Master Thesis

High-speed Data Acquisition and Processing for Transcranial Functional Ultrasound Brain Imaging

A.J. de Jong

TUDelft

High-speed Data Acquisition and Processing for Transcranial Functional Ultrasound Brain Imaging

Master Thesis

by

A. J. de Jong

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on June 29th at 10:00 AM

TUDelft cube
understanding the brain
with ultrasound

Erasmus MC
University Medical Center Rotterdam

Abstract

Functional ultrasound is by now a well-established technique in the neuroscience community to measure brain activity. However, the transcranial application of functional ultrasound on humans, with the exception of the acoustic windows in the skull, remains a huge challenge because of the properties of the cranial bone the acoustic waves will reflect, refract, attenuate or will cause aberration. Therefore, the signal-to-noise ratio (SNR) of the incoming signal is too low for transcranial functional ultrasound (TCfUS). This thesis tries to overcome the SNR problem of TCfUS with the design of a 64-channel ultrasound acquisition system that is focused on achieving the highest SNR possible. This is achieved by placing the analog front-end (AFE) chips as near to the transducer elements as possible and by oversampling the incoming signal with a factor of 25, a theoretical increase of 15 dB of the signal-to-quantization-noise ratio over the current state-of-the-art is estimated. The proposed design is a receive-only system where the transmission of the ultrasound pulses is carried out by a separate ultrasound system. The design splits the acquisition system into a front-end and a back-end subsystem, where the front-end system is implemented using four AFE58JD48 analog front-end chips from Texas Instruments. From here the data samples are transported over fiber optics to the back-end subsystem, which consists of a VCK190 FPGA board from Xilinx, where the samples are processed and/or transported to a workstation for storage. Because the system organization differentiates from more conventional research ultrasound systems, a trade-off is introduced between processing and throughput, resulting in three processing configurations for the FPGA with each a different focus: raw RF data sampling, real-time processing, and hardware processing. Due to time and resource constraints, no measurements and results are available on the SNR and decimation. However, a theoretical exploration has been done on the expandability of the number of channels which was found to be 192 channels for the VCK190.

Preface

I want to thank all of the people involved in making this thesis a reality. The journey toward this moment did not always go as planned, however, even with these hurdles down the road, I am glad to be at the finish line. This would of course not be possible if it wouldn't be for a lot of people in my life who made my career as a computer/electrical engineer possible.

First of all, my supervisor Stephan Wong, since without him I would have never ended up with this project in which I could put a lot of creativity and broadened my horizon into the medical world substantially. I would also like to thank my supervisors at the Erasmus MC, Christos Strydis and Pieter Kruizinga. Although at times it could be quite hectic, you were both there for me when I needed it and helped me find a way towards the final result.

I would also like to thank all other members of the lab at Erasmus MC, your daily presence and friendship helped a lot during Covid times. You were always available for a coffee or some small talk. Additionally, thank you Lennart for using your work for my cover page.

I would, above all, like to thank my parents, since without you I would not be standing at the point in life where I am now. You always supported me in my technical interests, starting from a young age with Legos, and later on, financially during my student career. Even though this study path could be quite hard at certain times, you were always there for me for mental support and a good hug, I love you both.

The Electrotechnische Vereeniging (ETV) has also played a big role in my years at the TU Delft and from my first year onwards you all feel like a second family. I feel very grateful for the opportunity for my board year and the activities and experiences I could enjoy during all of these years studying in Delft.

Of course, I would like to thank my friends, I feel like have made friendships for life during my years at the TU Delft and I couldn't be more thankful for that. For all the highs: drinking coffee or beers, study sessions, borrels, parties, but also always being there when I needed it most. You all helped me more than you could possibly know.

Finally, I would like to thank my good old trusty, right side of the Bell curve, MacBook Pro mid 2012 15", which supported me throughout all of my work during my study career at the TU Delft. You have never let me down.

A. J. de Jong
Rotterdam, June 2022

Contents

List of abbreviations

ADC Analog-to-digital Converter

AFE Analog Front-end

ASIC Application-specific integrated circuit

AWG Arbitrary Waveform Generator

BOLD Blood Oxygenation Level Dependent (imaging)

B-mode Brightness mode

CBV Cerebral Blood Volume

CIC cascaded integrator–comb (filter)

CLB Configurable Logic Block

CPU Central Processing Unit

CT Computed Tomography

CW Continuous Wave

DDR Double Data Rate

DMA Direct Memory Access

DSP Digital Signal Processing

ECC Error Correcting Code

FFT Fast Fourier Transform

fMRI functional Magnetic Resonance Imaging

FMC FPGA Mezzanine Card

FPGA Field Programmable Gate Array

fps frames per second

fUS functional ultrasound

Gbps Gigabit per second = 10^9 bits/s

GBps Gigabyte per second = 10^9 Byte/s

GPIO General-purpose input/output

GPU Graphical Processing Unit

HDL Hardware Description Language

HLS High-Level Synthesis

HSDC High speed data converter

IC Integrated Circuit

IP Intellectual Property

LUT Lookup Table

LVDS Low Voltage Differential Signaling

MPO Multi-fiber push on

MRI Magnetic Resonance Imaging

MSPS Megasamples per second = 10^6 samples/s

NoC Network on Chip

NVC Neurovascular Coupling

PCB Printed Circuit Board

PDI Power Doppler Image

PRF Pulse Repetition Frequency

PW Pulsed Wave

RAID Redundant Array of Independent Disks

RBC Red Blood Cell

SDRAM Synchronous Dynamic Random-Access Memory

SIMD Single Instruction Multiple Data

SNR Signal-to-Noise Ratio

SSD Solid-State (storage) Drive

TCfUS Transcranial functional ultrasound

uC Microcontroller

ULM Ultrasound Localization Microscopy

VLIW Very Long Instruction Word

VNA Vector Network Analyzer

<div align="right">

1

</div>

Introduction

In this chapter, an introduction is given to a new tool in neuroscience that can determine the active parts of the brain utilizing transcranial functional ultrasound. First, a small introduction to the topic is presented in order to come to a research question. Thereafter, the goal and methodology are described and the structure of the rest of the thesis is laid out.

1.1. Overview

Brain disorders and diseases are prevalent in today's society. In the Netherlands alone, more than one-fourth of the total healthcare budget is spend on brain disorders and diseases and one-fifth of the deaths in the Netherlands can be attributed to a brain disorder [80]. These disorders are also the leading cause of disability-adjusted life-years and the second leading cause of deaths globally [22]. It is therefore important to gain more insight into these brain diseases and disorders. Although a plethora of modalities are available for the imaging and inspection of the human body, the number of modalities that is applicable for the imaging of the brain is only a small subset thereof. For example, computed tomography (CT) and magnetic resonance imaging (MRI) are popular modalities for the detection of tumors and provide precision in the order of a millimeter. However, for a better understanding of our brain, we need to go to the core of the brain, i.e. the neurons and the interaction among all 85 billion of them [32]. The firing of a neuron in the human brain is a rather slow process (0.5-60 Hz) [7], however, to find a modality that 1) tracks this process over time, 2) for all individual neurons in the brain, and 3) to understand the interaction among the neurons, is a big challenge in the neuroscience community that is not solved until now. Technologies like implantable microelectrode arrays do provide a small insight into the firing of neurons in a specific region of the brain, however, invasive surgeries are required in order to come to these results which is a large burden for a test subject [71]. Therefore, the neuroscientific community prefers a noninvasive modality that can help study the workings of the brain and the interactions occurring in the brain.

A non invasive method to find active regions of the brain is found by utilizing proxies for brain activity. In functional MRI (fMRI), the oxygenation of the blood flowing to and from specific parts of the brain can be tracked over time using a technique called blood oxygenation level-dependent (BOLD) imaging. Due to the neurovascular coupling principle [6] the volume and ratio of oxygen in the blood can be correlated to brain activity. The brain activity in parts of the brain is then tracked over time and is correlated to specific tasks that a subject is performing. However, this imaging method comes with several disadvantages in terms of size, immobility, and cost. Experiments in fMRI machines can also be influenced by the loud noises of the machine and the limited room that is available in the machine, therefore, experiments with free-roaming subjects are out of the question. The fMRI technique shines in its ability to spatially accurately image a human brain in the range of mm or even 500 microns depending on the strength of the magnetic field. However, one big disadvantage of fMRI is that it is limited in terms of temporal resolution (3-0.1 s) [25].

Another modality that uses the blood as a proxy to locate activity in specific parts of the brain is tran-
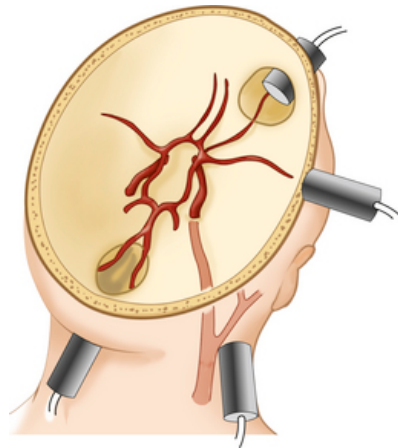
Figure 1.1: The locations of acoustic windows in the skull that allow for the functional ultrasound imaging of the blood vessels in the brain, which can be used as a proxy for measuring brain activity due to neurovascular coupling[19], from D'Andrea et al. [18]

scranial functional ultrasound (TCfUS). The inherent simplicity of the workings of ultrasound compared to that of fMRI and the size of the probe used in ultrasound are great advantages. With this imaging modality, it should be possible to do experiments with free-moving subjects at the fraction of the cost of fMRI. Functional ultrasound uses the Doppler effect and reflections of red blood cells (RBC) in order to measure the blood flow and blood volume in the brain. These metrics are then correlated to brain activity by using the same neurovascular coupling principle as used in fMRI [43]. TCfUS is proven to be more accurate in terms of spatial and temporal resolution in comparison to fMRI [44]. Additionally, TCfUS has proven to be a successful technique for whole-brain imaging in many neuroscientific studies using mice and rats [39, 43], as well as functional ultrasound during awake brain surgery [67]. The major hurdle in this technique at the moment, however, is the penetration of the skull bone with ultrasound waves. Therefore, TCfUS is limited by the utilization of the acoustic windows in the (human) skull [19, 20, 55]. The acoustic windows in the human skull are locations where the skull bone is thinner or nonexistent, the locations of these acoustic windows are depicted in Figure 1.1.

1.2. Motivation

In order to further advance neuroscientific research and provide tools in the clinic, it would therefore be logical to improve upon this TCfUS technique to provide *true* transcranial functional ultrasound in humans, where, true transcranial, means that the technique works at every arbitrary position on the skull instead of only at the acoustic windows. To overcome the biggest hurdle in TCfUS, the skull bone, it is necessary to focus on the receiving end of the system that receives the acoustic reflections from the RBCs. Due to the natural thickness and structure/material of the skull and the difference in impedance between the skull and the brain, a large portion of the waves will reflect, refract, attenuate, or will cause aberration, and therefore the waves cannot be focused correctly and the received signal contains relatively much noise and distortion [54]. Because of these effects caused by the skull bone, it is very difficult to distinguish the noise from the reflected ultrasound signal of the RBCs, that is needed for the imaging. Therefore, the signal-to-noise ratio (SNR) which captures the power of both the signal of interest and the noise, is one of the most important metrics and key to achieving true TCfUS. Subsequently, one of the big questions therefore within the ultrasound community is:

*Can we increase the signal-to-noise ratio of the ultrasound signal to allow for **true** transcranial functional ultrasound?*

This thesis focuses on tackling this exact question and will provide a hardware design of a research ultrasound acquisition system prototype that is focused on achieving the highest possible signal-to-noise ratio, in order to verify the feasibility of applying true transcranial functional ultrasound at any arbitrary position of the (human) skull. The design that will be presented will include several techniques to potentially increase the SNR, which are: 1) placing the analog front-end (AFE) chips as close to the probe

as possible, 2) oversampling the signal of interest with a factor of 25, 3) increasing the resolution of the AFE with 2 bits over the current state-of-the-art. On top of that, it is beneficial to have as many channels available since such a system can also allow for 3D imaging which can provide more insight into the vascular structure than its 2D counterpart.

In a research ultrasound system, the user needs to have access to raw RF samples of an ultrasound acquisition to explore the most ideal parameters in terms of signal processing. This means that the processing is not predetermined; on the contrary, the processing of the signals must remain flexible throughout the system, while at the same time the processing is expected to be near-real-time (<0.5 s) for operator positioning and time-critical applications. The design of the system also has to incorporate some important unique features to make the system competitive against fMRI. More specifically, it is important that compared to an fMRI machine, the spatial and temporal resolution are equal or better, the system has to be portable and the total cost has to be lower than $200.000, which forces the use of off-the-shelf-components.

1.3. Research Question

As briefly mentioned above, the fundamental question we seek to answer, starting with this thesis is: *can we achieve high-resolution transcranial imaging by using off-the-shelf components but with the objective of maximizing the SNR of the acquired signal?* But as mentioned above, maximizing the SNR depends on a number of signal-processing and system-organization aspects. The challenge is to show that such an approach – that is, an approach focusing on acquiring the highest quality signal and more importantly, the processing of the information this signal holds – is sufficient for achieving transcranial imaging in comparison to other approaches like the medical invasive approaches such as thinning of the skull and injecting contrast agents in the blood, or technical approaches using custom all-in-one transducers which combine build-in analog-to-digital converters into one application-specific circuit.

In this thesis, we make the first decisive step in answering this question by providing a detailed technical specification and prototype construction of a TCfUS system. Before committing to such a list of specifications, the actual design requirements of the system need to be carefully identified. This is not a trivial task since a number of antagonizing design goals are present in our envisioned system. For instance, improving SNR via the techniques mentioned in the previous section and by performing as much of the signal processing as close to the digitization stage conflicts with the wish to build a highly portable system as it impacts its physical dimensions, which eventually introduces heat-dissipation problems and, thus, power limitations.

In order to guide the rest of this work and to assess the outcomes at the end, below we break down this thesis goal into further sub-questions:

1. What is the status of related works with respect to transcranial functional ultrasound systems?

2. What are the design trade-offs involved in this system when optimizing for SNR?

3. What is the system organization that best serves these trade-offs?

4. What are the minimal technical specifications for a TCfUS system?

5. How do we guarantee that our system organization is feasible with off-the-shelf components?

6. How does a system prototype with these specifications look like and how can it be validated?

7. How does the designed prototype scale with regard to the number of channels?

Overall, this thesis will lay the foundation work and provide a roadmap for future works and future designs. This foundation consists of the requirements, specifications, and trade-offs for a high-SNR TCfUS system, which is the beginning and most difficult piece of every engineering problem.

1.4. Thesis Scope

The scope of this thesis then is the design and development of a novel research TCfUs system prototype that is: highly flexible for allowing experimentation, can remain portable such as the standard functional ultrasound technique, is expandable in terms of number of channels, whilst at the same time being near-real-time and only using off-the-shelf components. The verification of the SNR increase over the state-of-the-art can only be answered with a built prototype and real-life hardware, however this hardware was not or limited available during the project duration. This thesis will therefore focus on the implementation of state-of-the-art hardware in an architecture, the theoretical maximum performance and boundaries of the design, and how the system can expanded in terms of number of channels with the constraints following from the design decisions. Because of the maximum budget, and the limited time, only the use of off-the-shelf components are considered in the design, in the contrary to application specific integrated circuits (ASIC).

1.5. Thesis Overview

The remainder of the thesis is structured as follows:

Chapter 2 provides an introduction to the theoretical background of ultrasound and the derivatives thereof. Additionally, the limitations of transcranial ultrasound are discussed, as well as three techniques to improve upon the SNR problem that occurs due to the impedance mismatch and aberration of the ultrasound waves due to the the skull bone.

In **Chapter 3** a basic version of the receive structure of an ultrasound system is presented and compared to the state-of-the-art ultrasound research machines currently available on the market. The most important takeaways will serve as a foundation for the next chapter.

In **Chapter 4** a list of functional requirements is presented from which design decisions and specifications are derived. Some trade-offs have to be made between certain parameters in the specifications and three processing configurations emerge from these trade-offs. Every component in the receive system then follows with detailed specifications in each processing configuration.

In **Chapter 5** a design is presented for a high SNR TCfUS acquisition and processing system. Components are selected based on the specifications and the current state-of-the-art ultrasound system, after which an architecture for a 64 channel portable TCfUS system is presented. A preliminary study is conducted on the expansion of the number of channels in the design, based on the potential hardware components available since limited hardware was available during the thesis project. Thereafter the limiting factors are discussed and concluded upon.

In **Chapter 6** conclusions are drawn from the previous chapters and the proposed design for an as-high-as-possible SNR transcranial functional ultrasound system prototype is evaluated and further suggestions are made for future research to improve on the current work.

In **Appendix A** a design for an adapter printed circuit board is presented, which presents an opportunity to test both the selected analog front-end and the decimation filtering in the design. Additionally, this printed circuit board also incorporates a method to combine multiple transducer channels for compressed imaging. However, due to problems with the sample clock, the results of the experiments were limited.

2

Background

In this chapter, a theoretical background will be given in order to better understand the parameters, problems, and phenomena that are related to transcranial functional ultrasound (TCfUS). First, in Section 2.1 a basic introduction to brightness mode (B-mode) ultrasound will be given including the underlying physical phenomena and the processing required to create a B-mode image. Additionally, the basic components of an ultrasound system are discussed to provide a better picture of the hardware involved in an ultrasound system. In Section 2.2 the theoretical background for Doppler imaging and its derivative power Doppler imaging are explained which form the basis for functional ultrasound imaging discussed in the subsequent Section 2.3. Section 2.4 provides more information on the technical difficulties in imaging the human brain with ultrasound. Hereafter, Section 2.5 concludes that improving the signal-to-noise ratio (SNR) in the ultrasound machines is one of the ways to improve the quality of the signal in transcranial ultrasound. Therefore, three important SNR improvement techniques are discussed in the four subsequent sections. Section 2.6 discusses the use of plane waves and coherent compounding in order to increase the SNR, Section 2.7 discussed the use of contrast agents to increase the SNR, and Sections 2.8 and 2.9 discuss the technique of oversampling in decimation in order to decrease the quantization error and thus increasing the SNR. Hereafter, Section 2.10 will provide a short introduction and technical background on the workings of a field-programmable gate array (FPGA), which is one of the building blocks of a research ultrasound machine. Finally, Section 2.11 will conclude on the most important topics discussed in the chapter.

2.1. Ultrasound

Ultrasound is one of the most common noninvasive medical imaging modalities used today, because of the low cost [69]. Brightness mode (B-mode) ultrasound forms the foundation for ultrasound and due to the semiconductor revolution in the 60s and 70s ultrasound systems could be miniaturized which allowed prices to drop. In Figure 2.1 an overview of a conventional ultrasound system is given.

Ultrasound works on the principle of sending and receiving high frequency acoustic waves (1-20 MHz)[13]. These acoustic waves are excited by applying a voltage potential to a piezoelectric element, also called a transducer element, which vibrates and subsequently generates an acoustic wave, which insonifies the area of interest. An ultrasonic probe implements a series of these transducer elements to create an acoustic wavefront, the characteristics of these transducer elements determine the excitation frequency of the transmitted acoustic wave. Due to impedance differences between the different tissues in the area that are insonified, only part of the acoustic wave is reflected back. Some of the most important impedances used for the application of brain imaging are listed in Table 2.1. These reflected echos are then picked up by the same transducer elements that produced the acoustic wave. However, because the acoustic waves also undergo attenuation in the encountered media, the result is a small voltage potential which is then amplified and digitized by the analog front-end. With a processing technique called beamforming, these digitized signals can be reconstructed into an image presenting the geometrical information of the scatterers in the depth axis z and lateral axis x. Because the intensity of the pixel that is displayed corresponds to the signal value, and thus the amount of the signal reflected by

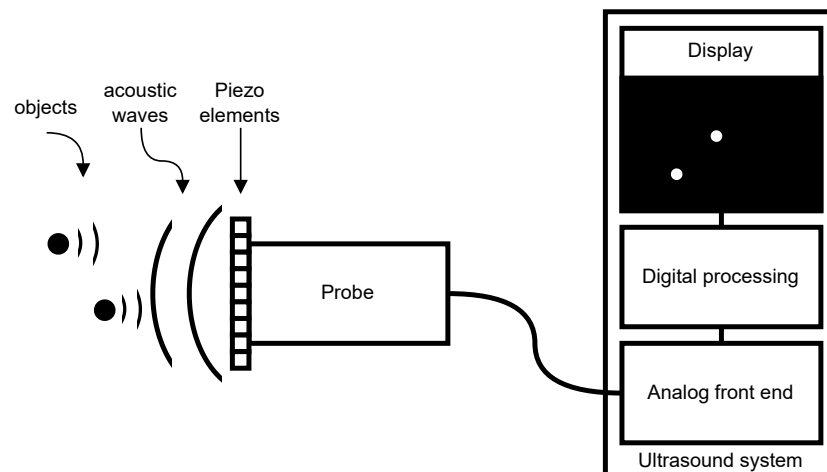a scatter, this mode is also called brightness mode or B-mode.

Figure 2.1: A functional overview of a conventional ultrasound system.

Medium	Phase Velocity c [m/s]	Impedance Z [Mrayl]	Atten. Coefficient α [Np/(cm·MHz)]
Air	333	0.0004	-
Water	1480	1.48	0.0002
Soft Tissue	1540	1.63	0.08
Blood	1580	1.67	0.02
Brain	1460	1.5	0.06
Cranial Bone	2770	4.8	2.5

Table 2.1: A list of tissues encountered in brain imaging and their characteristics embodied in common used metrics and units: phase velocity c, characteristic impedance Z and attenuation coefficient α, from [54].

2.1.1. Ultrasound System

In order to discuss the design of an ultrasound system, as proposed in Chapter 1, it is important to go into more detail on the individual components that make up an ultrasound system, because the architecture of an ultrasound system and the specifications of individual components eventually determine the overall performance of the system. In Figure 2.2 a schematic version of an ultrasound system is depicted. Each block in the diagram is responsible for a specific task. The components will be discussed following the signal path, from the transmission of an ultrasound wave to the received wavefront by the transducer to the processing and production of a B-mode image.

First, a wavefront has to be transmitted by the transducer elements. The shape, energy distribution, and thus focus of this wavefront are determined by the user. The firing order of the elements, that control the previously mentioned parameters, is handled by the TX beamformer, which loads the patterns from a pre-programmed memory. Subsequently, the pulses that are produced by the beamformer are amplified by the high-voltage pulser (HV pulser) and go through the RX/TX switch to the individual transducer elements to produce the desired acoustic wavefront. The reflected waves from the area of interest are then picked up by the transducer elements, go through the RX/TX switch, and are amplified and digitized by the analog front-end (AFE). It is important that the pulses that are sent out by the HV pulser do not end up at the input of the AFE since this will overload the inputs of the AFE and could potentially destroy it. Because of this reason the RX/TX switch is placed between these signal paths and is controlled to switch at the correct moments between transmission and receiving. Afterwards, the incoming signals are digitized by the analog front-end (AFE) the beamformer can process the signals from the individual transducer elements and produce a B-mode image, this will depend on the

technique that is used as can be read in Section 2.6. From here on further processing steps can be applied to the B-mode image(s), the result of which is subsequently stored on a drive or displayed to the operator, depending on the application and the environment.
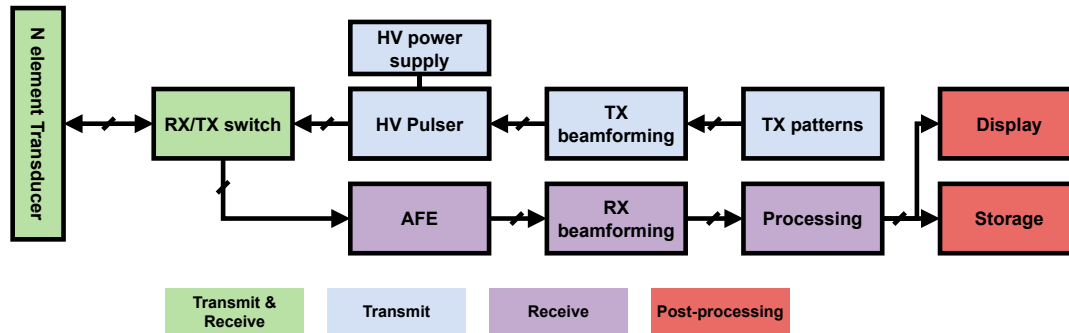

Figure 2.2: Block diagram displaying the basics of the signal flow inside an ultrasound system, each block indicates a specific component or function in the ultrasound machine, the paths indicate a bus of multiple signals, and the colors indicate the responsibility of each component/function in the signal chain.

2.2. Doppler Imaging

The Doppler effect is a commonly known phenomenon that is perceived as a change in frequency when a sound source moves towards or away from you, a common real-life example of this is an emergency vehicle, such as an ambulance, driving by while using its sirens. In 1962, Kato [37] published evidence that proved that the Doppler shift of an ultrasound signal was related to the velocity of red blood cells (RBCs) in blood, therefore laying the groundwork for a new application of ultrasound called Doppler imaging, which most importantly provides for a non-invasive measurement tool of the blood flow. Doppler imaging is often used as a cardiovascular imaging modality but can also be utilized for brain imaging, enabling scientists to map the vascular system of the brain. Power Doppler imaging which is a form of Doppler imaging forms the foundation of functional ultrasound, where the blood flow velocity from these images is coupled to neural activity; this is further discussed in Section 2.3. The current section will discuss the fundamentals of Doppler imaging.

