6G Distributed Massive MIMO Testbed for FR3

August Svensson
au5718sv-s@student.lu.se
Lukas Niklasson Lundqvist
lu3756ni-s@student.lu.se

Department of Electrical and Information Technology
Lund University

Supervisor: Liang Liu

Examiner: Pietro Andreani

June 22, 2026

Printed in Sweden
Tryckeriet i E-huset, Lund

Abstract

Research into distributed multiple-input multiple-output (D-MIMO) systems has gained increasing attention as a candidate technology for future 6G networks. However, evaluating such systems under realistic implementation constraints requires real-time testbeds that can capture effects such as synchronization, processing latency, hardware resource usage, and radio-frequency impairments.

This thesis further develops the Lund University Large Intelligent Surface (LuLIS) testbed. The testbed is a scalable, field programmable gate array (FPGA) based testbed with 16 distributed nodes. The work consists of two main contributions. First, a new distributed digital baseband architecture is developed to increase the supported uplink processing capability from four to 32 users. The architecture supports a distributed part of maximum-ratio combining (MRC) and zero-forcing (ZF) detection schemes. Second, a front-end design for operation in Frequency Range 3 (FR3).

The implemented digital baseband is evaluated in terms of functionality, latency, and FPGA resource utilization, while the analog front-end is characterized through measurements on an evaluation front-end. The results show that the baseband design is functional but that the implementation should be improved in a few ways, to reduce the use of BlockRAM resources on the FPGAs, and to reduce the latency in certain configurations. The evaluation front-end demonstrated performance similar to the theoretical estimates. The 16 channel front-end featured an improved design based on further research and learnings from the evaluation front-end but was not manufactured due to time constraints and potential for further improvement. Together, these contributions extend the LuLIS testbed toward larger-scale multi-user experiments and future upper-mid-band 6G research.

Popular Science Summary

Imagine a room full of people with a large speaker in the middle. Everyone in the room hears the same music, regardless of where they are standing. This is similar to how today's cellular base stations work: a single signal is broadcast over a wide area for many users to share.

Now imagine replacing that one large speaker with an array of small directional speakers. By cleverly controlling the speakers, sound waves corresponding to the music of each person can be directed in space. This is similar to what in wireless communications is called beamforming, and the system described is a close analogy to how current mobile communications are meant to work.

To improve this, the array of speakers can be split up and spread throughout the room. Each speaker array can more easily focus sound toward a specific person, allowing different people to listen to different music at the same time with less interference from one another.

With distributed multiple-input multiple-output (D-MIMO), the goal is to coordinate all of the speaker arrays in the room so they work together as a single system. The system tries to maximize the loudness and clarity of the music for each person while minimizing interference from everyone else's music.

Each speaker array also acts as a microphone array, listening to how the sound reaches different microphones in the room. By sharing this information with the other arrays, the system can continuously adjust which speakers transmit which sounds, adapting in real time as people move around the room.

It's easy to see that the more speaker arrays you have, the more precisely they can direct sound, the more people you can serve simultaneously. In the same way, D-MIMO uses many distributed antennas working together to increase both network capacity and user performance. This thesis focuses on accelerating the calculations necessary to focus the antennas and increasing the frequency range to allow for more users in the future.

Acknowledgments

First we'd both like extend out thanks to our supervisors Liang, Lina, and Vilgot for their valuable support and guidance throughout this spring. We also want to thank Dumitra who was an invaluable resource when Lina was away on her honeymoon.

The LuLIS testbed would not be possible without the funding of Swedish Foundation for Strategic Research (SSF) (Project CHI19-0001: Large Intelligent Surfaces - Architecture and Hardware), as well as support from the AMD University Program who supply the ZCU216 RFSoCs.

I'd like to thank my family and girlfriend Ida for always supporting me.
– Lukas

I especially want to thank Lena, Johnny and Tora.
– August

Table of Contents

List of Figures

List of Tables

List of Acronyms

ADC	analog to digital converter
BRAM	BlockRAM
BS	base station
DAC	digital to analog converter
DLD	downlink data
D-MIMO	distributed multiple-input multiple-output
DSP	digital signal processing
DUT	device under test
EIT	Department of Electrical and Information Technology
FF	flip-flop
FFT	fast fourier transform
FIFO	first-in first-out
FPGA	field programmable gate array
FR3	Frequency Range 3
HDL	hardware description language
IF	intermediate frequency
IFFT	inverse fast fourier transform
ILA	integrated logic analyzer
ISI	inter symbol interference
LIS	large intelligent surface
LNA	low noise amplifier
LO	local oscillator
LuLIS	Lund University large intelligent surface

LUT	look-up table
LUTRAM	look-up table RAM
MIMO	multiple-input and multiple-output
MMIC	Monolithic microwave integrated circuit
MRC	maximum-ratio combining
MU-MIMO	multi-user MIMO
OFDM	orthogonal frequency division multiplexing
PA	power amplifier
PCB	printed circuit board
PE	processing element
PLL	phase locked loop
RF	radio frequency
RX	receive
SMT	surface mount technology
SNR	signal to noise ratio
SSF	Swedish Foundation for Strategic Research
TDD	time division duplex
THT	through hole technology
TX	transmit
UE	user equipment
ULD	uplink data
ULP	uplink pilot
URAM	UltraRAM
ZF	zero-forcing

Chapter 1
Introduction

The demand for high-throughput and low-latency wireless communication continues to grow as society becomes increasingly interconnected. Emerging 6G systems are expected to enable new applications such as integrated sensing and communication, edge computing, and autonomous mobility [1]. To meet these demands, modern wireless systems rely on advanced multiplexing techniques to improve spectral efficiency and support large numbers of users. While time division duplex (TDD) and orthogonal frequency division multiplexing (OFDM) have formed the foundation of modern communication systems, increasing performance requirements have driven the adoption of spatial multiplexing techniques.

Since the introduction of 4G, multiple-input and multiple-output (MIMO) systems have become a key technology for increasing throughput and improving spectral efficiency [2]. A natural extension of conventional massive MIMO is D-MIMO, where base station (BS) antenna arrays are spatially distributed rather than co-located. By distributing the antennas, D-MIMO can improve coverage and increase macro-diversity. These advantages become increasingly relevant for upper-mid-band systems, where large antenna arrays and distributed deployments can provide the spatial resolution needed to separate users and support spatial multiplexing [3, 4].

The Lund University large intelligent surface (LuLIS) testbed is the world's first real-time D-MIMO testbed with 256 antennas [5]. The testbed uses field programmable gate array (FPGA) development boards as the basic building blocks. Real-time testbeds are important because they allow D-MIMO algorithms and hardware architectures to be evaluated under practical constraints that are difficult to capture fully in simulation, such as synchronization, processing latency, RF impairments, and resource limitations in the digital hardware.

The primary objective of this thesis project is to further develop the LuLIS testbed to support spatial multiplexing for a larger number of user equipments (UEs), while enabling operation in the lower part of Frequency Range 3 (FR3). The project consists of both analog and digital development. The analog part focuses on the design of a 16-channel front-end for operation at FR3 frequencies, while the digital part involves developing a new distributed digital baseband capable of supporting up to 32 UEs simultaneously.

The proposed front-end is intended to operate in the lower portion of FR3 range, between 7.4 and 8.5 GHz. Rather than targeting a specific band, the frequency range was selected to avoid future licensing constraints while maintain-

ing reasonable implementation cost and design complexity. The selected range is motivated by both the Swedish Frequency Plan [6] and international industry roadmaps. At the 2023 World Radiocommunication Conference (WRC-23), the International Telecommunication Union (ITU) identified the 7.125–8.4 GHz range as a so-called "golden band" due to its favorable propagation characteristics and broad spectrum availability across multiple regions. Figure 1.1 illustrates the spectrum availability identified by different organizations.

Compared to existing cellular frequency bands, the upper mid-band enables substantially wider bandwidth allocations, allowing significantly higher throughput. Although propagation losses increase with frequency, the path loss remains considerably lower than in FR2 systems operating between 24.25 and 71.0 GHz [7]. Consequently, upper mid-band systems can achieve improved coverage with fewer base stations while still benefiting from beamforming and D-MIMO techniques to compensate for propagation losses through antenna gain [7].

Implementing large-scale multi-antenna systems at these frequencies introduces significant hardware challenges. Each antenna element requires a dedicated radio frequency (RF) signal chain, and when implemented using discrete integrated circuits, careful layout becomes very important. Undesired coupling between channels through capacitive, conductive, and inductive mechanisms increases with frequency, as do transmission losses. As a result, the RF front-end design must balance performance, complexity, cost, and scalability.

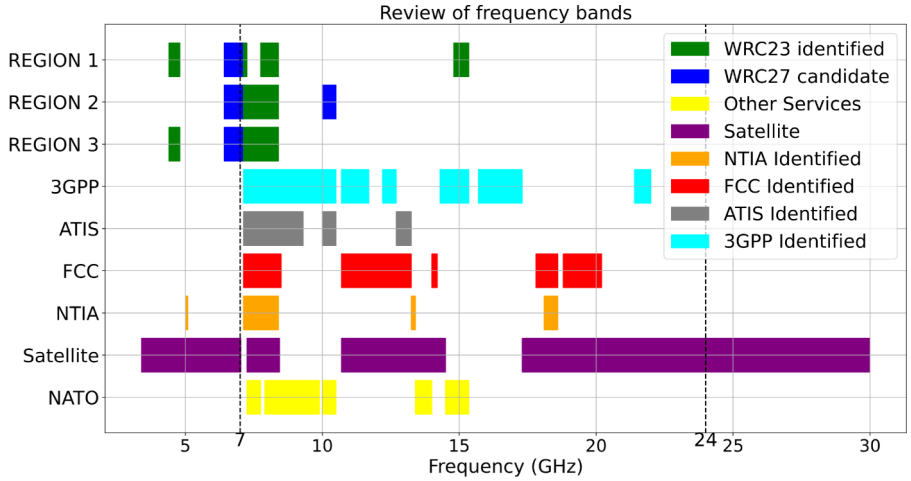


Figure 1.1: 6G spectrum availability [8].

Supporting 32 UEs is an important step toward making the LuLIS testbed useful for larger-scale multi-user MIMO (MU-MIMO) experiments. While the current four-UE configuration is sufficient for validating basic functionality, it represents a lightly loaded system compared to the 256 available antenna elements. Increasing the number of UEs allows the testbed to investigate scenarios where spatial multiplexing, inter user interference, and the choice of detection and precoding scheme have a larger impact on system performance. A target of 32 UEs still

maintains a favorable antenna-to-user ratio of 8:1, while creating a substantially more demanding and realistic processing workload. This makes it a suitable design target for evaluating both the scalability of the distributed architecture and the feasibility of real-time FPGA based processing.

1.1 Project requirements

- Implement high-throughput MIMO processing, enabling spatial multiplexing of up to 32 UEs.
- Enable operation in the upper mid-band frequency range (FR3), from 7.4 GHz to 8.5 GHz.
- Provide cost-effective and scalable frontend and digital signal processing (DSP) implementation solutions for large-scale deployment in the testbed.
- Integrate with the existing testbed with minimal modification to the overall LuLIS system.

The thesis project is based on requirements defined by the Department of Electrical and Information Technology (EIT). First, the proposed baseband improvements should extend support from the current four UEs to 32 UEs. Second, the system should support operation in the upper mid-band, specifically within 7.4–8.5 GHz, which is considered a strong candidate range for the final 6G specification. Third, the designs must be cost-effective and scalable, as the testbed comprises 256 RF chains. Finally, both the frontend and baseband improvements should integrate into the existing testbed with minimal modifications to other parts.

Due to the broad scope of the project requirements and the expertise needed, the project is split up into two parts. Author A. Svensson is responsible for the sub-project that focuses on the introduction of support for 32 UEs, while L. N. Lundqvist is responsible for the support for FR3. This partition is reflected in the structure of the report.

1.2 Previous Work

There has been plenty of research into MIMO systems in general. On the digital baseband side there is no shortage of relevant work [9, 10, 11, 12]. However, these works are not designed with LuLIS in mind, they are generally larger in scope, and most are designed for a centrally processed system. Nonetheless, we note that the use of systolic arrays (see section 2.3.1) for matrix multiplications is prominent.

For the analog front-end, examples of similar work can be found. But they are generally for evaluation purposes, and do not scale cost-wise in the way necessary for the LuLIS testbed, where 256 channels are needed [13, 14], and are not directly compatible with ZCU216 currently used. Generally analog front-end design for FR3 has not been done at this scale and other designs found have less than 16 channels.

1.3 Outline

The report initially, in chapter 2, goes over key wireless communications concepts before explaining the current LuLIS testbed in greater detail. Thereafter other relevant concepts are introduced. After that chapter 3 and 4 cover the design processes and methods for the digital and analog parts, respectively. Results are presented and discussed in chapter 5 and concluding remarks are given in chapter 6.

Chapter 2
Background & Theory

In this chapter we will create context that is important in order to understand the project. Initially, a few concepts specific to modern wireless communication will be introduced. Thereafter, armed with this knowledge, an overview of the LuLIS testbed will be given. Finally some complementary theory will be presented.

2.1 System

2.1.1 OFDM

In the latest two generation of wireless communications, OFDM has been the dominating transmission scheme and is going to be the primary frequency division multiplexing used in 6G as well. The reasons for using OFDM in MIMO systems are many: It divides the bandwidth into many subcarriers that can be treated individually with the linear precoding schemes. The subcarriers are orthogonal, which removes the need for guard bands and it can utilize fast fourier transform (FFT) and inverse fast fourier transform (IFFT) for efficient implementation.

The main principle of OFDM is splitting a large bandwidth into smaller subcarriers. By choosing the spacing such that each subcarrier falls on the trough of the neighbouring carriers, interference between subcarriers can be minimized. Assume that there is a a rectangular function

$$r(t) = \begin{cases} 1, & -\frac{\tau}{2} \leq t < \frac{\tau}{2} \\ 0, & \text{otherwise} \end{cases} , \tag{2.1}$$

that represents a subcarrier holding a value for τ amount of time. Its Fourier transform $X(f)$ is

$$F(x(t)) = X(f) = \tau \operatorname{sinc}(\tau f) \ . \tag{2.2}$$

By choosing the spacing between subcarriers to be $1/\tau$, we can see in Figure 2.1 that ideally, the magnitude of all subcarriers are zero except one. The important take away for this work is that the symbol time and subcarrier spacing are inversely proportional. If, for example, one would require more time in an OFDM frame the solution is probably to add symbols rather than change their time duration [15].

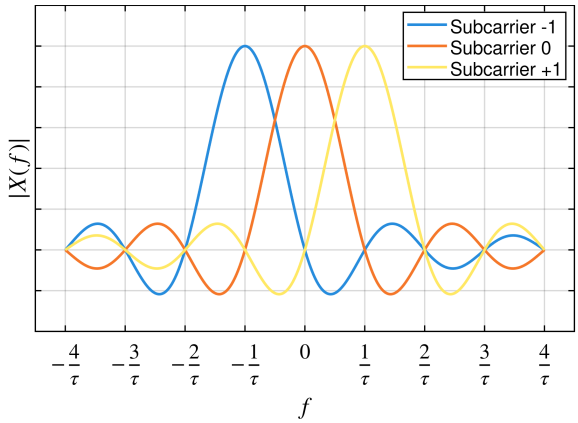


Figure 2.1: Example of $\frac{1}{\tau}$ spacing used in OFDM.

2.1.2 MIMO

Multi-antenna systems can greatly increase throughput and improve coverage hence the wide adoption of MIMO in nearly all modern wireless technology. MIMO is a broad term that covers many approaches that use the spatial dimension to increase throughput, robustness or coverage. MU-MIMO is for us the most interesting type, separating multiple users in space, see Figure 2.2.

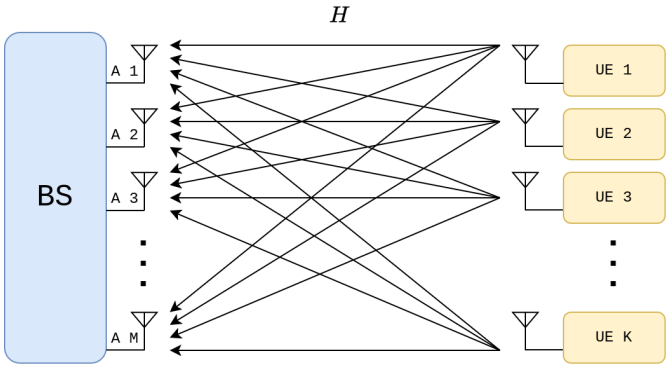


Figure 2.2: MU-MIMO diagram where a BS has M antennas, servicing K UEs.

2.1.3 Linear Detection and Precoding schemes

When the ratio between the number of antennas at the base station and number of UEs is large, linear detection and precoding schemes perform nearly as well as more computationally complex non-linear schemes [16]. The individual channels to each user become easier to separate when the number of antennas increase. These linear schemes also take advantage of the fact that the wireless channels

that distort signals between transmitter and receiver can be modeled with a linear filter $\mathbf{H}$ plus some Gaussian noise $\mathbf{n}$ [17],

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \qquad \mathbf{y}, \mathbf{n} \in \mathbb{C}^M, \quad \mathbf{s} \in \mathbb{C}^K, \quad \mathbf{H} \in \mathbb{C}^{M \times K}, \quad \mathbf{n} \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2). \tag{2.3}$$

There are two linear schemes that are of interest to us, zero-forcing (ZF) and maximum-ratio combining (MRC). The core principle for both methods is to approximate the channel matrix $\mathbf{H}$ by using predetermined pilots. Each UE transmits a known pilot value. The BS approximates $\mathbf{H}$ with $\hat{\mathbf{H}}$,

$$\hat{\mathbf{H}} = \mathbf{y}\mathbf{s}^* = \mathbf{H} + \mathbf{n}\mathbf{s}^* \,. \tag{2.4}$$

For MRC, the hermitian of the channel approximation $\hat{\mathbf{H}}^{\mathrm{H}}$ is used as detection matrix for uplink data (ULD) symbols and $\hat{\mathbf{H}}^*$ is used for precoding downlink data (DLD) symbols:

$$\mathbf{z} = \hat{\mathbf{H}}^H \mathbf{y}, \quad \mathbf{x} = \hat{\mathbf{H}}^* \mathbf{u} \,. \tag{2.5}$$

Similarly for ZF,

$$\mathbf{z} = \hat{\mathbf{H}}^\dagger \mathbf{y}, \quad \mathbf{x} = \hat{\mathbf{H}}^{\dagger \mathrm{T}} \mathbf{u} \,, \tag{2.6}$$

where $\mathbf{H}^\dagger$ is computed by

$$\mathbf{H}^\dagger = \mathbf{G}^{-1} \hat{\mathbf{H}}^{\mathrm{H}}, \quad \mathbf{G} = \hat{\mathbf{H}}^{\mathrm{H}} \hat{\mathbf{H}} \,. \tag{2.7}$$

MRC is simple to compute compared to ZF as the dreaded matrix inversion is avoided entirely. It works by combining the signals from each antenna in such a way that the signal to noise ratio (SNR) is maximized at the intended UE [17]. However, MRC does not account for inter-user interference, and in multi-user MIMO there are more than one receiving UE, which is often better served by a ZF scheme. Instead of maximizing SNR at the targeted user, the ZF algorithm attempts to cancel out interference to the other receivers in the cell. But by doing this ZF sacrifices some of the SNR at the target UE. The drawback of having to compute a matrix inversion is lessened by the fact that the gramian matrix $\mathbf{G} = \hat{\mathbf{H}}^{\mathrm{H}} \hat{\mathbf{H}}$ is diagonally dominant, which allows for certain approximative methods.

2.1.4 Time Division Duplex and OFDM Frames

TDD means that time is divided up in some way so that during certain time intervals one is receiving, and during others, one is transmitting. The downside is that one cannot receive and transmit at the same time, as that is prohibited. TDD is however incredibly well suited to the detection schemes discussed in the previous section, 2.1.3. The reciprocity of the channel is exploited so that the uplink pilot (ULP) is sufficient for both downlink and uplink communication. Each frame starts with a ULP being sent from the UEs. The BS estimates the channel $\hat{\mathbf{H}}$ and calculates $\mathbf{G}$. The UEs transition into sending data while the BS is inverting

Figure 2.3: An example OFDM frame. The frame includes one ULP, two ULD, and two DLD symbols.

$\mathbf{G}$. When the TDD mode of the system changes, the BS has been given enough time to be finished with computing $\mathbf{H}^{\dagger}$.

An example OFDM frame can be seen in Figure 2.3. The crosses represent guard symbols. A guard symbol is an empty symbol meant to protect ULD against inter symbol interference (ISI) from DLD and vice versa. They also give time for hardware to be configured for the next TDD mode. However, interference between symbols is not only an issue between TDD modes. Symbols can experience ISI from previous symbols due to delay in the channel, but instead of having a guard symbol between every normal symbol, the common approach is to extend the duration of each symbol with a so called cyclic prefix. The length of the cyclic prefix is such that all ISI appears on the prefix instead of on the symbol. Before the symbol is passed to the FFT, the prefix is removed, so that subcarrier orthogonality is maintained [15].

2.1.5 Coherence Bandwidth and Coherence Time

In the previous couple sections two techniques for detection and precoding were presented. One assumption is that the channel approximation that is based on the known ULP symbol stays valid for the entire remainder of the frame. This is only true if the time coherence T_c of the channel is longer than the total frame time. In the same way, another assumption is that the channel approximation is valid for all the subcarriers. This is only true if the total bandwidth of the subcarriers is smaller than the coherence bandwidth, B_c.

We will not go into the details of the physics behind coherence bandwidth and time. We will simply note what importance they have to us. Coherence time limits the possible length of the chosen frame, as by the end of it the channel approximation is outdated. Similarly, the coherence bandwidth limits the number of subcarriers we can assume experience the same channel. The consequence is that the schemes discussed in section 2.1.3 must be performed for each group of neighboring subcarriers such that the total bandwidth of the subcarriers does not exceed the coherence bandwidth [17].

Of course, both coherence bandwidth and time can vary over both time and frequency but we will not consider this. We will see them as upper bounds for the system as a whole.

2.1.6 Current LuLIS testbed

The LuLIS testbed is a research project funded by SSF [5], with the main objective of developing the processing architectures and hardware solutions to enable this type of distributed massive MIMO systems. The testbed is comprised of 16 panels

with 16 antennas each, see Figure 2.4. Each panel uses a ZCU216 RFSoC Evaluation Kit as its base. The ZCU216 is a high performance FPGA board specifically developed with RF use cases in mind. Most notably, the ZU49DR FPGA chip includes integrated RF-ADCs and RF-DACs. All the panels operate as one BS but in a distributed fashion.

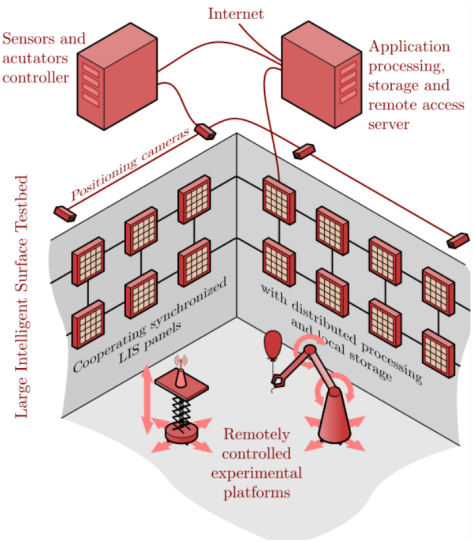


Figure 2.4: High level overview of LuLIS testbed.

OFDM is used as the transmission scheme of the testbed. A 50 MHz carrier is populated with 600 transmitted subcarriers. The subcarriers are spaced in frequency by intervals of 60 kHz, giving a symbol time of 16.67 μs, including the cyclic prefix that is 19.05 μs. The frame structure is as described in section 2.1.4. The current distributed digital baseband only supports four UEs. We summarize the specifications in Table 2.1 below.

Table 2.1: OFDM parameters used on the LuLIS testbed.