In Doppler imaging, the frequency shift that occurs from a moving object, such as a RBC, is measured. If the send and receive transducer elements are the same, the formula for the change in frequency (Doppler shift) produced by a moving scatter is given by [70, 13]:

$$f_D = f_R - f_T = -\frac{2v}{c_0} f_0 \cos \theta, \qquad (2.1)$$

where f_R is the frequency of the received signal, f_T is the frequency of the transmitted signal and thus equal to f_0 which is the center frequency of the probe, v is equal to the speed of the blood in the vessel, c_0 is equal to the speed of sound in the medium, and θ the angle of the ultrasound beam with respect to the insonified blood vessel. The negative sign indicates that the received frequency is less than the transmitted frequency meaning the flow of the scatterer is away from the source [13].

The invention of this technique, proven by Kato [37], and the properties of the RBC sparked a wave of new techniques. In newer techniques, however, pulses of acoustic signals are used instead of a Continuous-Wave (CW) that was originally used in the method proposed by Kato. As Jensen describes [35] in his book on the estimation of blood velocities, the general term Doppler system is used for systems that estimate blood velocities, however, pulsed wave ultrasound systems strictly speaking do not use the Doppler effect. Instead, pulsed wave (PW) ultrasound systems use the shift of position of the RBC to find the velocity.

In order to further explain the workings of Doppler ultrasound systems that use PW, it is important to explore the properties of the transmitted acoustic wave. The acoustic waves that are generated, are generally sent in short bursts because acoustic waves need time to travel to the area of interest,
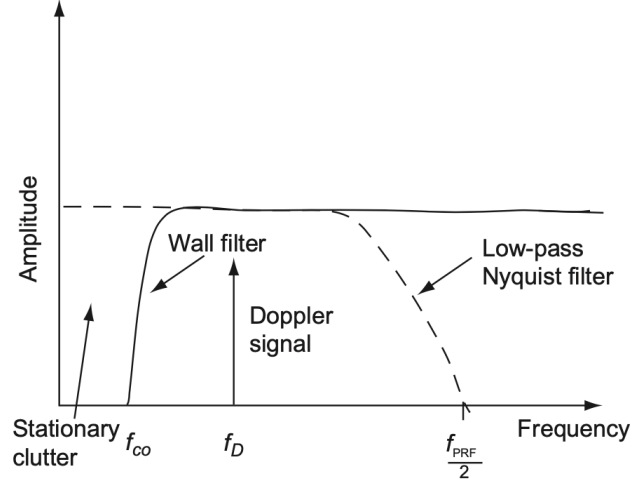
Figure 2.3: Frequency domain representation of the filter that extracts the Doppler signal $Z_F(x, z, t_i)$ from the input signal $Z(x, z, t_i)$, where f_{co} represents the cut-off frequency for the stationary clutter and $f_{PRF}/2$ representing the upper limit to prevent aliasing in the signal. From Szabo [70].

reflect and travel back. The pulse repetition frequency (PRF) is the number of pulses sent per unit time. Depending on the depth that needs to be imaged, the frequency that is used, and the length of the pulse, the PRF is selected. For medical ultrasound imaging the PRF is typically in the range of 1 to 10 kHz and is given by the following equation:

$$\text{PRF} = \frac{1}{t_{PRF}} = \frac{1}{t_{pls} + t_{rt}} = \frac{1}{\frac{r}{f_0} + \frac{2z}{c_0}}, \tag{2.2}$$

where t_{PRF} is equal to the time between each transmission, t_{pls} the duration of the pulse, t_{rt} the round trip time of the wave, z the imaging depth in meters, c_0 the speed of sound in the medium, r is equal to the number of pulse repetitions, and f_0 frequency of the generated acoustic wave.

If a consecutive set of B-mode images is acquired with a high enough PRF, each pixel in this consecutive set of images can be seen as a signal $Z(x, z, t_i)$ with $i = 1, ... N$ and N being the number of transmissions. Where $Z(x, z, t_i)$ is the complex signal after beamforming, this signal Z can then be filtered for the Doppler signal that represents the reflections of RBCs. This filter is a band-pass filter that removes the clutter in the lower frequency band and has a stop-band in at half the PRF to prevent aliasing of the Doppler signal, Figure 2.3 depicts the filter range in the frequency domain. Therefore, it is important that the PRF of the acquired signal is high enough in order for the Doppler signal not to alias.

Two parameters are commonly extracted from the Doppler signal with imaging techniques called: color flow imaging (CFI) which displays the axial blood velocity, and the power Doppler imaging which displays the mean intensity of the Doppler signal which is found to be equivalent to the cerebral blood volume (CBV) in each pixel [62, 61]. The CBV is a metric used for the estimation of neural activity in functional ultrasound and therefore the focus in this section on the power Doppler signal.

In Power Doppler imaging the filtered Doppler signal is integrated in order to acquire the power of the Doppler signal compared to the mean frequency shift [62] which is used in CFI. Therefore, in power Doppler, the directional information in the signal is lost, however with increased sensitivity of the image. The mean intensity $I(x, y)$ for a pixel in a Power Doppler Image (PDI) is calculated as

$$I(x, z) = \frac{1}{N} \sum_{i=1}^{N} |Z_F(x, z, t_i)|^2, \tag{2.3}$$

where the complex Doppler signal $Z_F(x, y, t_i)$ is the representation of the filtered B-mode signal at lateral position x, depth z, and sample t_i and N the number of consecutive images, typical ($N = 8, 16$) in

conventional Doppler ultrasound. The intensity I is commonly shown as a change in the range from red (low intensity) to yellow (high intensity), the fourth step in Figure 2.7 shows the result of a PDI. In Figure 2.4 a block diagram is depicted showing the Power Doppler processing pipeline where each number in the pipeline represents a processing step. Although Figure 2.2 specifies plane waves are used for transmission, the power Doppler imaging process for conventional line-by-line mode imaging described in this section is comparable. However, the line-by-line mode does not have the coherent compounding stage. The power Doppler imaging process that uses plane waves is described in more detail in Section 2.6.

Compared to CFI the sensitivity in power Doppler imaging is much larger and thus smaller veins can be detected, however, there are some drawbacks. First, Power Doppler is very susceptible to motion artifacts since the Doppler signal is influenced by all moving particles/objects and therefore it is difficult to distinguish between these artifacts and blood flow. This is partially solved by filtering the low frequencies with the clutter filter, however, this does impose a limit on the smallest veins, and thus slowest blood volume, that can be detected because this information is filtered out. The second drawback is that in conventional line-based ultrasound systems too many pulses are required to generate one B-mode image, and therefore the sensitivity of the signal I is too low for power Doppler imaging that can be applied for functional ultrasound imaging. To increase the sensitivity of I, the number of samples N per PDI needs to be increased, and therefore the PRF has to be increased which can be done using a technique called plane wave imaging which is discussed in Section 2.6.
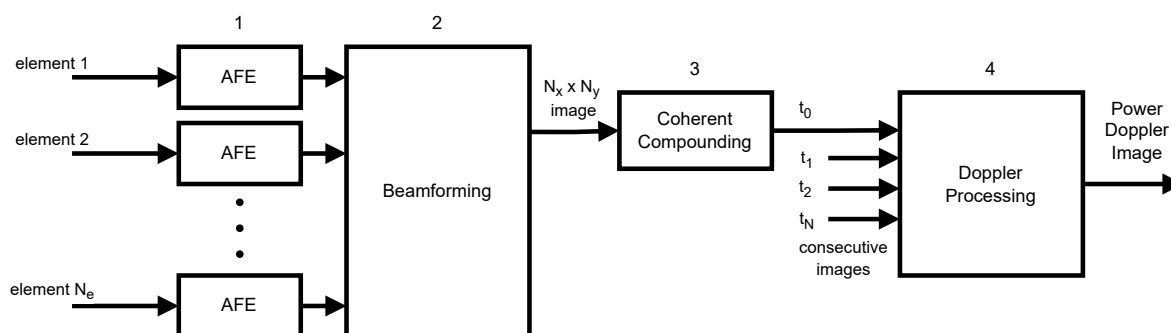

Figure 2.4: Block diagram of the signal processing pipeline used in functional ultrasound. 1) The reflected acoustic signals from the plane waves are coming in from each transducer element and a digitized by the analog front-end (AFE) 2) the raw RF data is beamformed into a $N_x \times N_y$ pixel image 3) multiple b-mode images are coherently compounded 4) Doppler processing is applied to a series of compounded B-mode images to form a Power Doppler Image (PDI)

2.3. Functional Ultrasound Imaging

Neurovascular coupling (NVC) is the mechanism in the brain that regulates the blood flow to the regions in which neurons are active, this response is termed functional hyperaemia [6]. The proper functioning of this NVC mechanism is critical, because with an inadequate supply of oxygen-rich blood to a region of the brain, then neurons may become injured or die, this is the leading cause of a stroke or Alzheimer's disease [6]. For a more comprehensive explanation, Martin [45] discusses the use of BOLD fMRI to understand the neurophysiological processes underlying NVC. Since power Doppler images are coupled to the blood volume in the vasculature of the brain these images can thus be linked to neuron activity using neurovascular coupling. In Mace et al. [43] the PDIs are further processed and the Doppler signal for each pixel is correlated to a known stimuli/task pattern that is provided to the subject. This technique is called functional Ultrasound (fUS) imaging. The outputs of this Pearson's product-moment correlation coefficient are called activation maps [43], an example of these activation maps is depicted in the last processing step (fUS) of Figure 2.7.

2.4. Transcranial Imaging

In this section, we discuss the fundamental limitations that prevent us from directly applying functional ultrasound on human subjects. It can be seen that transcranial functional ultrasound imaging is possi-

ble with mice and young rats [76]. However, this is only possible due to the thin skull of these animals. By utilizing the acoustic windows in the human skull and utilizing microbubbles as a contrast agent, Demené et al. [20] were able to map brain vasculature in humans using ultrafast ultrasound localization microscopy (ULM) to 25 μm. The goal of this thesis however is to provide transcranial functional imaging anywhere on the skull without any dependency on contrast agents or acoustic windows. Several problems, however, prevent this. The first problem occurs because of the impedance difference between the brain and the cranial bone is significant, as can be seen in Table 2.1. Due to this mismatch, a large part of the acoustic signal will be reflected or aberrated. The second problem is that due to the aberrations (complex changes due to the structure of the skull bone) the ultrasound waves do not focus well through skull bone, leading to no, or blurry images.

Equation 2.4 defines the transmissivity (T_I) and characterizes the intensity of the plane wave at the reflecting boundary [35, 54]:

$$T_I = \frac{I_t}{I_i} = \frac{I_t}{I_i} = T^2 \frac{Z_1}{Z_2} = \frac{4 Z_1 Z_2 \cos^2 \theta_i}{(Z_2 \cos \theta_i + Z_1 \cos \theta_t)^2}, \tag{2.4}$$

with I_t, I_i equal to acoustic intensities of the incident, and transmitted plane waves at the material interface, Z_1, Z_2 being the characteristic impedances of the two media, and θ_i, θ_t the angles of incidence and transmission at the material interface. With a plane wave assumed normal to the interface and taking the characteristic impedances from Table 2.1, it can be seen that $T_I = 0.76$ as also calculated by Pietrangelo in [55]. Since ultrasound encounters this boundary twice because of the inherit workings of the technique, a significant part of the energy of the acoustic wave has been lost.

The third problem is that, apart from the reflection of the acoustic wave, it can also be seen that the attenuation coefficient of the cranial bone is significant, therefore attenuating the part of the wave that does protrude through the skull significantly, approximately $10 - 20$dB per transmission [54], depending on the thickness of the cranial bone. For this reason, it is important to utilize a probe that produces acoustic waves at a low frequency (1.5 MHz), therefore allowing the acoustic wave to travel further with less attenuation, as can be seen in Figure 2.5. However a probe with a lower transmit frequency
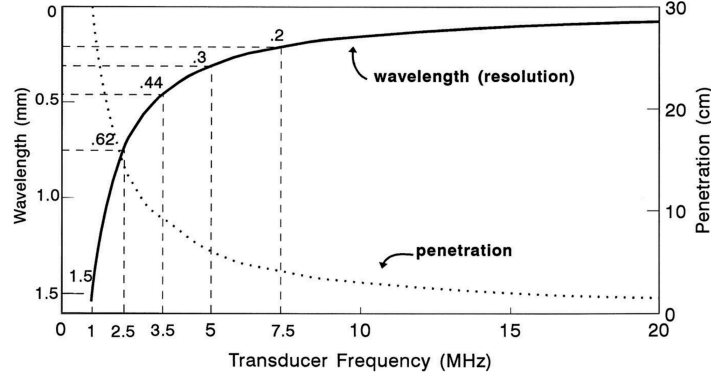
Figure 2.5: A graph depicting both the wavelength and the penetration depth of the acoustic wave that is produced by the probe, assuming the wave speed in soft tissue. As seen in the figure, these two parameters are inversely related and enforces a trade-off. Figure from [42].

also has its downsides, since the wavelength of the acoustic wave is inversely proportional to the axial resolution, defined by

$$\lambda = \frac{c_0}{f_c}, \tag{2.5}$$

where c_0 is the velocity in the medium (seen in Table 2.1) and f_c the center frequency of the transducer elements. Because center frequency f_c has to be lowered to increase the penetration and c_0 will not change, the wavelength λ increases and therefore decreasing the minimum observable distance between waves and lowering the axial resolution. An additional downside of using a lower transmit frequency can be derived from Equation 2.1, because the transmission frequency drops the Doppler shift will also decrease, therefore making the signal more difficult to distinguish from other low frequency signals as the clutter, therefore also decreasing sensitivity for CBV.

2.5. Signal-to-Noise Ratio and Ultrasound

From the parameters discussed in the previous section, it can be concluded that a large part of the acoustic wave transmitted in transcranial ultrasound imaging is reflected, attenuated, or aberrated during the insonification of the brain. Therefore, the waves that do propagate back are significantly attenuated and distorted compared to what is transmitted. This highlights the performance requirements of the ultrasound probe and the ultrasound system that amplifies and digitizes the received waves since any additional noise introduced here can significantly hurt the performance of the system as a whole. For this reason, it is important to introduce a metric that compares the power of the signal of interest and the noise power for a specific bandwidth. This signal-to-noise ratio (SNR) is defined as [13]

$$SNR = \frac{P_{signal}}{P_{noise}}, \qquad (2.6)$$

where P_{signal} is the power of the signal of interest and P_{noise} is the noise contributed from the transducer, coaxial cabling, amplification, digitization, background noise, and other sources. Increasing this parameter is of the utmost importance in order to create a system that can provide transcranial functional ultrasound and hence is the focus of this thesis. Therefore, it is important to discuss several techniques to achieve an increase in the SNR. An obvious solution would be to increase the power of the signal that is transmitted by the probe, therefore increasing the numerator of Equation 2.6. However increasing the transmit power has a limit, depending on the characteristics of the probe, breakdown voltage, power dissipation, and other factors. More important, since the attenuation of the ultrasound wave is caused by the transformation into heat there is a regulatory safety limit that cannot be exceeded [13]. There are, however, other ways to increase the signal-to-noise ratio without *just* increasing the transmit power of the probe. In the subsequent sections three techniques will be discussed: 1) Plane wave imaging which also solves the power Doppler sensitivity problem, 2) Contrast agents such as microbubbles, and 3) Oversampling and decimation.

2.6. Plane Wave Imaging

In conventional ultrasound systems, the Doppler signal is derived from a series of B-mode images with frame rates up to 40 frames per second [50]. For this reason, transcranial Doppler imaging of the brain through acoustic windows is limited to the (very) large arteries [19]. The Ultrafast Doppler (µDoppler) imaging technique proposed by Tanter and Fink [73] solved this issue by implementing plane wave imaging and coherent compounding. Allowing for frame rates up to ∼20kHz [44] and therefore enables the imaging of the small veins in a rat brain that has flow velocities that are much slower (25 mm/s) than the major arteries (50 cm/s) [44].

Compared to traditional line-by-line imaging, used in conventional ultrasound machines, plane wave imaging works by generating an unfocused plane wave utilizing all of the transducer elements. The angle of this plane wave can be determined by using beamforming. Instead of only reconstructing one line of the image per transmitted wave, the whole image is reconstructed from the received reflections. Since the beamforming is done individually for each pixel with the same data this process can be processed in parallel with each other. A visual comparison between line-by-line imaging and the plane wave imaging technique can be seen in Figure 2.6. A more in-depth visualization per step of the plane wave imaging process is depicted in Figure 2.7.

  With one single plane wave the sensitivity and the signal-to-noise ratio is worse than in a complete line-by-line mode image. However, by transmitting multiple plane waves, each with different angles and coherently compounding these b-mode images, the final image has a higher or equal SNR compared to line mode imaging [66, 50]. This comes with a significant benefit because compared to a line-by-line mode, the acquisition time can be reduced by a factor of 16 [8] and allows for frame rates up to 20,000 frames per seconds(fps) [44]. Because the number of samples per unit time N, available for the processing of a power Doppler image using Equation 2.3 the sensitivity of the pixel CBV also increases. Therefore, plane wave imaging/ultrafast Doppler has led to an important breakthrough in the application of ultrasound in functional brain imaging and allows for whole-brain imaging of the rat[44] or mice[39] brain. Together with microbubbles, discussed in the next section, ultrafast Doppler even allows subwavelength detection of blood vessels in patients, using ultrasound localization microscopy
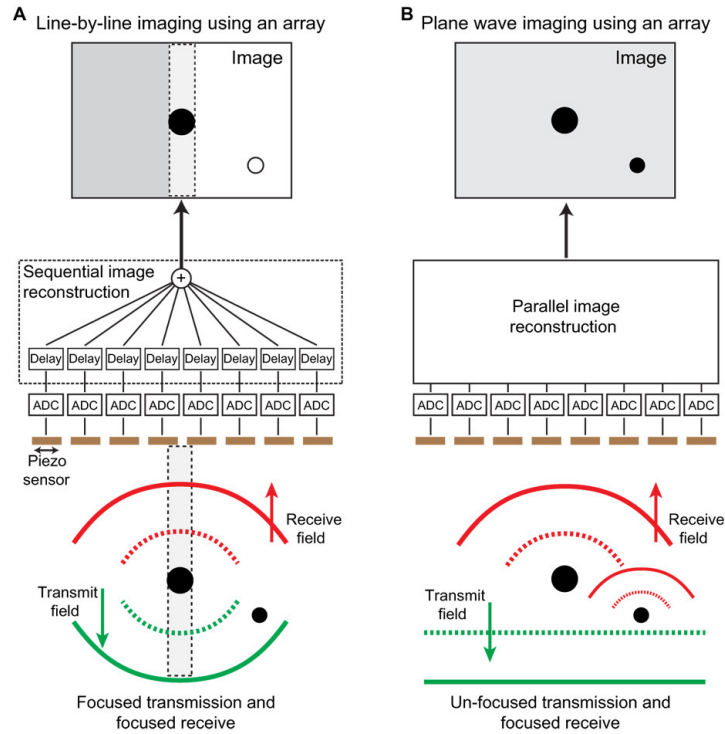
Figure 2.6: Difference between (a) conventional line-based ultrasound and (b) plane wave imaging, adapted from [40].

(ULM) and the acoustic temporal bone window [20].

2.7. Contrast Agents

Contrast agents such as microbubbles can be added to the blood to increase the contrast of the B-mode images. The microbubbles are gas-filled bubbles and the size of the bubbles is similar to RBCs. Therefore they also act as a scatterer and thus reflect the acoustic ultrasound waves. Because of the increase of reflections the signal power received from the arteries is also increased compared to using no microbubbles, therefore also increasing the SNR. However, a drawback of contrast agents is the need to ensure a constant concentration of microbubbles in the blood since the microbubbles quickly dissolve (5 min)[44]. Therefore, using contrast agents can be seen as a solution for increasing the signal-to-noise ratio, but this comes at the price of being a more invasive procedure whilst not being an effective solution over a longer time period. Because of this reason, the focus of this thesis is on a more permanent solution to increase SNR which is discussed in the following section.

2.8. Oversampling

Oversampling is a technique to increase SNR in a system where the bandwidth of the signal that is captured by that system is a multitude of the signal bandwidth of interest. The idea is that the noise that is introduced during the digitization process is spread out over a larger bandwidth, therefore when the signal is filtered and decimated only the bandwidth of interest is left and contains less noise power and therefore increasing the SNR. It is however important to remember that any sources of noise already present in the signal before it is presented to the analog-to-digital converter (ADC), from hardware or other phenomena that occur before the conversion are not affected by this technique. As may be obvious from the description of the oversampling technique, oversampling is only possible if the hardware of the right specifications is implemented. Therefore, in order to further explain this concept of oversampling, it is first important to understand the signal chain of the ADC because this is the hardware component that allows for this technique. Once the details of the signal chain are discussed and it is clear what causes the noise in the digitization process, the filtering and decimation to end up with the bandwidth of interest are discussed in Section 2.9.
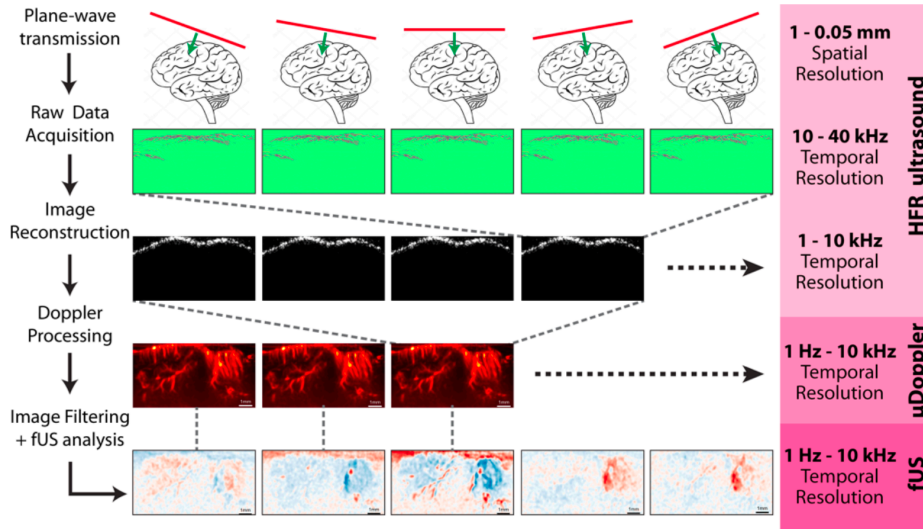
Figure 2.7: Illustration of the acquisition process of functional ultrasound. The maximum temporal and spatial resolution are stated per processing step.

The ADC, which is the working horse in the analog front-end (AFE), translates the analog signals into a digital value. The three main components for an analog-to-digital converter are depicted in Figure 2.8 and the three main components are: 1) filtering, 2) sampling and 3) quantization. The acoustic waves that are used in ultrasound have a center frequency in the range of 1-20 Mhz as mentioned in Section 2.1. However, depending on the specifications of the probe, the bandwidth of the signal picked up by the transducer elements is usually equal to 2.5 MHz distributed around the center frequency. The Nyquist–Shannon sampling theorem tells us that in order to prevent aliasing and thus to accurately represent the signal in the discretized domain, the sample frequency of the analog-to-digital converter has to be at least twice the frequency as the bandwidth of the signal that is presented at the input as:

$$B \leq 2 \cdot f_s, \tag{2.7}$$

where B is the bandwidth of the incoming signal and f_s is the sampling rate of the ADC. Because of these criteria, the low-pass filter, also called the anti-alias filter, is a critical component before the signal is sampled. One of the challenges for this filter is to find an implementation that has a very flat transfer function for the pass-band and is very steep for the stop-band without distorting the input signal that is then presented to the sampler. The sampler from Figure 2.8 discretizes the input signal at specific intervals in time determined by f_s. This discretized signal is now an analog signal sample and is presented at the input of the quantizer where it is quantized to a specific level and produces a digital sequence.

With a clear picture of the signal chain of the ADC, it is important to discuss the noise source affects the potential accuracy of the ADC. The accuracy of the quantization process is dependent on the step size of the levels of the quantizer and their linearity. Therefore a mathematical model of noise added in the quantizer, also called quantization noise, can be defined as [57]

$$x_q(n) = e_q(n) + x(n), \tag{2.8}$$

with $x(n)$ the input signal to the quantizer, $e_q(n)$ equal to the quantization noise in the system, and $x_q(n)$ the output of the system. From here we can estimate the power of the input signal $x(n)$ and the quantization noise $e_q(n)$ using a list of assumptions on the statistics and distribution of the quantization noise $e_q(n)$, which can be found in Proakis et al. [57]. In the end, this allows us to calculate the signal-to-quantization-noise ratio defined as [57]:

$$\text{SQNR(dB)} = 10 \log_{10} \frac{P_x}{P_{qn}} = 20 \log_{10} \frac{\sigma_x}{\sigma_e} = 6.02 N_Q + 16.81 - 20 \log_{10} \frac{R}{\sigma_x}, \tag{2.9}$$

with P_x equal power of the input signal, P_{qn} equal to the power of the quantization noise, σ_e, and σ_x the standard deviation of the input signal and the quantization error, N_Q equal to the number of bits to
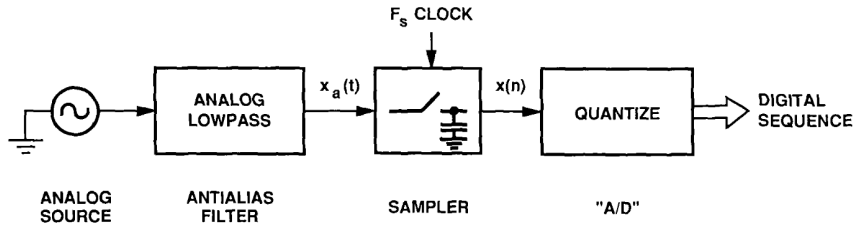
Figure 2.8: The three main components in the signal chain of an analog-to-digital converter, from Hauser [29].

represent the input value, and R equal to the input range in volt. By assuming the input signal $x(n)$ is sinusoidal, distributed over the entire range R, captured with a sample frequency of Df_s, and that the bandwidth of $x(n)$ is in the bandwidth of $f_s/2$ as seen in Figure 2.9, the quantization noise can be reduced by filtering and decimating $x_q(n)$ with a factor D. Therefore, the limits of integration for the quantization noise power now only spread over the bandwidth of the signal of interest. With the help of these assumptions and Equation 2.9, Hauser [29] states that the maximum SNR to be attained by an ADC using oversampling and decimation can be approximated with:

$$\text{SNR}_{\text{max}}(\text{dB}) \approx 6.02 N_Q + 1.76 + 10 \log_{10} D. \tag{2.10}$$

It is however important to note that Equation 2.9 and thus Equation 2.10 will only hold under a significant number of assumptions. Therefore, real-life performance will never reach these numbers, but it is important to take into account all the parameters and theoretical maxima to see what is possible for the design of an ultrasound system. From Equation 2.10 it can be seen that the SNR can be increased by increasing the number of bits for quantization N_Q and thus the resolution of the ADC, and by oversampling the input signal with a factor Df_s. Both parameters have one thing in common, and that is that the throughput resulting from the increase in resolution and sample frequency will also increase significantly if no decimation is applied. This will later pose a significant part of this thesis. For this reason, filtering and decimation are often directly applied after the digitization of the input signal. To provide a better insight into the filtering and decimation process Section 2.9 will further elaborate on this topic.
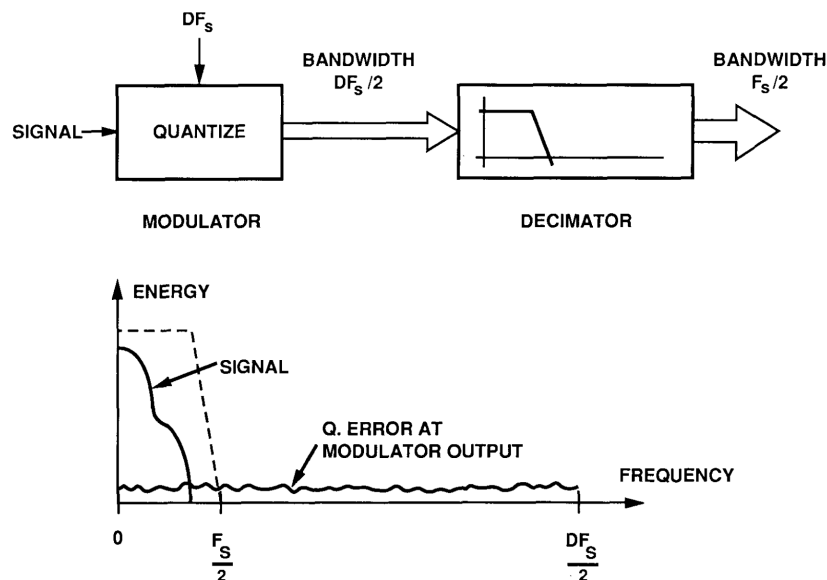

Figure 2.9: The illustration of the signal of interest and how oversampling and decimation can decrease the quantization noise power by decreasing the area of integration for the noise power. From Hauser [29].

2.9. Decimation

Decimation is a form of multirate signal processing, also called downsampling. In this process data samples from the input signal $x(n)$ are left out by resampling the signal at a factor of the original sample frequency f_s, this factor is called the decimation factor. Although in preceding sections of this chapter denote the decimation factor as D for convenience and illustration purposes, literature often denotes the decimation factor with M, therefore M is used in subsequent equations of the thesis to denote the decimation factor. If, however, downsampling is applied to an input signal $x(n)$ with frequency components higher than $f_s/2M$ the downsampling process will cause aliasing [57, 72]. Therefore, a low-pass filter or anti-aliasing filter is required for the decimation process. This filter will also filter out the quantization noise introduced by the ADC and is, therefore, a key component in the oversampling and decimation process. The decimation process can be seen as two cascaded blocks, the anti-aliasing filter, and the downsampling process. An illustration of the process is depicted in Figure 2.10. From Figure 2.10b
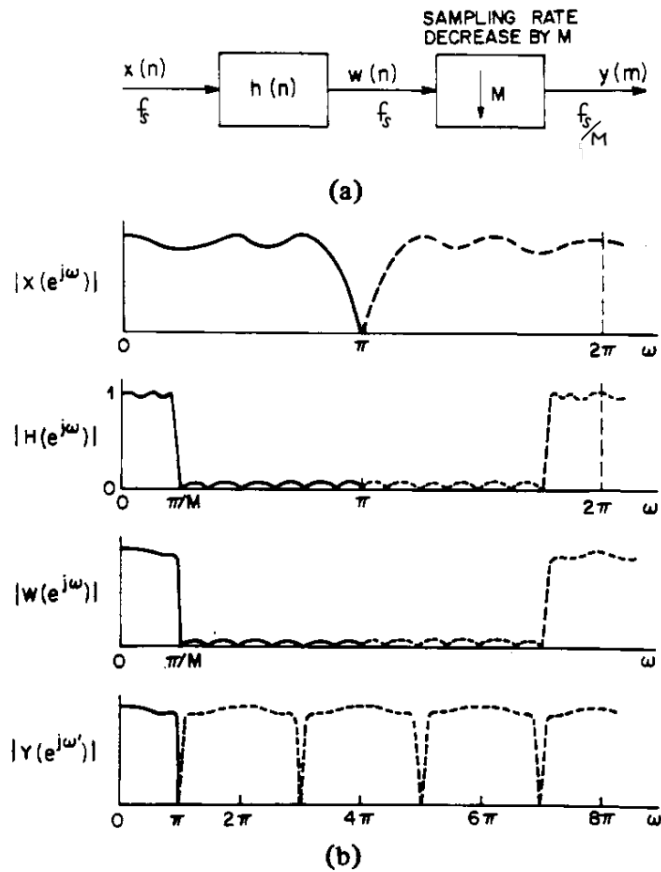
Figure 2.10: a) Structure of a decimation filter, b) typical spectra in the decimation filter, for decimation by a factor M, adapted from Crochiere [15].

it can be seen that the filter $h(n)$ is necessary to prevent aliasing of out-of-bounds frequency components into the band of interest $\pi/M = f_s/2M$. Therefore, the oversampling and decimation technique to reduce SNR is only effective if the filtering beforehand is accurate enough. An estimation can be done for the size of the filter with the assumption of parameters regarding the accuracy of the filter. The minimum number of taps N for a single-stage FIR filter can be estimated using the following formula [15, 81]:

$$N \approx \left((\log_{10} \delta_s) \left[a_1 \left(\log_{10} \delta_p \right)^2 + a_2 \left(\log_{10} \delta_p \right) + a_3 \right] \right.$$
$$\left. + a_4 \left(\log_{10} \delta_p \right)^2 + a_5 \left(\log_{10} \delta_p \right) + a_6 \right) f_s/\Delta f, \tag{2.11}$$

where σ_p is equal to the ripple in the passband from the ideal response, σ_s equal to the stopband ripple, $a_1 = 0.005309$, $a_2 = 0.07114$, $a_3 = -0.4761$, $a_4 = -0.00266$, $a_5 = -0.5941$, and $a_6 = -0.4278$. The

δf is equal the transition bandwidth of the filter and f_s is equal to the original oversampled sampling frequency.

In practice, however, more complicated multistage decimation filters, cascaded integrator–comb (CIC) filters, and polyphase filter structures are used to reduce one complex filter to a number of low complexity filters [57, 60]. This reduces the number of multiply and adds required and thus also reduces the amount of resources/area required while maintaining accuracy. For this reason, analog front-ends that include a polyphase filters decimation filter use the rule of thumb of $N = 16M$. Where N is equal to the length of the symmetrical filter and M is equal to the positive integer decimation factor.

2.10. Field Programmable Gate Arrays

Field Programmable Gate Arrays (FPGAs) are devices that are often used within research ultrasound systems because of the throughput and configurability they provide in comparison to general-purpose processors. Xilinx, one of the major FPGA manufacturers today describes an FPGA as follows: "Field Programmable Gate Arrays (FPGAs) are semiconductor devices that are based around a matrix of configurable logic blocks (CLBs) connected via programmable interconnects. FPGAs can be reprogrammed to desired application or functionality requirements after manufacturing. This feature distinguishes FPGAs from Application-Specific Integrated Circuits (ASICs), which are custom manufactured for specific design tasks."[84]. Because of the amount of configurability within an FPGA, custom designs for hardware logic, such as filters, fast Fourier transforms or other digital signal processing steps can be integrated with an FPGA. Because the FPGA can be reconfigured, the parameters or the functionality of the FPGA as a whole can be changed without replacing any hardware.

2.11. Conclusion

Ultrasound is a technique for non-invasive imaging of the human body and brain by using acoustic waves and the reflection of these waves on tissue boundaries which causes reflections. Doppler ultrasound imaging uses the Doppler effect to determine the blood velocity and volume and the cerebral blood volume in its turn can be correlated with brain activity using the mechanism of neurovascular coupling and this is called functional ultrasound. In small animals like mice and rats, it is possible to do transcranial functional ultrasound. However, the human skull reflects, aberrates, and absorbs the majority of the transmitted signal and therefore the signal reflected by the red blood cells is very weak compared to the transmitted signal. Therefore, it is currently only possible to do transcranial ultrasound using the acoustic windows within the human skull. In order to do transcranial imaging of the human brain, it is therefore critical to focus on the signal-to-noise ratio of the power of the signal of interest versus the power of the noise that is introduced. Since the thesis focus is to design an ultrasound system that is focused on reaching the highest possible SNR several techniques are discussed in order to improve the SNR in ultrasound imaging. 1) Ultrafast Doppler uses multiple plane waves at different angles to increase the sensitivity of the Doppler signal while also increasing/maintaining the SNR at framerates up to 10 kHz. This technique is the foundation enabling the required sensitivity for functional ultrasound. 2) Contrast agents are injected into the blood and increase the number of reflective objects besides the red blood cells already present therefore increasing the reflected signal amplitude. The major drawbacks of contrast agents are the short presence in the blood and the extra invasiveness introduced. 3) Oversampling and decimation are a combination of techniques in order to reduce the quantization noise that is introduced when the received signal is converted to the digital domain. This quantization noise can be spread over a broader spectrum by sampling the signal and bandwidth of interest with a significantly larger sample frequency. This however causes a significant data burden and by filtering and decimation, the amount of data can be reduced to only the signal/bandwidth of interest while reducing the quantization noise that is introduced. An important step in the decimation process is the filter that precedes it; without this filter, the decimation process would introduce aliasing artifacts into the bandwidth of interest and the quantization noise of the whole spectrum would still be present. FPGAs are a tool commonly used in ultrasound research machines due to their versatility and reconfigurability while performing application-specific tasks.

<div style="text-align: right; font-size: 3em;">3</div>

Related Work

With a general vision of the transcranial functional ultrasound system proposed in Chapter 1 and the most important technical background explained in Chapter 2 this chapter summarizes all related works in the field of transcranial ultrasound imaging, in order to present an overview of the state-of-the-art functional and/or transcranial ultrasound systems, that can subsequently aid the design of the envisioned high-SNR TCfUS system prototype.

In Section 3.1 the scope is defined of the state-of-the-art ultrasound systems that will be discussed. Additionally, a generalized structure of the receiving end of the ultrasound system is presented, which serves as a foundation in order to discuss all of the components of the systems using the same terminology. Following, Section 3.2 provides an overview of the specifications of the selected ultrasound systems and will discuss in detail all of these systems, including important or exclusive features. Afterwards, Section 3.3 discusses the most important parameters and features of the ultrasound systems and how they influence the design. This discussion serves as the foundation for the design decisions that have to be made from the requirements defined in the next chapter. Finally, Section 3.4 will conclude upon the most important takeaways of the chapter.

3.1. Scope

In the background chapter, Section 2.1, Figure 2.2, a general overview of the hardware architecture of an ultrasound machine is depicted. The scope of the systems that will be presented in the overview in Section 3.2, however, will be limited to the receive signal chain components. This includes the acquisition of the signal from the transducer to the final storage of the data. The reason to reduce the scope of the overview of the state-of-the-art ultrasound systems to only the receive signal chain is that the receiving end of the system is the place where the majority of the SNR gains can be achieved. In order to present an overview of ultrasound systems and compare the systems a generalized block diagram of the receive signal chain of a functional ultrasound system is depicted in Figure 3.1. In the subsequent paragraph, all the blocks [I-IV] and the interconnections (a-c) of Figure 3.1 will be discussed and will provide a platform in which to capture the state-of-the-art.

The [I] Transducer is responsible for transmitting and receiving the ultrasound waves to the target of interest, in our case, the brain. The design of the transducer is heavily dependent on the application, but the aim is a high-channel-count, high resolution, high sampling rate system, to achieve the highest possible image quality. Because of these parameters the transducer thus generates a large number of wideband (2.5 MHz) signals. In connection (a) the signals coming from the transducer are still analog, and thus susceptible to analog-noise phenomena in their transportation to the digitization stage [II]. The analog front-end (AFE) [II] contains all the analog signal conditioning circuitry (low noise analog amplifiers, filters, variable gain amplifiers, etc.) to condition the incoming ultrasound signals and an analog-to-digital converter (ADC) to digitize these conditioned incoming signals. Because of the number of signals that need to be digitized, this in its turn, creates several hundreds of Gbps in data. This poses a problem for the current high-speed interconnect that is used for connection (b). The

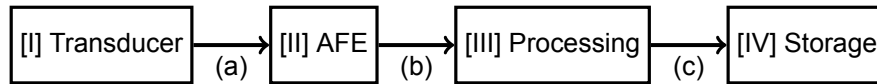<div style="text-align: center;">17</div>

Figure 3.1: A block diagram of the receive part of an ultrasound system.

processing [III] contains all the steps described in Section 2.2 in order to generate a power Doppler image. The processing is done in the digital domain for ease and flexibility of configuration and plays a critical role in the system because it reduces the data/information throughput. Therefore it is beneficial to implement the processing [III] as close to the digitization stage [II] as possible. From the processing stage [III], the signal can then be stored [IV], in practice, this is almost always a workstation. Depending on the amount of processing that is done, the throughput and size of the data in (c) varies from a couple of Gigabytes per second (Gbps) to the full throughput that is reached in (b).

3.2. State of the Art

Currently, there is a lot of work available on ultrasound systems, but the focus in this section will focus on the hardware and architecture of the current state-of-the-art research ultrasound systems. Since fUS and TCfUS are techniques that are still researched and not yet regularly implemented in the clinic these techniques will always require an ultrasound research system. Meaning that all the parameters of the system, such as the transmit waveform and receive processing are fully customizable. Another important feature of research systems is that raw RF data coming from the AFE [II] can be stored. So processing can be done afterwards in order to optimize the processing.

The following subsections will discuss the systems listed in Table 3.1 which contains an overview of state-of-the-art research systems. The systems listed in Table 3.1 are either general-purpose research ultrasound systems that can be or are used for functional ultrasound, or an ultrasound system specifically focused on transcranial ultrasound. Each following subsection will discuss one system and the generalized platform depicted in Figure 3.1 will help to capture the hardware features and architecture of the state-of-the-art designs, as well as their performance. The subsequent Section 3.3 will thereafter discuss the systems in a more generalized manner.

3.2.1. SARUS

The SARUS is a large but versatile ultrasound system that can handle up to 1024 channels. The system focuses on hardware processing of signals and researching the possibilities of ultrasound parameters. Therefore it features a full configuration of all the send and receive paths. The system however is not mobile and consists of multiple standard size 19-inch racks for conversion, processing, storage, and cooling. The system architecture splits the 1024 channels over 64 digital acquisition and ultrasound processing (DAUP) boards. These provide the transmit, digitization, and processing hardware and contain five FPGAs, each responsible for an individual processing task. The 64 DAUP boards are interconnected via 1 Gbps Ethernet and switches to 12 Linux servers where the data is stored. Therefore the conversion[II] and processing[III] are thus divided up per 16 channels where multiple of these boards are connected to multiple workstations [IV]. The input data is significantly reduced by the processing [III] from 13.44 Gbps to around 1 Gbps in order for connection (c) to keep up. Besides the processing[III] that can be done, the system also provides enough memory for the temporary storage of raw RF samples from the AFEs that can be offloaded over a larger time span.

3.2.2. ULA-OP 256

The ULA-OP 256 is a tabletop-sized ultrasound system that can accommodate probes up to 256 channels in both send and receive. It is unique in its hierarchical buildup of the system. It comprises front-end boards (FE) that provide the transmission digitization of 32 channels. Also included on the FE board is processing in the form of an FPGA and two digital signal processing (DSP) chips. A ring-like Serial RapidIO network connects the FE boards via a backplane and to the master control (MC) board. Which also contains an FPGA and a DSP chip. The processing [III] is therefore divided into two parts with a high-speed interconnect (RapidIO) in between. This system focuses heavily on hardware bases processing of the ultrasound signals as can be seen from the amount of FPGAs and DSP chips.

Because of this reason, the designers selected a relatively slow connection (c) to the workstation [IV]. Raw data from the ADCs still can be stored and accessed from the onboard memory (80 GB) however this is not real-time. To accommodate high element count transducers, the clock of multiple systems can be coupled for synchronous acquisition. Syncing can be provided from one master to four slaves. With two levels of hierarchy, giving a theoretical limit of syncing 21 systems and 5,376 channels. One workstation can be used for the control and offloading of data from multiple systems [47].

3.2.3. Aixplorer

A system that is not described Table 3.1 is the Aixplorer[68] (Supersonic Imagine, Aix-en-Provence, France) because the exact specifications of this system are not known publicly. However, it is known that this system laid the groundwork for software-based processing. Here the majority of the processing [III] is moved off the ultrasound system and to the workstation [IV]. A PCI Express bus provides the high throughput interconnect for this connection (c). The processing [III] is therefore split and the flexibility of the processing on the workstation allows for *ultrafast imaging* that was first proposed by Tanter et al. [73, 43]. The Aixplorer is also used for transcranial ultrafast ultrasound localization microscopy [20] on patients. Where the blood flow is measured using flow estimation techniques. The view however was limited by the acoustic windows. Multiple Aixplorers can also be combined [30, 31] to attach probes with a higher channel count such as 2D matrix probes to allow for 3D imaging.

3.2.4. UARP II & DiPhAS

The UARP II from the University of Leeds & the DiPhAS from the Fraunhofer Institute for Biomedical Engineering are both systems that are built to the 19-inch standard, with one system 3-4 standard units in height. The digitization [II] and processing [III] are split over boards of 16 (DiPhAS) and 8 (UARP II) channels. Via a backplane, this pre-processed data is then transported to a second processing step which offloads the data to a workstation [IV] via the PCI Express bus. Both systems combine a high sample rate and channel count with a high-throughput interconnect to a computer, and also provide raw RF samples even though memory is limited. Multiple DiPhAS systems can be combined in order to extend the number of channels up to 1024 for 2D matrix probes [33]. These four systems are connected to one workstation where further software-based processing [III] can be applied to the raw or pre-processed RF samples.

3.2.5. Verasonics Vantage 256

The Vantage 256 is a commercial research system produced by Verasonics [79] (Kirkland, Washington, USA). Just like the Aixplorer, this system is designed with software processing in mind. This is done in a Matlab environment on the workstation [IV], which provides the full configurability of the system, including transmit and receive parameters. Not too much is known about the internal hardware structure but from the approach and the interfaces it can be determined that there is a split in processing. The processing [III] is divided over the hardware acquisition part and the workstation [IV]. Because of the processing in the acquisition hardware, the throughput can be lowered enough to be able to fully utilize the maximum of PCI Express bus to the workstation. Because of this unique feature of configurable pre-processing using decimation filtering on the acquisition hardware, Verasonics can provide a platform for 256 channels at these rates and precision. This form of offering *quasi* raw RF samples to the user in near real-time ($\leq 0.5\mathrm{s}$) is unique especially at this form factor (47.6x28.0x48.9 cm). Eight systems can be combined to form a large acquisition system for a high channel count probe. An implementation where four systems (1024 channels) are used in parallel can be seen in [53, 87]. Here each system offloads its data and is controlled by, an individual workstation.

	SARUS [36]	ULA-OP 256 [10][11][9][47]	DiPhAS [59][58][33][23]	UARP II [77][49]	Verasonics Vantage 256 [79][78]	Lightprobe [26][27]	Proposed solution
Tx Channels	1024	256	256	128	256	64	-
Rx Channels	1024	256	256	128	256	64	64
Tx Voltage [V_{pp}]	0-200	0-200	0-160	0-200	2-190	0-160	-
Tx Freq [MHz]	1-30	1-20	0.3-20	0.5-15	0.5-20	0.625-20	-
Tx type	Lineair	Lineair	3-level	5-Level	3-level	3-level	-
max f_s [MHz]	70	78.125	80	80	62.5	65*	125
ADC resolution [bits]	12	12	12	12	14	12	16
RAM [GB]	256[†]	80	16	16	16	2	16
→ MB/channel	125	312.5	62.5	125	62.5	31.25	250
RAM type	DDR2	DDR3	DDR3	DDR3		DDR3	DDR4
PC interface	64x 1Gb/s Ethernet	USB 3.0	PCIe 2.0 x8	PCIe 3.0 x16	PCIe 3.0 x8	QSFP fiber Aurora 64b/66b	PCIe 4.0 x8
→ bandwidth [Gbps]	64	1.6	25	126**	52.8	26.4	126**
Raw RF data [Gbps]	860.16	240	245.76	122.88	224	24.96	128
sync multiple systems	-	21 (2 layers)	4	-	8	-	-

Table 3.1: Comparison between research-oriented ultrasound systems, adapted from [12], *interleaved sampling, **theoretical, [†]only memory of first FPGA after digitization

3.2.6. Lightprobe

The lightprobe is the work of Hager [26]. This work presents a system where stages [I], [II], and [III] are packaged into one enclosure called the Lightprobe, and stage [IV] is completely separated. However, since the protocol that is used for the interfacing on (c) the data can not directly be processed on the workstation [IV]. It can therefore be stated that this is again a form of splitting processing. However, here the interconnect (c) does provide enough throughput to transport the raw RF data samples in (near) real-time from the Lightprobe to the workstation. The focus of this work is on portability and power consumption while maintaining all the features of a research system in terms of access to the raw RF samples which provides for the software-based approach in processing. Processing can be done in hardware on the probe itself or on the workstation. The interface to the workstation, interconnect (c), can provide the bandwidth to send all the raw data from [II] back to the workstation [IV]. However, because of bandwidth limitations this connection (c) in this system, the number of channels is limited to 64. Because the system can provide a near real-time using interconnect (c) the memory that accompanies the processing part [III] on Lightprobe can be limited to 1 GB.

3.2.7. Pietrangelo

The work of Pietrangelo [55] is a phased-array prototype particularly built for transcranial imaging through the acoustic windows of the temporal bone. It is a portable system that incorporates a custom, low-frequency, 8x8 element transducer. In this system, the transducer [I], digitization [II], and processing [III] are combined into one system. All processing [III] is done on the system and because no high-speed link is available and storage on board is limited, there is no access to the raw RF data for future reference or offline software processing.

3.3. Discussion

With the individual state-of-the-art systems discussed, the following section will go further into a more generalized discussion, elaborating on the most important takeaways of these systems. The discussion is structured per subject and will provide more insight into how the design requirements and parameters in a research system affect each other.

3.3.1. Portability

The portability and size, of the systems that are discussed vary wildly, with the SARUS being the largest system and the system built by Pietrangelo [55] plus the Lightprobe [27] being the smallest. The system built by Pietrangelo proves that portability is a major advantage for transcranial imaging because the system can be directly attached to the subject. As opposed to cart-based systems, subjects are not bound to prolonged times of sitting still and the portability allows for free-roaming experiments and measurements. Portability in conventional systems such as the Vantage 256 [79] is achieved by utilizing a cable from the transducer to the analog front-end utilizing a bundle of analog micro coaxial cables (Figure 3.1 interconnect (a)). This type of cabling is implemented because the signals produced by the transducer elements are very weak and low in amplitude and thus more susceptible to analog-noise phenomena but is limited in terms of length. Therefore, the Lightprobe digitizes the signals as close to the probe as possible and utilizes a fiber optic connection at Figure 3.2 interconnect (c), which carries raw RF or pre-processed digital data over to a workstation. By placing the AFE [II] as close to the transducer as possible the Lightprobe attempts to reduce the analog-noise phenomena by eliminating this micro-coaxial cabling and by transporting the samples digitally, therefore reducing the possibility for errors. A major disadvantage of the system built by Pietrangelo and the Lightprobe is their limitation in the number of channels. Because if the whole design (including processing and digitization) is scaled down for portability, this negatively affects the power per unit area and therefore increases the amount of heat per unit area, complicating effective heat dissipation. We can, therefore, conclude that the number of channels and the amount of processing is inversely related to the degree of portability and power/heat dissipation. However, since portability is one of the key requirements of the system, power/heat limitations are therefore established. Therefore, the only parameters that can be changed, are the location and the parameters of the processing and digitization, and the number of channels.