Parameter	**Value**
Carrier bandwidth	50 MHz
Number of transmitted subcarriers	600
Subcarrier spacing	60 kHz
Useful symbol time	16.67 μs
Symbol time including cyclic prefix	19.05 μs
Frame structure	As described in Section 2.1.4

The linear schemes presented in section 2.1.3 are used in a similar way to how they would be used on a traditional MIMO BS. The channel estimation (2.4) is performed on each panel p and stored locally as $\hat{\mathbf{H}}_p$. Similarly each panel calculates the gramian $\mathbf{G}_p = \hat{\mathbf{H}}_p^{\mathrm{H}} \hat{\mathbf{H}}_p$ (2.7). $\mathbf{G}_p$ is transmitted to a central processing

node to be aggregated and inverted, but crucially $\hat{\mathbf{H}}_p$ is only used locally, reducing the required bandwidth between panels. The split between central and local computation is shown in (2.8) [18].

$$\begin{array}{cc} \text{Local:} & \text{Central:} \\ \hat{\mathbf{H}}_p = \mathbf{y}_p \mathbf{s}^* & \mathbf{G}^{-1} = (\sum_p \mathbf{G}_p)^{-1} \\ \mathbf{G}_p = \hat{\mathbf{H}}_p^{\mathrm{H}} \hat{\mathbf{H}}_p & \mathbf{z}_{\mathrm{MRC}} = \sum_p \mathbf{z}_{\mathrm{MRC}p} \\ \mathbf{z}_{\mathrm{MRC}p} = \mathbf{H}_p^{\mathrm{H}} \mathbf{y}_p & \mathbf{z}_{\mathrm{ZF}} = \mathbf{G}^{-1} \mathbf{z}_{\mathrm{MRC}} \\ \mathbf{x}_p = \hat{\mathbf{H}}_p^* \mathbf{v} & \mathbf{v} = \underbrace{(\mathbf{G}^{-1})^{\mathrm{T}} \mathbf{u}}_{\mathrm{ZF}} \text{ or } \underbrace{\mathbf{u}}_{\mathrm{MRC}} \end{array} \tag{2.8}$$

Figure 2.5 shows a closer look at the internals of one LuLIS panel. The highlighted blocks represent the parts that will be redesigned in this project. As you can see, each ZCU216 is connected to a custom 16 channel front-end that was designed for direct conversion using the RF-ADC and RF-DAC. The specifications are available in Tables 2.3 and 2.4 respectively, both the digital to analog converter (DAC) and analog to digital converter (ADC) have 14 bit resolution, but the effective number of bits is largely dependent on the clock source and its phase noise. The ADC has a maximum input power of 1 dBm but as stated in the comments/conditions column it features a programmable attenuator which allows for higher power levels and better dynamic range. The ADC sample rate is limited to 2.5 GSPS however it features a digital down converter allowing it to be used up to 6 GHz. The DAC is limited to 10 GSPS and the full-scale output power can be varied from -18.5 to 6.5 dBm, however the span decreases with frequency. The current front-end specifications can be seen in Table 2.2. Each front-end is connected to a custom 16 antenna array.

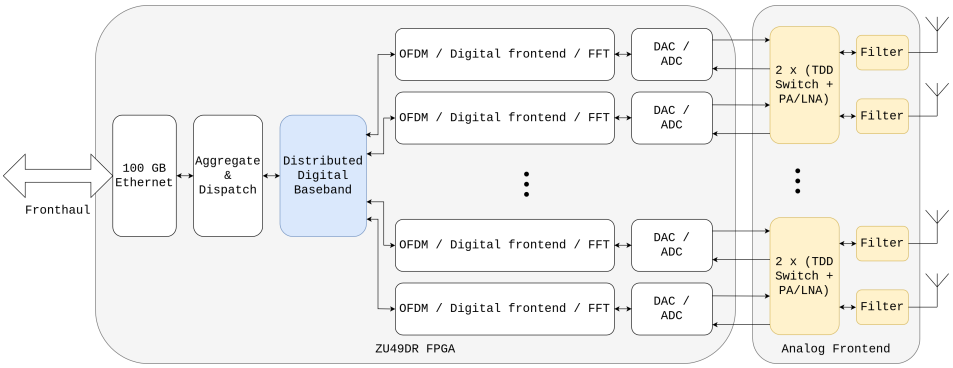


Figure 2.5: LuLIS panel overview.

Table 2.2: Current front-end specifications.

Transmit	
Gain [dB]	34.5
Max Input Power [dBm]	-9.5
Output Power Compression [dBm]	25
Frequency [MHz]	3700–3980
Receive	
Gain (high mode) [dB]	32.5
Gain (low mode) [dB]	13.5
Noise Figure [dB]	3.6
Frequency [MHz]	3700–3980

Table 2.3: ZU49DR RF-ADC specifications [19].

Parameter	Comment / Condition	Min	Typ	Max	Unit
Resolution	–	14	–	–	bits
Full-scale input	Input with 100 Ω on-die termination and DSA attenuation set to 0 dB	–	1	–	dBm
Digital attenuation range	–	0	–	27	dB
Attenuator step size	–	–	1	–	dB
Auto attenuation	Automatically set when amplitude overvoltage is asserted	–	15	–	dB
Analog input bandwidth	Full-power bandwidth, −3 dB	–	6	–	GHz
Return loss	Up to 4 GHz	–	−12	–	dB
Crosstalk isolation between channels	0–4 GHz	–	−75	–	dBc

Table 2.4: ZU49DR RF-DAC specifications [19].

Parameter	Comment / Condition	Min	Typ	Max	Unit
Resolution	–	14	–	–	bits
Sample rate	-2E, -2I, and -2LI speed grades without clock forwarding; datapath modes 2, 3, and 4	0.5	–	9.85	GS/s
Maximum output power	V_{DAC_AVTT} = 3.0 V, 100 Ω termination, signal frequency < 200 MHz	−18.5	–	6.5	dBm
Variable output power range	At 240 MHz	24	–	–	dB
	At 3500 MHz	20	–	–	dB
	At 4900 MHz	18	–	–	dB
	At 5900 MHz	17	–	–	dB
Analog bandwidth	Full-power bandwidth, −3 dB	–	6	–	GHz
Return loss, R_L	Up to 4 GHz	–	−12	–	dB
	Up to 6 GHz	–	−10	–	dB
Crosstalk isolation	F_{OUT} = 0–4 GHz	–	−75	–	dBc
	F_{OUT} = 0–6 GHz	–	−70	–	dBc

2.2 Analog Front-end

2.2.1 Superheterodyne

A superheterodyne is an architecture where a high frequency signal is down-converted to a lower intermediate frequency when receiving, and the intermediate frequency signal is up-converted to a high frequency signal when transmitting. The process of combining a low and high frequency signal is known as heterodyning.

The majority of the amplification is done at the intermediate frequency, this allows for much greater amplification due to the lower frequency and bandwidth required. In the superheterodyne architecture the intermediate frequency (IF) is fixed and the LO is swept to vary the frequency at which signals are received. A band-pass filter on the IF then removes unwanted frequencies. A filter on the RF side determines the range of frequencies that can be received or transmitted and can thus avoid images and other disturbances. An example of a superheterodyne receiver can be seen in Figure 2.6.

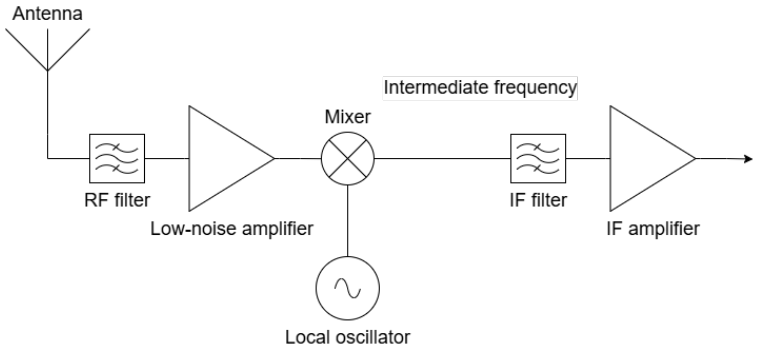


Figure 2.6: Block diagram of a basic superheterodyne receiver.

2.2.2 Noise Figure

Noise figure is a metric of how much the signal to noise ratio of a signal is degraded by a device or system.

$$F = \frac{SNR_i}{SNR_o} \tag{2.9}$$

where SNR_i and SNR_o are the input and output signal to noise ratios respectively.

$$NF = 10 \log F \tag{2.10}$$

Where NF is the noise figure expressed in decibels.

2.2.3 Friis formula

Friis formula is used to determine the noise figure of a multi-stage system.

$$F_{\text{total}} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \frac{F_4 - 1}{G_1 G_2 G_3} + \cdots + \frac{F_n - 1}{G_1 G_2 \cdots G_{n-1}} \tag{2.11}$$

Where F_1, G_1 is the noise figure and gain of the first stage in the system and so on. The significant take-away is that the noise figure is dominated by the first stage of the system and that a higher initial gain will suppress the noise-figure of the other stages.

2.2.4 Transmission lines

The characteristic impedance is critical in RF design as most RF devices are designed to have a characteristic impedance of 50 Ω. This is important as impedance mismatch between the load Z_L and the source impedance Z_o will result in reflections which reduces the power into the load.

A transmission line is a cable or structure designed specifically to propagate electromagnetic waves. The theory of transmission lines is closely related to characteristic impedance. An infinitesimal segment of a transmission line is represented as in Figure 2.7.

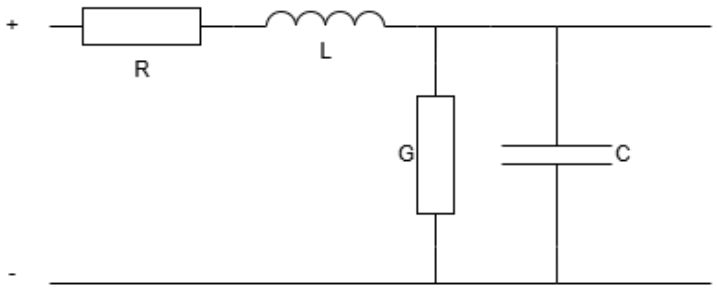


Figure 2.7: Transmission Line Model.

$$Z_o = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \approx \sqrt{\frac{L}{C}} \text{ for a loss-less line} \tag{2.12}$$

R, L and C represents the resistance, inductance and capacitance per unit length. G is the conductance of the dielectric per unit length. Z_o is the characteristic impedance of an infinitesimal segment of the transmission line [20].

Transmission lines on printed circuit board (PCB)s

On PCBs there are three major types of transmission lines with unique characteristics which will be shown below.

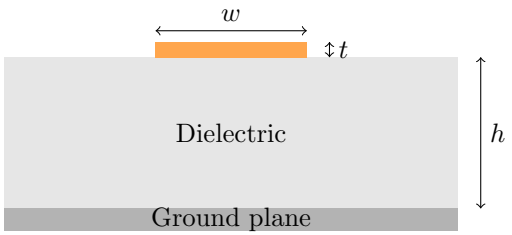


Figure 2.8: Illustration of a microstrip.

Microstrip is the simplest kind of transmission line. It consists of a conductor on an external layer over a ground plane and separated by a dielectric. Microstrip has slightly higher losses compared to a waveguide. It is also more susceptible to coupling and radiates more compared to stripline [21].

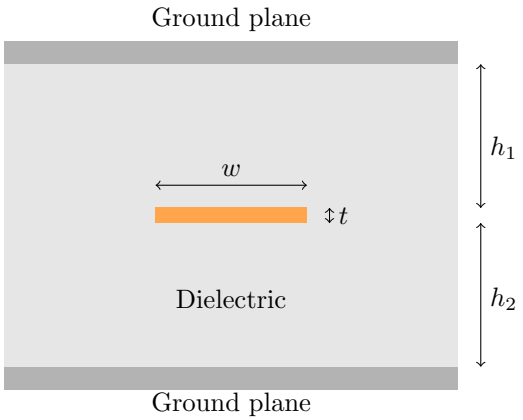


Figure 2.9: Illustration of a stripline.

Stripline is similar to microstrip, but the trace is on an internal layer and surrounded by ground planes on either side separated by a dielectric. The line being surrounded on top and bottom provides shielding and good isolation to other signals [21].

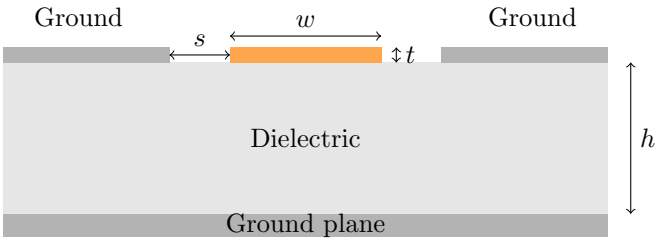


Figure 2.10: Illustration of a coplanar waveguide with ground (CPWG).

Coplanar waveguide with ground (CPWG) is similar to a microstrip line but it also has coplanar ground. CPWG is more sensitive to process variation in the PCB manufacturing process as the thickness of the copper layers affects the capacitance. The ground on the sides of the trace improves shielding and isolation as the field is more contained [22].

2.2.5 Loss in transmission lines

The dielectric loss stems from the losses of the electromagnetic wave traveling in the dielectric and the dielectric materials inherent dissipation of the wave's energy.

At high frequency the skin-effect becomes prominent and the current can no longer be assumed to be uniformly distributed in the conductor, instead it concentrates near the surface causing higher effective resistance and therefore losses.

2.2.6 Return & insertion loss

Return loss is a measure of how much power is reflected P_r relative to the incident power P_i as a result of impedance mismatch. Typically expressed in dB or as VSWR.

$$RL_{dB} = 10 \log_{10} \frac{P_i}{P_r} \tag{2.13}$$

$$VSWR = \frac{1 + 10^{-RL/20}}{1 - 10^{-RL/20}} \tag{2.14}$$

If expressed as above the return loss will always be a positive number as the incident power has to be greater than the reflected power, however the return loss is often expressed inversely as a negative number. In both cases a greater magnitude signify a better return loss and less reflection. An acceptable return loss is considered to be >10 dB and a good return loss is >15 dB.

Insertion loss is a measure of how signal power is affected by passing through an element. Defined as

$$IL_{dB} = 10 \log_{10} \frac{P_i}{P_o} \tag{2.15}$$

where P_i is the incident power and P_o is the power after the inserted element.

2.3 DSP Architecture

2.3.1 Systolic Array

The idea of systolic architectures is primarily to exploit the reuse of data in algorithms and thereby lowering the necessary memory or I/O bandwidth [23]. Multiplication of matrices (2.16) is a good example of such an algorithm.

$$c_{ij} = \sum_{k=0}^{M-1} a_{ik} b_{kj}, \qquad i \in [0 : N_a - 1], \quad j \in [0 : N_b - 1]. \tag{2.16}$$

Because each element in each matrix is multiplied by an entire vector of elements from the other matrix, it is clear that there is a high degree of reuse. Also, because the algorithm for matrix multiplication is regular it is possible to create a simple, regular topology of similar processing elements (PEs) to perform the calculation, a so called systolic array. Each PE in the systolic array multiplies and accumulates its' current values with its' previous result. When all the elements of the rows/columns have been processed the result is outputted and the accumulator is reset.

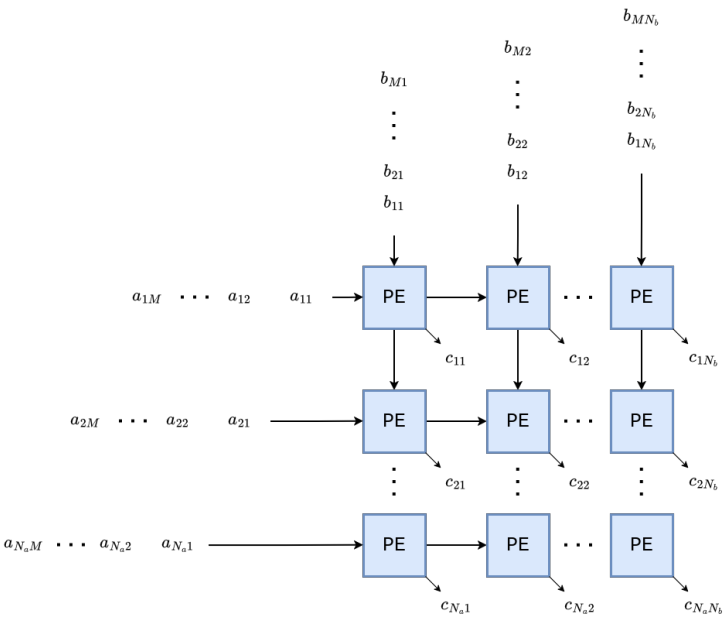


Figure 2.11: Model of output stationary systolic array for matrix multiplication.

Assuming the systolic array in Figure 2.11 is fully pipelined, we can calculate its throughput (TP), latency (L) and number of PEs (A).

$$TP_{systolic} = N_a N_b \,, \tag{2.17}$$

$$L_{systolic} = N_a + N_b + M - 2\,, \tag{2.18}$$

$$A_{systolic} = N_a N_b\,, \tag{2.19}$$

The example in Figure 2.11 is not the only possible configuration of a systolic array for matrix multiplication, other variations will balance performance and resource use differently [24].

Chapter 3
Digital Baseband Design

In this chapter, the design of the distributed digital baseband is presented in parts. Initially we will summarize constraints and assumptions, then take a closer look at the algorithms in question, and lastly the choice of architecture. In this and coming chapters, the use of $\hat{}$ to indicate an approximation will be foregone for ease of reading. Similarly, for the remainder of the report, when mentioning coherence bandwidth, B_c, it's described in terms of the amount of subcarriers, with a subcarrier spacing of 60 kHz, for which the channel is flat.

3.1 Constraints and Assumptions

We will make some assumptions and apply some constraints regarding the choices of coherence bandwidth B_c, the number of UEs K, and the number of total data symbols per frame D. This is done to manage the scope of the design whilst being flexible enough for research purposes.

Firstly, we will constrain the number of data symbols D to be no fewer than 4 and no larger than 16. With $D < 4$ the latency demands on the inversion of the Gramian G will very likely be too high, even for a small K. With $D > 16$ the coherence time of the channel might be insufficient.

Secondly, the chosen coherence bandwidth, B_c, is physically constrained by the actual coherence bandwidth of the channel. This value is unknown, but for implementation purposes we constrain the choice of the coherence bandwidth to be between 4 and 16. Also, the choice of coherence bandwidth must divide the number of data carrying subcarriers (600). That leaves us with the set $B_c \in \{4, 5, 6, 8, 10, 12, 15\}$. Thirdly, to meet the goal of supporting up to 32 UEs, $K \in [1 : 32]$. We summarize in Table 3.1.

Table 3.1: Design parameter constraints.

Parameter	Description	Allowed values
K	Number of UEs	$[1 : 32]$
D	Number of data symbols	$[4 : 16]$
B_c	Coherence bandwidth	$\{4, 5, 6, 8, 10, 12, 15\}$

Naturally, the design is also constrained by the available resources. Each dis-

tributed node consists of one ZCU216 RFSoC development board, with a XCZU49 FPGA. The most critical resources are the BlockRAM (BRAM) and DSP slices. In Table 3.2 the total number, and percentage that is used by other IPs can be seen. That is, the utilization of all other subsystems on the FPGA. Table 3.2 also includes other resource types: look-up tables (LUTs), look-up table RAMs (LUTRAMs), and flip-flops (FFs).

Table 3.2: XCZU49 FPGA resource overview.

Resource set	LUT	LUTRAM	FF	BRAM	DSP
FPGA total	425280	213600	850560	1080	4272
Used	19%	5%	13%	30%	4%

Another important constraint is the bit width accepted by the DSP slices. The Ultrascale+ family of FPGAs uses DSP48E2 slices that include one two's complement 27x18 bit multiplier [25]. The result is that if a bit width of more than 18 bits is needed the required resources will double. For the scope of this project, the bit width will be kept at 16 bits with the possibility for future alteration if it is necessary.

The future implementation of other subsystems will also require resources and generally increase congestion during implementation. It is therefore important that the design uses a proportionate amount of resources to allow for future work.

The efficiency and performance in terms of power consumption are not prioritized for the design. It extends the scope and implementation time significantly and is of greater interest if one is doing an ASIC tape-out of the design.

Lastly, the digital front-end on the FPGA is using a clock signal generated by the integrated ADC. This clock has a period of $T_{clk} = 6.5\,\text{ns}$. The aim is to use the same clock for this new design, mainly to avoid having to deal with different clock domains.

3.2 Algorithm Analysis

The operations that the accelerator shall perform are the local parts of the linear processing schemes in (2.8). Each type of symbol, apart from the guard symbols, is associated with some computations. For ULP the first calculation is the channel approximation $\mathbf{H}_p$ (2.4), as well as the calculation of the partial gramian $\mathbf{G}_p$ (2.7). Both ULD and DLD symbols require multiplications with either $\mathbf{H}_p^{\mathrm{H}}$ or $\mathbf{H}_p$ respectively. We summarize this in Table 3.3 below.

3.2.1 Throughput Analysis

With the testbed supporting a total of 600 active subcarriers (Table 2.1), an approximation for the total number of complex multiplications per frame can be found:

Table 3.3: Distributed operations for each symbol.

Symbol type	Operations	Sizes
ULP	$\mathbf{H}_p = \mathbf{y}_p\mathbf{s}^*$	(M x 1) by (1 x K)
	$\mathbf{G}_p = \mathbf{H}_p^{\mathrm{H}}\mathbf{H}_p$	(K x M) by (M x K)
ULD	$\mathbf{z}_{\mathrm{MRC}p} = \mathbf{H}_p^{\mathrm{H}}\mathbf{y}_p$	(K x M) by (M x 1)
DLD	$\mathbf{x}_{\mathrm{ZF}p} = \mathbf{H}_p^*\mathbf{v}$	(M x K) by (K x 1)

$$\underbrace{\#ULP \cdot 600M + \frac{1}{2}\frac{600}{B_c}MK^2}_{\text{ULP}} + \underbrace{\#ULD \cdot 600MK}_{\text{ULD}} + \underbrace{\#DLD \cdot 600MK}_{\text{DLD}}, \tag{3.1}$$

where M is the number of antennas, K is the number of users, and B_c is the chosen coherence bandwidth. Note

- that the complexity of the calculation of $\mathbf{G}_p$ is quadratic in K,
- that $\mathbf{G}_p$ symmetry reduces the needed multiplications be roughly half,
- that $\mathbf{G}_p$ does not grow linearly with the number of symbols like the other operations.

The calculation of $\mathbf{G}_p$ is instead performed once for each subcarrier group, i.e. for all subcarriers that share the same channel approximation. Similarly, an expression for the total time of the frame can be made based on the predetermined OFDM symbol time $T_{symbol} = 19.05\,\mu$s:

$$\underbrace{\#ULP \cdot T_{symbol}}_{\text{ULP}} + \underbrace{\#ULD \cdot T_{symbol}}_{\text{ULD}} + \underbrace{\#DLD \cdot T_{symbol}}_{\text{DLD}} + \underbrace{2 \cdot T_{symbol}}_{\text{guard symbols}}, \tag{3.2}$$

With the expressions (3.1) and (3.2), and the constraints on K, B_c, and the number of data symbols ($D = \#ULP + \#DLD$), we find that the maximum required throughput in terms of complex multiplications per second is

$$ComplexMultPerSecond = 13.2\,\mathrm{Gcmult/s} \tag{3.3}$$

Assuming that the already mentioned clock frequency of 154 MHz is used in the design (see section 3.1), that corresponds to 85.7 cmult/cc. This maxima is observed when all variables are maximized. We can conclude that the maximum throughput requirement of the design is approximately 85.7 cmult/cc. However, there are latency demands to consider.

3.2.2 Latency Analysis

There are two major latency demands that need to be met for the design. The first is that there is a hard deadline on the DLD symbols. There is a finite time from

receiving the centrally processed $\mathbf{v}$ until $\mathbf{x}_{\mathrm{ZF}p}$ has to be passed to the digital front-end. The second requirement is related. In order for the central processing node to be able to perform the central part of the zero forcing algorithm, it must first invert the gramian $\mathbf{G}$. Inversion of matrices is a notoriously computationally demanding operation and it is therefore crucial to give the central node as much time as possible while keeping in mind the resource requirement of the implementation.