3.3.2. Channel Count

For a research system to operate in near real-time, the system has to be able to process the data produced by the AFEs [II] fast enough. Additionally, interconnect (b), which transports the data from the AFEs [II] to the processing stage [III] has to keep up. Because high channel count systems produce more data and thus more throughput, systems like the ULA-OP 256 [10, 11, 9, 47] and the SARUS [36] process the data [III] near the physical location of digitization [II]. However, if the user wants access to the raw RF data of the AFEs [II] these systems lack a high-speed interface between the processing stage [II] and storage [IV]. Therefore, these systems store the raw RF data on memory available next to the physical location of processing [III], which requires enough temporary fast memory to provide access to raw RF samples on the workstation. However, by temporarily storing the sampled data the specification of a near real-time system is abandoned. From these systems, it can thus be seen that the number of channels implemented in a system influences the throughput of the interconnect (b) and (c), depending on if the user wants access to the raw RF samples or processed data. This choice in processing then influences the ability for real-time imaging. The Vantage 256, on the other hand, implements a trade-off between processing and throughput by applying decimation filtering to the raw RF samples. This processing step causes a significant reduction in data throughput while maintaining the information from the selected bandwidth of interest and increasing the SNR as can be read in Section 2.8. Although the data samples that are available after the filtering and decimation are not the *true* raw RF samples from the AFE, the most important information is available for the subsequent (software-based) processing steps. Therefore, it can be concluded that the number of channels in a system is coupled to the throughput and thus interconnect. Where the throughput can be alleviated by either processing or temporary storage. A second technique to increase the number of channels can be seen in some systems that provide the option for the use of a multiplexer within the probe. With this technique, the probe contains more channels than a given system can handle. But, by multiplexing the channels of the probe and accessing each set of channels over time the system can still sample all of the channels, though be it at the comprise of sampling speed. However, to reduce additional noise in the system, high-voltage multiplexing of channels is generally avoided [86].

3.3.3. Processing

The system that is envisioned in Chapter 1 is one that is optimized for reaching the highest SNR, but at the same time also remains a research system. It is therefore important to take into account that the processing in the system has to remain flexible in order for parameters and different signal processing techniques to be explored. This flexibility can also be seen in the ULA-OP 256 and the SARUS systems, where all the FPGAs and/or DSP chips are fully configurable by the user, depending on the application. Additionally, it can also be seen that the amount of processing power that is envisioned for near real-time power-Doppler imaging (50 TFlops [73]) or power-Doppler imaging using A-matrix compressive imaging [40] is significant in terms of compute. For this reason, systems like the Aixplorer and the Vantage 256 offload this processing to a workstation and utilize graphical processing units (GPUs) for the processing. To better understand the design trade-offs that this processing flexibility and compute density imposes, we consider here the two design extremes: in the first case, all processing is packed at the portable system right after digitization, which maintains a high SNR but creates dramatic problems for compute density, design scalability and heat dissipation, and is impossible for the compute that is required for TCfUS. In the second case, the signal is digitized near the probe but all processing is postponed until the "far-side" workstation. Scalability, portability, and heat dissipation are improved over the first case but the connection throughput becomes intractable due to the sheer volume of raw data that needs to be transferred. Therefore, it is impossible to transport the raw RF data of a high channel count system (1024) over the current day available digital interconnect. Clearly, a compromise between these two extremes needs to be found. As seen in the overview of the state-of-the-art systems, conventional systems like the ULA-OP 256 and SARUS combine the digitization [II] and processing [III] together, but as seen in the previous Section 3.2.5, the Vantage 256 only applies a preliminary amount of pre-processing to the data before transferring the data to a workstation where the majority of the processing and thus compute can be executed. Therefore, data is reduced enough to transport the data to a workstation and the compute is offloaded to a location with ample compute resources. The Lightprobe combines the same decimation principle as the Verasonics in order to reduce the throughput enough to send the data over a QSFP fiber connection to the workstation for software-based processing. Therefore, the hardware integrated into the probe does not
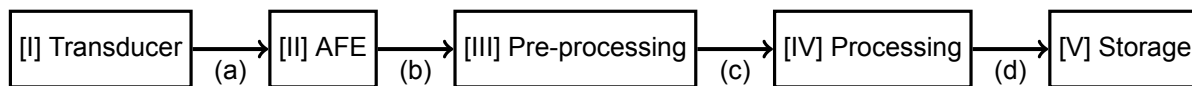
Figure 3.2: An updated block diagram of the receive part of an ultrasound system. The third stage is termed pre-processing because data is prepared and pre-empts further processing. What exactly is encompassed in the pre-processing is further elaborated upon in Chapter 4.

carry the full processing load and the processing stage [III] is split into 1) pre-processing consisting of decimation filtering and 2) the remaining computational intensive part of the fUS processing pipeline, including beamforming and the Doppler processing. In order to aid the following discussion, Figure 3.1 is updated to accommodate this split in processing in Figure 3.2.

Splitting the processing into two separate locations has its advantages and disadvantages. An advantage is that the pre-processing alleviates the throughput of (Figure 3.2) interconnection (c) and therefore allows for an increased number of channels that can potentially be accommodated in the probe section of the system. However, because the processing has to remain flexible and raw RF samples should also be available to the user if required this increase in the number of channels is impossible since throughput becomes too high. Therefore, building a high-channel-count system without any pre-processing would not be feasible with the current-day interconnect technology without breaking the portability constraint. So either the range of flexibility of processing in the system has to be fixed or the number of channels has to be fixed in order to design the system that is envisioned.

It can also be seen that there is a different approach to the processing between the systems. Most of the newer systems, such as the Aixplorer, Verasonics, and Lightprobe, choose the software-based processing approach. Here the raw RF or pre-processed data is transported to a workstation, as quickly as possible. On the workstation, all of the computational intensive processing is executed on the central processing unit (CPU) and graphical processing unit (GPU). The advantage here is that the parameters of the processing pipeline can be quickly adjusted and optimized, however, this technique only became possible due to the recent development and efficiency of GPUs. A more traditional approach is hardware-based processing, implemented in systems such as the SARUS and the ULA-OP 256. In this approach, the processing is executed on FPGAs. FPGAs can be very efficient in executing custom and/or complex processing tasks, however, the problem is that designing an efficient configuration for an FPGA can be very hard, time-consuming and parameters cannot be changed easily.

3.3.4. Channel expansion

From Figure 3.2 we can also see how the modularity in systems like the ULA-OP 256, UARP II, and DiPhAS are achieved. The systems include multiple front-end circuit boards, that subsequently include one or multiple AFEs [II] and preliminary processing [III] before the data is transported via a backplane to a central circuit board, which performs further processing and/or transports the data to a workstation. This modularity makes sure that systems can easily be extended in the number of channels, for example, by expanding the maximum number of front-end boards that can be connected to the central board. Therefore, it is important to remember that, to explore a high-channel-count system, provisions have to be made in the hardware architecture of the system. Most systems are thus designed as such that for a fixed number of channels the raw RF samples are available for software-based processing on a workstation but are expandable in terms of the number of channels if additional pre-processing is applied before the interconnection to the main processing part.

Additional scaling can be achieved by scaling out to multiple ultrasound acquisition systems, as also seen implemented by Verasonics, DiPhAS, and the ULA-OP 256. In this way, the required number of channels for the high-channel-count 2D transducers is reached. To run multiple systems in parallel, the synchronization of the core clocks of the systems is a strict requirement for beamforming and Doppler processing as seen in systems such as the ULA-OP 256 [47]. Another consideration to be taken into account is that a server/workstation can only handle a limited number of systems before maximum interface and storage bandwidths are reached, as seen in the DiPhAS system, which is limited to four systems [33].

3.4. Conclusion

In this chapter, a general description of the receiving end of an ultrasound system is given and depicted in Figure 3.2. Recent work on the design and use of ultrasound systems for functional and transcranial functional ultrasound are discussed. Several key factors influence the design that is envisioned. The physical location where digitization and processing take place influences the requirements on the interconnects between the AFE and the processing and between the processing and storage. Additionally, for a portable system, it can be concluded that the processing is inversely related to the degree of portability and thus power/heat dissipation. From the design of the system by Pietrangelo [55] and the Lightprobe [26] it can be seen that noise can be reduced if the proximity of the transducer to the analog-to-digital converter is reduced. Additionally, Xu et al. [86] advise not to introduce multiplexing to reduce additional noise in the receive path of the ultrasound signal. Therefore, if portability is a key design goal, only the number of channels and physical location of processing can be parameterized. Because of the computational load, part of the processing has to be physically separated from the digitization and transducer part due to power and heat limitations. Processing can be software-based, where the data is directly transported to a workstation, which provides all of the processing capabilities, or a system can be hardware-based, where the data is processed in the system using FPGAs. In a hardware-based processing approach the signals can be processed more efficiently, however, this comes at the cost of ease of configurability. Some ultrasound systems use the concept of oversampling and filtering and decimation to increase the SNR of the ultrasound signal. However, without any pre-processing like filtering and decimation the oversampling causes the interconnect to reach the throughput limits of the current-day interconnect technology. Therefore, the processing stage can be split into two parts – pre-processing and processing – and several trade-offs are introduced: 1) a fixed amount of pre-processing can achieve a higher channel count at the cost of information loss. 2) A design that is flexible for the type and parameters of the processing pipeline and thus research but has a limited number of channels because of the throughput that is reached if no pre-processing is applied and the user requires access to the raw RF samples. 3) if memory is added to a system, near real-time constraints cannot easily be satisfied with the addition that the temporary storage also needs to be placed close to the processing stage to retain signal integrity and timing.

4

Design Specifications

This chapter describes how the previous chapters can be taken into account in order to set up the design requirements and subsequently, the design specifications for the envisioned high-SNR TCfUS system prototype from Chapter 1. In Figure 4.1 an overview is depicted of how the preceding chapters provide the necessary information in order to formulate the functional requirements and design specifications, and thus further narrowing down the details of the to-be-designed system. From Figure 4.1 it can be seen that both Chapter 1 and Chapter 2 serve as the foundation for the formulation of the functional requirements in Section 4.1. Chapter 1 provides the vision and motivation for the system and serves as a basis of the requirements for the system, at the same time, Chapter 2 provides us with the theoretical background information. The functional requirements from Section 4.1, together with the related work
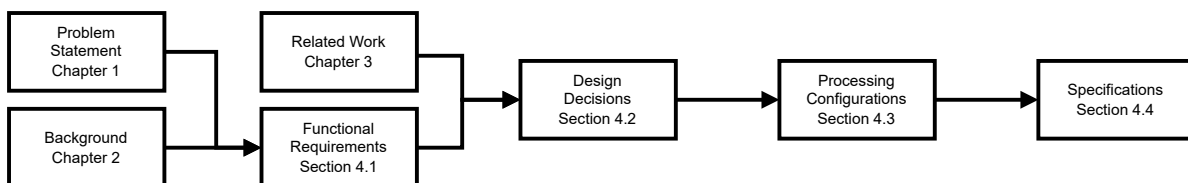

Figure 4.1: Chapter 4 content navigation tree.

from Chapter 3 provide a framework from which design decisions are made in Section 4.2. In the design decisions section, the requirements of the system are narrowed down to a system that is feasible to design, the related works chapter here helps to establish this feasibility aspect. In this section, it becomes clear that not all requirements can be met, or not all requirements can be met at the same time. Therefore, Section 4.3 presents an overview of three processing configurations in order to target three envisioned target applications and their related requirements. Section 4.4 first provides an overview of the system resulting from the previous sections in Figure 3.2. This overview contains all of the components of the to-be-designed system and from here, each component and the connection between the components will be discussed per processing configuration. Section 4.5 discusses the limitations in the system resulting from the previous sections. Finally, in Section 4.6 conclusions are drawn that can aid the subsequent implementation of the system in Chapter 5.

4.1. Functional Requirements

In order to realize a system that can explore the boundaries of the signal-to-noise ratio for transcranial functional ultrasound, a more detailed description is needed of how such a system would look like. With the help of Chapter 1 and 2, a list of functional requirements is compiled to further drive design decisions and specifications.

1. **Requirement 1: High signal-to-noise ratio**
   The main objective of the system is to increase the SNR to the highest point possible, to overcome the problem of the reflection, attenuation, and aberration of the skull, in order to explore the

transcranial imaging of the brain at every point on the skull and thus not to be dependent on the acoustic windows. At this moment it is not known what order of magnitude of SNR improvements are possible, therefore, all design decisions should be made in order to maximize the SNR.

2. **Requirement 2: High spatial and temporal resolution**
A high temporal resolution is critical in order to perform functional ultrasound as seen from Section 2.2 and 2.6. Otherwise, the SNR will be too low in order to spot the small vessels in the brain. In combination with the specifications of the probe, that is already designed and produced, frame rates of 5-10 kHz. In order to be competitive with fMRI, the spatial resolution should at least be on par, and thus be in the order of 1 mm.

3. **Requirement 3: High channel count**
The system will require a high number of channels (1024) to support the option for volumetric imaging. Because a 2D probe requires the squared number of channels compared to a linear probe this extra dimension can provide powerful insights into tracking blood vessel structure in comparison to the conventional two-dimensional images. 1024 channels are chosen because multiple works in Chapter 3 utilize 1024 channel 2D probes for volumetric imaging.

4. **Requirement 4: Flexible processing**
To explore and optimize the parameters and SNR, the signal processing that is used in functional ultrasound has to be flexible and configurable. This flexibility is necessary to further explore and optimize the parameters and type of signal processing that is used in functional ultrasound and SNR can be further increased. Therefore, direct access to the raw RF samples from the transducer should be available to the user with no processing applied. However, the user should also have the opportunity to create a fully integrated processing pipeline on the system.

5. **Requirement 5: Real-time operation**
A near-real-time (<0.5 s lag) operation of the system is essential feedback for the operator for the positioning of the probe. Additionally, in time-sensitive and time-critical environments this requirement is essential.

6. **Requirement 6: High portability**
The system should be portable so it can be used on subjects that do not have to remain perfectly still. Therefore, providing the possibility of conducting experiments with free-roaming subjects.

7. **Requirement 7: Cost-effectiveness**
The system prototype has to be built under $100.000 in order to be competitive against $3M fMRI machine. Therefore, the system should be built with commercially available components.

4.2. Design decisions

In the previous section, the requirements for the envisioned TCfUS system were presented, in the following section, the implication of these functional requirements is discussed with the knowledge from Chapter 2 and the lessons learned from related work in Section 3.4. These implications are weighed against the functional requirements and decisions are taken on which requirements are kept and which requirements have to compromise, it is therefore decided which parameters in the system are fixed and which parameters are left variable.

1. **Decision 1:** In order to create a high SNR system, several techniques from Chapter 3 are implemented in the prototype. First of all, the signal from the transducer will be digitized as soon as possible to reduce noise. Secondly, a state-of-the-art analog front-end is selected in terms of specifications such as SNR, sample frequency, and resolution to allow for the oversampling of the incoming signal. This decision also increases the possibilities of a high-temporal resolution system. Third, because the system is solely focused on SNR, any component that can introduce noise is omitted. This means no multiplexing of the incoming signal but also any other kind of switching, more specifically, the switching between transmit and receive. By designing a receive-only system, the receive path of the system contains a minimum number of components and thus also a minimum amount of noise sources. The transmit side of the system is handled by a Verasonics Vantage 256 system. How these two systems interact and work together can be found later on in Section 5.1.

2. **Decision 2:** From the conclusion of Chapter 3 it can be seen that multiple factors influence the throughput of the interconnections between the blocks sketched in Figure 3.2. These factors also correlate with the requirements listed in Section 4.1. These factors are: flexibility of processing, real-time operation, high channel count, and the decision for high-performance AFEs in design decision 1. If all of these factors are taken into consideration in one design. This poses a problem since state-of-the-art interconnects simply cannot provide that throughput as shown in the following example. As an example, we would like a real-time system that provides the raw RF data, following requirement 4. Given 1024 channels with a sampling rate of 125 Megasamples per second (MSPS) at 16 bits per sample (highest specification component available; see Section 5.2) this would result in $1024 \times 125 \cdot 10^6 \times 16 = 2$ Tbps = 256 GBps. Since there is no realistic way to store and transport this volume of data in real-time whilst also staying portable, some compromises have to be made.

3. **Decision 3:** Processing can reduce the data rate which can make throughput rates realistic for current technology. However, this comes at the cost of an increase in compute and therefore, also power demand. Because of the limited space, due to the portability requirement 6, the power budget is limited and heat limitations play a role. Therefore, a split is introduced in processing as suggested in Section 3.3. With this split part of the processing can be offloaded to a location that does not have any limitations in power and heat. However, the range of flexibility in processing will eventually determine the amount of pre-processing in the first stage. This range of flexibility in processing is necessary because a research system is envisioned and therefore signal-processing parameters are still unclear and need to be explored. Therefore, the user should have access to all the raw RF samples from the AFE. This means there is no processing occurring in the first part of the split design and all the possible information is retained. This scenario provides the worst case when it comes to the throughput of the connection to the second part of the split design. Therefore, this worst-case scenario determines the number of channels in the system, since it limits the maximum throughput that can be produced. Another option would be to have a minimum amount of pre-processing to expand the number of channels, as also seen in the Verasonics Vantage in Section 3.2.5, however, this comes at a cost of loss of information. Since it is unknown how important that information can be and because the focus of this system is on parameter and signal processing exploration, the number of channels is fixed to 64. This number of channels is chosen because it is the minimum to do any kind of beamforming for conventional ultrasound imaging. Additionally, 64 channels also allow for low-channel-count 8x8 2D probes or 64 channel linear probes while maintaining manageable throughput rates of $64 \times 125 \cdot 10^6 \times 16 = 128$ Gbps = 16 GBps with raw RF samples. From Chapter 3 it can be seen that other portable research systems such as LightProbe [27], Pietrangelo [55] and Xu [86] implement a similar number of channels.

4. **Decision 4:** Even though raw RF samples of 64 channels can be transported with current-day interconnect technology does not mean the system is real-time because of potential memory bottlenecks further down the line. Therefore, it is important to take a look at the rest of the cases that are envisioned in the flexibility of processing. Since in these cases, there is pre-processing taking place and thus the volume of data and throughput are reduced the real-time requirement of 5 is possible. However, a trade-off has to be made in the pre-processing between how much and which information is discarded. This is limited by the fact that the power budget is limited, as described in decision 3, and therefore the processing hardware and thus amount of pre-processing are also limited. How the pre-processing is configured and accounts for these requirements presented in requirements 4, 5, and 6 and the trade-off is described in this paragraph. Section 4.3 introduces three scenarios for each different targeted application, each with a specific amount of pre-processing.

5. **Decision 5:** As discussed in decision 3 the system will contain 64 channels, however, by making provisions in the selection of the hardware of the system it would be possible to extend the number of channels and thus going for the latter option suggested in decision 3 with a minimum amount of pre-processing. However, the maximum throughput of the interconnected in the system is still fixed and thus expansion in the number of channels is correlated to how much reduction of data the pre-processing provides.
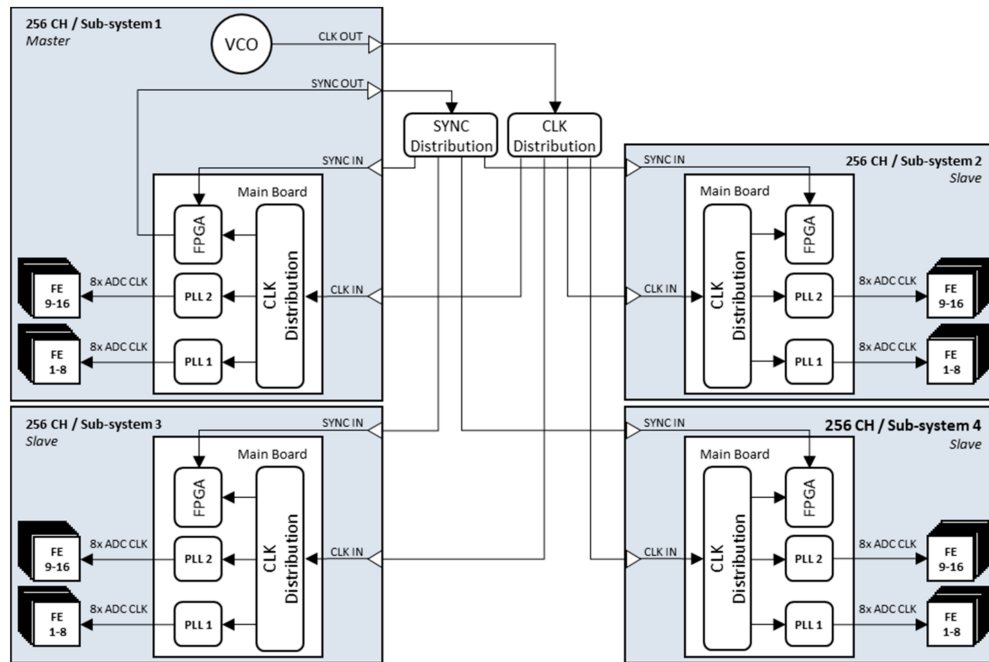
Figure 4.2: Block diagram providing an overview of the synchronization of multiple system clocks of four connected DiPhAS systems, from Hewener et al. [33].

Another form of scaling to increase the number of channels can be accomplished by scaling out to multiple ultrasound systems. As also seen in Section 3.2.4 and Table 3.1. To utilize multiple ultrasound systems in parallel does however require provisional hardware features such as the synchronization of multiple clocks. Otherwise, the machines will not sample in-phase and beam-forming and Doppler processing (which are necessary for TCfUS) are impossible. An example of the scaling out to multiple systems can be seen in Figure 4.2.

6. **Decision 6:** To keep development time short and costs low, the system should be built from off-the-shelf parts and commercially available components. The use of evaluation kits can be predominantly used for prototyping.

4.3. Processing Configurations

From the design decisions in the previous section follows a 64 channel, high-performance ADC, near real-time acquisition system that can store raw RF data samples. However, even though the channel count has been reduced to 64 channels, it is impossible to adhere to all of these requirements at once, since the raw RF data that is produced still is 128 Gbps (taking into account the best AFEs currently available), without taking into account correctional overhead for the transmission of the data. Therefore, processing and storing this data in real-time on a workstation at this throughput is not possible, as can be seen when deriving the specifications in the next Section 4.4. Because of these reasons, three system configurations are presented in which a trade-off is made between the requirements of a real-time system vs. a system that can store raw RF data samples. In each configuration, the processing is applied differently since – with processing – a reduction in data rate can be made. In the processing stages, a trade-off has to be made between the throughput and the precision of the data. The function and the requirements of the three different configurations are as follows:

1. **Configuration 1: Relay system**
   Raw pre-beamformed RF data is relayed to a workstation for storage. No processing is applied to the data in the system and all processing is done afterwards on a workstation. Because of the constraints of the interconnect, this can not happen in real-time. Therefore data has to be temporarily stored in the system. From Chapter 3 it can be seen that locating memory near the processing stage and the type and size of the memory are the most important parameters. Then

a trade-off has to be made between the time required for the recording of the samples to the temporary memory and the time needed to offload the data from the temporary memory to the workstation. These times depend on the maximum throughput and the size of the temporary memory plus the maximum throughput to the workstation.

2. **Configuration 2: Real-time system**
   This system is focused on the real-time requirement 5 whilst maintaining as much raw data as possible so processing can be done in software at the workstation/storage part. Minimal information is lost during the processing. The RF data is then further processed in software on the host PC. The pre-processing that is done has to be matched to the interconnect.

3. **Configuration 3: Hardware-processing system**
   This configuration incorporates a hardware-processing approach with a rudimentary form of the processing pipeline envisioned in Chapter 2 to provide near-real-time feedback to the operator for the positioning of the probe. In parallel to this, the data is sent to the workstation for storage for a more comprehensive approach to beamforming called A-matrix beamforming [40]. This stored data can then be processed at a later moment in time due to the compute complexity. The most important factor in this processing configuration is that there are enough resources available on the hardware side of the back-end system and that the data can be transferred to the workstation at the same time.

To better illustrate the problem of the throughput in each configuration, Figure 4.3 provides an overview of how the throughput changes at each point in the pipeline proposed in Figure 3.2. Since different configurations will each have a different amount of processing applied to the incoming data, the reduction of data throughput in a component can be seen as a downward slope. During the transfer of the data over the interconnects, the throughput will remain fixed.
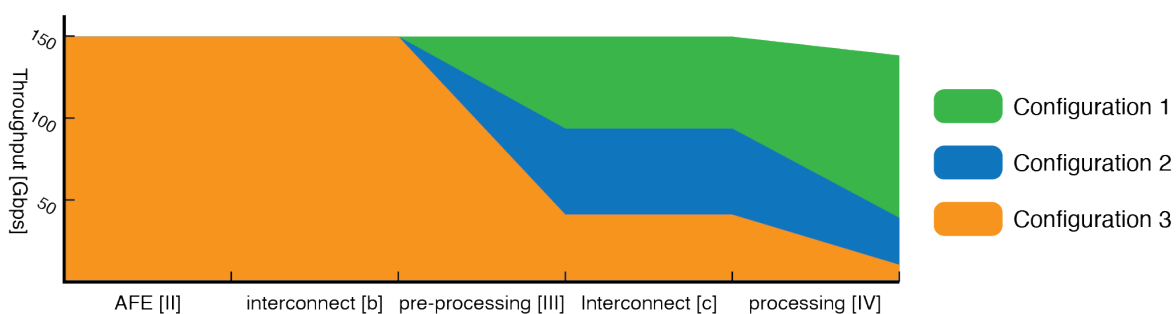

Figure 4.3: Visualization of throughput and processing per configuration and per connection and component

4.4. Specifications

Since the functional requirements and the design decisions resulting from those have been discussed, the block diagram sketched in Figure 3.2 can now be adapted to fit the design goals of the system. Subsequently the updated Figure 4.4 is further explained. Hereafter, specifications of the components [I-V] and interconnections (a-d) can be determined and elaborated upon in the following subsections.