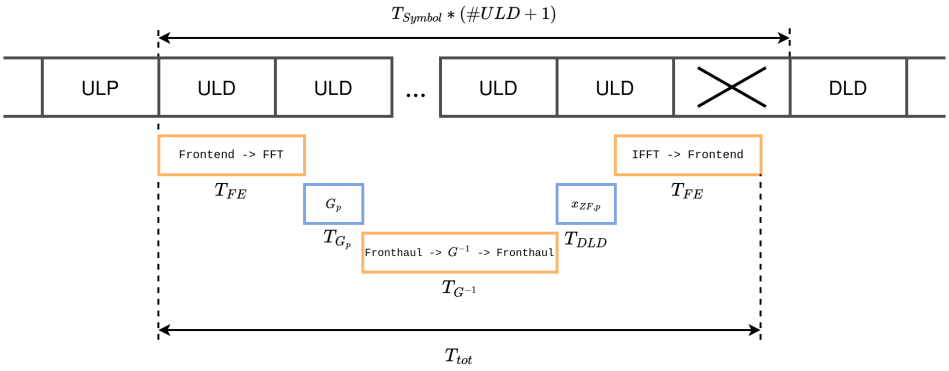


Figure 3.1: Illustration of delay model.

Since the LuLIS project is still at an early stage, the central node is not yet capable of inverting matrices in real time, and the system has not yet been configured for downlink communication. Therefore, the analysis aims to give a good indication rather than a full picture of the delays in the system. In Figure 3.1 you find a diagram of the delays that we consider. T_{FE} is the delay associated with the digital front-end and FFT, $T_{\mathbf{G}^{-1}}$ is the time it takes for $\mathbf{G}_p$ to be centrally aggregated, $\mathbf{G}^{-1}$ to be calculated, and $\mathbf{v} = \mathbf{G}^{-1}\mathbf{x}$ to be distributed. T_{G_p} and T_{DLD} are the delays that arise from the latency of the design. The delays in the analog front-end are not considered as they are small in relation to the other delays. Also, observe that it is likely the case that both ULD and DLD data are received simultaneously by the design. In such cases DLD should be prioritized as ULD has no deadline. T_{FE} is known to be $25\,\mu$s [18], $T_{G^{-1}}$ is unknown but based on the model presented in [18], we can approximate

$$T_{G^{-1}} \approx 2 \times 0.87\,\mu\mathrm{s} \times (P-1) + \frac{T_{clk}}{2.25} K^3\,, \tag{3.4}$$

where $P = 16$ is the number of panels, and $T_{clk} = 6.5\,\mathrm{ns}$. The total delay T_{tot} in Figure 3.1 is

$$T_{tot} = 2T_{FE} + T_{G_p} + T_{DLD} + T_{G^{-1}}\,. \tag{3.5}$$

Let us now investigate the T_{DLD} delay. We define the latency to be the number of clock cycles from when the last $\mathbf{v}$ vector is received until the last $\mathbf{x}_{\mathrm{ZF}p}$ is outputted. To not incur any latency from lack of momentary throughput, the design would have to be able to process $max(KM) = 512\,\mathrm{cmult/cc}$, significantly more than $85.7\,\mathrm{cmult/cc}$. The case for T_{G_p} is similar. Because the channel approximation $\mathbf{H}_p = \mathbf{y}_p\mathbf{s}^*$ must be spread over multiple pilot symbols for $B_c < K$, the

same can be done for $\mathbf{G}_p = \mathbf{H}_p^{\mathrm{H}}\mathbf{H}_p$ by tiling the multiplication. The worst case is then $\max(KM) = 512\,\mathrm{cmult/cc}$, since each vector of $\mathbf{H}_p$ must both be multiplied with it's own hermetian, and the hermetian of one vector for each other UE. We can calculate the delay caused by buffering, for a certain maximum throughput tp, in the following way,

$$T_{DLD} = 600T_{clk}\max\left(\frac{16K}{tp} - 1,\, 0\right), \qquad T_{G_p} = 600T_{clk}\max\left(\frac{16K}{tp} - 1,\, 0\right). \tag{3.6}$$

In Table 3.4 the total T_{tot} is tabulated for different values of tp and K, and below, the fraction of T_{tot} that comes from buffering can be seen. Dividing T_{tot} with T_{symbol} we find the number of $\#ULD$ needed in each case:

$$\min(\#ULD) = \lfloor \frac{T_{tot}}{T_{symbol}} \rfloor, \tag{3.7}$$

which can be found furthest down in Table 3.4.

Table 3.4: Total delay, fraction caused by buffering, and number of needed ULD symbols for different number of UEs K, and throughput tp.

$\mathbf{T_{tot}}$, $[\mu s]$				
$K\backslash tp$	85.7	128	256	512
8	81.42	77.58	77.58	77.58
16	103.43	95.73	87.93	87.93
32	209.56	194.16	178.56	170.76

$(\mathbf{T_{G_p}} + \mathbf{T_{DLD}})/\mathbf{T_{tot}}$				
$K\backslash tp$	85.7	128	256	512
8	2% + 2%			
16	7% + 7%	4% + 4%		
32	8% + 8%	6% + 6%	2% + 2%	

$\min(\#\mathbf{ULD})$				
$K\backslash tp$	85.7	128	256	512
8	4	4	4	4
16	5	4	4	4
32	11	10	9	8

We can see that there is a large difference in momentary throughput and the average throughput requirement. There are a few things of note. Firstly, for $tp \geq 256$ the delay contribution from buffering is dominated by $T_{G^{-1}}$. Secondly, when K is small the constraint on $\#ULD$ is not significant. Lastly, all tp choices

are possible if one chooses to include more ULD symbols than DLD, the issue is that this reduces the flexibility of the whole system. We can conclude that whatever throughput is chosen, the number of users will in some cases effect the possible frame structures.

3.3 Architecture

We will now discuss the main considerations that went into the architecture of the design as well as a description of the data-flow with this architecture. In Figure 3.2 you find a high level block diagram of the final design.

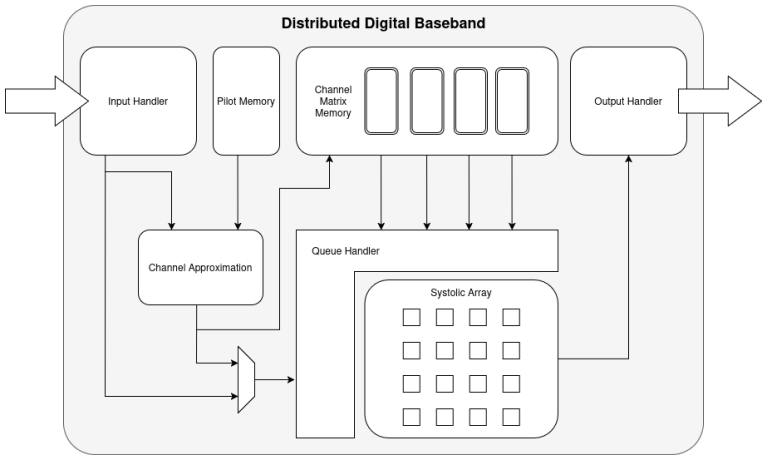


Figure 3.2: Overview of distributed digital baseband.

3.3.1 Architecture Choice

Observing the operations that the design shall perform it is clear that most of them can be viewed as matrix multiplications. For the ULD and DLD cases, each channel matrix $\mathbf{H}_p^{\mathrm{H}}$ or $\mathbf{H}_p^*$, is shared by B_c vectors that are in the same subcarrier group. As mentioned in section 2.3.1, systolic arrays are well suited for this. Accepting a small latency increase (2.18), we gain high throughput and low memory bandwidth.

3.3.2 Systolic Array Dimensioning

The question then becomes what size of systolic array to choose. To match the momentary throughput mentioned in section 3.2.2, a 32×16 array would be needed, the downside being a very high use of DSP slices. Alternatively, a 10×10 version could be used, satisfying the average throughput requirement, but the T_{G_p} and T_{DLD} delays would be on the order of tens of microseconds even when K is low. In the end, a compromise was chosen: $\mathbf{16 \times 16}$, giving a maximum throughput of 256 cmult/cc. The delay introduced is noticeable, but significantly less than in the 10×10 case. Looking at the minimum number of ULD symbols needed (in

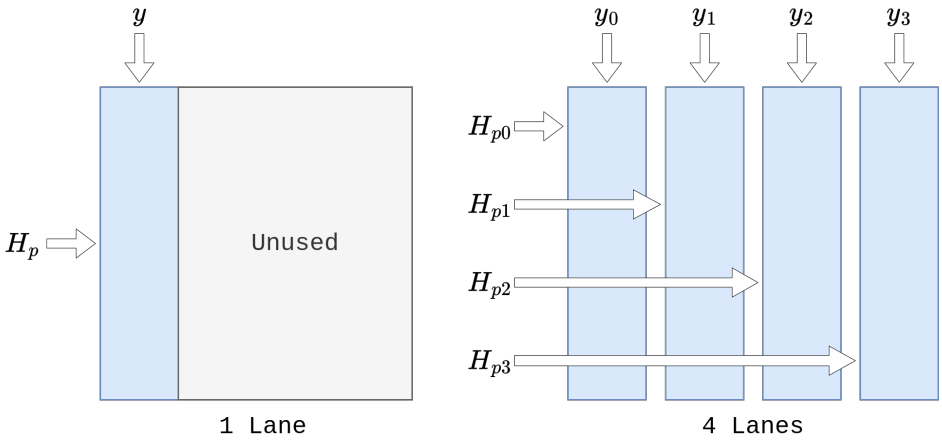


Figure 3.3: Illustration of the difference between having one lane vs. four lanes in the systolic array.

Table 3.4), a further increase in the size of the systolic array gives diminishing returns. The size also works well with the constraints set previously, $M = 16$ and $max(B_c) = 15$. Tiling is then only needed for $K > 16$, which simplifies the design from a control and dataflow perspective.

3.3.3 Configurability

However, we need a way of handling when B_c is chosen as one of the lower values, e.g. 4. When K is small, the throughput requirement is lower and we can get away with a low utilization of the systolic array. But when B_c is small, the throughput requirement is the same, but utilization will decrease.

Let us look at the example in Figure 3.3, where $K = 16$ and $B_c = 4$. Without modifications, the maximum throughput for a ULD operation ($\mathbf{z}_{\mathrm{MRC}p} = \mathbf{H}_p^{\mathrm{H}}\mathbf{y}_p$) is $16 \times 4 = 64\,\mathrm{cmult/cc}$. This is because there are only four subcarriers that share the same $\mathbf{H}_p$. The solution we chose is to split the systolic array into 4 lanes that can receive different side inputs, see Figure 3.3. This way the systolic array can operate as four 16×4 arrays, two 8×16, or one 16×16. The reason to not split it "horizontally" as well is that B_c only constrains one of the matrices in each operation, never both.

There will still be a slight decrease in throughput for some choices B_c, causing a few additional microseconds of buffering delay. The worst ones being when B_c is five and ten because the column utilization is at its lowest ($\sim 63\%$). One could split the array even more to remedy this, but seeing as $min(B_c) = 4$, even more granularity would bring more problems than gains. The bloating of the control and data-flow overhead was deemed to be too large compared to the latency reduction.

After the channel matrix estimations are made in the beginning of each frame, they must be stored in some way to be accessed at a later time. Just as the systolic array is split into four lanes, the channel matrix memory is split into four banks.

Each bank stores complete channel matrix vectors, stored in such a way that they are easily read in the order they are needed in the systolic array. All four banks can be read from simultaneously, allowing for the throughput mentioned previously.

3.3.4 Data-flow

This section describes the data flow through the design for each supported symbol type. The overall architecture is illustrated in the block diagram in Figure 3.2.

The one operation in Table 3.3 that is not a matrix multiplication is the channel approximation: $\mathbf{H}_p = \mathbf{y}_p\mathbf{s}^*$. It is computed by a separate, dedicated scalar multiplication block, which in Figure 3.2 is named *Channel Approximation*. The $\mathbf{s}^*$ values are loaded into the dedicated *Pilot Memory* during initialization.

Uplink Pilot (ULP)

The pilot symbol is received by the Input Handler and each of the 600 subcarriers are passed to the channel approximation block. The resulting channel estimation matrices $\mathbf{H}_p$ are routed both to the Queue Handler which pushes them into the systolic array to compute $\mathbf{G}_p$, as well as the memory. If there are multiple pilot symbols, the channel matrices from previous pilot symbols are simultaneously read from the memory as they are also needed to compute $\mathbf{G}_p$. The resulting tiles of the $\mathbf{G}_p$ matrix are fed to the Output Handler. The Output Handler temporarily stores and repackages the tiled inputs into a complete $\mathbf{G}_p$ matrix. Remember that $\mathbf{G}_p$ is symmetrical, all tiles don't need to be computed. Once all pilots have been processed, the complete gramian is outputted.

Uplink Data (ULD) and Downlink Data (DLD)

The data symbol arrives at the Input Handler. From there the vectors are forwarded to the Queue Handler. The Queue Handler is also transferred the channel matrices from the memory. Once the systolic array has performed the calculation the resulting vectors are passed to the Output Handler, which outputs them.

3.4 Implementation

The design was implemented with a Scala based language called Chisel. Chisel is an open source alternative to traditional hardware description languages (HDLs) like Verilog and VHDL. The advantage of using something like Chisel is that the underlying Scala infrastructure is used, which can make Chisel very powerful in regards to for example code reuse and configurability. Chisel also has an integrated simulation engine that was used to verify the design. The chisel code can be compiled to Verilog, it was in this way imported to the Vivado design tool for implementation.

Chapter 4
Front-end Design

This chapter describes the design process of the 2 channel evaluation front-end as well as the 16 channel front-end and the associated LO splitter board. First the method used for determining the architecture and components is presented, which is followed by the frequency plan after which the architectures are evaluated and compared. Lastly a breakdown of the 2 channel evaluation front-end, the 16 channel front-end and the LO splitter design is presented. The 16 channel front-end refers to the board containing the 16 RF paths and associated circuitry which connects to the ZCU216 RFSoC Evaluation Kit. The LO splitter is a stand-alone board consisting of power splitters to split an incoming local oscillator signal and distribute it to all mixers on the front-end board through SMA connectors. The report does not cover the local oscillator generation or the distribution to multiple distribution boards.

4.1 Method

Initially, the previous front-end was examined to identify components and ideas that could be re-used and reference material for the RFSoC ZCU216 was examined to understand the features of the RF-ADC and RF-DAC. Thereafter different front-end architectures were studied and compared. Available components were evaluated based on performance, cost and implementation complexity, based on this calculations were done to estimate the performance and cost of different implementations. Once the front-end itself was defined, the surrounding circuitry such as voltage regulators followed a similar process. A two channel evaluation front-end was designed and manufactured to evaluate the performance and identify problems before finalizing the design of the 16 channel front-end. In parallel, a frequency plan was developed to determine filtering requirements and spur performance.

4.2 Frequency plan

The existing system has a carrier frequency of 3.84 GHz and an analog bandwidth of 280 MHz therefore this was used as the starting point for the frequency plan. As the frequency range of interest is between 7.4 and 8.5 GHz, to maintain a fixed IF such as in a superheterodyne architecture, the LO frequency would be

between 3.56 and 4.66 GHz which unfortunately overlaps with the IF. This is an issue because typically mixers do not have very good LO to IF isolation. In other words the LO would be present on the IF as a large spur. Additionally, the 2 x LO spur will overlap with the RF range. Sweeping the LO also requires the FPGA to interface with an RF synthesizer which adds complexity thus it was preferable to keep the LO fixed. This also means that the LO can be generated by a stand-alone device and the wide bandwidth of the RFSoC can be used to shift the carrier frequency.

The frequency plan was based on the spur amplitude table available in the mixer datasheet[26] and the equations to calculate the frequencies of the spurs according to equations (4.1) for IF spurs and (4.2) for RF spurs. In addition, the image frequency was calculated according to (4.3), to evaluate whether an image reject mixer would be necessary.

The largest spurs were assumed to remain at the same multiple of the LO, RF and IF frequencies as in the mixer datasheet, since spurs are the result of leakage, intermodulation or distortion and largely influenced by the type of mixer.

$$f_{SPUR} = |(M \cdot f_{RF}) - (N \cdot f_{LO})| \tag{4.1}$$

$$f_{SPUR} = |N \cdot f_{LO} \pm M \cdot f_{IF}| \tag{4.2}$$

$$f_{IM} = f_{RF} - 2f_{LO} \text{ for } f_{LO} < f_{RF} \tag{4.3}$$

The initial frequency plan based on a 3840 MHz intermediate frequency can be seen in Table 4.1. As seen in the spur Table 4.2, the only viable frequency range is 8210 to 8495 MHz as the spurs will otherwise be nearly impossible to filter with a passive filter due to their proximity to the desired frequencies. And despite this there are certain frequencies with large magnitude spurs which are still difficult to filter.

Table 4.1: Frequency plan for a 3840 MHz IF, with a fixed LO.

Frequency Plan				
PTS frequencies [MHz]	7427.1–7568	7581.1–7722	7900–8185	8210–8495
LO frequency [MHz]	–	–	–	4510
Analog IF center frequency [MHz]	–	–	–	3840
Analog IF bandwidth [MHz]	–	–	–	280
Min RF frequency [MHz]	–	–	–	8210
Max RF frequency [MHz]	–	–	–	8490

Table 4.2: Spur frequency tables for 3840 MHz IF. The first and second value for a given M and N represent the spur bounds. Red cells indicate dominant spurs by intensity. Green cell is the desired frequency.

(a) Down-conversion IF output spur frequencies.

M	N											
	0		1		2		3		4		5	
0	–	–	4510	4510	9020	9020	13530	13530	18040	18040	22550	22550
1	8210	8490	3700	3980	810	530	5320	5040	9830	9550	14340	14060
2	16420	16980	11910	12470	7400	7960	2890	3450	1620	1060	6130	5570

(b) Up-conversion RF output spur frequencies.

M	N											
	0		1		2		3		4		5	
0	–	–	4510	4510	9020	9020	13530	13530	18040	18040	22550	22550
1	3700	3980	810	530	5320	5040	9830	9550	14340	14060	18850	18570
1	3700	3980	8210	8490	12720	13000	17230	17510	21740	22020	26250	26530
2	7400	7960	2890	3450	1620	1060	6130	5570	10640	10080	15150	14590
2	7400	7960	11910	12470	16420	16980	20930	21490	25440	26000	29950	30510

For the 16 channel front-end, the frequency planning was revisited to improve spur performance and better utilize the bandwidth of the RFSoC. The result was the frequencies in Table 4.3, the spur table up to $N = 5$ and $M = 2$ can be seen in Table 4.4. This frequency plan contains some in-band spurs, however all are very low magnitude for a double balanced passive mixer (<50 dBc). For the RF the $M = 2$, $N = 2$ and $M = 2$, $N = 1$ spurs are the most problematic, overlapping with parts of the frequency range but both have very low magnitude. For IF the important spurs are the $M = 2$, $N = 2$ and $M = 2$, $N = 3$ both with low magnitude.

Table 4.3: Frequency plan for 1860 MHz IF, utilizing the wide RF-SoC bandwidth to cover the entire frequency range with a fixed LO.

Frequency Plan				
PTS frequencies [MHz]	7427.1–7568	7581.1–7722	7900–8185	8210–8495
LO frequency [MHz]	6100	6100	6100	6100
Analog IF center frequency [MHz]	1860	1860	1860	1860
Analog IF bandwidth [MHz]	1100	1100	1100	1100
Min RF frequency [MHz]	7410	7410	7410	7410
Max RF frequency [MHz]	8510	8510	8510	8510

Table 4.4: Spur frequency tables for 1.86 GHz IF. The first and second value for a given M and N represent the spur bounds. Red cells indicate dominant spurs by intensity. Green cell is the desired frequency.

(a) Down-conversion IF output spur frequencies.

M	**N**											
	0		1		2		3		4		5	
0	–	–	6100	6100	12200	12200	18300	18300	24400	24400	30500	30500
1	7410	8510	1310	2410	4790	3690	10890	9790	16990	15890	23090	21990
2	14820	17020	8720	10920	2620	4820	3480	1280	9580	7380	15680	13480

(b) Up-conversion RF output spur frequencies.

M	**N**											
	0		1		2		3		4		5	
0	–	–	6100	6100	12200	12200	18300	18300	24400	24400	30500	30500
1	1310	2410	4790	3690	10890	9790	16990	15890	23090	21990	29190	28090
1	1310	2410	7410	8510	13510	14610	19610	20710	25710	26810	31810	32910
2	2620	4820	3480	1280	9580	7380	15680	13480	21780	19580	27880	25680
2	2620	4820	8720	10920	14820	17020	20920	23120	27020	29220	33120	35320

Finally due to the high IF the image frequencies are located far from the band of interest and attenuated significantly by the RF filter. For down-conversion the image frequency is located at

$$f_{IM} = f_{RF} - 2f_{LO} \text{ for } f_{LO} < f_{RF}\,,$$
$$3690\,\text{MHz} \leq f_{IM} \leq 4790\,\text{MHz}.$$

4.3 Architecture

The RF path can be divided into three main stages: the RF stage, the mixer stage, and the IF stage, each containing separate transmit and receive paths. Although these blocks may be combined in several alternative ways, certain configurations are omitted for clarity. The purpose of this section is to identify suitable architectures for both the evaluation front-end and the 16-channel front-end based on the previously defined requirements. The considered architectures are first presented and analyzed quantitatively, followed by a numerical comparison using representative components, as described in section 4.3.5.

4.3.1 Dual mixers

A traditional receiver/transmitter front-end features dual mixers, one for the receive path and one for the transmit path. This offers the lowest possible insertion loss. However, in a TDD system both mixers are never used simultaneously, as the mixer is quite expensive a single mixer architecture was preferred to save cost.

A dual mixer architecture can be seen in Figure 4.1. A switch is used instead of a splitter to distribute the LO signal to the respective mixer since a switch has lower insertion loss resulting in a lower LO power requirement and does not add significant cost.

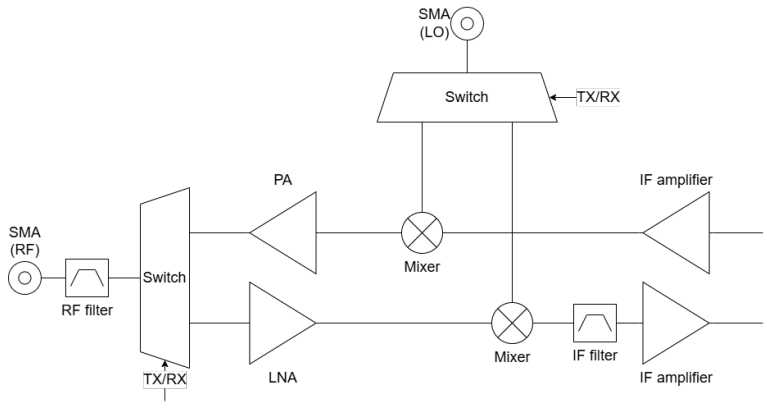


Figure 4.1: Architecture with dual mixers, IF amplifier on both transmit and receive paths.

4.3.2 Driver amplifier

A pre-amplifier stage before the power amplifier (PA) (also referred to as a driver amplifier), such as in Figure 4.2 was considered. This is the typical arrangement in most front-ends as it avoids mixer compression and has better spur performance, although it is typically more expensive than amplification at the IF since the amplifier has to handle the RF directly.

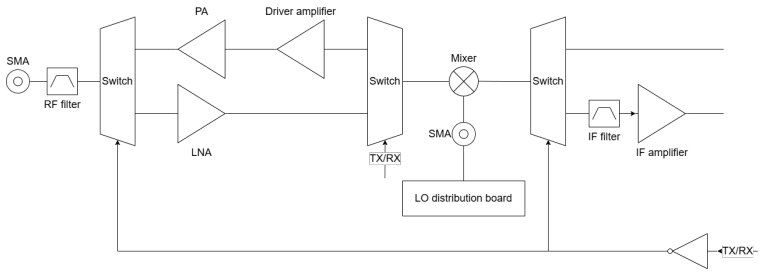


Figure 4.2: Architecture with a driver amplifier before the PA.

4.3.3 TX IF amplifier

An architecture using intermediate frequency amplifiers in both the transmit and receive paths is shown in Figure 4.5. In most front-end designs, IF amplification is applied only in the receive path, since transmit-path amplification can cause mixer compression and increased spur levels due to the higher mixer input power [27]. Its main advantage over the driver-amplifier architecture is lower cost.

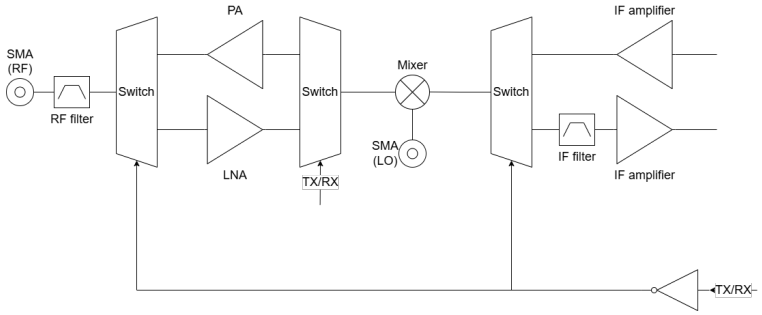


Figure 4.3: Architecture with IF amplifier on the transmit path.

4.3.4 Integrated front-end

An architecture with an integrated front-end which has the TDD switch, PA and low noise amplifier (LNA) built-in was also considered. The front-end ICs currently available at these frequencies are quite expensive (and few) as the upper mid-band is not widely used in commercial applications. The integration saves space which would be advantageous in a MIMO system but despite the integration, negative bias and supply sequencing was required, increasing the complexity particularly for the 16 channel front-end. The architecture can be seen in Figure 4.4.

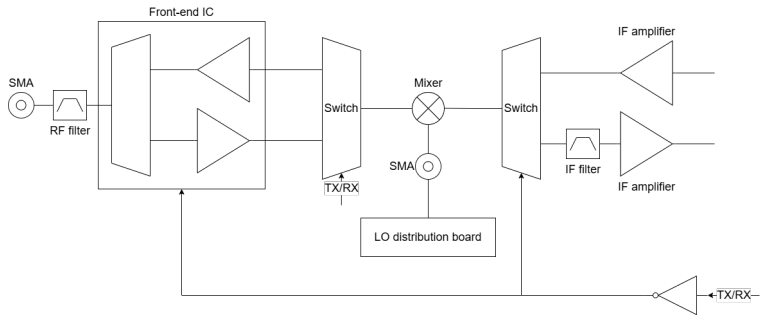


Figure 4.4: Integrated front-end, single mixer.

4.3.5 Cost & performance comparison

A theoretical performance as well as cost comparison between the architectures can be seen in Table 4.5. In section 4.4.1 the component selection used for the performance estimation is described. However, for most practical purposes the switches, mixer and filters can be assumed to be generic, with an insertion loss of 2 - 2.5 dB for the filters, approximately 1 dB for the switch and 9 dB for the mixer.

Table 4.5: Front-end performance and cost comparison.

	TX IF amplifier	**Dual mixer**	**Integrated front-end**	**Driver amplifier**
Cost [SEK]	1549	1899	2885	1679
		Receive		
Noise figure [dB]	3.8	3.5	5.2	3.8
Gain [dB]	23.05	25.35	18.70	23.05
		Transmit		
Noise figure [dB]	4.1	3.6	3.5	14.8
Gain [dB]	22.05	24.35	27.20	23.55
P1dB [dBm]	17.85	17.85	21.00	17.85
Max Input Power [dBm]	-4.2	-6.5	-6.2	-5.70

The maximum input power is defined as the input power to the front-end required for the PA to reach compression (P1dB), and the noise figure is calculated according to equation 2.11, accounting for all components in the signal path assuming perfect matching and no PCB losses. For the transmit comparison it is important to consider the fact that different input powers will appear very different at the mixer input, this affects the spur performance. Additionally, allowing for higher transmit input power can benefit the overall system SNR, since reducing the full-scale of the RF-DAC will reduce the SNR as the noise floor is essentially constant although noise figure of the transmit path is generally not considered an

important aspect in front-end design.

As can be seen in Table 4.5 the differences are not significant, but in certain areas the different architectures have unique advantages. The integrated front-end module has a higher transmit gain and better P1dB than the PA used for the discrete architectures. Although it is nearly twice the cost per channel. At which point the discrete architectures could use a PA with a better P1dB if a more complicated implementation is acceptable. The dual mixer architecture has slightly better receive noise figure than the other architectures but the noise figure is still dominated by the RF filter and TDD switch. Using dual mixers could be done with any of the architectures but the cost is prohibitive and offers little practical benefit to anything but receive gain. Initially the transmit (TX) IF amplifier architecture seems like the best alternative, however considering that the mixer input power will be 13.1 dBm compared to -7.9 dBm in the driver amplifier architecture and the fact that third-order intermodulation spur amplitude increases by 3 dB per 1 dB input power [28][27] the spur performance is theoretically $(13.1 + 7.9) \cdot 3 - (13.1 + 7.9) = 42$ dBc worse compared to the driver amplifier architecture. Despite this the TX IF amplifier architecture was chosen for the evaluation front-end as a result of somewhat poor planning. For the 16 channel front-end the driver amplifier architecture in 4.2 with additional filtering was chosen due to the superior spur performance and wider bandwidth as shown in section 4.2.

4.4 Design of the 2 channel evaluation front-end

In an attempt to design the evaluation front-end to be compatible with the testbed without any changes to the baseband an IF of 3840 MHz was used. Unfortunately due to the timeline of the project, the frequency plan was done in parallel and it was not yet clear that this had significantly worse spur performance.

as the frequency plan was done in parallel with the design. As mentioned previously the architecture chosen featured an intermediate frequency amplifier on the transmit path which is expected to worsen spur performance. Despite this the design and RF component selection is presented in this section due to the similarities with the improved design featured in section 4.5.

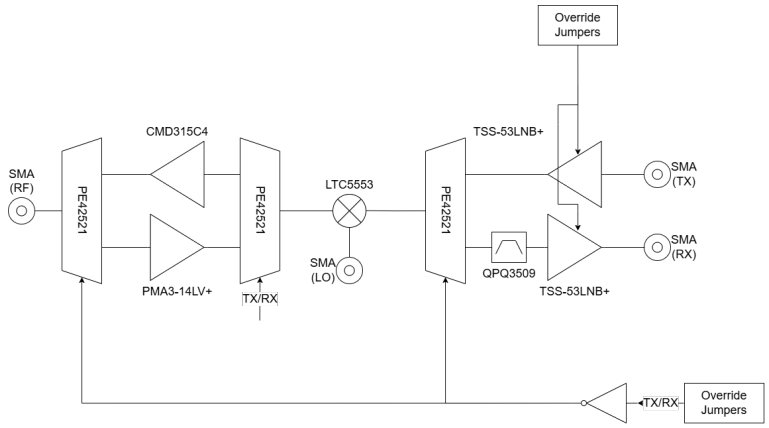


Figure 4.5: Evaluation front-end block diagram.

4.4.1 RF components

Switch

To realize the time-domain duplexing and allow for a single mixer, multiple RF switches are required. The frequency range should cover both the IF and the RF with good return as well as insertion loss. It should be capable of handling the full output power continuously when transmitting. The switch chosen was the PE42521 from pSemi as it exceeds all the requirements, only requires a single 5V supply and is fairly inexpensive. Additionally it has a fairly reasonable ESD rating which could prevent damage to the rest of the front-end as it is the first active component in the signal path.

Mixer

The mixer chosen was the LTC5553 from Analog Devices, it features 3 - 20 GHz RF bandwidth, a built in LO amplifier and wide IF bandwidth with good return loss at suitable IF frequencies, both of which greatly simplify the design of a MIMO front-end. It requires a 3.3 V for the LO amplifier but is otherwise a traditional double-balanced passive mixer. The downside is that it is fairly expensive and as a passive mixer it has some conversion loss, up to 9 dB at the frequencies of interest, detailed specifications are available in Table 4.6. Initially the lower cost but similar LTC5549 was considered although unfortunately it was unavailable from major distributors.

Table 4.6: LTC5553 Mixer Specifications [26].

Parameter	Min	Typ	Max	Units
LO Frequency Range		1 to 20		GHz
RF Frequency Range		3 to 20		GHz
IF Frequency Range		500 to 9000		MHz
RF Return Loss		> 9		dB
LO Input Return Loss		> 10		dB
LO Input Power	-6	0	6	dBm

Low-noise amplifier

The LNA chosen was the PMA3-14LV+ from Mini-Circuits it features a very low noise figure and reasonably high gain at a competitive cost, see Table 4.7. It requires some external biasing circuitry which can affect the performance once implemented but there is a comprehensive reference layout available. There is an alternative which does not require any external biasing namely the PMA3-5153+ but it is more expensive and availability is poor as it was recently released.

Table 4.7: PMA3-14LV+ Specifications at 8 GHz [29].

Parameter	Min	Typ	Max	Units
Frequency Range	0.05		10	GHz
Gain (8 GHz)	20.9	23.3		dB
Input Return Loss (8 GHz)		21		dB
Output Return Loss (8 GHz)		14		dB
Isolation (0.05–10 GHz)		25		dB
Output P1dB (8 GHz)		+18.3		dBm
Output IP3 (8 GHz)		+27.8		dBm
Noise Figure (8 GHz)		1.3		dB
Device Operating Voltage	+4.75	+5.0	+5.25	V
Device Operating Current		59		mA

Intermediate frequency amplifier

The IF amplifier chosen was the TSS-53LNB+ it operates between 0.5 - 5 GHz and features a bypass mode which was a requirement set by EIT, this allows for greater dynamic range. In Table 4.8 detailed specifications for the device at 4 GHz can be found. It requires no external circuitry on the RF path which simplifies the implementation.

Table 4.8: TSS-53LNB+ Specifications at 4 GHz [30].

Parameter	**Min**	**Typ**	**Max**	**Units**
Supply Voltage	4.8	5.0	5.2	V
Amplifier Mode, 4 GHz				
Gain	-	19.5	-	dB
Noise Figure	-	1.6	-	dB
Output P1dB	-	20.2	-	dBm
Output IP3	–	33.4	–	dBm
Input Return Loss	-	14.5	-	dB
Output Return Loss	-	10.7	–	dB
Supply Current	–	82	105	mA
Bypass Mode, 4 GHz				
Insertion Loss	-	0.9	-	dB
Input Return Loss	-	14.5	-	dB
Output Return Loss	-	14.0	-	dB
Supply Current	-	2	-	mA

Filters

The RF filter is a band-pass which is supposed to prevent unwanted frequencies from reaching the front-end when receiving and attenuate spurs from the front-end when transmitting. The evaluation board did not include an RF filter to evaluate mixer spur performance. Similarly to the RF filter, the IF filter is used to remove spurs and out-of-band signals but in this case from the down-converted signal, the IF filter chosen for the evaluation board was the QPQ3509 from Qorvo. It is a band-pass filter with very sharp filtering and 40 dB of attenuation (outside the pass-band) up to 10 GHz. This filter was used as the RF filter on the previous board.

SMA connectors

Although SMA connectors themselves can be used at high frequencies, the interface to the PCB affects the impedance [31]. The choice was made to go with vertical hybrid through hole technology (THT)/surface mount technology (SMT) SMA connectors from SAMTEC such as on the previous front-end namely the SMA-J-P-H-ST-MT1, where the ground connections are through hole whilst the center RF pin is surface mount, typically horizontal board edge connectors would be used as these present less of an impedance discontinuity but due to the mechanical constraints and difficulty laying out all the channels side by side on the full-scale board vertical connectors were preferred despite worse performance. Solder-less, also known as screw-mount SMAs were also considered. These are typically better than the SMT alternative but cost significantly more and require manual assembly.

Power amplifier

A comparison of all amplifiers considered and their biasing/sequencing requirements can be seen in Table 4.9.

Table 4.9: Comparison of PAs. (NB) indicates need for negative bias, (S) need for sequencing.

Part Number	Gain [dB]	NF [dB]	P1dB [dBm]	Cost [kr]	Freq. [GHz]	Supply [V]
CMD170P4	30	–	28	748	7.5–9	7
AVA-6183MPS+	20.9	6.3	26.6	859	6–18	6
HMC1082LP4E	22	–	24	726	5.5–18	5 (NB)
HMC633LC4	30	10	23	713	5.5–17	5 (NB)(S)
ADTR1107	22	9	23	2045	6–18	5, ±3.3 (NB)(S)
CMD315C4	19.5	5.5	21	629	4–10	5

Ultimately the power amplifier chosen was the CMD315C4 from Qorvo, despite being marketed as a driver amplifier it is suitable as a power amplifier in this case as a 1 dB compression point of 20 dBm or higher (100 mW) was deemed acceptable. In contrast to many other amplifiers with similar specifications, it does not require any sequencing or external biasing components and only a single 5V supply which greatly simplifies implementation. The availability of the amplifiers considered varies greatly and influenced the choice as well due to the large quantity required for the testbed. The specifications for the CMD315C4 be seen in Table 4.10.

Table 4.10: Qorvo CMD315C4 Specifications [32].

Parameter	Min	Typ	Max	Units
Frequency Range		4–10		GHz
Gain		19.5		dB
Noise Figure		5.5		dB
Input Return Loss		10		dB
Output Return Loss		15		dB
Output P1dB		21		dBm
Output IP3		33		dBm
Supply Current		143		mA

An alternative which could be considered if a higher P1dB is desired is the CMD170P4 also from Qorvo, due to the higher gain it allows for a weaker driver amplifier for example the Mini-Circuits AVA-183A+, which is similarly priced to the PMA3-14LV+ and requires no external biasing. The downside is that the CMD170P4 and the driver amplifier would require different supply voltages.

4.4.2 Layout

On the Evaluation board all the RF traces are coplanar waveguides with ground (CPWG). Using the built-in impedance calculator in Kicad and the manufacturer stack-up from Figure 4.6 the trace width chosen was 0.35 mm, and the gap 0.35 mm. Additionally for the high-frequency traces the solder-mask was removed as it improves matching and reduces insertion loss slightly [31].

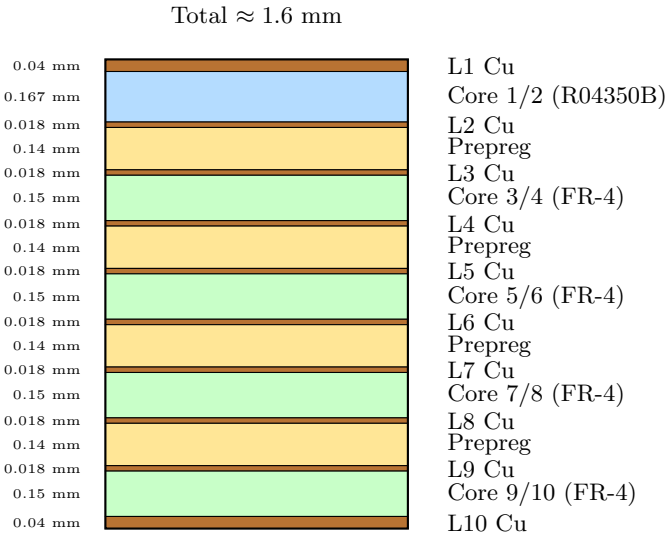


Figure 4.6: Evaluation board stack-up.

4.5 Design of the 16 channel front-end

The design for the 16 channel front-end consists of a backplane and 16 front-end modules. The backplane contains the power circuitry, LO ports, IO and control logic. The backplane connects to the RFSoC through the two SAMTEC LPAM connectors on the bottom side. This section covers firstly the design of the backplane and then the design of the front-end modules.

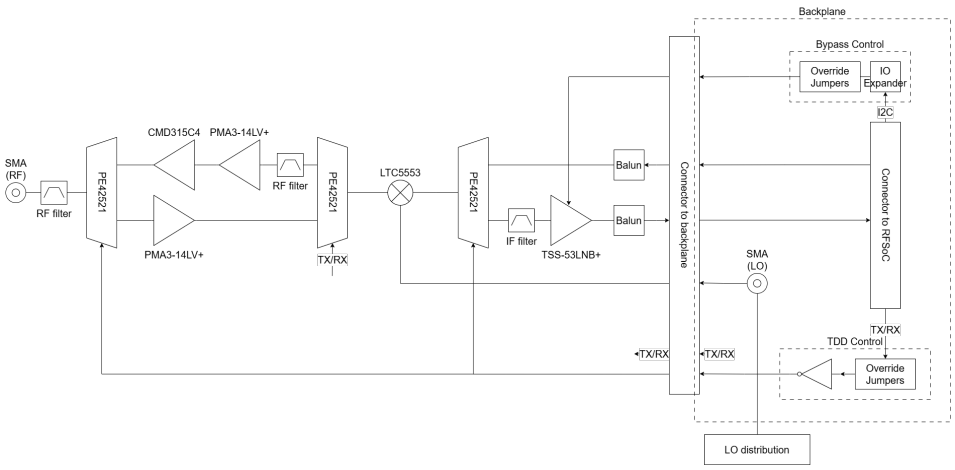


Figure 4.7: Overview of front-end module and backplane interface.

4.5.1 Backplane

Voltage regulator

As mentioned previously, a very important characteristic for the power supply in an RF system is noise, because of this the LT8625SIV from Analog Devices was chosen. It has a very low output noise and is suitable for RF applications [33]. To further improve the noise a larger inductor and more output capacitance was also added. The output voltage is set using a single resistor for output voltages under 6 V using the built-in precision current source. An important factor to consider for the output capacitance is the derating curve of ceramic capacitors, specifically X7R, X5R and similar style dielectrics. The capacitance rapidly decreases with applied DC voltage at the rated voltage of the capacitor as much as 80 % of the capacitance is lost [34]. Therefore additional capacitors were added compared to the reference design.

The LT8625SIV is capable of supplying up to 8A which is limited by the operating temperature and the thermal dissipation of the device. To avoid thermal issues and reduce power supply coupled noise separate regulators are used for the TX and receive (RX) components, the common components (e.g RF switches and other IO) are supplied from the RX supply. It is also possible to turn off the 5VTX from the FPGA in order to save power. However it has some start-up time as the output capacitance is quite large which could be a problem for quick TDD operation.

Input filter

To avoid conducted emissions from the switching regulator or front-end a simple input filter was designed, using a ferrite, a large electrolytic capacitor and a ceramic capacitor for high frequency attenuation, one of which is in series with a resistor to adjust the ESR to dampen the resonance. The filter with dampening is shown

in Figure 4.8. A comparison of the frequency responses of the dampened and undampened filter is shown in Figure 4.9

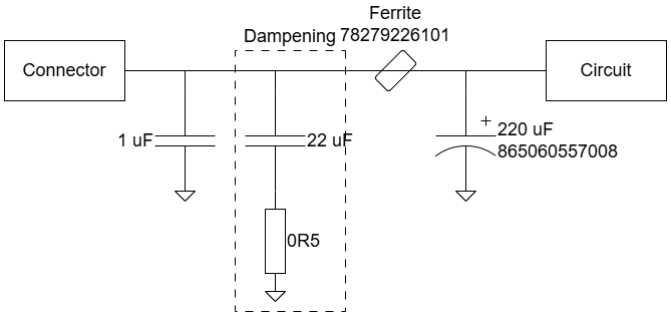


Figure 4.8: Input filter schematic.

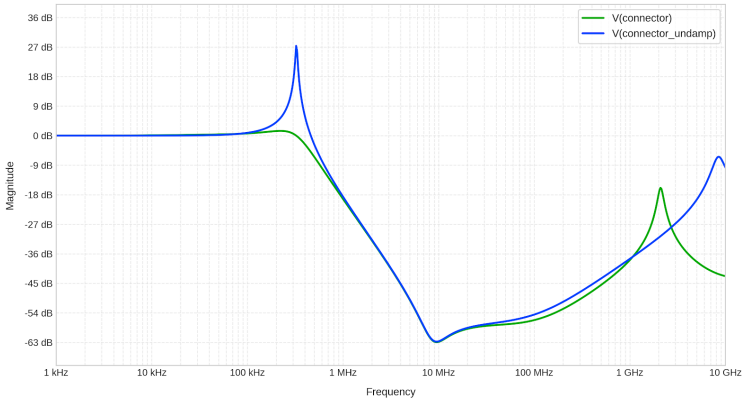


Figure 4.9: Input filter simulation including parasitics for the capacitors and ferrite. AC source is on the circuit side and the result is at the connector. Blue trace is without the dampening section and green is with dampening.

Control interface

To ensure backwards compatibility with the previous front-end board a similar jumper based control interface was implemented and the GPIO pins for communication between FPGAs and TDD were kept the same, the bypass mode of the IF amplifiers was designed to be controlled over the I2C interface available from the FPGA, this allows for more convenient routing of the signals as the interface to the connector is serial and therefore interferes significantly less than the otherwise parallel interface of the GPIO pins.

Low drop-out regulator

An LDO is used to generate the local IO voltage for interfacing with the FPGAs I2C and digital IO pins, so that the IO expander which controls the IF amplifier bypass can use 5 V logic as required by the amplifiers. The specific model in this case does not have any significant effect on the performance and the load current is very low. The selected LDO was the AP2127K-ADJ from Diodes Incorporated. It comes in a standard footprint which is pin compatible with LDOs from many other vendors and has good availability.

Layout

Selecting the correct PCB material and finding an appropriate stackup is more difficult than it first appears. The losses of a 200 mm stripline in FR4 with a dielectric thickness of 0.308mm between the ground planes will have a combined dielectric and conductor loss of over 7 dB which is a significant loss of power. To reduce the conductor losses a material with a lower dielectric constant can be selected, this allows for increased trace width for the same dielectric thickness. To reduce the dielectric losses a material with lower loss tangent (or lower dielectric constant, although it only scales with the square root of the dielectric constant) has to be used [35], FR4 has a loss tangent of approximately 0.02 whereas typical high-frequency dielectrics have an order of magnitude lower loss tangent.

The possible stack-ups were discussed with the PCB manufacturer and reviewed to ensure that it did not result in performance degradation. First an all FR4 stack-up was considered due to the low cost and similarity with the previous design, unfortunately due to the module design the layout included long traces which would result in signficany losses as mentioned earlier. In a traditional system where the IF is much lower this would be less of an issue as both conductor loss and dielectric loss scale with frequency. Anyways changing the stack-up was preferable from a layout perspective and the PCB manufacturer suggested the stack-up seen in Figure 4.10 for the full-scale board, using a lower cost alternative to RO4350B namely S7163H which results in approximately 3 dB losses for the longest traces compared to the previous 7 dB due to the much lower loss tangent and lower dielectric constant.

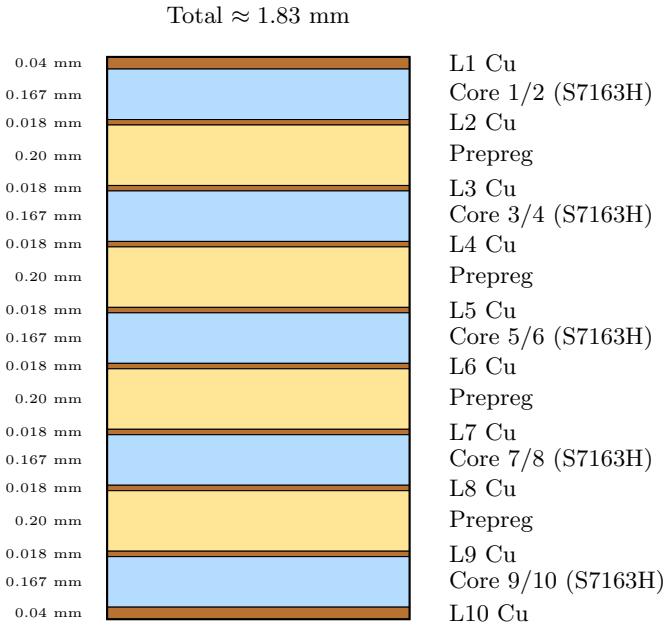


Figure 4.10: Backplane stack-up.

For the differential stripline Saturn PCB Toolkit was used to calculate the trace widths based on the manufacturer stack-up as seen in Figure 4.10. The conductor width and spacing was experimented with until the differential impedance closely matched 100 Ω, the characteristic impedance closely matched 50 Ω and was relatively resistant to possible process variations in dielectric constant and conductor spacing. The result was a trace width of 0.2 mm and spacing of 0.4 mm. The via impedance was also calculated using Saturn PCB Toolkit a 0.6 mm pad, 0.3 mm hole via with a plane opening of 1 mm on all layers as seen in Figure 4.11 gave an impedance close to 50 Ω. It is important to note that these calculators can be inaccurate when it comes to via impedance and ideally via structures should be simulated using an EM solver [36]. The vias are not back-drilled creating stubs at different lengths depending on the layer transitions. Below the SMA connectors the ground fill is removed to improve impedance matching at high frequencies.