4.4.1. Overview

In Figure 4.4, an updated overview of the system is depicted. A distinct split has been introduced in the system based on the design decisions made in the previous section. This split into two distinct subsystems follows from the first design decision that states that signals are digitized as close to the transducer as possible. In Chapter 3 it was seen that digitization and processing are often combined to decrease throughput and increase expandability in channel count by repeating the design, both contributing to requirement 3. The processing is further stipulated by the fact that the place of processing has to be flexible following requirement 4 and essential for the interconnect handling. However, the portability requirement dictates that, from a design perspective, any additional hardware such as processing hardware should be minimized due to power and heat constraints. Therefore the majority of the

processing should be moved to another location, without relaxed power or heat constraints, with ample resources for compute, temporary storage, and close proximity to the workstation that can provide high-speed interconnect and storage. The result of these decisions is a separation into two subsystems with a critical link (c) in between. The role of two subsystems and the interconnection between the subsystems will now be further clarified:
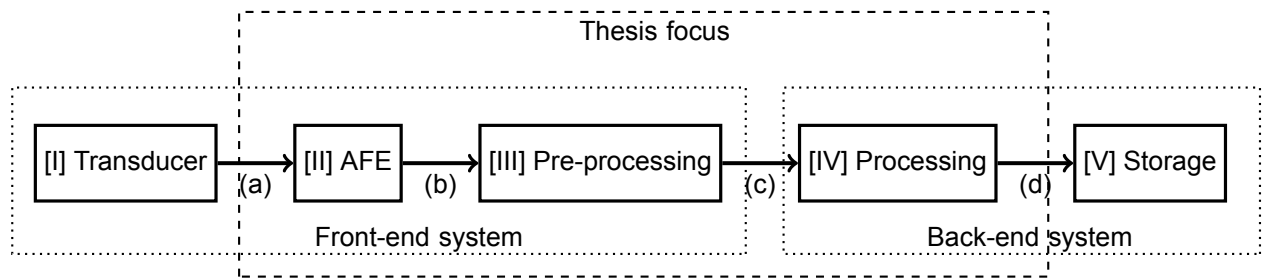

Figure 4.4: An updated block diagram of the receive part of an ultrasound system including proximity of components (black, dotted) and the scope of the components included in the design in this thesis (red, dashed)

- **Subsystem 1: Front-end**
  The front-end system contains the transducer and the digitization, including all the amplification and/or analog signal preparation associated with this. The subsystem also contains a pre-processing stage. The most important part of digital pre-processing is that the interconnect between the front-end and the back-end subsystem can handle the throughput being generated at the front-end. The type of pre-processing is minimal because of hardware constraints.

- **Interconnect (c)**
  This connection connects the front-end to the back-end. Since the portability requirement has been shifted to the front-end system, a long cable (>2m) is needed to transport the data. The throughput that can be attained in this connection is critical for the performance limits of the system.

- **Subsystem 2: Back-end**
  The back-end system contains the majority of the processing and storage. How this subsystem provides the processing depends on the configuration that is used described in Section 4.3. For the back-end, we must find the hardware that can adapt the high-speed interface from the front-end to an interface to the workstation for configurations 1 and 2 but at the same time can also provide the processing capabilities for the hardware-based processing described in configuration 3. For the further processing of raw RF samples using software-based processing, as described in configurations 1 and 2, a workstation with hardware such as high-end GPUs is also required.

With a clear understanding of the functionality of the system and the subsystems, the specifications of the system can be explored. Subsections 4.4.2-4.4.9 describe the functionality and specifications, per component, and per configuration, as outlined in Section 4.3. Because the specifications of each component and interconnect influences the next, the system is discussed in the same order as a received ultrasound signal propagates through the system.

Since the design of a transducer is another expertise and discipline on its own and because workstations are general purpose and can be configured/assembled to handle the target application without any custom hardware, the outline and scope of this thesis is limited to the specifications of components [II]-[IV] and the connections between them. It is however impossible to design such a system without any specifications on the transducer [I] or the storage [IV]. Therefore, these components are discussed but not in the same detail as components [II]-[IV].

4.4.2. Transducer [I]
A custom receive transducer is designed in the department for the to-be-designed system. The receive transducer will be composed of a concave 64-element round probe with a hole in the middle. This middle hole will allow room for the transmit probe. To meet the specification for the imaging depth and

the SNR, the insonification frequency of the receive probe is chosen to be 1.5 MHz with a bandwidth of 2 MHz. The low center frequency helps with the penetration of the skull as described in Chapter 2. This will diminish the axial resolution to an order of a millimeter. To compensate for the low channel count, the size of the probe elements is increased so therefore the aperture size and thus received energy remains the same. This increase in element size will however also decrease the spatial resolution. But with the incorporation of sparse imaging [28] and by using a coded mask for the transmission probe [40], a targeted resolution of 0.5 to 1 mm is foreseen with an imaging depth up to 100 mm and a reconstruction volume of 30x30x30 mm. Since the application will stay the same for each configuration the transducer will be the same for all configurations. The interconnections to the transducer are however considered universal and the transducer can thus be swapped in the future if the application and/or specifications change.

Since the focus of this thesis is on the acquisition system and not on the probe design further details and discussion on the geometry and the design of the probe is considered out of scope for this thesis. Nevertheless, Chapter 5 provides a depiction of the designed probe in Figure 5.2.

4.4.3. Interconnect (a)
The interconnect between the transducer and the AFE should be as short as possible to minimize noise, as mentioned in the first design decision. In this short interconnection, it is still important to focus on noise reduction. Therefore shielded micro-coaxial cabling or ground separated flat flexible cables are used to transport the signals to the analog-to-digital converter. Although this interconnect will not change in all of the configurations, the last option of utilizing flat flexible cables does provide an opportunity for an easy change of the transducer or expansion in the number of channels if more ADCs can be added.

4.4.4. Analog Front-end [II]
If chosen carefully, the specifications of the analog front-end will decide the majority of the SNR increase of the system. Therefore, no compromises are taken and the same AFE will be used in all three processing configurations. As seen in Chapter 3, an AFE does more than just the digitization of the incoming signals. Depending on the AFE, an amount of analog conditioning circuitry is present, as well as application-specific digital processing after digitization. From design decision 1 follows that a high-performance AFE is necessary for reaching to reach the highest attainable SNR. Therefore, for a high SNR ultrasound-specific analog front-end, the following specifications are important:

- signal-to-noise ratio (SNR) [dBFS]
- resolution [bits]
- sample frequency [Hz]
- application-specific circuitry
- channel count
- power [W]
- price [$]

We will now go into more detail on these specifications. The signal-to-noise ratio is one of the most important specifications for the ultrasound system since this determines the sensitivity of the system. Moehring et al. state a minimum SNR of 78 dB at the point of digitization [48] for transcranial Doppler ultrasound with a single element transducer. However, a multi-channel system is foreseen and the same or a higher SNR is required for all the channels. From Chapter 2 it can be seen that SNR for quantization in the ADC is also a function of resolution and sample frequency. Combining this with the oversampling theorem from Section 2.8. An approximation can be made on the maximum attainable signal-to-quantization-noise ratio (SQNR) with the specifications of current commercially available products with $f_s = 125 \cdot 10^6$ and $N = 16$, that is,

$$\text{SNR}_{\text{max}} \approx 6.02 N_Q + 1.76 + 10 \log_{10}\left(\frac{f_s}{2 \cdot B_{probe}}\right)$$

$$\approx 6.02 \cdot 16 + 1.76 + 10 \log_{10} \left(\frac{125 \cdot 10^6}{2 \cdot 2.5 \cdot 10^6} \right) \approx 112.06 \text{dB}$$

This is under the assumptions of the background chapter and an oversampling ratio of 25 times. Where the bandwidth B_{probe} of the signal from the probe is equal to 2.5 MHz. This approximated SNR is still theoretical, commercially available implementations will likely have lower specifications. Thus it is best to aim for an AFE with a minimum of 16 bits accuracy, (2 bits more than current state-of-the-art in ultrasound systems) and the highest available sampling frequency for oversampling. Compared to the current state-of-the-art of the Verasonics Vantage 256 with 14 bits of resolution and a sample frequency of 62.5 MHz, this is a theoretical increase of 15.05dB in SQNR.

It is also critical that the AFE contains ultrasound-specific hardware features, such as a variable input gain to scale the signal to the full input resolution of the ADC. As well as a programmable gain in order to implement time gain compensation. This programmable gain can be accomplished in hardware with an amplifier or later on, in the processing stage. However, having the option of doing it in hardware is beneficial for parameter exploration.

Another common feature of ultrasound-specific AFEs is the ability for internal down-conversion, digital filtering, and decimation. This form of pre-processing [III] is done on the same integrated circuit (IC) and thus can be done early on in the pipeline without additional hardware. This could be beneficial for a) throughput and interconnect b)power dissipation/heat, thus portability c) flexibility in processing

The number of channels available on an AFE determines two factors. At first, the physical size of the IC and thus how much of these ICs can be efficiently integrated onto a PCB. Secondly, the power consumed by the AFE. In general, commercial AFEs are packaged per 4, 8, or 16 channels, where a higher channel count AFE is more efficient in power consumption per channel. This increased power efficiency and channel count is important for the portability requirement of the front-end subsystem since space and power are limited. It is however important to keep in mind that a focus on power reduction does not come at a penalty in SNR.

Following requirement 7 it is important to reduce the price of the system as much as possible. If any significant price reductions can be made without sacrificing SNR, the selection of the cheaper option is evident. Otherwise, the best performance in SNR always has priority, since this is the main goal of the system.

Synchronization and accuracy of the sampling clock for multiple AFE chips is an important specification in TCfUS systems since beamforming would not be possible without synchronization of the sampling clocks and with phase noise, and the time domain equivalent, jitter on the sampling clock, Doppler processing would not be possible. Since beamforming and Doppler processing form the foundations of functional ultrasound, phase shift and the jitter on the main clock are the main performance limitations for the processing of the signals. Therefore, it is important to design a system that provides an equal-length, clock-distribution network and jitter of a maximum of 90 ps [63].

4.4.5. Interconnect (b)

For all configurations, the interconnect depends on the amount of pre-processing and where this is executed. If any pre-processing such as decimation and/or downconversion are done in the AFE itself this connection can be regarded as an internal connection on the IC. Otherwise, if no processing is done the raw RF samples need to be forwarded like configuration 1. In combination with the specifications of the AFE, this will result in a high throughput. Assuming the parameters estimated in the previous paragraph the total throughput from all the channels is equal to

$$T_{sys} = f_s \cdot N_{ch} \cdot N_Q = 125 \cdot 10^6 \cdot 64 \cdot 16 = 128 \text{Gbps}$$

and the data produced by one channel is equal to

$$T_{ch} = f_s \cdot N_Q = 125 \cdot 10^6 \cdot 16 = 2 \text{Gbps}$$

How this interconnect and thus interface is implemented depends entirely on the choice and manufacturer of the AFE. At the moment two standardized interfaces dominate the market: low voltage

differential signaling (LVDS) and the JEDEC standard JESD204 with version C being the latest revision.

In LVDS the data of one channel is often transported over one pair. If any downconversion is done, a separate differential pair can be utilized for the in-phase and quadrature-phase components of the signal. This does result in $2 \cdot N_{ch}$ number of differential pairs per AFE. With these high data rates, it is also good to clarify that the bandwidth per LVDS pair is in practice limited to approximately 1 Gbps [21].

In the JESD204 standard, the data of multiple channels is serialized and transported over a smaller number of high-speed connections called lanes. In version B of the standard, the bandwidth per lane is equal to 12.5 Gbps. JESD allows for the synchronization of multiple devices (i.e. AFEs) using additional signaling. A drawback of the JESD standard is the extra latency that is introduced with the serializing of the data. The JESD204 standard uses 8B/10B encoding for easy clock recovery and DC balance [41]. This does however increase the total throughput of the system to $B_{sys} = 128 \cdot \frac{10}{8} = 160 \text{Gbps}$. For both LVDS and JESD204 the connections are routed using differential pairs on a printed circuit board (PCB) for signal integrity.

4.4.6. Pre-processing [III]

Depending on the configuration the pre-processing stage would need to do the following:

- **Configuration 1:** No processing, all the raw data is preserved. Consequently, the data throughput rate stays the same, approximately 128 Gbps or 160 Gbps, depending on the communication standard used in interconnect (b).

- **Configuration 2:** Filtering and decimation is a critical step in the processing pipeline presented in Chapter 2. Although not depicted in Figure 2.4, the filtering and decimation processing step is critical to reducing the data throughput before the data is transported to the beamforming stage and subsequent processing steps, handled in the back-end subsystem. Pre-processing is limited to these two elements because the processing is basic, and is the same for every channel. For elements later in the pipeline, the complexity becomes larger but also the flexibility constraints weigh higher. For example one might opt for a completely different beamforming processing approach, this has a lot of impact on the reconfiguration of the processing side. However, the filtering and decimation can remain the same. The same approach can be seen in the Vantage 256 by Verasonics in Section 3.2.5. Additionally, the filtering and decimation stage can always be bypassed for Configuration 1.

  In configuration 2 the pre-processing is tuned for maximum retention of information and ideally, processing should be kept at a minimum while reaching maximum throughput. Because of the parameter exploration functionality, the pre-processing (and thus filtering and decimation) have to be fully configurable. From Section 4.4.4 follows that the maximum decimation factor $M = 25$, with the specifications for the AFE and the transducer, is estimated in the same chapter. The decimation block should thus be configurable in the range of $M \in [1, 25]$. As can be read in Chapter 2 the effectiveness of the oversampling and decimation is completely dependent on the filtering that precedes it. The filtering component preceding the decimation block has a minimum number of taps is equal to $N = 16M$ when a symmetric polyphase FIR filter is used. Longer filters are possible with more precision but they also introduce a longer time delay, but depending on the implementation this is negligible compared to the 0.5 s and is in the ns range.

  Hardware implementation for this high-speed digital processing has to meet the requirements stated in the previous paragraph but also the other requirements of the front-end system (i.e. portability). High-speed digital processing is typically handled by: 1) application-specific integrated circuits (ASIC) or 2) field programmable gate arrays (FPGA), because of the data throughput these devices can deliver in comparison to von Neumann architecture based processors on this scale. ASICs however are out of the question since the design should only consist of commercially available components.

  Because of the full configurability of the function of an FPGA and the ability to handle high-speed data signals at a low a relatively low power, this is an ideal candidate for pre-processing. FPGAs

come in different shapes and sizes depending on their capabilities. In order to estimate the type and family of FPGA that is required for our application, some key specifications of the FPGA are listed.

1. **Transceivers** For high-speed interconnect e.g. connections with a throughput higher than 2.5 Gbps an FPGA has transceivers that are optimized for transmitting or receiving. The physical number of transceivers available on the FPGA determines the throughput for interconnect (b) and (c). Estimated from Section 4.4.4 assuming 16 JESD lanes total coming from the AFEs equals 16 transceivers, each handling 10 Gbps. In order to handle configuration 1, the same number of transceivers has to be assumed for the output, totaling 32 transceivers.

2. **LUTs & Flip-Flops** The look-up tables (LUTs) and the registers that store data (Flip-Flops) form the heart of the logic that is implemented.

3. **DSP slices** The digital signal processing (DSP) slice in the Xilinx architecture houses a broad set of features that can be utilized for signal processing. With complex features such as a hardware multiply and accumulate unit, pattern detection, and multiplexing of wide data buses. The number of these slices is important for the implementation of the FIR filters of each channel. The number of DSP slices per channel for an FIR filter can be estimated as:

$$N_{DSP/ch} = \frac{N_{taps}}{2} \cdot \frac{f_s}{f_{clk}} \cdot \frac{1}{M}$$

With N_{taps} equal to the length of the FIR filter, f_s equal to the sampling frequency of the incoming signal, f_{clk} equal to the clocking frequency of the filter, and M equal to the decimation rate.

4. **BRAMs** BRAMs can be used to store filter coefficients for the FIR filter and in multiple configurations, this filter and decimator can be used.

With these five building blocks, all the digital logic in the FPGA is constructed. An FPGA design can be separated into different parts where each block is responsible for a specific task, these blocks are called IP-blocks. In the application of pre-processing these blocks would be: decoding of the incoming data bus, a block that does both filtering and decimation in one and encoding the data to send the data to the back-end subsystem. Table 4.1 contains a resource utilization estimate of these IP-blocks assuming a Xilinx family architecture.

IP-block	LUTs	Flip-Flops	DSP slices	BRAM blocks*
JESD204B 4L RX	2360	3596	0	0
1ch 102 tap FIR, D=5	358	923	11	0
64ch 102 tap FIR, D=5	22912	59072	704	0
JESD204B 4L TX	1942	1674	0	0

Table 4.1: Estimate of resources based on Xilinx architecture. *BRAMs estimated as 1×36k BRAM and 0.5×18k BRAM

In between the custom ASIC and the FPGAs, there is also a third option. There are also application-specific analog front-ends available that provide the option for the processing steps of digital filtering and decimation. In comparison to an FPGA implementation, these AFEs are configurable in terms of filtering and decimation characteristics but their specification range is fixed. This option will be further explored in Chapter 5.

• **Configuration 3:** In the third configuration the same pre-processing is applied to the signal as in configuration 2. Because of the data reduction, this presents opportunities in terms of channel expansion, if there is hardware to accommodate this. The demands for processing however may change in the future with new insights and therefore the flexibility exists to also negate pre-processing entirely and shift all the processing to the hardware of the back-end subsystem.

4.4.7. Interconnect (c)

With interconnection (c) the front-end subsystem and the back-end subsystem are connected. Without this connection, the front-end subsystem would not be portable even though it might have a small form factor. Therefore, an important specification of this connection is the length which is approximately 2-5 meters. Depending on the configuration of the system, the data throughput in this connection is equal to either the raw RF generated equal to 128 Gbps, or 5-64 Gbps if processing like decimation by a factor of 2-25 is applied, as described in the previous subsection. Together, length and throughput are conflicting factors when it comes to high-speed interconnect. In coaxial cables, with the increase of the frequency of the signal comes an increase in attenuation per unit distance. Because of the high frequencies and thus fast switching characteristics the dielectric and thus capacitive losses take over and the signal-to-noise ratio of the signal decreases with distance and frequency. Following the Shannon–Hartley theorem the channel capacity is proportional to the bandwidth of the channel and the logarithm of SNR. Even when using differential pair signaling, a way to reduce the noise on a transmission line, the distances and throughput that are envisioned cannot be met. For this reason, long-distance high-speed interconnect is often transported via fiber optics. Fiber optics do not suffer from these impedance problems and can carry very wideband signals into the gigahertz range without an increase of attenuation as compared to copper cabling and avoiding electromagnetic interference. The added benefit is that multiple fibers can be used alongside one another without introducing crosstalk due to electromagnetic coupling. A drawback of using fiber optics however is the time that is needed for the conversion from the electrical domain to the optical domain and back. However, for small distances (≤10m) this is equal to around 3 ns.

4.4.8. Processing [IV] & Interconnect (d)

The digital data from the front-end subsystem now arrives at the back-end subsystem. The goal of the processing component is to create a power Doppler image (PDI). How this happens and where the signal processing takes place depends on the configuration. The purpose of the processing component, however, is more than only processing of the pipeline described in Figure 2.2. Since the processing component should also provide a hardware interface to which the front-end interface and the workstation can connect to. Since both of these connections need to be high-speed interfaces, a sensible choice would be to use an FPGA to handle these tasks as also described in Section 4.4.8. Because of the different configurations described in Section 4.3, the processing component consists, in terms of hardware, of an FPGA and a workstation. Thus depending on the configuration, the FPGA is either leveraged for processing or the processing is situated on the workstation and the FPGA only acts as an adapter of interfaces.

Therefore, each configuration and use case is described in the following three sections, so an estimation of resources and specifications can be made.

- **Configuration 1:** In this configuration the raw RF data from interface (c) from the front-end subsystem to the workstation. No processing is done on the FPGA. Therefore, the FPGA acts as an adapter of interfaces. To interface with the workstation, which can also be seen as interconnect (d), a limited number of options are available. PCI express is a straightforward choice since this is an interface that can handle high throughput data (31.508 GBps for gen 4.0 x16) at low latency (300 ns [52]). High-speed network interfaces are also an option but these interfaces are eventually also converted into PCI express to communicate with the workstation. Therefore, the former option which communicates directly with the workstation is preferred.

  In contrast to the highest current standard of PCI Express gen 5.0 x16 and the FPGA which incorporate hardware for this standard, Xilinx does not produce evaluation boards incorporating the gen 4.0 x16 standard and is limited to gen 4.0 x8 or gen 3.0 x16 providing the same maximum throughput of 15.754 GBps. An estimation of the raw RF data samples from the front-end is approximately 128 Gbps which is higher than the PCIe 4.0 x8 interface of 15.754 GBps = 124.23 Gbps [52]. And this is not taking into account the 128b/130b encoding, which means a 1.53% overhead in data rate. In Chapter 5 this bottleneck will be discussed more detail.

  Since the PCIe interface cannot keep up with the input data rate, the data has to be temporarily stored. Since FPGAs only tend to have small amounts of on-chip storage (34 Mb) [82], off-

chip storage such as Double Data Rate Synchronous Dynamic Random-Access Memory (DDR SDRAM) has to be incorporated into the design including a memory controller. Therefore, a structure like in Figure 4.5 has to be implemented. Because of the mismatch in interface data rates, this also means the system has to switch between storing data on the DDR4 memory and offloading the data to the workstation. Therefore, continuous recordings are not possible. The length of the recording depends on the size and speed of the memory and the memory controller.

As soon as the raw RF data is transferred to the main memory of the workstation processing can be applied to the data. As seen in Chapter 3 high-end GPUs are leveraged to parallelize the processing tasks of the processing pipeline described in Section 2.2. An estimation is that with plane wave compounding real-time B-mode imaging will require approximately 50 Tflop/s [73]. An assumption is that processing can not be done in real-time, since a top-of-the-line GPU like the Nvidia A6000 has 38.7 TFLOPs of single-precision performance [51]. Therefore the raw RF data samples have to be buffered in the system memory of the workstation, other strategies for storage will become apparent in Section 4.4.9.

An estimation of the resources needed on the FPGA by acting as an adapter of interfaces is shown in Table 4.2.

IP-block	LUTs	Flip-Flops	DSP slices	BRAM blocks*
JESD204C 4L RX	2,360	3,596	0	0
JESD204C 4L RX x4	9,440	14,384	0	0
AXI DMA s2mm**	6,386	10,460	0	20
PCIe 3.0 x16 MM	33,565	26,101	0	45
PCIe 3.0 x16 QDMA	89,726	78,880	0	99

Table 4.2: Resource estimation configuration 1 on a Versal Core AI platform by Xilinx, memory controllers are integrated hardware blocks on the Versal architecture, *BRAMs estimated as 1×36k BRAM and 0.5×18k BRAM, **based on Xilinx Ultrascale Plus architecture

- **Configuration 2:** In this configuration the FPGA also acts as an interposer of interfaces. However, in configuration 2 the throughput has been significantly reduced by the pre-processing stage, and therefore any buffering and thus latency can be reduced to a minimum. The data is transferred from the PCIe interface of the FPGA directly to the main memory of the workstation. After which the same high-end GPUs as in configuration 1 can now process the data further into power Doppler images in near-real-time (≤0.5s).

- **Configuration 3:** In this configuration of the system, the FPGA handles the processing of the blocks 2-4 of the processing pipeline described in Chapter 2, Figure 2.4. Because the processing is done in hardware on the FPGA, this allows for near-real-time performance and rudimentary feedback to the operator for the positioning of the probe via an external monitor. In parallel to this, the data can be sent to the workstation for a more comprehensive approach to beamforming
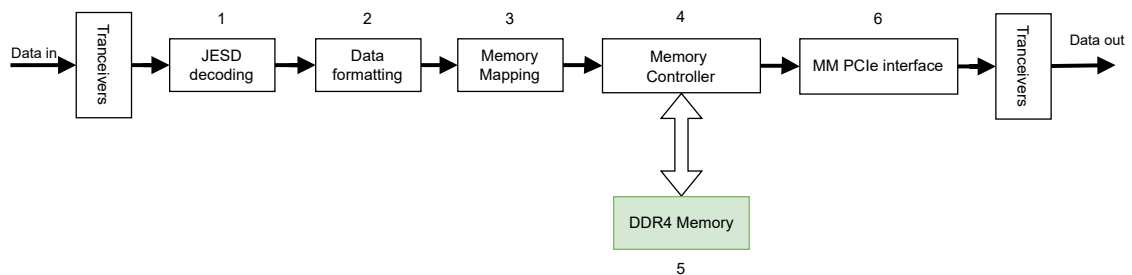

Figure 4.5: Block diagram illustrating the flow of data in the FPGA in configuration 1

Supported Channels	LUTs		Flip-flops		DSP slices		BRAM blocks	
64	120,473	49.7%	108,111	22.3%	139	7.2%	588	49.0%
32 × 32	206,040	85.0%	167,256	34.5%	139	7.2%	1166	97.1%

Table 4.3: Adapted from [34], based on Kintex UltraScale KU040 implementation results

called A-matrix beamforming [40]. However, the computational complexity of this technique is too high, and therefore it is not feasible to implement this technique on the FPGA and the processing can be done at a later moment in time, distributed over a cluster with sufficient resources. The data is transferred to the workstation in the same way as in processing configuration 2.