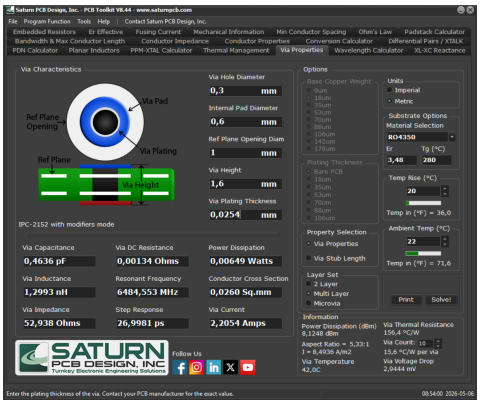
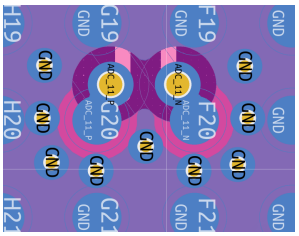


Figure 4.11: Via impedance calculation in PCB Toolkit.

The routing of the signals into the SAMTEC LPAM connector is based on the whitepaper from SAMTEC regarding their Analog over Array™ technology which can be seen in Figure 4.12a but with the via size for the signals from PCB toolkit. The adapted design on the backplane can be seen in Figure 4.12b.

(a) Analog over Array™ reference. [37]



(b) Implemented design.

Figure 4.12: LPAM Connector Differential Pair Routing.

The routing between the modules and the RFSoC connector is done to maximize the distance between adjacent ADC and DAC signals to reduce crosstalk. Additionally, the differential pair skew is tuned so that the positive and negative traces are the same length.

4.5.2 Front-end module

Table 4.11: Estimated performance of the FR3 front-end.

Cost [SEK]	1679
Transmit	
Gain [dB]	21.0
Max Input Power (PA compression) [dBm]	-4.1
Output Power Compression [dBm]	17.4
Expected Worst Spur [dBc]	< -45
Frequency [MHz]	7410–8510
Receive	
Gain [dB]	22.6
Gain (bypass mode) [dB]	2.65
Noise Figure [dB]	3.8
Max Input Power (IF amplifier compression) [dBm]	-3.1
Expected Worst Spur [dBc]	< -50
Frequency [MHz]	7410–8510

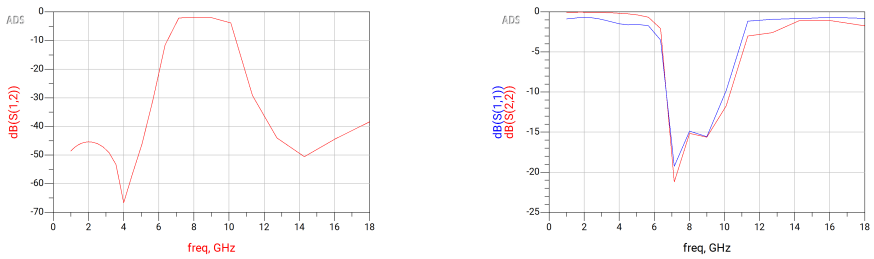
Filters & balun

The lower IF determined for the 16 channel front-end creates different filtering requirements compared to the evaluation front-end. The RF filter was redesigned using separate high-pass and low-pass filters to create a band-pass filter suitable for the entire frequency range, the S-parameters for the combined filter were simulated in ADS, the result can be seen in Figure 4.14. A new IF filter and balun were chosen for the lower frequency and wider bandwidth IF. A filter before the amplifiers on the transmit path (referred to as the TX filter) was also included, it is the same design as the RF filter using a discrete high and low-pass filter to further improve spur performance and allows for slightly higher input power and therefore nearly no performance penalty. The filters used can be seen in Table 4.12.

Table 4.12: Filter and balun selection.

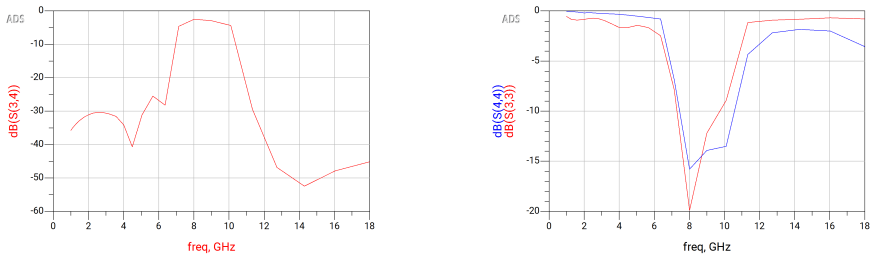
Type	Part Number
IF filter	BFCN-1801+
RF LP filter	LFCW-8400+
RF HP filter	HFCW-6600+
TX LP filter	LFCW-8400+
TX HP filter	HFCW-6600+
Balun	NCS2-33+

Originally the HFCN-7150+ was considered for the high-pass filter although the lower cut-off frequency HFCW-6600+ was preferred, sacrificing some out-of-band rejection (particularly for the 1 x LO spur) for a better in-band return loss and insertion loss, a comparison of the pass-band insertion loss can be seen in Figure 4.15



(a) Insertion loss. **(b)** Return loss.

Figure 4.13: S-parameter simulation of combining LFCW-8400+ low-pass filter and HFCW-6600+ high-pass filter to create a suitable band-pass filter for the 7.4 GHz to 8.5 GHz band.



(a) Insertion loss. **(b)** Return loss.

Figure 4.14: S-parameter simulation of combining LFCW-8400+ low-pass filter and HFCN-7150+ high-pass filter to create a suitable band-pass filter for the 7.4 GHz to 8.5 GHz band.

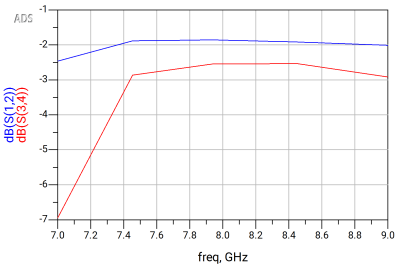


Figure 4.15: Pass-band comparison between the HFCW-6600+ and HFCN-7150+ in combination with the LFCN-8400+.

Another idea which was considered was the use of band-pass filters designed for Ultra Wide-Band (UWB), as the frequency range overlaps significantly with the lower range of the upper-midband. The specific filter considered was the Abracon AFII-LC-0048. Unfortunately when using an IF of 1.86 GHz the 1 x LO spur becomes problematic as 6 GHz is within the pass-band. However this spur could be filtered by using a TX high-pass filter at the expense of worse attenuation of higher frequency spurs.

Driver amplifier

The amplifier used as the driver amplifier is the same part as the LNA on the receive path. It is very cost-effective, the design is proven and the gain is suitable for the CMD315C4 to reach compression with reasonable input power to the front-end. As mentioned earlier a possible alternative suitable for a higher gain PA is the Mini-Circuits AVA-183A+.

Layout

The module features CPWG, microstrip and differential microstrip. The impedances were calculated for the 4 layer stackup in Figure 4.16, resulting in a 0.3 mm trace width and 0.15 mm gap for the CPWG trace, 0.3 mm trace width and 0.25 mm spacing for the differential microstrip and lastly 0.35 mm trace width for the microstrip. To minimize discontinuities in controlled impedance traces, minimal footprints for the 0402 and 0201 components were designed to fit within the controlled impedance trace to the extent possible. Reducing the footprint size is worse for manufacturability therefore the minimal footprints were only used for components in the RF path. A smaller shielding cage was utilized so that the module could be made smaller. The edge-card connector was designed according to the specifications for a SAMTEC HSEC8 connector. Below the SMA connectors the ground fill is removed to improve impedance matching at high frequencies.

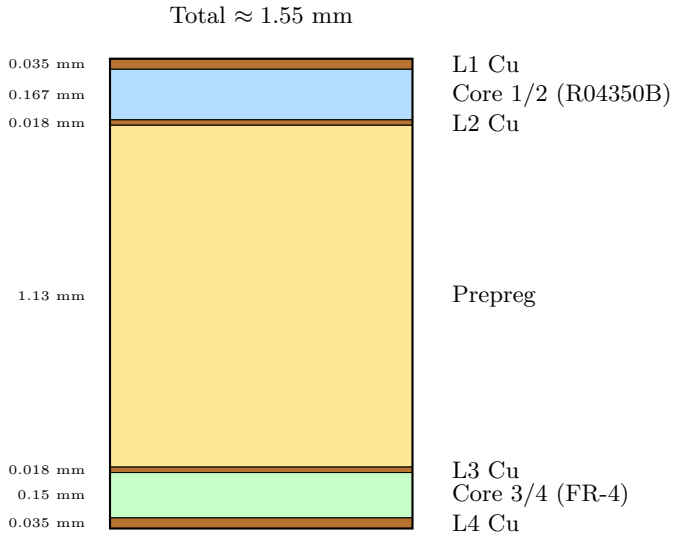


Figure 4.16: Front-end module stack-up.

4.6 Design of LO splitter

To split the local oscillator to all 16 mixers of each front-end board, another board consisting of tree-like distribution using Monolithic microwave integrated circuit (MMIC) splitters was designed, see Figure 4.17. From frequency planning the LO frequency is 6100 MHz. The decision to keep this board separate was partly to reduce the overall footprint of the already large and complex front-end board and also to avoid further complicating the routing by additional RF signals which would have to run across the board more or less perpendicular to the other signals. Using splitters comes at a cost of at least 3 dB per stage as the signal power is divided into two (and the splitter has some insertion loss). To get the necessary 16 outputs per board 4 stages of splitting is required. The loss from splitting is 12 dB and the insertion loss is between 2.8 and 4.4 dB therefore the total loss is between 14.8 and 16.4 dB, splitters have some phase imbalance but more importantly does not significantly affect the phase noise. The mixers require 0 dBm LO drive thus the signal power to the distribution board needs to be at least 16.4 dB to ensure proper functionality of all mixers. Distributed MIMO also requires that the LO signal is distributed between all front-ends. Practical implementation of this in a real-world scenario is out of the scope of this thesis as it is assumed each distribution board will have the required input power.

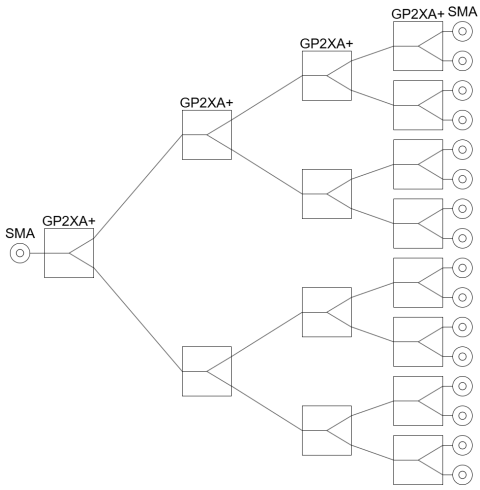


Figure 4.17: Overview of the LO splitter.

4.6.1 Power splitter

The choice of power splitter was made based on cost, bandwidth and isolation. The isolation is important to avoid potential spurious tones from the mixers interfering through the LO port. Phase imbalance is not important, as it is compensated for by the channel estimation since it is constant in time. Bandwidth needs to cover frequencies up to 6.1 GHz as determined by the frequency planning and preferrably with low insertion loss and good return loss. The GP2XA+ from Mini-Circuits fulfills all these requirements with a 32 dB isolation, 2800 to 6200 MHz frequency range, in a compact QFN package.

4.6.2 Layout

The LO splitter board uses the same stack-up as the front-end modules and the same CPWG parameters. Below the SMA connectors the ground fill is removed to improve impedance matching at high frequencies.

Chapter 5
Results & Discussion

5.1 Digital Baseband

In this section the results of the digital implementation will be presented and discussed. Due to time constraint, the support for DLD symbols has not been fully implemented as of writing this report. Because downlink had not been implemented on the rest of the large intelligent surface (LIS) system, it was lowest priority. The design was validated through a mix of targeted and constrained random simulation. It was simulated for all combinations of K, B_c, and $\#ULD$ for multiple frames, while the data was randomly seeded. The outputs where compared to a reference model.

5.1.1 Results

The distributed digital baseband was integrated on the FPGAs of the LuLIS system and in Figure 5.1 a diagram of the resulting constellations can be seen. The measurements where done for 4 UEs. The measurement was taken with the LuLIS-panels distributed in a 8×4 m arena. The antenna panels were pointed inward toward the 4 UEs. UE antennas were separated by no more than 10 cm.

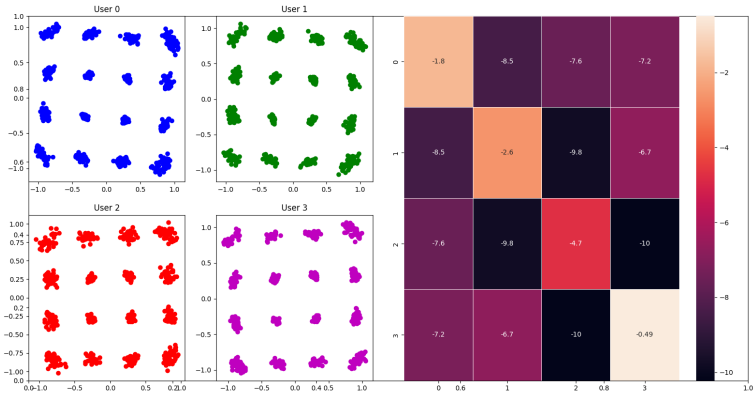


Figure 5.1: Measured constellation diagrams and average gramian matrix magnitude $10\log\left(|\bar{\mathbf{G}}|\right)$ in dB. 12 LuLIS panels where used.

The resource utilization of the design is tabulated in Table 5.1. The values were extracted from the Vivado design tool after Implementation. No integrated logic analyzer (ILA) was included in the design. Notably the percentage of BRAM that are used is quite high, which will be discussed further in section 5.1.2.

Table 5.1: XCZU49 FPGA resource utilization on distributed nodes.

Resource set	LUT	LUTRAM	FF	BRAM	DSP
FPGA total	425280	213600	850560	1080	4272
Used	39%	5.7%	26%	88%	30%

Figure 5.2 plots the latency, T_{G_p}, as defined in subsection 3.2.2, for all combinations of K and B_c. The use of the same color for different values of B_c represents that they are computed with the same lane configuration: blue means four 4×16 lanes, orange means two 4×16 lanes, and green means one 16×16 lane.

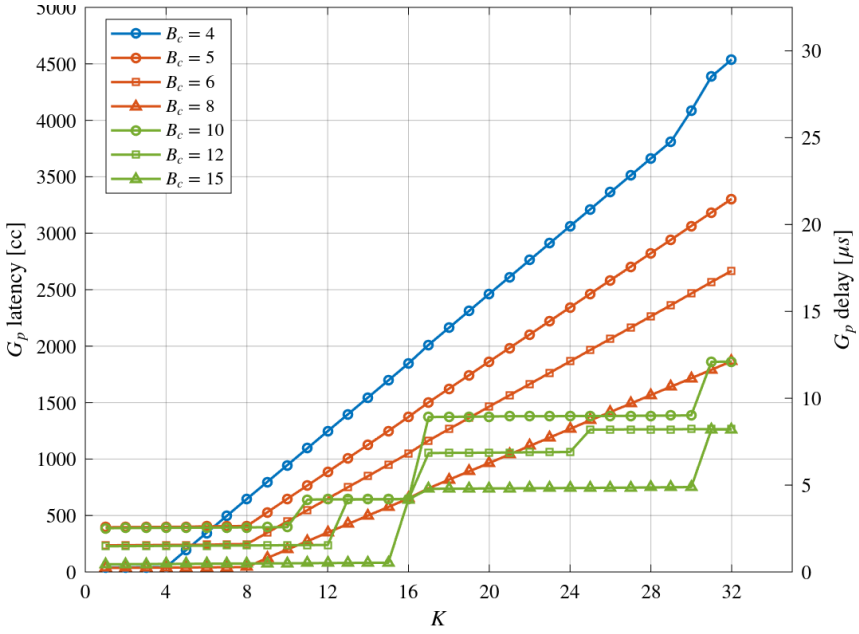


Figure 5.2: T_{G_p} for all combinations of K and B_c.

5.1.2 Discussion

The constellations of Figure 5.1 indicate that the IP block is performing the distributed operations correctly. There is a slight phase shift relative to the frame of reference. Notably, each UE has a different phase shift. Moving the UE antennas has a negligible effect. This slight phase shift is also present when the testbed is using the previous design, see [5].

The main diagonal of the gramian matrix shows how the transmissions from each UEs are scaled. We expect the magnitude of the major diagonal to be much larger than the off-diagonal magnitude, which is what we see. The low off diagonal magnitude is what reduces the interference between UEs.

The most critical resource of the FPGA is the BRAM. Table 5.1 shows that 88% of BRAM blocks on the fabric are utilized. This high utilization has been causing timing closure problems. Because the placer engine gets very constrained in where it can place the BRAM blocks, it must choose worse options, making it more likely that the routing agent cannot meet the timing demands.

There are a few reasons for this high BRAM utilization. The first is that first-in first-outs (FIFOs) mapped to BRAM are used to manage the data-flow of the design. The BRAM that instantiated as FIFOs make up 46% of the total usage. Some of these FIFOs can be reduced in depth to reduce overall use.

All BRAM FIFOs have been manually instantiated with a depth of 1024 and a width of 32, and are then cascaded and put in parallel to create the sizes needed. The reason for the manual instantiation is that we were unsuccessful in getting the Vivado synthesis tool to do this mapping implicitly from the Verilog generated by the Chisel compiler. By changing the option of depth and width to 512 and 64

respectively, a meaningful reduction in BRAM use can be expected.

Another 12% of BRAM utilization is occupied by the channel matrix memory. In this case, another opportunity presents itself. The XCZU49 includes 80 UltraRAM (URAM) blocks that are currently not utilized at all. Moving the memory module to URAM would require 32 URAM blocks and would free up those 12% of the total BRAM.

Another issue relates to the latency of the gramian calculation, shown in Figure 5.2. When choosing the size of the systolic array for this design, a buffer delay of a few microseconds was considered acceptable. In the graph, we see that for $B_c \leq 8$, there is a linear increase with K that makes the worst case latency much higher than desired.

We can also observe a step-like latency increase for $B_c > 8$. This latency comes from the choice of granularity we made during the design phase. We can see that the highest latency is achieved when $B_c = 10$, which aligns with the analysis made in section 3.3.3.

The linear latency increase stems from the FIFO-based tile reordering that is done in the Output Handler, see Figure 3.2. Every time a tile of length B_c arrives to be processed, it must wait for $K - B_c$ values to be dequeued from the storage FIFO. For $B_c > 8$, the throughput of the systolic array is lower than that of the tile reorderer, therefore these values do not experience the same effect. The reason why tile reordering was implemented this way was time constraint. The solution is simple and was very fast to implement. Because DLD was already outside the scope of implementation, the increase in T_{G_p} does not affect the short term capability of the IP block, but will need to be addressed in the near future.

5.2 Evaluation front-end

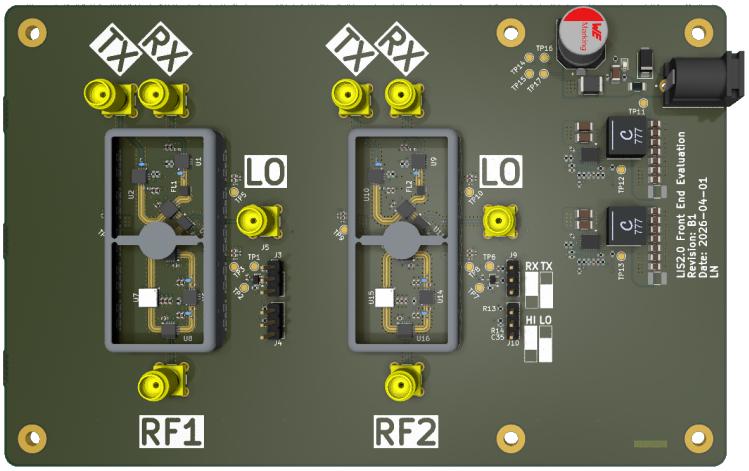


Figure 5.3: Render of the evaluation front-end.

5.2.1 Schematic

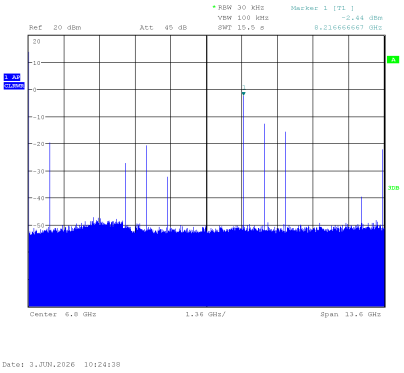
The schematic is available as appendix A.

5.2.2 Measurements

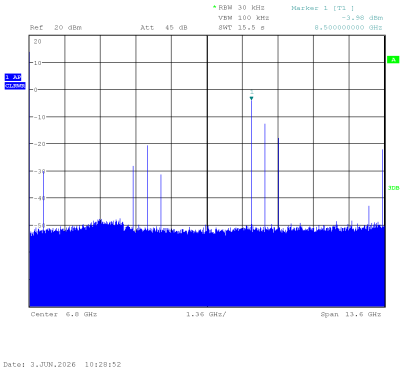
Transmit

To measure the transmit gain, the TX SMA was connected to the R&S SMIQ 06B RF signal generator, likewise the LO port was connected to the Agilent E8257D RF signal generator. The RF SMA connector was connected to the R&S FSU50 Spectrum Analyzer. The IF signal frequency was set to 3980 MHz, and the power set to −25 dBm. The LO frequency was set to 4510 MHz and 0 dBm. The spectrum from the analyzer was saved. The measurement was repeated with an IF of 3700 MHz.

Thereafter the LO was set to 6100 MHz and 0 dBm. The measurement was repeated with an IF of 1310 MHz and 2410 MHz respectively.

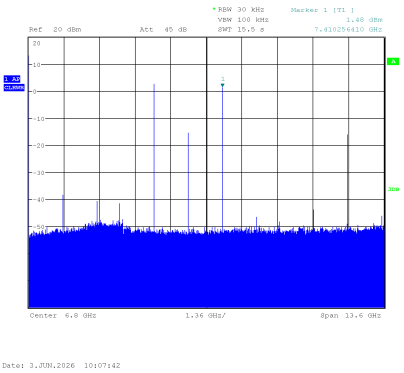


(a) 3700 MHz IF with 4510 MHz LO output spectrum.

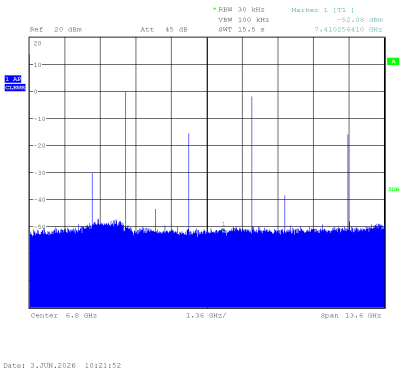


(b) 3980 MHz IF with 4510 MHz LO output spectrum.

Figure 5.4: Up-conversion (TX) spectrum for 3700 MHz and 3980 MHz IF with 4510 MHz LO.



(a) 1310 MHz IF with 6100 MHz LO output spectrum.



(b) 2410 MHz IF with 6100 MHz LO output spectrum.

Figure 5.5: Up-conversion (TX) spectrum for 1310 MHz and 2410 MHz IF with 6100 MHz LO.

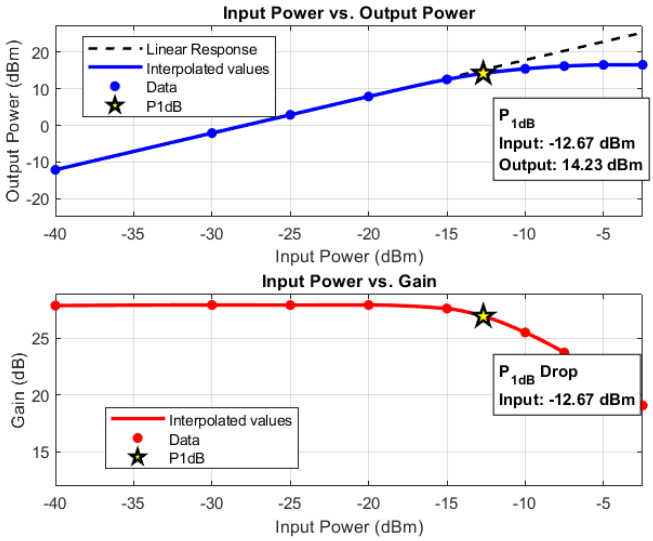


Figure 5.6: Up-conversion output power and gain against input power for IF 1860 MHz, and LO 6100 MHz.