In order to get an estimation for the resources of the beamformer and next blocks in the processing pipeline, we take a look at current work in ultrasound beamforming [4, 34]. In this work, an ultrasound image processing pipeline is designed for 64 channels including TCG, apodization, delay and sum beamforming, log compression, and scan conversion. This design can be scaled up to 1024 channels or more depending on the resources available on the FPGA. An overview of the resources used in this design can be seen in Table 4.3. From Table 4.3 it can be seen that the design scales non-exponentially in terms of resources with an exponential increase in the number of channels. The number of DSP slices used in the design even remains equal.

With a clear overview of the purpose and resource estimation of all of the three configurations, some final remarks on the specifications of the processing component.

It is important to remember that the goal is to design a research ultrasound system. Therefore the flexibility of the processing (pipeline) and exploration of the parameters has to be present in the design and thus the amount of resources on the FPGA has to be ample to provide for this flexibility. Table 4.3 and 4.2 provide a good starting point with a minimum amount of resources. But it is best to find an FPGA that can easily accommodate these resources and provides enough resources for expansion.

With the last two configurations, the number of transceivers available on the FPGA is important. Since a fixed number of the transceivers is assigned to the PCIe interface to the workstation the rest of the transceivers are available for the connection to the front-end system. Because configurations 2 and 3 provide a significant data reduction in the pre-processing, the data throughput and thus also the number of transceivers required is diminished. This allows for a potential scalable design in terms of the number of channels. So apart from the minimum number of transceivers set by configuration one, a high number of transceivers is beneficial for such a scalable design. This does, however, has to be taken into account in the choice of FPGA but also in the board that is chosen where the FPGA resides on since the transceivers have to be brought out to connectors or points that can be connected to the front-end subsystem.

The design can be scaled out further if need be to include more channels. Here multiple back-end subsystems will be required. A tactic that is also seen in several systems in Chapter 3. Since the PCIe bus has become very congested with traffic and system memory has a limited capacity in (sequential) write speeds. A scaled-out version of the system has to be distributed over multiple workstations. And as mentioned in Section 3.3 the synchronization of these systems is critical in order to enable power Doppler processing, therefore, the FPGA has to provide support for external clocks as an input.

4.4.9. Storage [V]
Storage is critical for the processing of the data at a later point in time, or also called offline processing. Therefore, in configuration 1, we are at a crossroads. Either data can directly be processed but because of the volume of data, this can not occur in real-time. Or the data has to be stored without any processing. The latter case poses a new problem. The data can be stored in the main memory if DDR4 or DDR5 memory is used [16]. The main memory of desktop-grade workstations is often limited to 128 GB, server-grade CPUs can provide an outcome and go up to 2 TB of main memory but come

at the cost of requiring more expensive error checking code (ECC) RAM. However, with this amount of memory potentially storing up to

$$t_{rec} = \frac{N_{storage}}{N_{ch} \cdot N_Q \cdot f_s} = \frac{2 \cdot 10^{12} \cdot 8}{64 \cdot 16 \cdot 125 \cdot 10^6} = 125\text{s}$$

of raw RF samples in recording, without taking into account the overhead of a file system, operating system, and other factors. Where $N_{storage}$ equals the total number of bits available for storing data samples.

The data can also be directly written to a solid-state drive using the DMA and PCIe. But the performance of a top-of-the-line PCIe 4.0 SSD can only handle a throughput up to 6,200 MBps of sequential writes [38] and therefore does not suit this purpose unless a redundant array of independent disks (RAID) is introduced. This RAID array, however, has to be managed by the FPGA and the workstation in order to also do processing on the data or distribute the data. This problem is out of the scope of the thesis and will be left for future work.

For configuration 2 and 3 the data rate has been diminished enough for real-time writing due to the processing and or pre-processing.
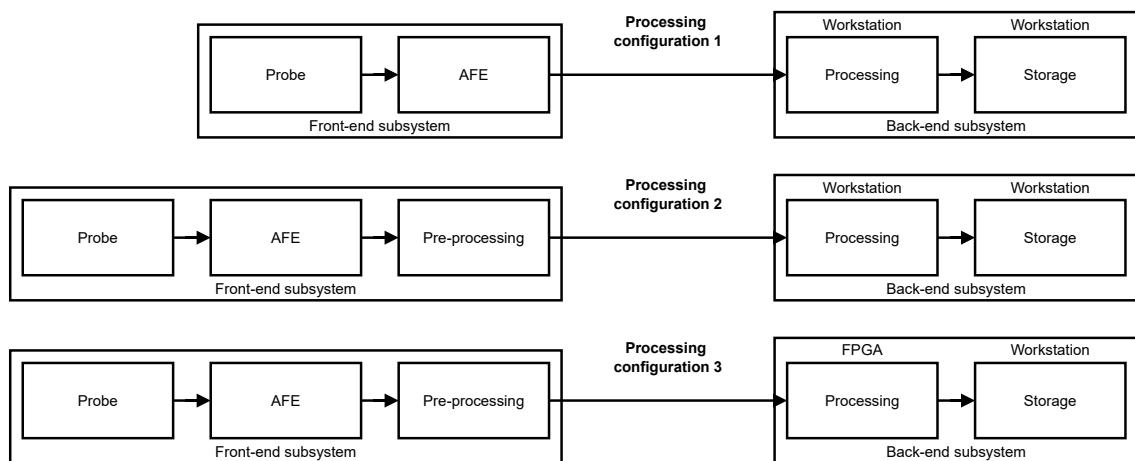

Figure 4.6: An overview of the integration of the components in the TCfUS acquisition system in the three different processing configurations.

4.5. Discussion
Theoretically, the best SNR in a system can be achieved using the most precise AFEs placed directly near the transducer elements. However, in the specifications that are derived in this chapter, some limitations are introduced. Due to the maximum specifications of current commercially marketed AFEs, the number of AFEs that can realistically fit in the front-end subsystem is limited because of the power and heat constraints due to portability. Therefore, the number of channels is limited to only 64. However, this is compensated by 1) the fact that the AFE is situated close to the probe 2) the oversampling factor of 25x 3) 2 additional bits compared to the state of the art in ultrasound 4) the fact that no transmit electronics or multiplexing is introduced into the receive circuitry. Since the parameters in the probe design are set (such as the element size and the number of elements) which form the basis of the trade-off from axial and lateral resolution vs. the SNR and the penetration depth of the signals. With traditional techniques, the spatial resolution would therefore come down to 1 mm and a temporal resolution to 900 fps (imaging depth 5 cm and 16 angles). However, with the custom-designed receive probe and the use of a coded mask for the transmission probe [40] a spatial resolution of 0.5 mm is foreseen.

One of the other limitations in the system is the throughput. In order to store all the raw RF data

for future processing, described by configuration 1, the data that is produced causes a significant challenge for storage and interconnect. Additionally, with every solution, some latency is introduced. In the JEDEC JESD204 standard data is serialized, sent over a high-speed interconnect, and deserialized. With the distances that are envisioned the data has to be transported via fiber optics and thus has to be converted from the electrical domain to the light domain and back introducing approximately 3 ns. With the long filters which are required the decimation process in pre-processing using a $25 \times 16 = 400$ tap filter a latency of $\frac{400}{125 \cdot 10^6} = 3.2 \mu s$ is introduced. After which the FPGA also introduces latency with the scheduling and mapping of the data and the transfer of the data via DMA and PCIe to the workstation.

In the software-based approach of configuration 2, the majority of the processing is done in software and determined by high-level languages. The clear benefit of this approach is that it is very flexible for ultrasound researchers to change and tweak the algorithms and processing they are using to develop better ultrasound imaging techniques. In configuration 3 the processing is handled by an FPGA which has a clear benefit in processing time, but configuring an FPGA is not done in high-level languages but in hardware description language (HDL). High-level synthesis (HLS) is an upcoming trend where high-level languages can be used in FPGAs. However, this needs a framework already in place to incorporate IP-blocks written in this way. Thus, to incorporate both approaches an architecture on the FPGA needs to be incorporated that can stream the raw RF data to the workstation whilst also being flexible enough to incorporate processing IP-blocks. Created from scratch using HDL or using HLS.

The current design is not expandable in terms of the number of channels if processing configuration 1 is used. Since the sampling frequency and accuracy are specified, we reach the maximum of the buses such as PCIe 4.0. Therefore, with this configuration, the design can be scaled out and thus multiplied several times however it is important to then provide clock synchronization between systems. Since in configuration 2 and 3 the throughput is significantly reduced this can provide options for expandability in the number of channels.

4.6. Conclusion

Requirements are set up from which design decisions have been drawn. The design requirements of the system are limited to a receive-only system with 64 channels and split into 1) a front-end subsystem containing the AFEs and pre-processing and 2) a back-end system containing the processing and storage. This split is made to relieve power and heat constraints off the front-end system. The interconnect between these subsystems is realized with a fiber optic link due to length and throughput specifications. The system can be configured into three modes of operation also called processing configurations to provide different functionality to the user. Ranging from non-real-time raw RF data storage for exploration of processing parameters to a near real-time solution for direct feedback to the operator. Depending on the processing configuration described in Section 4.3, the configuration inside the FPGA will change, as does the throughput to the workstation. However, the hardware used for all the processing configurations has to remain the same. Processing plays a large role in the design of the system to balance throughput and accuracy, and therefore the envisioned processing pipeline is split up into a pre-processing part containing filtering and decimation, and a processing part that contains beamforming followed by subsequent processing steps for Doppler processing. Depending on the configuration, the secondary processing stage in the back-end subsystem is executed either on the workstation or the FPGA that handles the signals coming from the front-end subsystem. In either case, the incoming samples from the FPGA are transferred to the workstation and buffered in the system memory, to be later stored in a storage raid of solid-state drives. It is critical that the system is designed in the same way the data flows when receiving an ultrasound pulse since the specification and interface of every component influences the next component.

<div align="right">

5

</div>

<div align="right">

System Design

</div>

In this chapter, an architecture is presented based on the findings of Chapter 3 and 4. The end goal is to design a single hardware solution that adheres to all of the specifications of Section 4.4.

First, in Section 5.1, an overview is provided on how the acquisition system fits in the envisioned application setup, including the Verasonics system for providing the transmit signals. Thereafter, in Section 5.2, the components for the envisioned acquisition system can be selected. The architecture and implementation of these components are then presented in Section 5.3. Section 5.3 will also include any extra hardware that is introduced in order to realize the implementation that is envisioned. It is important to note that any work regarding the design, and the place-and-routing of all the printed circuit boards (PCB) presented in this chapter is carried out by the skilled researchers and technicians at the TU Delft, Department of Microelectronics, Electronic Instrumentation Laboratory. The selection of the hardware and the architecture behind the system is the main focus of this thesis. With the complete implementation described, more detail can be provided on the configuration of the FPGA in Section 5.4. However, since the parameters and specifications of the processing are still unknown at the moment of this writing, the resources and architecture in the FPGA can only be approximated at a high level. Therefore, Section 5.5 will provide a theoretical approach for an expansion of the number of channels in the designed system, since increasing the number of channels that are available to the user can increase the image quality, and foremost allows for volumetric imaging. The focus of Section 5.5 is on the hardware selection, matching of interfaces, and pre-processing using decimation filtering, but not on the scaling of processing implementation on the FPGA. The most important limitations in this chapter are then discussed in Section 5.6 and finally Section 5.7 concludes on the main findings in this chapter.

5.1. Overview

Because design decision 1 states that the main focus of this system is maximizing the SNR, any multiplexing and/or transmitting hardware is omitted. Therefore, a secondary system is required to provide these transmit capabilities, since the ultrasound modality is physically dependent on this interaction of transmitting and receiving sound waves. Subsequently, in this chapter when the term *acquisition system* is used, it will refer to the receive system that is to be designed, otherwise transmit system or Verasonics will be used. The secondary ultrasound system that will provide transmit is the Verasonics Vantage 256. In Figure 5.1, the setup including both ultrasound systems can be seen.

Each system is connected to a separate probe, where the receive probe is a custom-designed probe described in Section 4.4.2 and the transmit probe is a commercially available 2D probe. For a better understanding of how these two probes geometrically relate to each other, see Figure 5.2. The concave round receive probe provides a recess in the middle which allows room for the transmit probe. The split in the receive and transmit path continues further. The transmit probe is directly attached to the Verasonics Vantage 256 using microcoaxial cabling. But the receiver probe is connected as soon as possible to the analog front-end ICs. After the digitized raw RF samples are transported via fiber optics
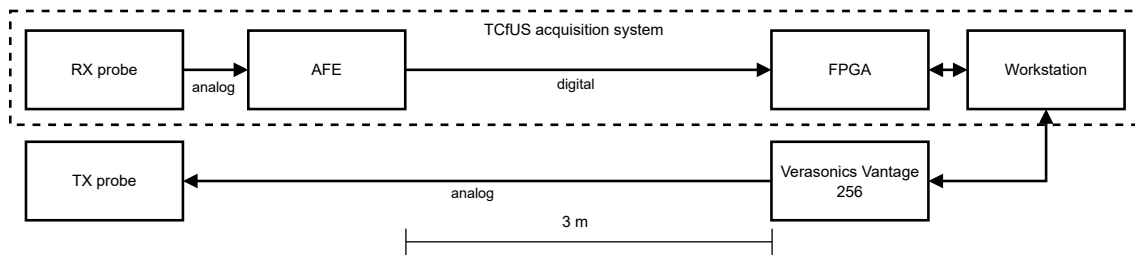
Figure 5.1: Setup of the TCfUS acquisition system, that designed in this chapter, and the interaction with the Vantage 256 by Verasonics.

to the back-end subsystem. Both ultrasound systems communicate with the same workstation, which provides control and management to both systems. To synchronize, the transmit and the receive systems are both connected using a trigger signal which is synchronized to the clock of the receive system which is also shared by the core clock of the Verasonics. The implementation of this synchronization can be read in Section 5.3.2.
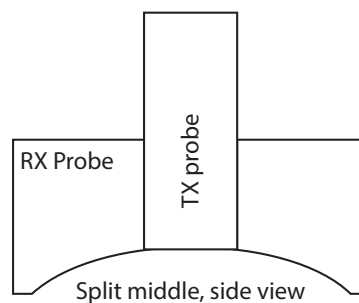

Figure 5.2: A cross-sectional view in the x plane of he receive and transmit probe. The round concave receive probe has a recess in the middle which allows room for the transmit probe.

5.2. Component Selection

With a clear view of the intended setup and the specifications, the hardware components of the TC-fUS acquisition system can be selected. Since the hardware of the acquisition system is critical for the performance, and thus the SNR, no shortcuts can be permitted. However, it is also important to keep in mind the interconnectivity of all of the components. Therefore, a systematic approach is made based on the interdependence of the components and thus the selection order of the components is predetermined. The selection starts with the components of the front-end subsystem, which sets the exact throughput specifications for the rest of the system and consists of: the transducer[I], analog front-end[II], and digital pre-processing[III], as described in Section 4.4.1. Since the transducer is already custom-made, there is no selection to be made here. Therefore, the AFE selection will be discussed first in Section 5.2.1, because of the associated throughput and interface and the dependence of subsequent components on this selection. In this section, it will also become apparent that, with the features of current top-of-the-line AFEs, no extra hardware is necessary for pre-processing. Secondly, the throughput generated by the front-end subsystem has to be matched to the back-end subsystem and therefore the FPGA is the next hardware component to be selected. Because integrating an FPGA on a custom-designed PCB is a time-intensive job and requires a lot of expertise, an off-the-shelf evaluation board is chosen in Section 5.2.2. With the components selected for both subsystems, the interface between the two subsystems, interconnect (c), can be selected in Section 5.3.3. Specifics on the integration of all of the components will be discussed in the next section, Section 5.3.

5.2.1. Analog Front End (AFE)

From Chapter 4, two main conclusions were drawn for the selection of an AFE for a transcranial functional ultrasound acquisition system. The first conclusion was, to build an as high as possible signal-to-noise ratio system the main focus of the specifications of the AFE should be the SNR, resolution,

and sample rate. Therefore, the analog front-end plays a crucial role in the architecture of the design because of the two assumptions made in Section 3.3, specifically: positioning the AFEs close to the transducer and selecting an AFE with the best specifications available with the intention of using over-sampling to increase the SNR.

The second conclusion was the portability requirement, which manifests itself in the specifications of the number of channels per AFE, the power each AFE requires, and the interface, where the interface is also the result of the specifications of the first conclusion. Both of these conclusions are taken into account when selecting the AFE in this section.

Analog front-end ICs come in very different shapes and sizes. There are, for example, FPGAs with built-in ADCs and analog front-ends such as the Xilinx Zynq RFSoC series. This would be quite ideal for our situation given the specifications and the ability to do pre-processing in the front-end subsystem with the flexibility of changing the pre-processing steps. However, because of the limited number of channels per RFSoC device, size, power consumption, and price, it is not suited for the application that is envisioned. Therefore, the selection of the analog front-end ICs has been narrowed down to three ultrasound-specific analog front-end ICs, the specifications of which are listed in Table 5.1. Consequently a comparison is made between the three AFEs listed in Table 5.1 and the specifications stated in Section 4.4.4. After the comparison one of the AFEs is selected and this choice is further substantiated.

Specification	AD9671[2]	MAX2088[46]	AFE58JD48[1]
Channels	8	16	16
Resolution [bit]	14	14	16
Sample rate [MSPS]	125	65	125
ADC SNR [dBFS]	69*	81**	80***
Interface	JESD204B	LVDS	JESD204B/LVDS
Power [W]	1.2	2.45	2.24
Price [$]	90	178	373,22

Table 5.1: Comparison between top ultrasound AFEs on the market from commercial suppliers. *Full channel characteristic, f_{in} = 5 MHz at −12 dBFS, VGAIN = −1.6 V, ** f_{in} = 5 MHz and 2 MHz bandwidth and only for ADC, ***ADC Idle-Channel SNR 16-Bit, 125-MSPS Mode

The minimum of SNR at the point of digitization is stated in the specifications (Sec. 4.4.4) at 78 dB [48]. Together with the principle of oversampling an increase in SNR is expected, depending on the over-sampling ratio. Therefore the AFE58JD48 (Texas Instruments, Dallas, Texas, USA) best fits these requirements. With almost double the sampling frequency of the MAX2088 (Maxim Integrated, San Jose, California, USA) and two extra bits in resolution. The MAX2088 does promise 1 dB extra in SNR at the point of the ADC. However, it is difficult to compare these three chips on this metric since every AFE is tested with different setups and circumstances.

High channel count AFEs are in general more energy efficient compares to using multiple AFEs with a lower channel count. The AFE58JD48 has a clear advantage in this aspect over the other two AFEs, with the amount of power used per channel being 0.14 W/ch. Combined with the fact that higher channel count AFEs can provide for more compact designs.

One of the most important specifications that are not captured in Table 5.1, is the analog signal conditioning and digital pre-processing executed AFE. All of the above mentioned include analog signal conditioning features specific to ultrasound including low noise amplifiers, variable gain amplifiers, anti-aliasing filters, selectable active input impedance, and time gain compensation. When it comes to digital pre-processing the specifications of the AFEs begin to differentiate from each other. Both the AD9671 and MAX2088 provide a $16 \times$ M tap decimation filter with programmable decimation between $2 \leq M \leq 32$ decimation with a coefficient size of 14 bits. The AFE85JD48 can provide decimation in the range of $2 \leq M \leq 64$ with the extra benefit of also providing fractions in between integers of:

0.25, 0.5 and 0.75. Depending on, if and which fraction is used, the filter is $16 \times M$, $32 \times M$ or $64 \times M$ with filter coefficients of 14 bits. The coefficients used by the filter can be default values generated on-board or the coefficients can be configured using the SPI bus. By taking the specifications of the hardware-implemented decimation filtering into consideration and by comparing it to the specifications, it can be concluded it is not necessary to implement any further hardware for pre-processing, since these chips provide both the precision and the reconfigurability at a lower power usage and complexity compared to an additional FPGA.

In high-speed ADCs and AFEs, low voltage differential signaling (LVDS) is the most common interface to communicate data to an FPGA. One LVDS pair per channel is then utilized to transport the signal to the next device, often an FPGA. However, as mentioned in Chapter 4 the throughput per differential pair is in practice limited to ~1 Gbps [21]. Therefore with the increase in resolution and sampling frequency seen in the AFEs in Table 5.1 the output data rate will increase above this 1 Gbps per channel threshold. Therefore it can be seen that the AD9671 and the AFE58JD48 utilize a JESD204B interface. Here the data of multiple channels are aggregated into one high-speed interface that consists of multiple lanes where in version B of the standard each lane can provide a throughput of 12.5 Gbps[21, 88]. The AFE58JD48 still also supports LVDS, however, is then limited to 80 MSPS instead of the full 125 MSPS. JESD204B has a clear advantage when it comes to the implementation of the interface between the front-end and the back-end subsystem as will become clear in Subsection 5.3.3.

In terms of pricing, the AD9671 (Analog Devices, Wilmington, Massachusetts, USA) is the clear winner. With a factor of two in regard to the amount of dollars per channel to the AFE58JD48 and almost equal to the MAX2088. However, by taking the most expensive chip, the AFE58JD48 to create a 64 channel system approximately $1500 is needed. Therefore, the costs for the analog front-end would only be a fraction of the cost for a whole ultrasound system which is typically priced at $100,000 for a commercially available unit.

By taking all of the above factors into consideration the AFE58JD48 from Texas Instruments is deemed the best candidate for the application of TCfUS that is envisioned and is therefore selected as the AFE to be used in the ultrasound acquisition system. Since the AFE58JD48 is the best performing AFE in terms of sampling rate and resolution, equal or better compared to the rest of the AFEs, has the best power per channel performance and channel count is the most comprehensive in terms of digital pre-processing and provides both an LVDS and a JESD204B interface. When it comes to the price, this is where the only disadvantage comes into play, however, this disadvantage can be considered negligible compared to other components selected later on in this chapter.

5.2.2. FPGA and FPGA Board

The FPGA in the system is best captured in the component of processing [V]. However, as Section4.4.6 describes the FPGA is not always used for processing but also as an adapter of interfaces to get the data to a workstation for further processing and storage. Because of the processing configurations described in Section 4.3, the configuration of the FPGA changes per processing configuration, as does the throughput to and from the FPGA. However, the hardware used in the back-end subsystem in all these processing configurations has to remain the same, as concluded in Section 4.6. Therefore, an FPGA has to be selected that can support all three configurations described in Section 4.4.6, with room to spare for future parameter and processing explorations. The FPGA needs one hardware setup for all the processing configurations. However, in order to reduce the design time and costs, the obvious choice for implementing and developing with an FPGA is with an evaluation board. Either from the FPGA manufacturer or a third party. With the choice for an evaluation board, the options become more limited in terms of interfaces, memory, etc. Therefore, a trade-off has to be made between the defined specifications and the specifications of the available FPGA boards. In this thesis, the focus is to realize a hardware architecture for the back-end that can be used in all processing configurations while using an FPGA evaluation board.

In terms of evaluation boards, both Xilinx/AMD and Altera/Intel provide an extensive lineup of FP-GAs and FPGA evaluation boards in the class that is applicable for the envisioned system. However, since a majority of the produced IP and hardware at the lab is based on the Xilinx platform, Xilinx is the

favorable vendor. Xilinx in particular has an extensive lineup of evaluation boards that provides a large number of options in terms of high-speed connectivity. In Table 5.2 four evaluation boards are listed based on the specifications of Section 4.4.6, all of these evaluation boards are manufactured by Xilinx. Third-party boards are left out of scope due to the often limited support. A fifth column is added specifying the minimum resources required by the FPGA based on the specification of the components listed in Section 4.4. The specifications can then be compared against the available boards consequently the FPGA can be selected.

| FPGA family | Virtex 7 | Kintex UltraScale | Virtex Ultrascale+ | Versal AI Core | Minimum |
Evaluation board	VC707	KCU105	VCU118	VCK190	specifications
Transceivers FMC 1	8 GTX	8 GTH	24 GTY	12 GTY	16 total
Transceivers FMC 2	8 GTX	1 GTH	-	12 GTY	
→ Gbps/Transceiver[†]	12.5	16.3	32.75	32.75	12.5
PCI-E	gen 2.0 x8	gen 3.0 x8	gen 3.0 x16	gen 4.0 x8	gen 4.0 x16
→ bandwidth [Gbps][†]	32	63.016	126.032	126.032	252.064
I/O pins	700	520	832	770	
Logic cells	485,760	530,000	2,586,000	1,968,000	418,428
DSP slices	2,800	1,920	6,840	1,968	843
AI Engines	-	-	-	400	-
BRAM [Mb]	37	21.1	75.9	34	20.98
RAM	1GB DDR3*	2GB DDR4**	5GB** DDR4	8GB* DDR4	16 GB DDR4
				8GB** LPDDR4	
price [$]	3,495.00	2,995.00	8,394.00	11,995.00	

Table 5.2: Table comparing Xilinx FPGA evaluation platforms. [†]theoretical maximum, *SODIMM expandable, **non expandable

1. **Transceivers** In order to process high-speed serial links an FPGA uses transceivers to handle the physical interconnects from the front-end system and to the workstation. The number of transceivers required on the FPGA is therefore determined by the number of AFEs, the number of JESD204B lanes each AFE uses, and the version plus type of the PCI Express connection to the workstation.
Since there are four AFEs required to achieve 64 channels, the FPGA should at least have $N_{AFE} \times L = 4 \times 4 = 16$ transceivers for the AFEs, where N_{AFE} is the total number of AFEs and L is the number of lanes per JESD204B link. Four JESD204B lanes are the minimum number of lanes for the AFE58JD48 to send over all the raw data while also incorporating the 20% overhead introduced by the 8b/10b encoding. The total throughput per lane is then equal to 10 Gbps per lane and totaling 160 Gbps for all four AFEs. Besides the number of transceivers, the physical implementation of the interface on the evaluation board is just as important. The VITA 57.1 and 57.4 standard, also called the FPGA Mezzanine Card (FMC) provides a universal connector for transporting a wide variety of high- and low-speed input and output signals. Most FPGA evaluation boards, therefore, utilize one or more FMC connectors with transceivers routed to this connector for future expansion possibilities. These FMC connectors will be used to connect to the front-end subsystem since there is no other standardized interface that can provide 16 transceivers that are required to provide the specified throughout. In the first two rows of Table 5.2 the number of transceivers and type of transceiver are specified per FMC connector. The third column then provides more information on the maximum throughput of each type of transceiver.
For the connection to the workstation, the number of transceivers is determined by the PCIe standard and version that is implemented into the evaluation board. In order to incorporate a PCIe 4.0 x16 on an FPGA a minimum of 16 transceivers is required. Since PCI Express 4.0 x16 uses 16 lanes to transport data at 15.754 Gbps per lane. Unfortunately, there are no evaluation boards offered by Xilinx that support the full-width PCIe 4.0 x16 on a board level. The VCK190, for example, internally does provide a hard IP for PCI Express, including the 4.0 x16 specification using the QDMA subsystem.

2. **DSP slices** The number of DSP slices required on the FPGA is determined by the processing and multiplexing of high bandwidth streams executed on the FPGA. Since processing configuration 3 will predominantly use the most resources in terms of DSPs for the processing pipeline that is envisioned, the resources estimate from Table 4.3 and 4.1 are summed. All of the FPGAs selected in Table 5.2 have enough DSP slices available to accommodate the minimum amount of resources

for the estimated design. However, there is a difference between the architectural implementation of the DSP slices. For example, the VC707 implements the DSP48E1, the VCU118 the DSP48E2, and the VCK190 the DSP58. Even though this generation difference is always a superset of the previous generation, the DSP58 slice can do a single-precision floating-point multiplication and addition, to produce both the floating-point product and sum [83], whilst this was distributed over multiple DSP slices in previous generations.

3. **LUTs & Flip-flops** The number of lookup tables and flip-flops in the design is predominantly determined by the complexity and the number of IP-blocks that are integrated into the design. With more complex beamforming strategies and processing steps, this number will automatically also increase. However, it is difficult to estimate the exact number of LUTs and flip-flops in a design without actual IP-blocks that can be synthesized. Therefore, it is a safe strategy to overcompensate these resources in order to provide freedom in the design of the FPGA configuration.

4. **BRAMs** The number of BRAMs required in the FPGA configuration is predominantly dependent on the demand for any temporary storage in the configuration. Thus for processing configuration 1 the number of BRAMs could play a significant factor. Because the interface to the workstation has a lower throughput than the data coming from the AFEs, the raw RF samples cannot be transferred to the workstation in real-time. Therefore, the RAM and BRAMs are needed to alleviate this problem. BRAMs will then be predominantly used in memory controllers such as seen in Table 4.2. It can also be seen in configuration 3 that this is one of the limiting factors in scaling up the beamformer to an increasing number of channels. However, the storage of weights in beamforming can also be offset for compute at the cost of more DSP slices, LUTs, and flip-flops [26].

5. **Expandable memory** One of the most important factors on the FPGA board, besides the FPGA itself is the size of the off-chip memory, which in practice in the current market is DDR4 DRAM. As seen from Section 4.4.6 the size of the RAM determines how many samples can be temporarily held in storage. This is especially an important factor for processing configuration 1, because of the disparity in throughput between the interfaces. Evaluation boards come with soldered DRAM or with DRAM modules that are expandable using the SODIMM standard. The problem with DRAM memory that is soldered on the evaluation board, is that it cannot be expanded later on, even though the parameters and processing strategy of the system are not completely known at the moment of writing. Therefore, the VCK190 with 16GB of RAM, equal to one second of raw RF data samples in this 64 channel system is a clear favorite compared to the other options in Table 5.2. There is also a version of the Virtex Ultrascale+ family that contains HBM memory that is stacked on top of the FPGA die. However, the evaluation board of this family does not suit the needs of the TCfUS system that is envisioned and is therefore left out of the comparison.

With all of the specifications mentioned above taken into account and looking at the evaluation boards in Table 5.2, it becomes clear not every specification can be satisfied. The VCU118 compared to the VCK190 provides in theory a better platform when we look at the number of logic cells, DSP slices, and BRAMs, and it is on par with the number of transceivers brought out to an FMC connector(s). However, the VCU118 is significantly hampered in the amount of DRAM that is available to the FPGA, adding to that is the fact that the RAM is also not expandable. In that regard, the VCK190 has a significant advantage with first the total RAM available, second the expandability, and third the hardware IP memory controllers on board of the FPGA. With this 16GB of RAM, the system can at least buffer a full second of raw RF recording, without taking into account any other components that need storage or smart offloading strategies. The added benefit of dividing the transceivers over two FMC connectors instead of one is that the number of commercially available that fully utilize the number of transceivers of the VITA 57.4 standard is very limited. An important feature that cannot be absent on the FPGA board is the possibility for external input and output of clocking signals. Although clock recovery is possible with the JESD204B standard, a reference clock that is in synchronization with the reference clocks of the AFEs provides more stability in the system. All of the evaluation boards mentioned in Table 5.2 provide external clocking interfaces for both input or output, with the VCK190 for example also including an external clocking IC for extra settings such as clock division and phase shifting. Additional hardware will be necessary in order to synchronize all the AFEs, the FPGA, and the other devices in the system, this will be discussed in Section 5.3.2. Both the VCK190 and the VCU118 support 100G networking capabilities using the QSFP28 standard, but this is an option not currently used in

the design since the throughput over the interfaces is lower than that of the PCIe link. Also, the PCIe link is more flexible in terms of configuration and usage, but this comes at the cost of complexity. The advantage of the QSFP28 standard is that it can be used immediately for long-distance interconnect as opposed to PCIe which needs to be located on the motherboard of the workstation. The distance of the interconnect using this QSFP28 standard depends on the networking transceivers and fiber optic cable that is used. An example using a portable ultrasound probe and the lower bandwidth QSFP standard is the Lightprobe [27], more details can be found in related work Section 3.2.6. The VCK190, however, has three added benefits over all the other boards, these are:

1. The VCK190 has two onboard dual-core ARM processors: one real-time Dual-Core Arm® Cortex®-R5F and one general-purpose Dual-Core Arm® Cortex®-A72 processor. Both of these processors can be utilized for managing the on-board processing configuration, configuration parameters of the AFEs, and additional hardware that is described later in Section 5.3.2. The user can communicate with these processors via the FPGA and the ethernet standard. A soft processor like the MicroBlaze™ is also an option for the FPGAs in the lineup, however, a soft-core processor like this would utilize valuable resources of the FPGA fabric that could otherwise be used for processing logic or other tasks.

2. The second major unique advantage of the VCK190 is the hardware-implemented network on chip (NoC), integrating the AMBA AXI protocol[5]. The AXI protocol allows for a memory-mapped bus communicating to multiple IP-blocks on that bus or a streaming version of the protocol called AXI Stream which is a one-to-one link. Both of these protocols have a very high probability of being used in the FPGA and it is, therefore, beneficial to have a dedicated NoC available for these data buses and streams. The NoC also provides a direct interface between the programmable logic and the hardware integrated IPs on the FPGA such as the memory controller or PCI Express blocks. This NoC and communication to the integrated IPs are beneficial for data throughput when moving the ultrasound data to the workstation compared to older non-NoC FPGA architectures.

3. The third advantage in the Versal AI Core architecture is the 400 *AI Engines*. These AI Engines are a VLIW (Very Long Instruction Word), SIMD (Single Instruction Multiple Data) vector processor optimized for advanced signal processing applications [85]. Advanced signal processing workloads include beamforming, FFTs, and filters. An example of the beamforming and other steps in the plane wave ultrasound processing pipeline using the Versal ACAP architecture can be found in Corradi et al. [14]. The processors can reach frequencies up to 1.3GHz enabling very efficient, high throughput, and low latency functions. The AI Engines are structured into a 2D array and therefore allow for a very scalable solution. Data is moved between the AI Engines via the shared memory of the direct neighboring AI Engine or via the NoC that provides movement horizontally and vertically and is directly connected to the programmable logic and the hardware integrated IPs.

In terms of pricing, the VCK190 is the most expensive board in the lineup. However, this comes with the added benefit that the newest features and standards are integrated into this FPGA. The goal of this system is to provide a platform that can explore all the parameters and processing options/configurations. Since only one FPGA board is required for the number of channels of the envisioned system, the price is reasonable for the freedom of exploration it provides. For more cost and resources optimized designs more details must be known on the processing aspect of the design which is not part of this thesis. However, a small exploration in terms of channel exploration considering only the hardware requirements and implementing full custom boards is discussed in Section 5.5.

From the above-mentioned discussion, it can be concluded that the VCK190 is the best suited, off-the-shelf evaluation board from Xilinx for the application of TCfUS with the specifications and constraints mentioned in Section 4.4.

5.2.3. Interconnect (c)
The interface that is available on the AFE affects the implementation of the interface between the front-end and the back-end. As seen in Section 5.2.1 and 4.4.5, by implementing an AFE IC with the envisioned specifications of sampling frequency and resolution implementing LVDS will not be a viable option due to the throughput per LVDS pair. On top of that implementing LVDS would result

in at least 64 differential pairs equaling 128 signals to be transported over a cable. For this reason, JESD204B is implemented in the design, because each AFE will only need 4 JESD lanes per link, therefore totaling $4 \times 4 = 16$ lanes, equaling 32 signals to be transported. As has become apparent from Section 4.4.7, the combination of the amount of data that the AFEs produce and the distance that has to be bridged between the front-end and the back-end causes the interconnection between the two to be implemented using fiber optics since both of these factors are not a problem with this transport modality. Optical transceiver modules, that enable this interconnect, are predominantly designed for high-speed long-distance network interconnects, and use the SFP, QSFP, QSFP28 standards. Depending on the standard transceivers modules incorporate two or eight fibers, where half of the fibers are for transmitting and the other half are for receiving. To connect one AFE to the back-end system four transmit fibers would be enough, however, the system is designed with channel expandability in mind, and therefore a solution with extended capabilities and a smaller form factor has to be found. Samtec (New Albany, Indiana, USA) provides these capabilities with the FireFly™ Micro Flyover System™ [65]. One FireFly module can handle up to 28 Gbps per fiber and includes designs integrating 4 or 12 fiber channels. The specific module in the FireFly Micro Flyover System line-up for the transmit at the AFE side is the ECUO-T12-16. This selected model includes 12 channels and thus fibers over which the JESD204B data is transported to the back-end subsystem. The integration of this module is described in more detail in Section 5.3.3.

5.2.4. Workstation

In this thesis, the main focus is on the hardware design of the TCfUS acquisition system, and therefore the workstation is regarded as out of scope. However, it is important to match the interfaces between the FPGA and the workstation and that the workstation meets all the specifications from Section 4.4.9. Therefore, a workstation is selected that has a server-class motherboard and CPU. This allows for: first, multiple lanes of ECC memory, which is important for proving a large and reliable amount of 1 TB of DRAM storage. Second, multiple PCIe 4.0 x16 root complexes, which is important to achieve the full bandwidth of the CPU to the acquisition system, the Verasonics, and other devices. Third, a high core count CPU that will help with a reduction in execution time for processing configuration 1 and 2. In addition, an enterprise-grade GPU is also added for the same processing purpose. Lastly, the workstation is also outfitted with a high capacity, high-speed solid-state storage RAID array of 40 TB for permanent storage of raw and/or processed data.

5.3. Hardware Architecture

In this section, the hardware architecture for the TCfUS acquisition system is presented including the selected components from the previous section. In Figure 5.3 an overview of the design is presented. First, the general architecture of the design will be discussed and the considerations that are taken into account. Secondly, in the subsequent subsections, each individual PCB in the design will be discussed in more detail including design decisions and considerations. The architecture of the acquisition system is split into two distinct parts, as proposed in Section 4.4.1: the front-end subsystem containing the probe, digitization and pre-processing, and the back-end subsystem containing the FPGA and workstation for processing and storage.

The front-end subsystem consists of four individual printed circuit boards, containing the selected AFE58JD48 and FireFly module. These four PCBs are called the AFE PCBs and are situated as close to the receive probe as possible, in order to reduce noise and distortion. From the receive probe the signals of 16 transducer elements are connected to each AFE PCB. In addition to this AFE PCB, a separate clocking and management PCB is introduced which is situated near the AFE PCBs. The clocking and management PCB provides the AFEs with a phase-matched clock signal for the sampling of the ADC and the JESD204B interface. Furthermore, the PCB also provides power and configuration settings for the clocking IC on the board itself and AFE PCBs, for the AFE58JD48 and the FireFly module. This management function is implemented by an ARM Cortex-M7 microcontroller (uC) because if each IC was individually controlled by the FPGA there would run an enormous amount of wiring from the back-end to the front-end, which would be detrimental to the portability of the system. This interconnection for the configuration data is now provided by just one USB cable with the added benefit of also proving the power for the front-end subsystem.

The back-end subsystem consists of three components: first, the FPGA, which acts as a processing medium and as an interface adapter between the front-end subsystem, and second, the workstation. The VCK190 is connected to the workstation as a PCIe card and uses the PCIe 4.0 x8 standard. Because the selected FPGA board, the Versal AI Core Series VCK190 Evaluation Kit, does not have any way to directly connect FireFly modules, a third component called the FireFly FMC PCB is introduced. The FireFly FMC PCB provides this FireFly module connectivity but also additional features such as external input for clocking, triggering, and a GPIO interface.

The front-end and the back-end subsystem are connected by utilizing the Samtec FireFly Micro Fly-over System for the physical layer between the subsystems. The FireFly modules, which provide the optical fiber interconnect, use custom patched Multi-fiber push on (MPO) fiber cables to connect two AFE PCBs from the front-end to one FireFly FMC PCB at the back-end subsystem. These custom patched cables are 3 meters in length and because of the patch, they are in a Y shape, with 2.8 m, the majority of the cable, being the stem of the Y. The higher layers of the interface are handled by the JESD204B standard [3, 21] implemented on the AFE58JD48, where each AFE58JD48 utilizes four lanes per JESD204B link.

The following subsections will each discuss one of the PCBs present in the design and will elaborate in detail on the design decisions and considerations.
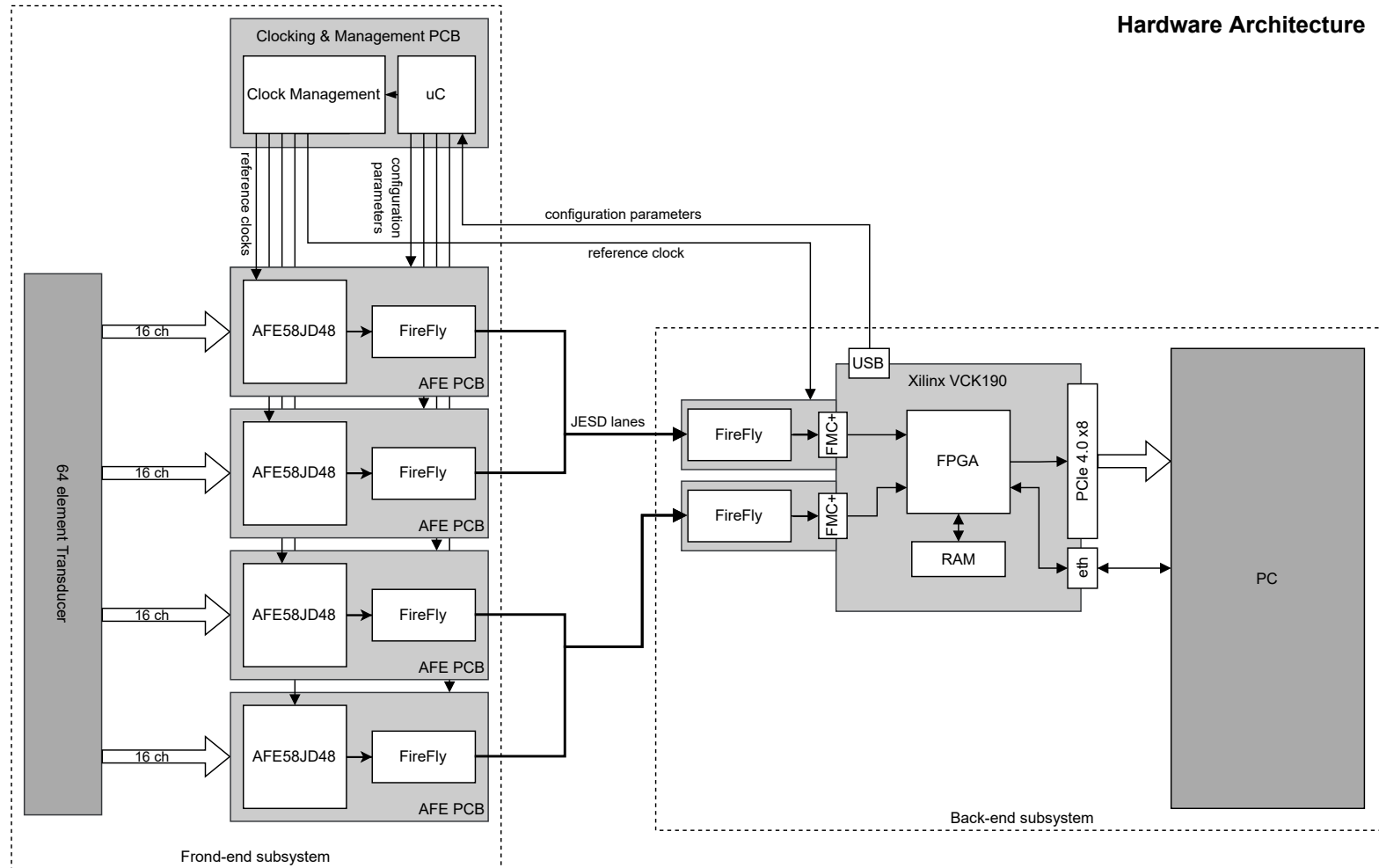
Hardware Architecture

Figure 5.3: Block diagram of the hardware architecture of the TCfUS acquisition system. Each individual PCB is described in Section 5.3.1-5.3.3. For more detail on the interconnect between the front-end and the back-end subsystem, see Section 5.3.3 and Figure 5.8.

5.3.1. AFE PCB

To create a compact implementation of the front-end subsystem, it would be beneficial to implement multiple AFEs on one PCB. However, since this is the first prototype of the TCfUS acquisition system, the decision is made to integrate each AFE onto an individual PCB, since integrating one IC on a PCB reduces the complexity of the PCBs and increases possibilities for debugging. Another factor that influenced this decision was the geometry of the receive and transmit probe. Because the receive probe is round with a recess in the middle for the transmit probe, the four AFE PCBs are positioned in a square around the transmit probe. This is further illustrated in Figure 5.4. A design that would contain circuit boards in the same plane as the receive probe would first need to take a circular shape but second of all would also need to incorporate a hole in the middle for the transmit probe. Therefore, a design has been made that contains the four AFE PCBs in a rectangular shape and perpendicular to the receive probe, so there is more flexibility for the design of the PCB in terms of area and geometry. All of the four AFE boards are identical which is beneficial for the routing of the clocking circuits to the AFE58JD48 ICs since the traces to each AFE need to be of equal length to avoid phase differences between the sampling clocks of the AFE. The clocking signals, configuration for the AFE, configuration signals for the FireFly, and power are provided to the AFE PCB using a board-to-board connector to the clocking & management PCB. To connect the receive probe to the AFE PCB flat flex connectors are used, with a ground conductor separating each signal conductor. By incorporating these connectors, instead of soldering the signal wires from transducer elements directly on the PCB, allows the design to easily accept other probes or signals. A schematic overview of the AFE PCB can be found in Figure 5.5. The FireFly module integrated on the AFE can convert up to 12 differential pairs at throughputs up to 16.1 Gbps. However, only 4 of these differential pairs are used at 10 Gbps per lane since each AFE is situated on its individual PCB.
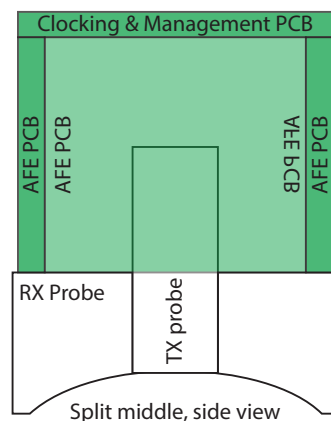

Figure 5.4: An illustration of the cross-section down the middle of the design to illustrate the receive and transmit probe and their geometric relation to the AFE PCBs and the clocking and management PCB
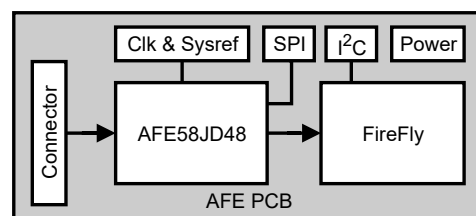

Figure 5.5: A schematic overview of the AFE PCB. Including a flat flex connector input for signals from the receive probe and a board-to-board connector to provide: 1) input clock plus SYSREF 2) an SPI plus I^{2}C interface for the configuration of the AFE58JD48 and the FireFly and 3) power to the board.

5.3.2. Clocking & Management PCB

The clocking and management PCB has multiple roles in the front-end subsystem. It provides clocking distribution to all the devices in the acquisition system, configuration parameters for all the ICs in the front-end subsystem, and distributes power over all of the components in the front-end subsystem. A schematic overview of the components on the clocking and management PCB is shown in Figure 5.6.

In order to accomplish beamforming and power Doppler, as intended by the envisioned TCfUS acquisition system, it is critical to provide an accurate and synchronized sampling clock to all of the AFEs. This clock needs to be jitter-free and distributed over the AFEs without introducing any phase errors. For this reason, the clocking & management PCB manages all the clocking generation and distribution in the acquisition system. In order to provide the distribution over the four AFE, a specialized IC is selected which supports multiple input clocks and contains support for JESD204B by providing besides a reference clock also a SYSREF signal for each device that is connected. The SYSREF is a signal derived from the reference clock and is an important feature of the JESD204B subclass 1 standard because it allows for deterministic latency and synchronization of the lanes between devices. In total the selected LMK04826 from Texas Instruments can support up to 7 JESD204B devices.
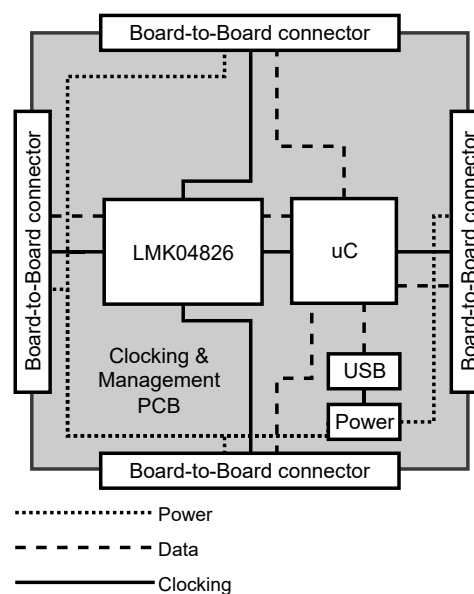

Figure 5.6: A schematic overview of the clocking & management PCB. The LMK024826 provides all the necessary sample clocks and SYSREFs to the AFE PCBs. In the prototype system this PCB also contains a large circular hole in the middle to allow the cable and transmit probe through as can also be seen in Figure 5.4.

The implementation of the clocking distribution network in the TCfUS acquisition system can be seen in Figure 5.7. A 125 MHz oscillator provides a reference clock to the LMK04826, which then uses PLLs, dividers, and multiplexers to create a 125 MHz clock and SYSREF for each AFE PCB, the FPGA (VCK190), and the Verasonics. Although the JESD204B interface provides excellent clock recovery possibilities, and thus a reference clock for the FPGA should not be needed, we do not want to be dependent on this feature and therefore the decision is made to still incorporate both the reference clock and SYSREF signals to the FPGA. By having both signals also available on the FPGA also provides us with better debugging possibilities. The Verasonics is responsible for the transmission of the ultrasound waves. To initiate an acquisition sequence, the Verasonics system will generate a trigger for both the FPGA and the AFEs. Because the trigger generation in the Verasonics is linked to the core clock of the system, the synchronization of the clocks between the systems is required. The FPGA can also trigger the AFEs and the Verasonics, however, the trigger input refresh rate of the Verasonics is limited to 15.625 MHz (a factor 16 of its core clock) and is therefore not a valid option. To connect the reference clock and SYSREF to the Verasonics and the FPGA a Cat 7 S/FTP cable is used because of the shielding and differential pair capabilities. To connect the reference clock and SYSREF to the AFEs length matched traces are routed to the board-to-board connectors going to the AFE PCBs to ensure phase alignment between the reference clocks. However, because the physical length of the

interconnect to the AFEs is by definition smaller, and thus not equal to the interconnect running to the FPGA and the Verasonics, a phase difference in these clocks will occur, meaning they are not aligned. The LMK04826, however, has multiple features available to mitigate this problem in the form of digital and analog delay. A SYNC~ which is a more rudimentary form of synchronization between JESD204B lanes, also needed for JESD204B subclass 1, runs between the FPGA and the AFE and is handled by the extra GPIO on the custom FireFly FMC PCB. The clocking circuit on the clocking & management PCB also allows an external clock as an input for the LMK04826, for debugging, or for example, the Verasonics, which provides a clock out port for its master clock, can be connected.
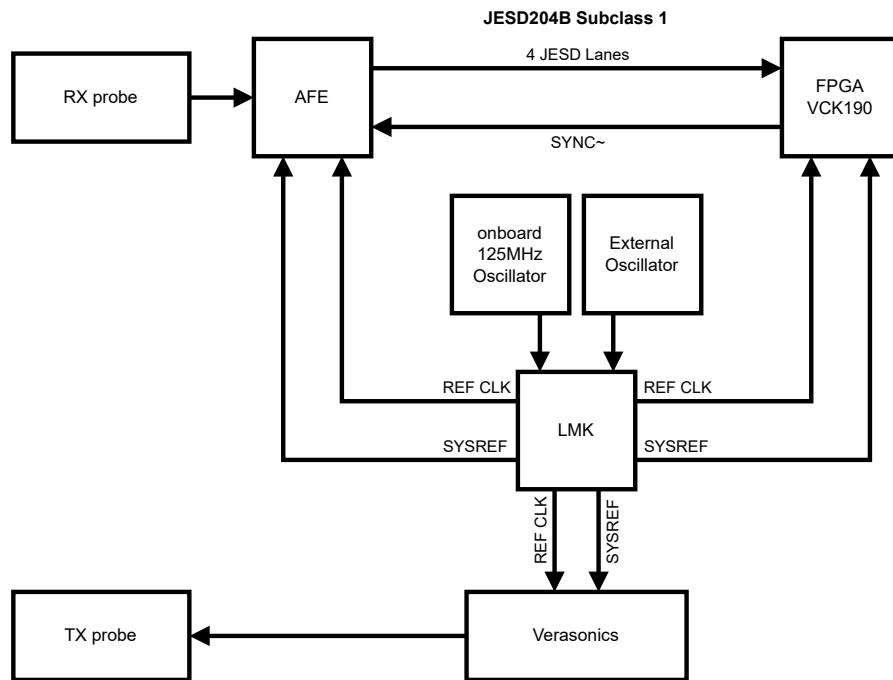
Figure 5.7: Clocking overview of the system. JESD204B subclass 1 is used particularly for the synchronization of the four AFEs, therefore requiring both the SYNC~ and SYSREF signals.