From this it is evident that the compression point is reached earlier than expected, the compression point is only 14 dBm, however the cable loss is not accounted for and therefore the actual value slightly better (around 1 – 2 dB), although it is still lower than the estimated 19.8 dBm. The gain is around 28 dB, which is very close to the estimated 28.8 dB (based on typical values from the datasheets).

The LO power was also swept to evaluate the effect on spur performance and conversion loss. The mixer datasheet states that the LO power should be between ±6 dBm. The result can be seen in Figure 5.7.

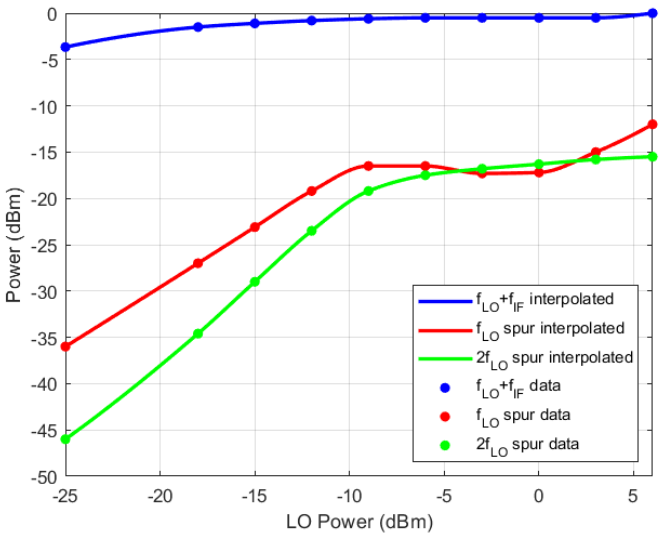
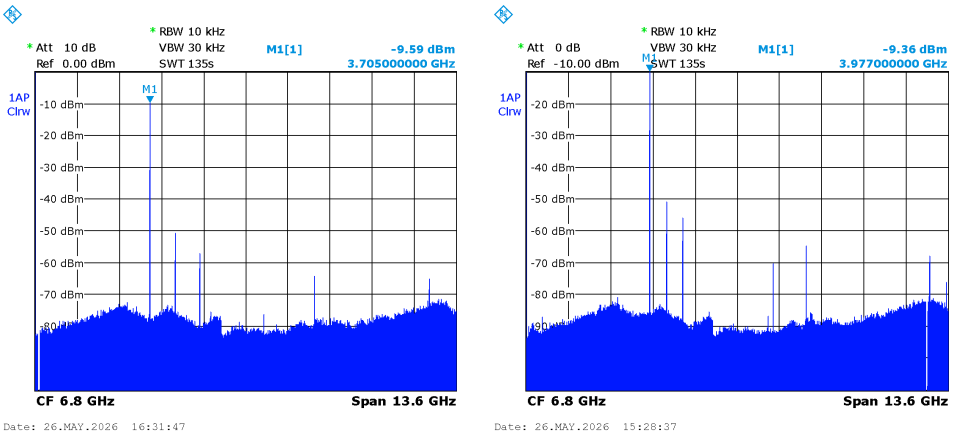


Figure 5.7: LO input power against desired tone power and LO harmonic power.

From the graph it can be seen that the mixer allows for a very wide range of LO powers without significant performance degradation. Below −12 dB the LO amplifier inside the mixer seems to no longer be able to keep the LO power constant and the spur power drops, resulting in increased intermodulation products and conversion loss.

Receive

The receive gain was measured similarly to the transmit gain. The R&S SMIQ 06B RF signal generator was connected to the LO port. The signal frequency set to 4510 MHz, 0 dBm. The R&S ZVL Network Analyzer was connected to the RX port. The Agilent E8257D RF signal generator was connected to the RF port and the frequency was set to 8210 MHz and 8490 MHz respectively, the spectrums are shown in Figure 5.8.



(a) Spectrum from front-end with 8210 MHz RF input and 4510 MHz LO.

(b) Spectrum from front-end with 8490 MHz RF input and 4510 MHz LO.

Figure 5.8: Down-conversion (RX) spectrum for 8210 MHz and 8490 MHz RF at −30 dBm.

For the receive path the spur performance is acceptable, arguably quite good. The −30 dBm signal is down-converted and amplified to −10 dBm. When accounting for the cable loss between the Agilent E8257D and front-end the gain is approximately 23 dB, and in theory it is 26 dB.

The input power was swept from −20 dBm to 5 dBm (0 dBm for the normal mode) at an RF of 8410 MHz. The total output power was calculated by summing the tones and the gain was calculated by taking the difference between the input power and total output power, for both the normal mode and the bypass mode. The result can be seen in figure 5.9

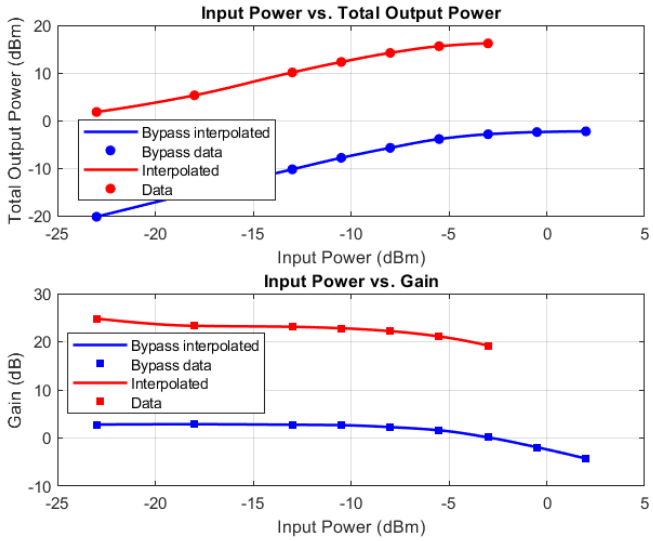
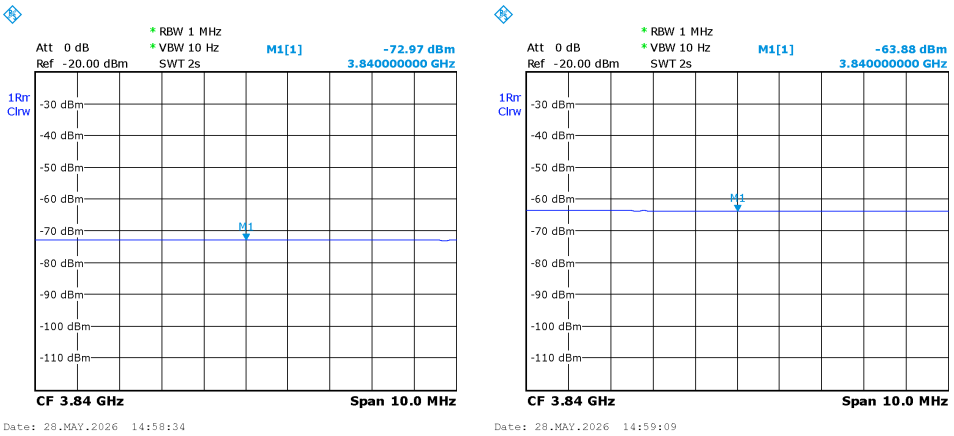


Figure 5.9: Down-conversion output power and gain against input power. for RF 8410 MHz, and LO 4510 MHz.

From figure 5.9 the receive path is working mostly as expected, although the gain is slightly reduced which can be partially by transmission line losses and cable losses. The above plot does not compensate for the cable loss of approximately 3 dB. The input compression point is approximately -2.5 dBm limited by the mixer or LNA. It is difficult to isolate exactly which part limits the compression point as both are similar. It is evident that the IF amplifier is not limiting the compression point as it is unchanged with the amplifier in bypass mode.

The noise figure was measured using the Y-factor method. The RF SMA connector was connected to the Agilent 346B Noise Source, the LO was connected to the SMIQ 06B signal generator set to 4510 MHz, 0 dBm. The RX SMA connector was connected to the Mini-Circuits ZJL-6G amplifier to lower the noise floor of the spectrum analyzer. The amplifier was connected to the R&S ZVL network analyzer set to spectrum analyzer mode.



(a) Noise-floor when front-end is enabled and noise source is disabled

(b) Noise-floor when front-end is enabled and noise source is enabled

Figure 5.10: Spectrums used for determining the front-end noise figure.

$$F = F_{DUT} + \frac{F_A - 1}{G_{DUT}} + \frac{F_{SA} - 1}{G_{DUT} G_A}$$
$$F_{DUT} = F - (\frac{F_A - 1}{G_{DUT}} + \frac{F_{SA} - 1}{G_{DUT} G_A})$$
$$F_{DUT} \approx F$$
$$NF_{DUT} = 10 \log F$$

The noise figure measurement is described by the equation above, where F_{DUT}, F_A and F_{SA} are the noise factors of the device under test (DUT) (the evaluation front-end), amplifier and spectrum analyzer respectively. As G_{DUT} is known to be at least 20 dB it is reasonable to assume that $(\frac{F_A - 1}{G_{DUT}} + \frac{F_{SA} - 1}{G_{DUT} G_A}) \approx 0$. Then using the equation

$$NF = 10 \log_{10} (\frac{10^{ENR/10}}{10^{Y/10} - 1})$$

where ENR is the excess noise ratio, as the evaluation front-end does not have an RF filter the mixer converts the noise from both the image frequency and the desired frequency, doubling the noise power therefore 3 dB has to be subtracted from the ENR table of the noise source, $ENR = 15.37 - 3 = 12.37$ dB. and Y is the measured noise floor increase as seen in Figure 5.10 ($Y = 72.97 - 63.88 = 9.09$ dBm) after enabling the noise source. This gives

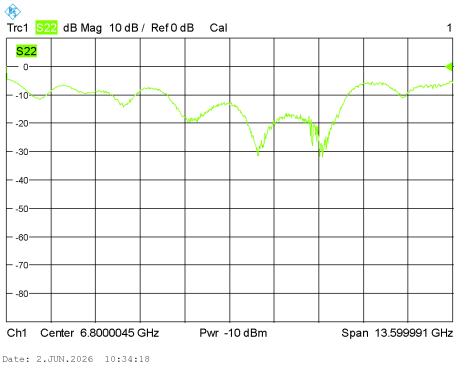
$$NF = 10 \log \frac{10^{12.37/10}}{10^{9.09/10} - 1} = 3.85 \text{ dB}$$

which is quite close to the theoretical value of 3.1 dB. The actual noise figure is without a doubt worse than the theoretical value but the increase can be partially

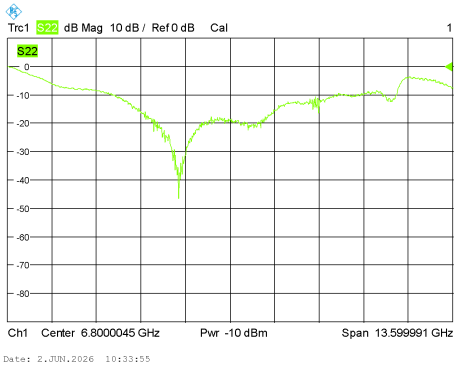
explained by the fact that there is some cable loss between the noise source and the front-end which directly affects the noise figure.

Return loss & cross-talk

The return loss of the SMA connector in combination with the front-end was measured with the front-end in both receive and transmit mode to validate the matching of the SMAs and the traces to the components. The return loss can be seen in figure 5.11.



(a) Combined SMA and front-end return loss receive

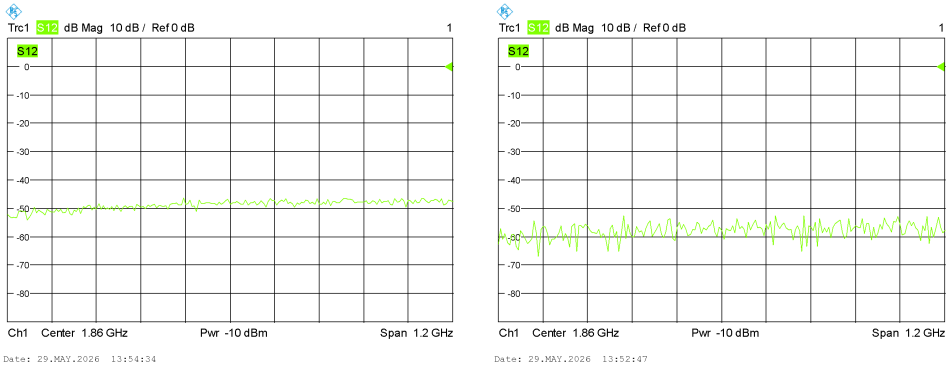


(b) Combined SMA and front-end return loss transmit

Figure 5.11: Combined SMA and front-end return loss for transmit and receive

As shown above the SMA and front-end has acceptable matching of >10 dB return loss for the entire frequency range of interest.

In addition to the evaluation front-end PCB it also featured a break-away test piece to study the effect of reducing the distance between the striplines on the full-scale board from 6 mm to 3 mm which would save some board space. The measurement was done using the test piece with the SMAs connected at opposite ends with the other end terminated.



(a) S12 with 3 mm spacing **(b)** S12 with 6 mm spacing

Figure 5.12: S12 for the 3 mm and 6 mm spaced striplines

From Figure 5.12 it can be seen that the 3 mm spaced striplines have around 10 dB more cross-talk, which is not insignificant.

5.3 Backplane & front-end modules

A design and layout of the 16 channel front-end was completed and a render of what the board would look like can be seen in Figure 5.14. A close-up of the front-end module can be seen in Figure 5.13. The board was not manufactured due to delays with the evaluation front-end and subsequent issues. During the course of the work and through the evaluation front end several opportunities to further improve the design in terms of cost, size, and overall efficiency were identified, as discussed in section 5.4. These improvements should be implemented before the 16-channel front-end is manufactured and integrated into the testbed. Nevertheless, the proposed architecture fulfills the requirements established at the beginning of the project, and the estimated performance is satisfactory.

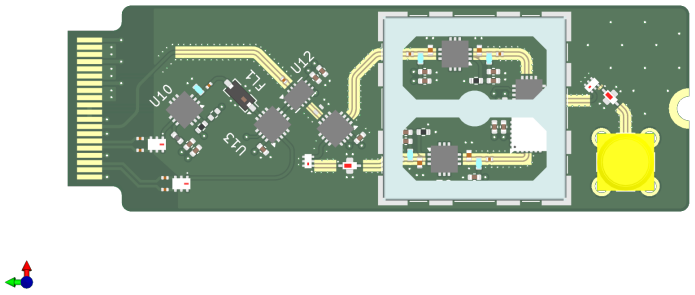


Figure 5.13: Render of a front-end module, containing 1 channel.

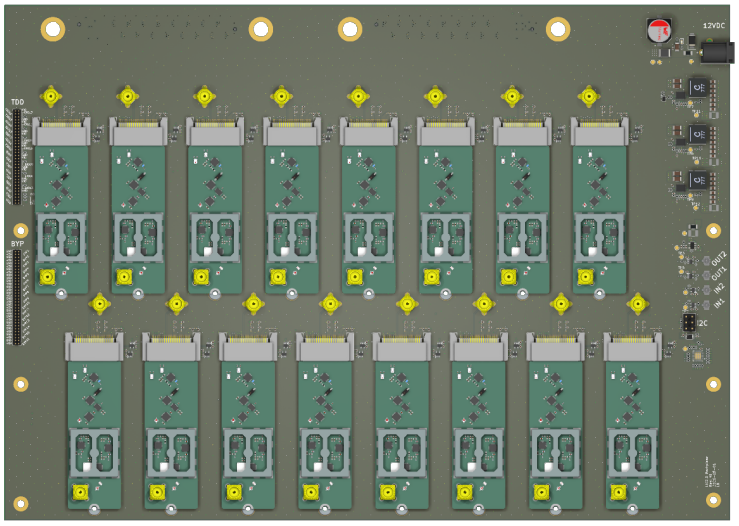


Figure 5.14: Render of the backplane populated with 16 front-end modules.

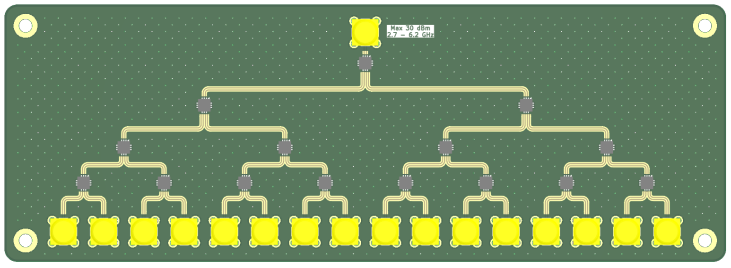


Figure 5.15: Render of the LO splitter board.

5.3.1 Schematic

The schematics for the front-end module and backplane are available in appendix C. and B. respectively.

5.4 Future Work

5.4.1 Distributed Digital Baseband

The future work regarding the distributed digital baseband should be evaluated with future research in mind. The next big step for the LuLIS project is system wide implementation of DLD. For the distributed digital baseband this entails

three things. Unsurprisingly, the first thing is to finish the implementation of DLD. Secondly, an effort to minimize BRAM use, as discussed in section 5.1.2, should be made. Thirdly, the tile reordering should be redesigned to decrease the T_{G_p} delay. The reason to first finalize the implementation of DLD is that it may change some minimum FIFO depths in the design. When redesigning the tile reordering, using URAMs could be considered here as well, as it would free up even more BRAM.

After this the direction is less clear. It depends on what research questions need to be answered. One feature that was not prioritized for this project is enabling the change of B_c, K, $\#ULD$ and $\#DLD$ during runtime. They are currently set by writing to certain AXI-Lite registers during initialization. Allowing for them to change every frame would allow for research into frame selection and the like.

5.4.2 Testbed LO distribution

For the testbed the LO signal needs to be distributed to all the LO splitter boards, due to the inefficiency of distributing a high frequency signal between all the distribution boards, a local phase locked loop (PLL) could and likely should be used to multiply a lower distribution frequency (such as the existing 10 MHz reference clock) at each distribution board as shown by LaCaille et. al (2019) [38]. The SINR increases very little beyond 32 antenna elements per PLL (antennas assumed to be on the same array). There seems to be no equivalent research in the distributed MIMO case, however the general ideas should be applicable. At 16 elements per PLL as would be the case here, the performance is acceptable.

5.4.3 Backplane & front-end module design

An architecture where the channels are individual modules is very flexible and continued development could integrate the LO splitters and/or LO synthesis onto the backplane, if the LO is to be fed externally from vertical SMAs perhaps frequency doublers should be considered on the front-end modules to reduce the insertion loss. The modules could be designed to use edge-mount SMAs if the modules are used in a vertical configuration or designed similarly to traditional PCIe cards such as in Figure 5.16. Edge-mounted SMAs have superior return loss compared to the hybrid SMT/THT SMAs used in the current design [31]. The redesign could significantly reduce the size of the backplane which could make it possible to use an all FR4 stack-up (due to shorter traces) and thus save significant cost.

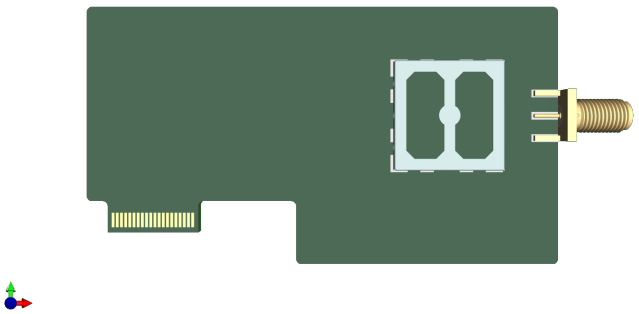


Figure 5.16: Render of how the modules could be designed for denser vertical integration and allow for an all FR4 backplane.

Instead of using SMT filters, the filters could be designed using trace geometry and simulated to obtain the desired filter characteristics. This allows for greater flexibility (with cut-off frequencies, possibly improving performance) and reduces dependence on the availability of many different filters as well as reducing cost. Additionally, simulating the via structures and SMA connector pad in an EM field solver such as Ansys HFSS could improve return and insertion loss.

Chapter 6
Conclusion

6.1 Digital Baseband

The digital baseband part of this thesis project focused on extending the LuLIS testbed toward support for a larger number of UEs by replacing parts of the existing digital baseband with a more capable FPGA implementation. The main computational tasks were identified as channel estimation, Gramian matrix calculation, and the matrix-vector multiplications required for uplink and downlink data processing. Based on the throughput and latency analysis, a systolic-array-based architecture was selected as a suitable compromise between resource utilization and processing latency.

A 16×16 systolic array was chosen to provide a peak throughput of 256 complex multiplications per clock cycle while avoiding the excessive DSP usage of a full 32×16 implementation. To improve utilization with lower coherence bandwidth, the array was divided into configurable lanes. The design was implemented in Chisel and verified using integrated simulation tools before integration into the existing LuLIS FPGA design.

The implemented distributed baseband was successfully integrated and evaluated on the LuLIS testbed. Measurements with four UEs showed that the design can perform the intended channel estimation and uplink processing in the real system. However, the full downlink data path was not completed within the scope of this thesis, mainly because downlink support was not yet available in the rest of the LuLIS system and when time began to run out it was the lowest priority.

Resource utilization showed memory usage is the main limitation. In particular, the high BRAM utilization caused timing-closure difficulties and limits the margin available for future additions. This indicates that further work should focus on reducing memory pressure, for example by moving suitable blocks to URAM. Overall, the work demonstrates that a dedicated systolic array accelerator is a viable approach for scaling the distributed digital baseband beyond the previous implementation, while also identifying the memory mapping and DLD as the most important areas for immediate continued development.

6.2 Front-end

The front-end portion of the thesis focused on designing a front-end to extend the frequency range of the LuLIS testbed to FR3 (upper midband). Due to the frequency requirements, a superheterodyne topology was selected as the basis for the design, which necessitated the development of a frequency plan. This was followed by an architectural comparison in which component placement, selection, and expected performance were evaluated. Based on this analysis, an evaluation front-end was designed and manufactured. Through the evaluation board a significant portion of the schematic and layout design was validated, with performance closely matching theoretical expectations, however the lack of filtering and the transmit path IF amplifier resulted in worse spur performance. Ultimately, a 16-channel front-end was designed with an improved architecture based on further research and the outcomes of the evaluation front-end. However the 16 channel front-end was not manufactured due to delays with the evaluation front-end and as opportunities to further improve the design in terms of cost, size, and overall efficiency were identified. These improvements should be implemented before the 16-channel front-end is manufactured and integrated into the testbed. Nevertheless, the proposed design should fulfill the requirements initially set.

Components suitable for the upper midband are expensive at the time of writing, as few devices are specifically designed for the 7.125–8.5 GHz frequency range. Consequently, many of the selected components are wideband devices, which are generally more expensive and offer compromised performance compared to application-specific alternatives. The cost was further increased by the high IF, since most low-cost mixers do not provide well-matched IF ports at frequencies above a few hundred MHz. The high IF, together with layout considerations, also influenced the substrate choice. In the future further cost reduction will be possible.

Chapter 7

References

[1] *IMT towards 2030 and beyond (IMT-2030)*. en-US. URL: https://www.itu.int:443/en/ITU-R/study-groups/rsg5/rwp5d/imt-2030/Pages/default.aspx (visited on 05/27/2026).

[2] ETSI. *LTE; Feasibility study for Further Advancements for E-UTRA (LTE-Advanced)*. Tech. rep. ETSI TR 136 912. European Telecommunications Standards Institute, Apr. 2022.

[3] Yingjie Xu et al. "Spatial Separation of Closely-Spaced Users in Measured Distributed Massive MIMO Channels". In: *2024 IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. ISSN: 1948-3252. Sept. 2024, pp. 381–385. DOI: 10.1109/SPAWC60668.2024.10694595. URL: https://ieeexplore.ieee.org/document/10694595 (visited on 05/27/2026).

[4] Emil Björnson et al. "Enabling 6G Performance in the Upper Mid-Band by Transitioning From Massive to Gigantic MIMO". In: *IEEE Open Journal of the Communications Society* 6 (2025). arXiv:2407.05630 [cs.IT], pp. 5450–5463. ISSN: 2644-125X. DOI: 10.1109/OJCOMS.2025.3576931. URL: http://arxiv.org/abs/2407.05630 (visited on 05/27/2026).

[5] Dumitra Iancu et al. *A Scalable 256-Antenna Distributed MIMO Testbed with Real-Time Fully Digital Beamforming*. Version 2. 2026. arXiv: 2605.02388 [eess.SP]. URL: https://arxiv.org/abs/2605.02388.

[6] Post och telestyrelsen (PTS). *Den svenska frekvensplanen*. 2026. URL: https://frekvensplanen.pts.se/ (visited on 04/08/2026).

[7] Gary Xu et al. *Upper Mid-Band Spectrum for 6G: Opportunities and Key Enablers*. Accessed: 2026-05-22. Samsung Research. 2025. URL: https://research.samsung.com/blog/Upper-Mid-Band-Spectrum-for-6G-Opportunities-and-Key-Enablers.

[8] Dipankar Shakya et al. "Comprehensive FR1(C) and FR3 Lower and Upper Mid-Band Propagation and Material Penetration Loss Measurements and Channel Models in Indoor Environment for 5G and 6G". In: *IEEE Open Journal of the Communications Society* 5 (2024), pp. 5192–5218. ISSN: 2644-125X. DOI: 10.1109/OJCOMS.2024.3431686. URL: https://ieeexplore.ieee.org/document/10605910 (visited on 05/30/2026).

[9] Mohammad Attari et al. "An Application Specific Vector Processor for Efficient Massive MIMO Processing". In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 69.9 (Sept. 2022), pp. 3804–3815. ISSN: 1558-0806. DOI: 10.1109/TCSI.2022.3182483. URL: https://ieeexplore.ieee.org/document/9801536/ (visited on 02/04/2026).

[10] Guiqiang Peng et al. "A 1.58 Gbps/W 0.40 Gbps/mm2 ASIC Implementation of MMSE Detection for 128\times 8 64 -QAM Massive MIMO in 65 nm CMOS". In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 65.5 (May 2018), pp. 1717–1730. ISSN: 1558-0806. DOI: 10.1109/TCSI.2017.2754282. URL: https://ieeexplore.ieee.org/document/8051120/ (visited on 02/04/2026).

[11] Hemanth Prabhu et al. "A 1070 pJ/b 169 Mb/s Quad-core Digital Baseband SoC for Distributed and Cooperative Massive MIMO in 28 nm FD-SOI". In: *2021 Symposium on VLSI Circuits*. ISSN: 2158-5636. June 2021, pp. 1–2. DOI: 10.23919/VLSICircuits52068.2021.9492455. URL: https://ieeexplore.ieee.org/document/9492455/ (visited on 02/04/2026).

[12] Guiqiang Peng et al. "A 2.92-Gb/s/W and 0.43-Gb/s/MG Flexible and Scalable CGRA-Based Baseband Processor for Massive MIMO Detection". In: *IEEE Journal of Solid-State Circuits* 55.2 (Feb. 2020), pp. 505–519. ISSN: 1558-173X. DOI: 10.1109/JSSC.2019.2952839. URL: https://ieeexplore.ieee.org/document/8911207/ (visited on 02/04/2026).

[13] Zhou, Ziqi and Wang, Jingwei. *RF system for mmWave massive MIMO*. eng. Student Paper. 2018.

[14] *ADMX7103-EBZ Analog Devices | Mouser*. en-eu. URL: https://eu.mouser.com/ProductDetail/Analog-Devices/ADMX7103-EBZ?qs=Q%2F5vwRdkjWJyd%252B5LILaTyw%3D%3D (visited on 05/28/2026).

[15] NR Explained. *OFDM Fundamentals*. 2026. URL: https://www.nrexplained.com/basics/ofdm (visited on 05/08/2026).

[16] Fredrik Rusek et al. “Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays”. In: *IEEE Signal Processing Magazine* 30.1 (Jan. 2013), pp. 40–60. ISSN: 1558-0792. DOI: 10.1109/MSP.2011.2178495. URL: https://ieeexplore.ieee.org/document/6375940/ (visited on 02/04/2026).

[17] Emil Björnson, Jakob Hoydis, and Luca Sanguinetti. “Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency”. In: *Foundations and Trends in Signal Processing* 11.3–4 (2017), pp. 154–655. DOI: 10.1561/2000000093. URL: https://github.com/emilbjornson/massivemimobook.

[18] Lina Tinnerberg et al. “Spectrum Efficiency and Processing Latency Trade-offs in Panel-Based LIS”. In: *2025 59th Asilomar Conference on Signals, Systems, and Computers*. ISSN: 2576-2303. Oct. 2025, pp. 519–524. DOI: 10.1109/IEEECONF67917.2025.11443936. URL: https://ieeexplore.ieee.org/abstract/document/11443936 (visited on 06/11/2026).

[19] *Zynq UltraScale+ RFSoC Data Sheet: DC and AC Switching Characteristics*. DS926. v1.14. AMD. Dec. 2024.

[20] Lars Sundström, Göran Jönsson, and Henrik Börjeson. *Radio Electronics*. English. Electrical and Information Technology, Lund University, Jan. 2014.

[21] Analog Devices. *Introduction to Common Printed Circuit Transmission Lines*. 2003. URL: https://www.analog.com/en/resources/technical-articles/introduction-to-common-printed-circuit-transmission-lines.html (visited on 04/08/2026).

[22] John Coonrod. *Comparing Microstrip and Grounded Coplanar Waveguides*. Accessed: 2026-05-20. Rogers Corporation, 2020. URL: https://www.rogerscorp.com/blog/2020/comparing-microstrip-and-grounded-coplanar-waveguides.

[23] Kung. “Why systolic architectures?” In: *Computer* 15.1 (Jan. 1982), pp. 37–46. ISSN: 1558-0814. DOI: 10.1109/MC.1982.1653825. URL: https://ieeexplore.ieee.org/document/1653825/ (visited on 02/04/2026).

[24] C. R. Wan and D. J. Evans. “Nineteen ways of systolic matrix multiplication”. en. In: *International Journal of Computer Mathematics* 68.1-2 (Jan. 1998), pp. 39–69. ISSN: 0020-7160, 1029-0265. DOI: 10.1080/00207169808804678. URL: http://www.tandfonline.com/doi/abs/10.1080/00207169808804678 (visited on 02/04/2026).

[25] “UltraScale Architecture DSP Slice User Guide”. en. In: (2021).

[26] *LTC5553 Datasheet*. Analog Devices. 2017.

[27] Ferenc Marki and Christopher Marki. "Mixer basics primer". In: *Marki Microwave* 116 (2010).

[28] Steve Arar. *Understanding Intermodulation Distortion and the Third-Order Intercept Point in RF Systems*. All About Circuits, May 2025. URL: https://www.allaboutcircuits.com/technical-articles/understanding-intermodulation-distortion-and-the-third-order-intercept-point-in-rf-systems/ (visited on 05/09/2026).

[29] *PMA3-14LV+ Datasheet*. Rev. B. Mini-Circuits. Aug. 2025.

[30] *TSS-53LNB+ Datasheet*. Rev. C. Mini-Circuits. Aug. 2025.

[31] *Coaxial PCB Connector PCB-Transmission Line Design Guide*. ANE012e. Wurth Elektronik. 2024.

[32] *CMD315C4 Datasheet*. Rev. B. Qorvo. June 2024.

[33] Analog Devices. *LT8625S*. Version Rev. C. 2025. URL: https://www.analog.com/media/en/technical-documentation/data-sheets/lt8625s.pdf.

[34] Cadence. *An Introduction to Ceramic Capacitor Derating Voltage*. 2022. URL: https://resources.pcb.cadence.com/blog/2022-an-introduction-to-ceramic-capacitor-derating-voltage (visited on 04/23/2026).

[35] John Coonrod et al. "High-Performance PCB Materials for Next-Generation Electronics: Packages, Driving Signal Integrity, Power Handling, and Reliability". In: *Proceedings of the IMAPS 2025 Symposium*. Sept. 2025.

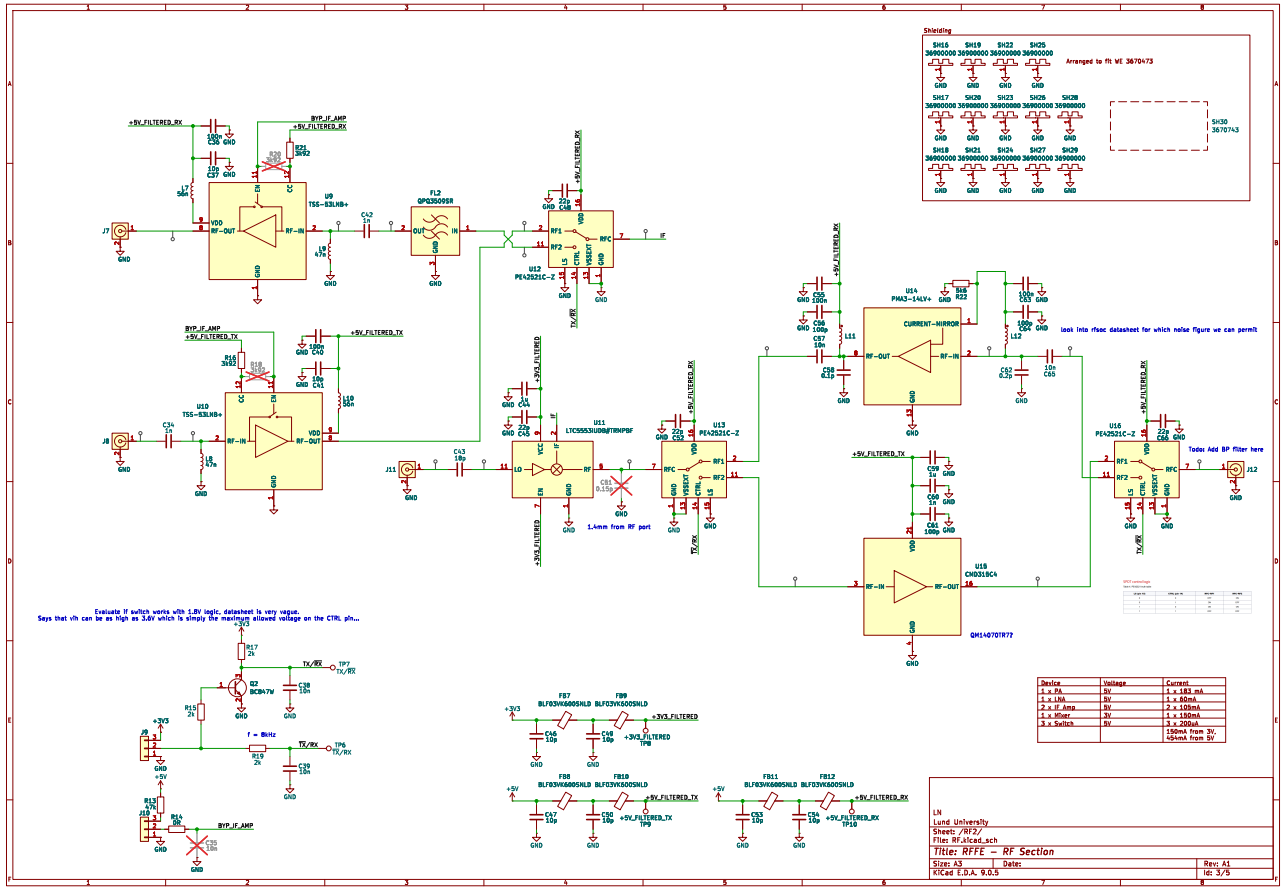
[36] Zida Li et al. "The Via Impedance Simulation and Control in PCB Manufacture for 112Gb/s". In: *2024 19th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT)*. 2024, pp. 215–218. DOI: 10.1109/IMPACT63555.2024.10818967.

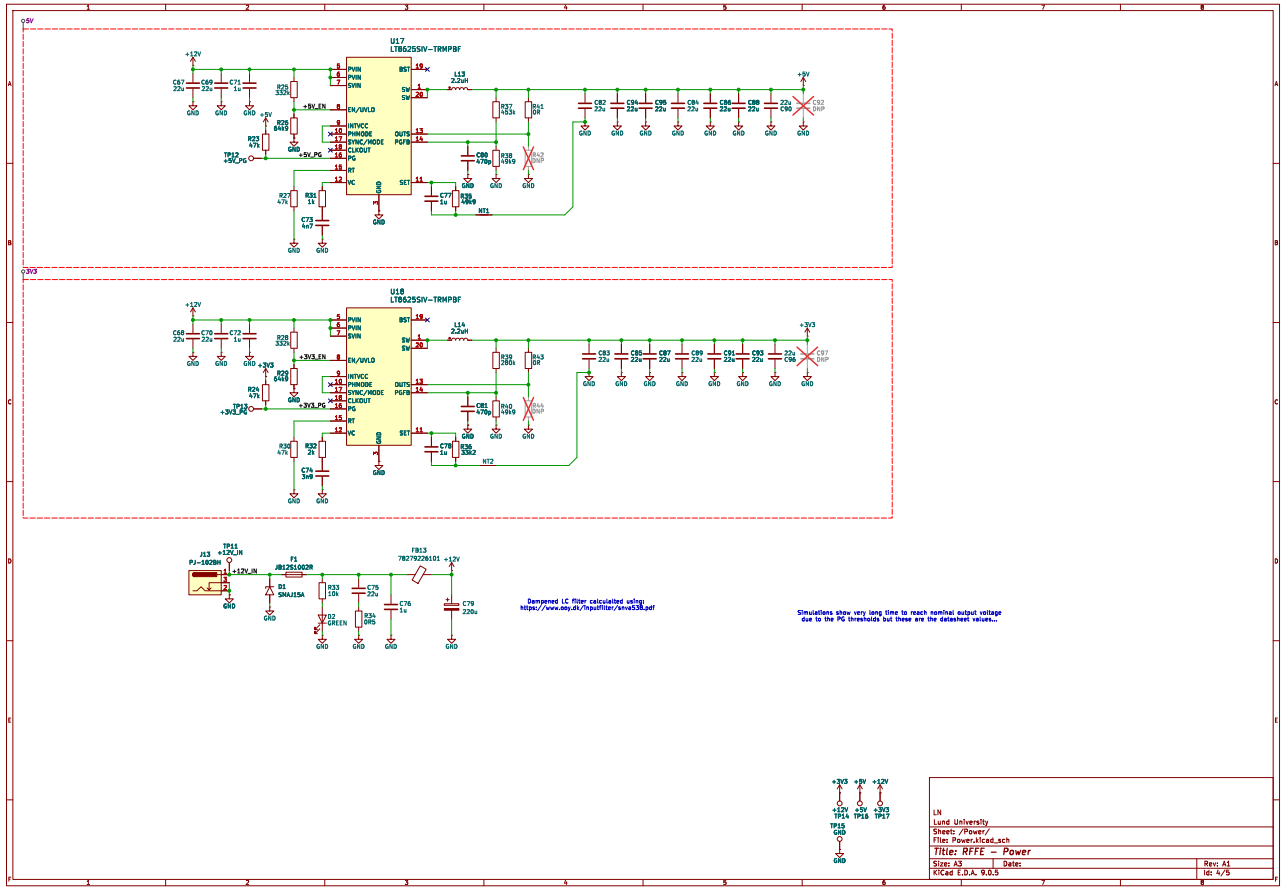
[37] *Samtec Array Connectors for Mixed-signal multi-channel RFSoCs up to 8 GHz Whitepaper*. SAMTEC. 2023.

[38] Greg LaCaille et al. *Optimizing the LO Distribution Architecture of mm-Wave Massive MIMO Receivers*. 2019. arXiv: 1911.01339 [eess.SP]. URL: https://arxiv.org/abs/1911.01339.

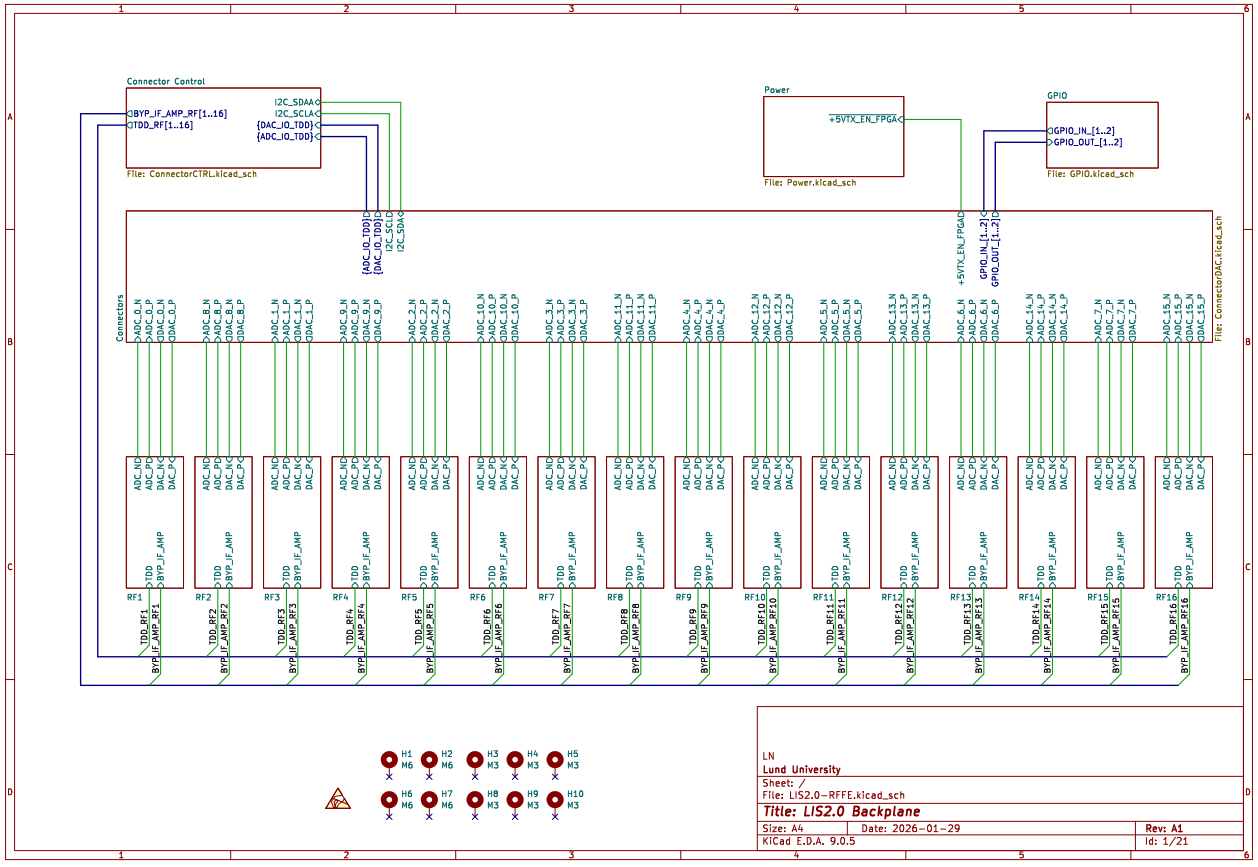
Appendix

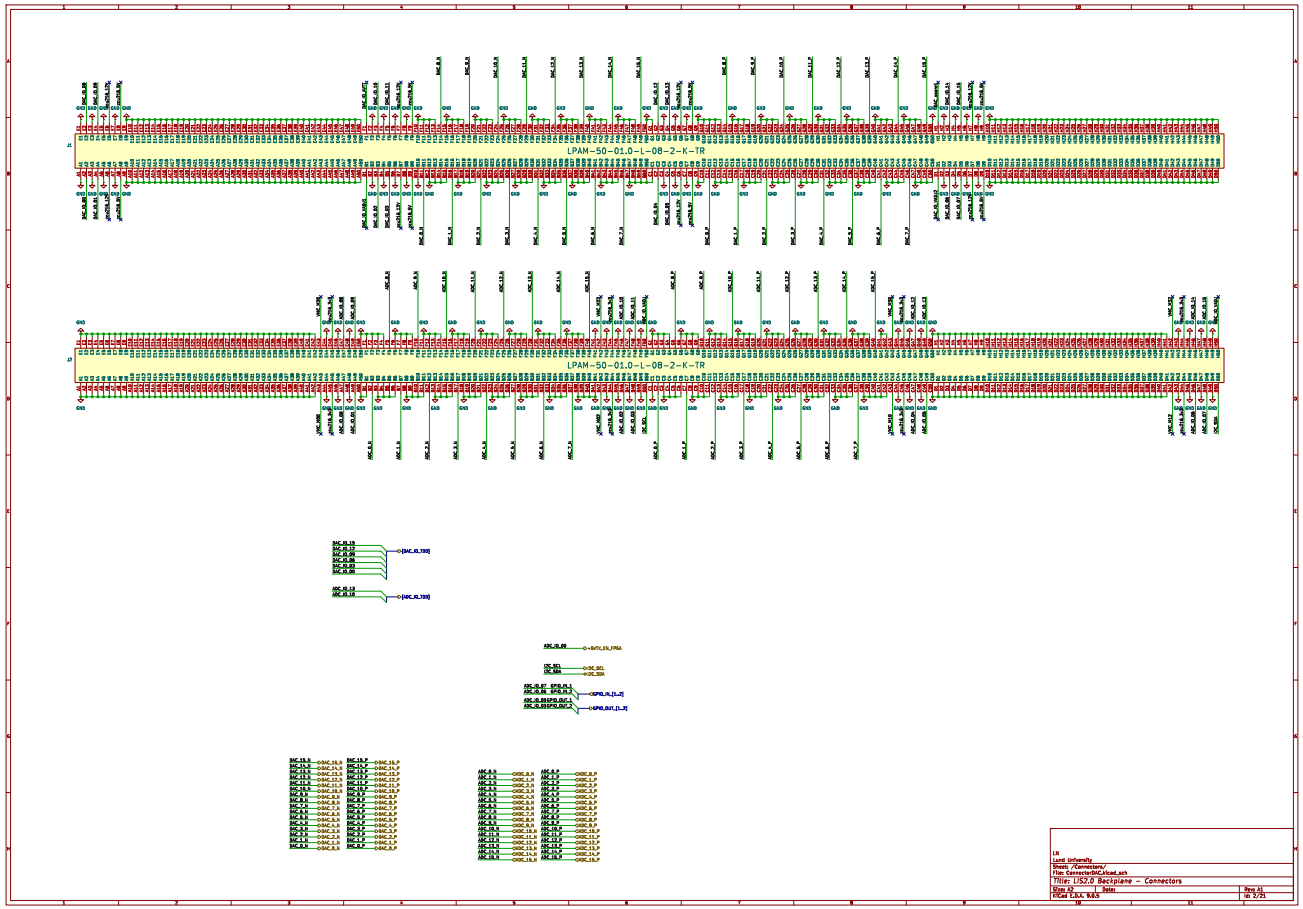
A. Evaluation Front-end Schematic


Shielding
Arranged to fit WE 3670473
SH30 3670743
U9 TSS-83LNB+
FL2 QPQ350095R
U12 PE42521C-Z
U8 TSS-83LNB+
U11 LTC5553IUD#TRMPBF
U13 PE42521C-Z
U14 PMA3-14LV+
U16 PE42521C-Z
U18 CHX31BC4
QM14070TR7Z
look into rfsoc datasheet for which noise figure we can permit
Todos: Add BP filter here
1.4mm from RF port
Evaluate if switch works with 1.8V logic, datasheet is very vague.
Says that vih can be as high as 3.6V which is simply the maximum allowed voltage on the CTRL pin...
f = 8kHz
Q2 BCB847W
FB7 BLF03VK6005HLD
FB9 BLF03VK6005HLD
FB8 BLF03VK6005HLD
FB10 BLF03VK6005HLD
FB11 BLF03VK6005HLD
FB12 BLF03VK6005HLD
Device | Voltage | Current
1 x PA | 5V | 1 x 183 mA
1 x LNA | 5V | 1 x 60mA
2 x IF Amp | 5V | 2 x 108mA
1 x Mixer | 3V | 1 x 150mA
3 x Switch | 5V | 3 x 200uA
| | 150mA from 3V, 40mA from 5V
LN
Lund University
Sheets: /RF2/
Files: RF.kicad_sch
Title: RFFE - RF Section
Size: A3 | Date: | Rev: A1
KiCad E.D.A. 9.0.5 | Id: 3/5

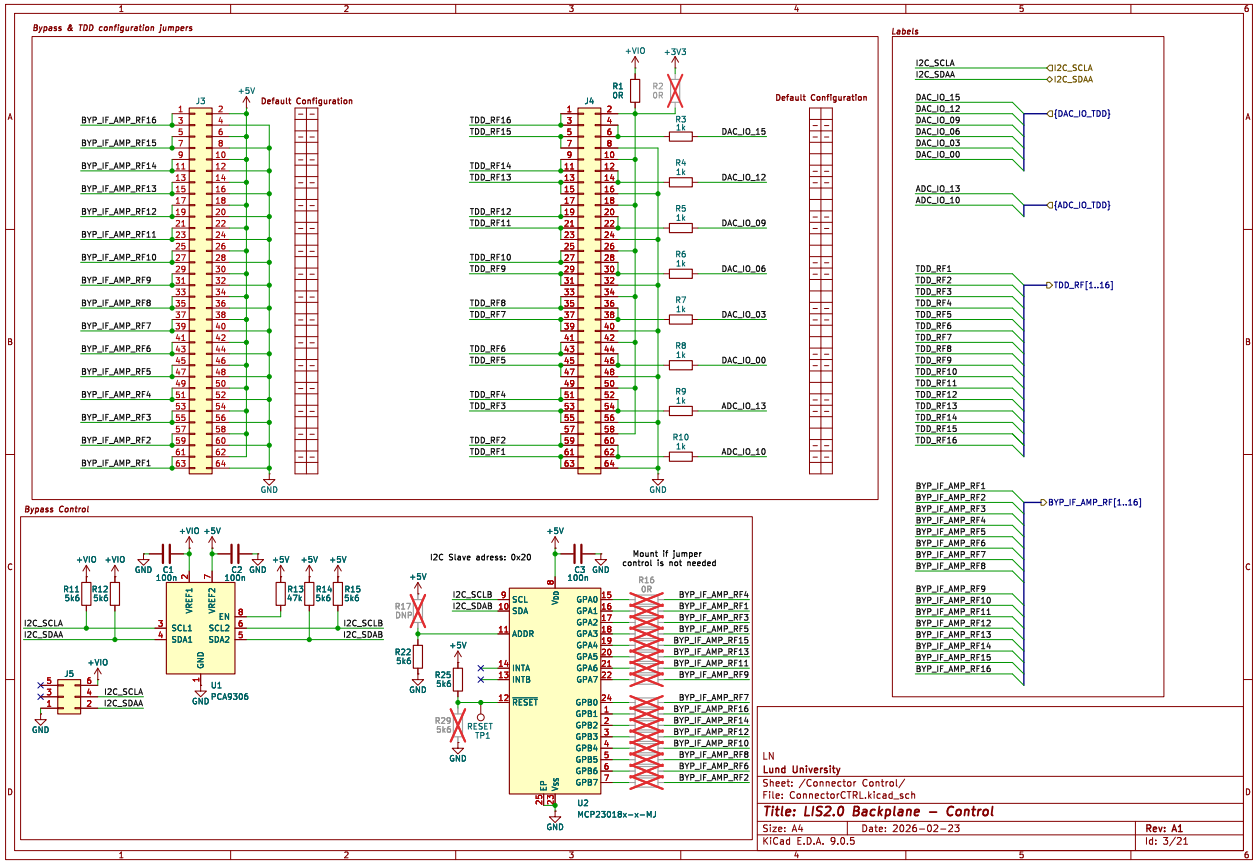

5V
U17
LTB625SIV-TRMPBF
PVIN
PVIN
SVIN
EN/UVLO
INTVCC
PHMODE
SYNC/MODE
CLKOUT
PG
RT
VC
BST
SW
SW
OUTS
PGFB
SET
GND
+12V
+5V_EN
+5V
+5V_PG
TP12
+5V_PG
L13
2.2uH
+9V
NT1
3V3
U18
LTB625SIV-TRMPBF
+3V3_EN
+3V3
+3V3_PG
TP13
+3V3_PG
L14
2.2uH
+3V3
NT2
J13
PJ-102BH
TP11
+12V_IN
F1
MB1261002R
D1
SMAJ15A
D2
GREEN
FB13
78279226101
+12V
Dampened LC filter calculated using:
https://www.ecy.st/inputfilter/snva538.pdf
Simulations show very long time to reach nominal output voltage
due to the PG thresholds but these are the datasheet values...
TP14
TP16
TP17
TP15
LN
Lund University
Sheets: /Power/
Files: Power.kicad_sch
Title: RFFE - Power
Size: A3
Date:
Rev: A1
KiCad E.D.A. 9.0.5
Id: 4/5

B. Backplane Schematic

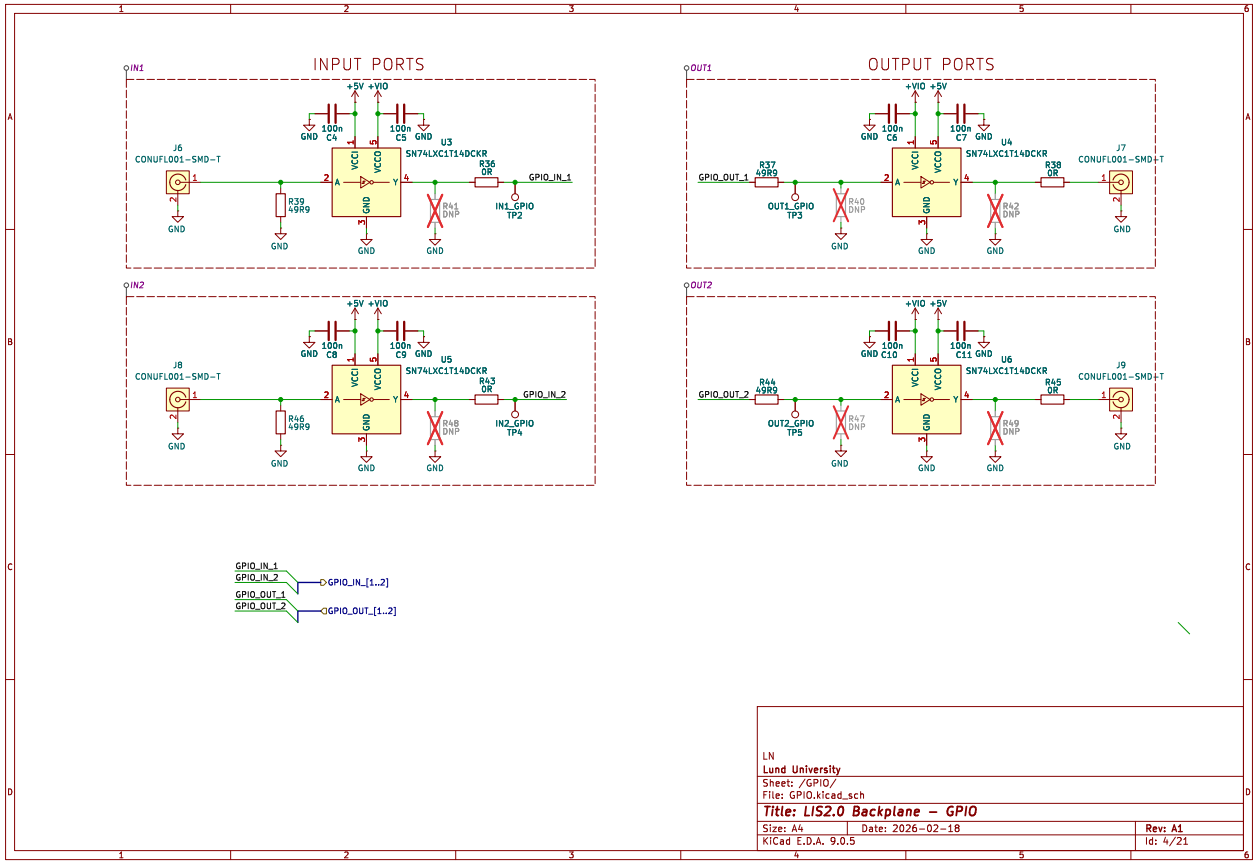

Connector Control
BYP_IF_AMP_RF[1..16]
TDD_RF[1..16]
I2C_SDAA
I2C_SCLA
{DAC_IO_TDD}
{ADC_IO_TDD}
File: ConnectorCTRL.kicad_sch
Power
+5VTX_EN_FPGA
File: Power.kicad_sch
GPIO
GPIO_IN_[1..2]
GPIO_OUT_[1..2]
File: GPIO.kicad_sch
Connectors
File: ConnectorDAC.kicad_sch
RF1
RF2
RF3
RF4
RF5
RF6
RF7
RF8
RF9
RF10
RF11
RF12
RF13
RF14
RF15
RF16
H1 M6
H2 M6
H3 M3
H4 M3
H5 M3
H6 M6
H7 M6
H8 M3
H9 M3
H10 M3
LN
Lund University
Sheet: /
File: LIS2.0-RFFE.kicad_sch
Title: LIS2.0 Backplane
Size: A4
Date: 2026-01-29
Rev: A1
KiCad E.D.A. 9.0.5
Id: 1/21



LN
Level Unfamiliy
Sheets /Connectors/
File: ConnectorBAC.kicad_sch
Title: LIS2.0 Backplane — Connectors
Size: A2 | Date: | Rev: A1
KiCad E.D.A. 9.0.5 | Id: 2/21

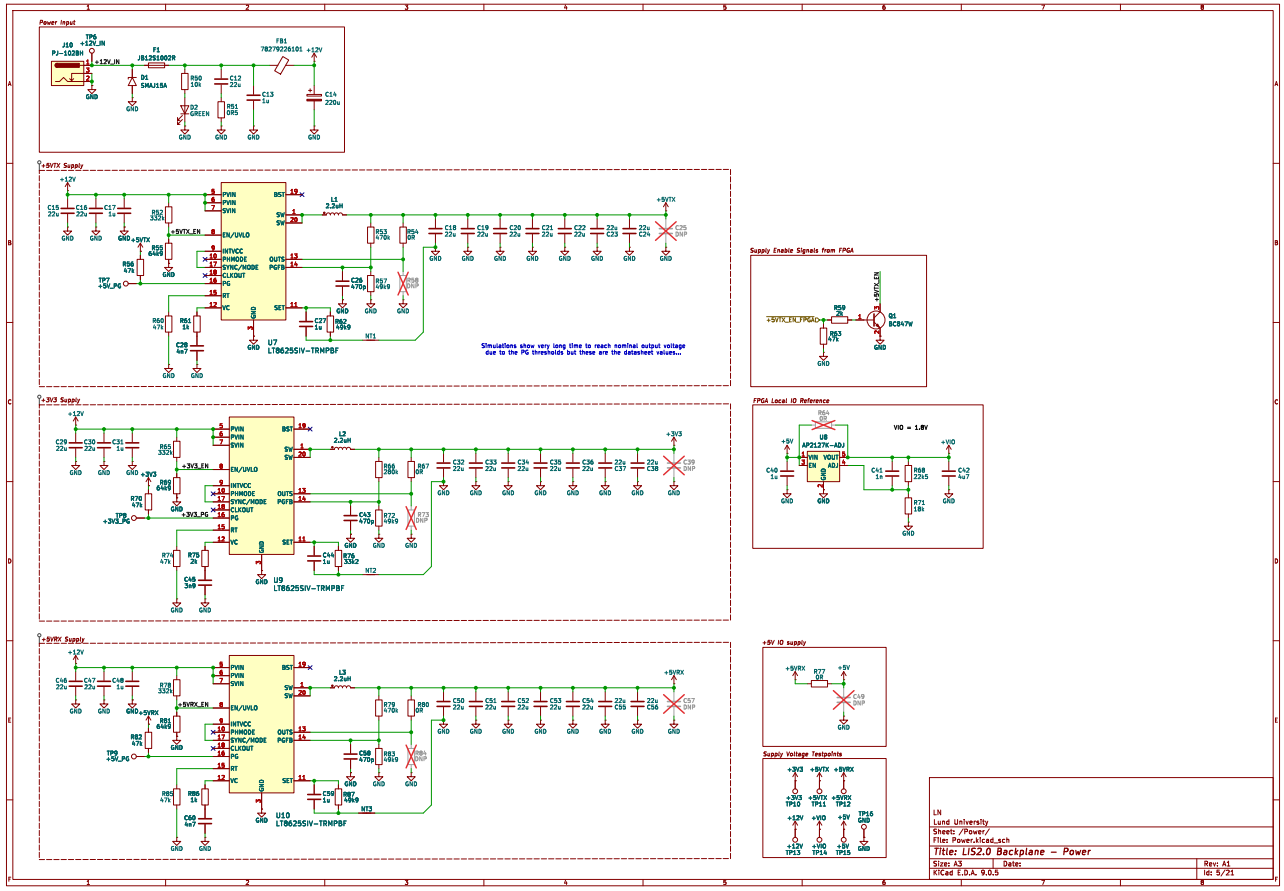


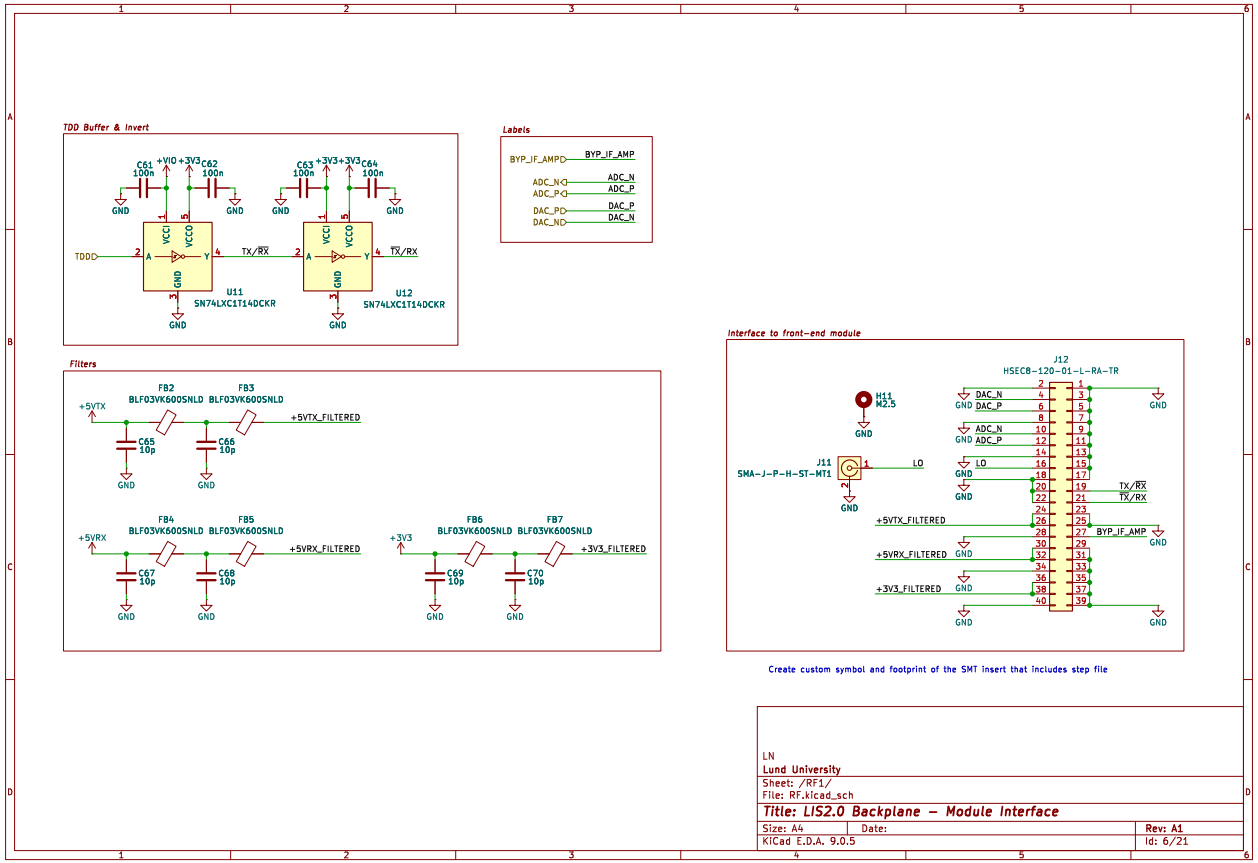
LN
Lund University
Sheet: /Connector Control/
File: ConnectorCTRL.kicad_sch
Title: LIS2.0 Backplane – Control
Size: A4 | Date: 2026-02-23 | Rev: A1
KiCad E.D.A. 9.0.5 | Id: 3/21



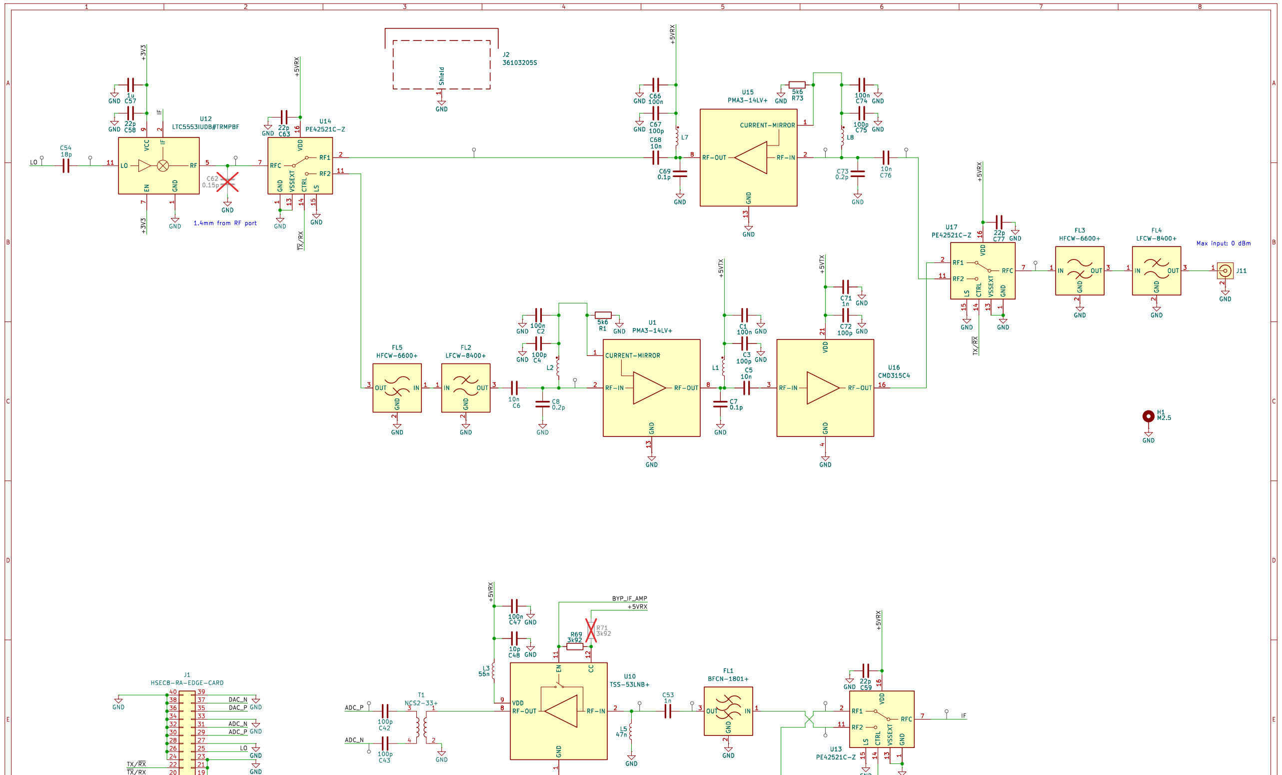
LN	
Lund University	
Sheet: /GPIO/	
File: GPIO.kicad_sch	
Title: LIS2.0 Backplane – GPIO	
Size: A4	Date: 2026-02-18
KiCad E.D.A. 9.0.5	Id: 4/21

Rev: A1


Power Input
J10 PJ-102BH
TP6 +12V_IN
F1 JBS251002R
FB1 78279226101
D1 SMAJ15A
D2 GREEN
R50 10k
R51 0R5
C12 22u
C13 1u
C14 220u
+5VTX Supply
U7 LT8625SIV-TRMPBF
L1 2.2uH
R52 332k
R53 470k
R54 0R
R55 649k
R56 47k
R57 4.99k
R58 DNP
R59 DNP
R60 47k
R61 1k
R62 4.99k
C15 22u
C16 22u
C17 1u
C18 22u
C19 22u
C20 22u
C21 22u
C22 22u
C23 22u
C24 22u
C25 DNP
C26 470p
C27 1u
C28 4n7
TP7 +5V_PG
+5VTX_EN
+5VTX
NT1
Simulations show very long time to reach nominal output voltage due to the PG thresholds but these are the datasheet values...
Supply Enable Signals from FPGA
+5VTX_EN_FPGA
R59 2k
R63 47k
Q1 BC847W
+3V3 Supply
U9 LT8625SIV-TRMPBF
L2 2.2uH
R65 332k
R66 380k
R67 0R
R68 649k
R69 649k
R70 47k
R72 4.99k
R73 DNP
R74 47k
R75 2k
R76 33k2
R77 3k48
C29 22u
C30 22u
C31 1u
C32 22u
C33 22u
C34 22u
C35 22u
C36 22u
C37 22u
C38 22u
C39 DNP
C43 470p
C44 1u
C45 3n9
TP8 +3V3_PG
+3V3_EN
+3V3
NT2
FPGA Local IO Reference
U8 AP2127K-ADJ
R64 0R
R88 22k5
R71 10k
C40 1u
C41 1u
C42 4u7
VIO = 1.8V
+5VRX Supply
U10 LT8625SIV-TRMPBF
L3 2.2uH
R78 332k
R79 470k
R80 0R
R81 649k
R82 47k
R83 4.99k
R84 DNP
R85 47k
R86 1k
R87 4.99k
C46 22u
C47 22u
C48 1u
C50 22u
C51 22u
C52 22u
C53 22u
C54 22u
C55 22u
C56 22u
C57 DNP
C58 470p
C59 1u
C60 4n7
TP9 +5V_PG
+5VRX_EN
+5VRX
NT3
+5V IO supply
R77 0R
C49 DNP
Supply Voltage Testpoints
TP10 +3V3
TP11 +5VTX
TP12 +5VRX
TP13 +12V
TP14 +VIO
TP15 +5V
TP16 GND
LN
Lund University
Sheets: /Power/
Files: Power.kicad_sch
Title: LIS2.0 Backplane – Power
Size: A3
Date:
Rev: A1
KiCad E.D.A. 9.0.5
Id: 5/21


TDD Buffer & Invert
C61 100n
+VIO
+3V3
C62 100n
C63 100n
+3V3
+3V3
C64 100n
GND
TDDD
VCCI
VCCO
GND
A
Y
TX/RX
U11
SN74LXC1T14DCKR
U12
SN74LXC1T14DCKR
Labels
BYP_IF_AMP
ADC_N
ADC_P
DAC_P
DAC_N
Filters
+5VTX
FB2
BLF03VK600SNLD
FB3
BLF03VK600SNLD
+5VTX_FILTERED
C65 10p
C66 10p
+5VRX
FB4
BLF03VK600SNLD
FB5
BLF03VK600SNLD
+5VRX_FILTERED
C67 10p
C68 10p
+3V3
FB6
BLF03VK600SNLD
FB7
BLF03VK600SNLD
+3V3_FILTERED
C69 10p
C70 10p
Interface to front-end module
J12
HSEC8-120-01-L-RA-TR
H11
M2.5
J11
SMA-J-P-H-ST-MT1
LO
DAC_N
DAC_P
ADC_N
ADC_P
TX/RX
BYP_IF_AMP
+5VTX_FILTERED
+5VRX_FILTERED
+3V3_FILTERED
Create custom symbol and footprint of the SMT insert that includes step file
LN
Lund University
Sheet: /RF1/
File: RF.kicad_sch
Title: LIS2.0 Backplane – Module Interface
Size: A4
Date:
Rev: A1
KiCad E.D.A. 9.0.5
Id: 6/21

C. Front-end Module Schematic



LN
Lund University
Sheet: /
File: LIS2.0-RFFE-EDGE-CARD.kicad_sch

Title: LIS2.0 – Front-end module

Size: A3 Date: 2026-05-03 Rev: A1
KiCad E.D.A. 9.0.5 Id: 1/1