The management part of the PCB consists of a Teensy 4.0 [56] microcontroller board. The microcontroller is controlled by the FPGA and receives commands via the USB interface for the configuration parameters of the AFE58JD48, LMK04826, and the FireFly modules. The option for the microcontroller was implemented because the individual control of all of the devices in this first of the system is required, however connecting the devices directly to the FPGA would require a significant number of wires, which would not benefit the portability requirement of the design. The AFE58JD48 ICs and the LMK04826 are configured using the SPI communication bus. The FireFly modules are configured using I^2C communication bus. Besides these buses, each device also has additional GPIO controlled by the Teensy 4.0 including resets and other specific features that need to be controlled or are beneficial for debugging. All of these control signals are transported via the board-to-board connectors as seen in Figure 5.6.

The power to the front-end subsystem is provided using the USB interface available on the clocking & management PCB. This is the same USB interface that is used for the configuration of the individual components in the front-end subsystem. By using this solution only one cable is run to the FPGA for the configuration and power of the front-end subsystem minimizing cabling. The front-end subsystem can also be powered externally by using an external power supply or battery for debugging purposes.

5.3.3. FireFly FMC PCB
In order to convert the JESD204B data coming from the front-end system back to the electrical domain, another FireFly module is implemented. Since these modules cannot be directly connected to the VCK190 Evaluation kit a separate PCB will be needed to provide this connection. The VCK190

implements 12 high-speed GTY transceivers per VITA 57.4 FMC+ connector. Samtec offers a development kit for the FireFly modules implementing a 14 Gbps transmit and a 14 Gbps receive module from which at both modules 10 of the 12 differential pairs are brought out to the VITA 57.1 FMC connector. Since the VITA 57 standard is backward compatible this 14 Gbps FireFly FMC development kit [64] can be used to connect two AFE boards to the VCK190 by implementing a custom patched fiber optic MPO cable. The decision is, however, made to create a custom version of this development board in order to allow for extra in and output signals, since the options for this on the VCK190 are limited. The customized version of the 14 Gbps FireFly FMC development kit is called the FireFly FMC PCB and also contains general purpose input and outputs (GPIO), clocking inputs and outputs, and triggering in and outputs for the FPGA. Especially the clock input here is important since the development board from Samtec connected the reference clock of the GTY transceivers directly to a clocking IC without the option for connecting an external clock. However, the intention of this design is to connect the reference clock (and SYSREF) from the clocking & management PCB to the reference clock-in of the GTY transceivers. The GPIO connector is used for the Sync~ signal and other debugging purposes and future-proofing of the design.

To get a better understanding of the interconnect between the front-end and the back-end subsystem, Figure 5.8 provides an overview of two AFE PCBs that connect to one FireFly FMC PCB. Two of these interconnections make up the system as seen in Figure 5.3. Each FireFly module on the AFE PCB terminates the fibers into a 12 fiber MPO connector. The output of two 12 fiber MPO connectors are then transported via a custom patched fiber optic Y cable into one 24 fiber MPO connector. This 24 fiber MPO connector then feeds 12 of the fibers to the receive module and 12 of the fibers to the transmit module. The custom patched cable brings the 2×4 JESD lanes to the receive module where the data is converted back to the electrical domain and brought out to the GTY transceivers on the FMC+ connector of the VCK190. The custom patched cable is 3 meters long with the stem of the Y around 2,8 meters in order to reduce the number of cables to increase portability.
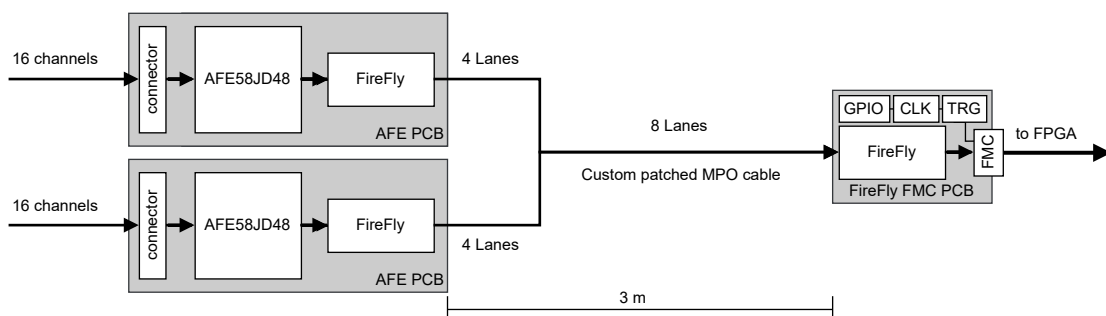

Figure 5.8: An overview of the fiber optic connection between the front-end consisting here out of 2 AFE boards and the custom FireFly 14 Gbps FMC module. Also highlighted is the custom patched MPO cable.

5.4. FPGA Configuration

In this section, the configuration of the FPGA will be discussed. Because the exact processing configuration and parameters are still uncertain, and hardware for the system is not available, a study of the FPGA configuration is done at the IP-block level, and neither implementation results nor exact resources are available. From the previous chapter and Section 5.2.2 it is known that the FPGA fulfills two purposes. At first, the FPGA acts as an interface adapter between the AFEs and the workstation. This is regarding processing configuration 1 and 2. The main difference between the processing configurations is that in the first occasion the throughput of the interface to the workstation can not keep up with the throughput that is being generated and thus ingested by the FPGA from the front-end subsystem. Secondly, the FPGA can also be utilized for the actual processing of ultrasound images. In processing configuration 3, a rudimentary form of a beamformer and power Doppler imaging will be executed on the FPGA. A resource estimation for 64 channels for this kind of processing can be found in Section 4.4.6. In parallel to this, the data can be sent to the workstation for a more comprehensive approach to beamforming called A-matrix beamforming [40]. However, the computational complexity

of this technique is too high, and therefore it is not feasible to implement this technique on the FPGA. Therefore basic image processing pipeline is implemented on the FPGA to provide near-real-time feedback to the operator on the positioning of the probe.

In Figure 5.9 a global overview of the signal flow in the FPGA is presented. Each number in Figure 5.9 represent an IP-block in the configuration. Data will come in at the transceivers on the left-hand side of the figure from the front-end subsystem and the AFEs. The GTY transceivers in the VCK190 feature hardware support for the decoding of the 8b/10b encoded JESD204B signal, from here the data is streamed to IP-block 1 which takes care of the higher layers of the JESD204B interface such as frame alignment between the JESD lanes, descrambling and deframing of the JESD data. The data from the AFEs will then be provided to the next IP-block as an AXI4 stream. From here on there is split between the three processing configurations.
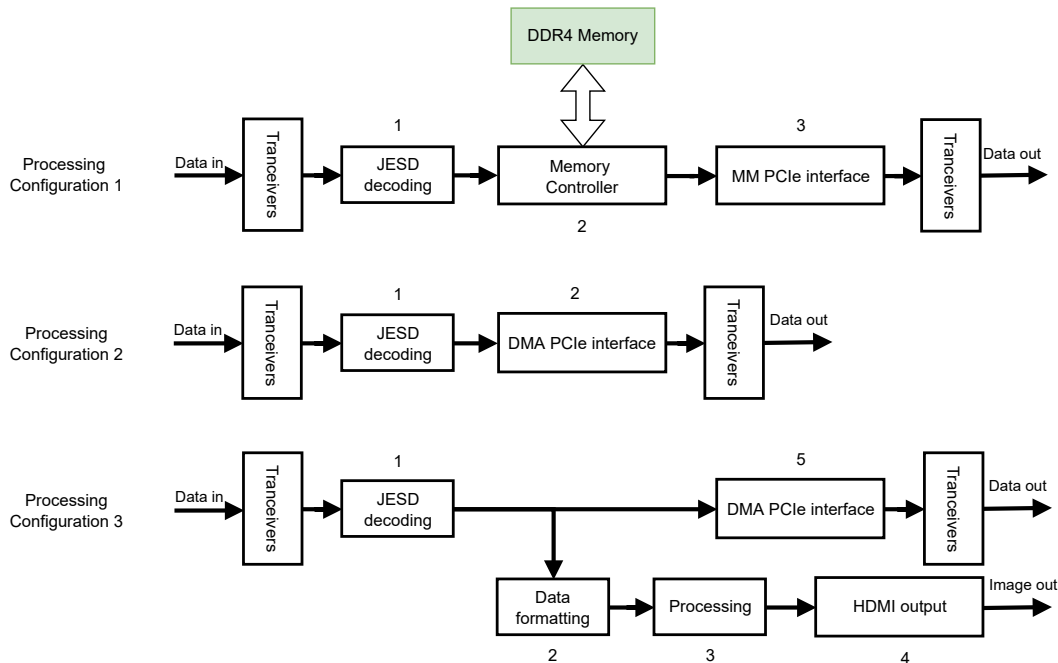
Figure 5.9: A schematic overview of the FPGA configuration for all of the three defined processing configurations. Each number in the figure represents an IP-block, with the transceivers regarded as hardware components of the FPGA.

- **Processing Configuration 1:** In the first processing configuration, the data from the analog front-end is sent to the hardware memory controller of the VC1902 via the AXI NoC. The memory controller receivers instructions for the scatter gather engine from one of the ARM processors. The Arm processor coordinates the interleaving read and write operation to the DRAM memory located on the VCK190. From the memory controller, the data is sent via an AXI4 memory mapped interface to the memory mapped PCIe interface (3). The same ARM processor will then use DMA and write data to the system memory of the workstation. In this configuration, the ARM processor in the VC1902 has an important task and orchestrates the reading and writing from memory on the FPGA and the writing to the workstation. Because of the discrepancy in throughput between interfaces the ARM processor will need to halt the ingestion of any new data from the front-end subsystem when the onboard memory has reached its maximum.

- **Processing Configuration 2:** The second processing configuration is a more straightforward approach. In this configuration, there is no discrepancy in throughput between the interfaces since the data has been reduced by the decimation filters on the front-end subsystem. Therefore, the data can be immediately sent to the main memory of the workstation. The memory mapping and interaction between the workstation is handled by the driver on the workstation which communicates to the ARM processor running on the VC1902 which handles the memory mapping.

- **Processing Configuration 3:** The third processing configuration is by far the most complicated since it involves both sending the data to the workstation which would be the same as in configuration 2 and locally processing the data into power Doppler images and providing these to the HDMI output of the VCK190.

  The data coming from the JESD decoding block (1) will be streamed to both the data formatting block(2) and forwarded to the DMA PCIe interface (5) which will forward the data to the workstation in the same way as the processing configuration 2. The data formatting IP-block (2) takes care of splitting the AXI4 stream coming from the JESD decoding IP into individual AXI4 streams per channel of the AFE. However, the AFE can contain multiple configuration parameters for the formatting of the data such as compression of multiple input channels, resolution, truncation, etc. Because of the number of configuration parameters, the data formatting IP (2) is configurable by an AXI4 lite interface so the configuration of this IP-block matches that of the configuration of the AFE.

  With the individual data from each channel now split into its own individual stream the data can then be fed into the processing IP (3) where the steps from the processing pipeline of Section 2.2 will be followed to provide a power Doppler image to the HDMI output block(4). The HDMI output block(4) will format the image such that it can be displayed on a standard monitor for guiding the operator.

Unfortunately, there are still a large number of unknowns in the processing configurations for the FPGA. However, due to the limited time available, the focus of the thesis is directed to the hardware design of the system and the configuration of the FPGA will be seen as future work.

5.5. Channel Expandability

Because the design that is presented in the previous sections has not been realized into hardware, and thus not much can be set about how successful the implementation is, this section will provide an insight into the expandability of the design regarding the number of channels. With an increase of the number of channels, we can increase the lateral resolution, depending on the probe specifications. Additionally, an increased number of channels could allow for volumetric imaging. The design that has been presented in the previous sections focuses heavily on SNR and flexible processing for parameter exploration. However, Section 3.3 and design decision 3 provide the insight that the flexibility in processing, which is heavily stipulated in the current design, can be traded in for an increased number of channels. With the hardware provision made for such an expansion of channels, according to design decision 5, together with a minimum amount of pre-processing (decimation) in the AFE chips, this expansion can be accomplished. Two channel expansion scenarios will be investigated.

1. How can the current design of the acquisition system be scaled to increase the number of channels and what are the hardware limitations? At what point do we need to scale out to more than one FPGA board?

2. If in a fully customized design a custom FPGA board could be introduced, is the VC1902 of the VCK190 still the most applicable FPGA in the Xilinx lineup? And to how many channels would the design expand?

5.5.1. Scaling Channels with Current Design

In the current design, the number of channels is limited by the maximum number of transceivers on the FMC connector and the FireFly interconnect. The VITA 57.1 standard that is implemented by the custom FireFly FMC board supports up to 10 transceivers and thus a maximum of 10 of the 12 channels for receive and transmit of the FireFly module are utilized. The FMC+ Standard increases the number of multi-gigabit interfaces from 10 to 24. However, on the VCK190 only 12 transceivers per FMC+ connector are connected. Since each AFE utilizes 4 JESD204 lanes, in order to transport all the raw RF data, the maximum number AFEs can be brought up to

$$N_{AFE} = \frac{N_{txrx}}{L} = \frac{24}{4} = 6 \text{ AFEs,}$$

equaling

$$N_{ch} \times N_{AFE} = 16 \times 6 = 96 \text{ channels,}$$

where N_{AFE} is the number of AFE ICs in the system, N_{txrx} is the number of transceivers brought out to the FMC+ connectors on the VCK190, L is the number of JESD204B lanes per AFE58JD48 and N_{ch} is the number of channels per AFE58JD48. However, having to connect six AFEs to the VCK190 has multiple hardware implications. First, the custom FireFly FMC board has to be revised in order to incorporate all of the 12 channels that are brought out by the VCK190 to the FMC+ connector. Second, new custom-patched MPO cables have to assembled. Third, all of the AFE boards require a synchronized sampling clock from the clocking and management PCB. Since the current setup is designed such that the AFE PCBs and the clocking and management PCBs are connected using board-to-board connectors, both have to be revised. Fourth, because in this design all the raw RF samples can still be stored or forwarded to the workstation the maximum recording length, in seconds, will now be reduced by 2/3, since the RAM on the VCK190 has remained the same. Because of the number of PCBs that have to be redesigned to expand the system up to six AFEs and because this expansion is still incorporating the raw RF sample requirement of design decision 3, this is not seen as a practical solution. Since the limitation is the FPGA board, it does not make sense to scale out to multiple systems.

5.5.2. Channel expansion with Revised Front-end PCBs

Since the previous section proves that with the current design and PCBs, it is not deemed feasible to expand the number of channels without a major redesign of PCBs. Therefore, a redesign of these PCBs is proposed based on the limitations of the current design. However, the feasibility of this redesign implementation cannot be verified without the lessons learned from the current design, in particular on the heat and power management of the front-end system. The redesign will incorporate three AFE chips on one PCB in order to fully utilize the transmit capabilities of the 12 multi-gigabit channels on the FireFly module, thus making the design more power-efficient than only using four of the twelve channels per AFE in the current design. Integrating three AFE ICs on one PCB also makes for a more compact design, which is advantageous for portability. A schematic overview of this AFE PCB and the interconnect to the back-end subsystem of the FPGA can be seen in Figure 5.10. Two of these redesigned 48 channel AFE PCBs could be connected to the VCK190 via two redesigned FireFly FMC PCBs, totaling 96 channels. In this proposed design the requirement of raw RF sampling is still possible. However, by also introducing a minimum amount of pre-processing (decimation filtering) in the analog front-ends, the total throughput from the AFEs can be reduced and therefore the number of JESD lanes can be reduced to only two lanes per AFE. This means the number of 48 channel AFE PCBs that can be connected to the VCK190 can be doubled if the same interconnect solution is used in the current design. Specifically, the custom patched fiber optic MPO cables. In Figure 5.11 an overview of this proposed solution is shown, the number of channels in this design would add up to a total of 192, tripling the capacity of the current design proposed in Section 5.3. However, this potential increase of the number of channels comes at the cost of a minimal amount of pre-processing. The implementation of this pre-processing, also seen in processing configuration 2, will use filtering and decimation to reduce the data throughput. In order to calculate the minimum decimation factor M that is needed to drop from four to two JESD lanes, the following formula will be used

$$T_L = \frac{f_s \cdot N_{ch} \cdot N_Q \cdot 1.25}{L \cdot M} \tag{5.1}$$

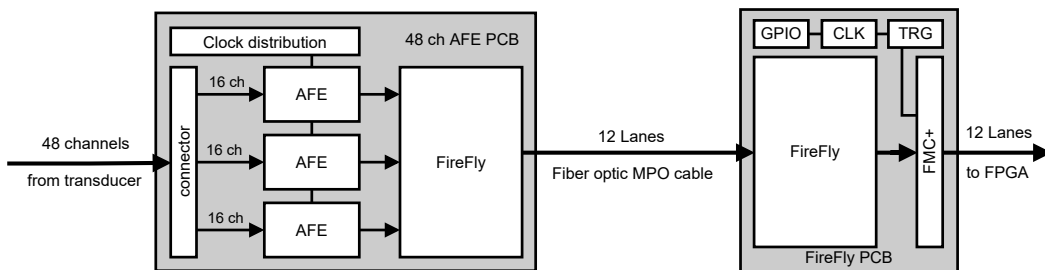Figure 5.10: Compared to the original design seen in Figure 5.8 the AFE PCB now contains 3 AFEs and a clock distribution chip for distributing the sample clock to the AFEs, integration of this IC decreases the complexity for the clocking and management PCB. The FireFly module is now also fully utilized in terms of number of channels. This design also incorporates the proposed changes on the custom FireFly FMC PCB from Section 5.5.1.
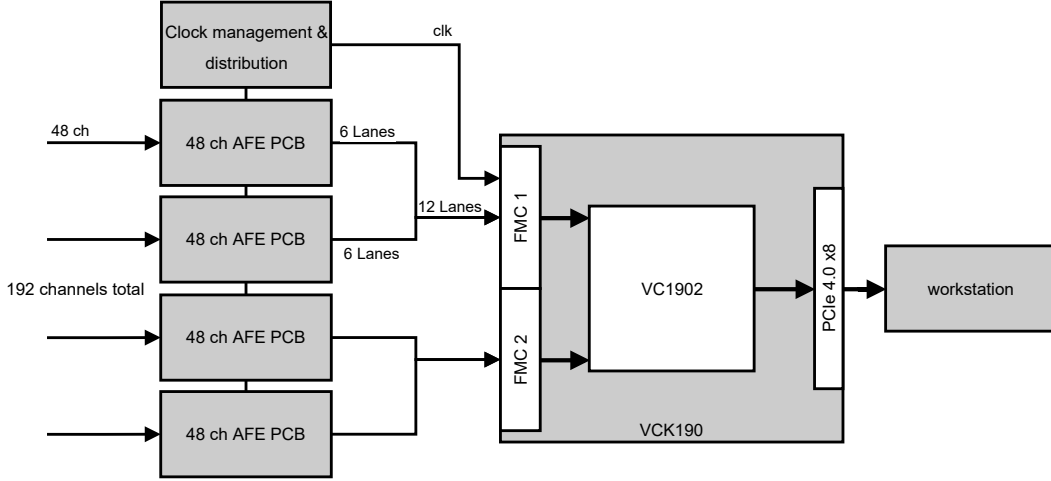
Figure 5.11: The interconnect of the revised AFE boards with the incorporation of pre-processing in the form of filtering and decimation. The minimum decimation factor in this design is $M = 2$, therefore requiring only 2 JESD204B lanes per link. The same strategy, applied in the original design of using a custom patched MPO cables to connect two AFE boards to one custom FireFly FMC+ board is also implemented in this design.

The factor of 1.25 in Equation 5.1 is due to the 8b/10b encoding in the JESD204B link. By transforming Equation 5.1 and filling in the known variables: the number of lanes per JESD204B link $L = 2$, maximum bandwidth per lane in a JESD204B link $T_L = 12.8$ Gbps and the sampling frequency plus resolution of the AFE, $f_s = 125$ MSPS and $N_Q = 16$ bits, the minimum decimation factor of the AFE58JD48 becomes

$$M = \frac{f_s \cdot N_{ch} \cdot N_Q \cdot 1.25}{M \cdot T_L} = \frac{125 \cdot 10^6 \cdot 16 \cdot 16 \cdot 1.25}{12.8 \cdot 10^9 \cdot 2} = 1.6. \tag{5.2}$$

The AFE58JD48, however, can only increment the decimation factor with integers and quarter fractions of integers, therefore making the minimum decimation factor $M = 1.75$. In order not to immediately jump to the absolute limitations of the JESD204B specification and as a precaution, the maximum throughput per lane of the JESD204B link is set at $T_L = 10$ Gbps. Therefore, making the minimum decimation factor $M = 2$. Because with an increase in throughput, the room for error in the tolerances and specifications of the PCB becomes more critical. There is, however, another factor that has to be taken into account that is related to the PCIe interface. The maximum bandwidth of the PCIe 4.0 x8 interface of the VCK190 is equal to $B_{PCIe} = 126$ Gbps. When taking into account a safety margin and including 128b/130b encoding, realistic real-world throughput would be equal to approximately $B_{PCIe} = 110$ Gbps. However, even with a decimation factor of $M = 2$ the 192 channels together still produce

$$T_{sys} = \frac{f_s \cdot N_{ch} \cdot N_Q}{M} = \frac{125 \cdot 10^6 \cdot 192 \cdot 16}{2} = 192 \text{Gbps}$$

therefore, it is not realistic in this situation that data is transferred to the workstation in real-time, as also has become apparent in processing configuration 1. As a consequence, the decimation factor M has to be increased even more, if the system has run in near real-time, as seen in processing configuration 2 or 3. The decimation factor then has to be equal to

$$M = \frac{f_s \cdot N_{ch} \cdot N_Q}{T_{PCIe}} = \frac{125 \cdot 10^6 \cdot 192 \cdot 16}{110 \cdot 10^6} = 3.49.$$

With the introduction of a decimation factor of $M = 3.5$, part of the data of the incoming signal will be lost, but because the bandwidth of the transducer that is specified in Section 4.4.2 is known, decimation can be applied without loss of information in the bandwidth of interest due to the oversampling that is applied and the Nyquist–Shannon sampling theorem. Therefore, the maximum bandwidth of the signal after decimation is equal to

$$B = \frac{f_s}{2 \cdot M} = \frac{125 \cdot 10^6}{2 \cdot 3.5} = 17.86 \text{MHz}.$$

The AFE58JD48 meets the specifications of the filter preceding the decimation with regard to the number of coefficients integrated into the decimation block, as seen in Section 5.2.1. In addition to that, the specifications of the decimation filter of the AFE58JD48 even beat the specifications of the Verasonics in terms of filter length. However, the coefficient length utilized in the Verasonics is not known [78]. It is, however, important to verify the parameters of the filter coefficients and the decimation factor with the designed transducer and designed acquisition system in order to make sure results are satisfying to expand to this channel expandable hardware design that is so reliant on this feature of the AFE58JD48. However, since this design has not been fully built into real-life hardware, some preliminary testing can be done utilizing the AFE58JD48EVM, the evaluation module of the selected AFE, and the TSW14J56, the appropriate FPGA sold by Texas instruments for testing with this evaluation board. In Figure 5.12 the measurement setup is presented for the characterization and verification of the decimation filter. The SDG2082X is an arbitrary waveform generator by Siglent that can generate waveforms up to 80 MHz. This is connected to one channel of the AFE58JD48EVM via a coaxial cable. The AFE58JD48EVM and the TSW14J56 are connected via a Vita 57.1 board-to-board mezzanine connector over which the JESD204B lanes are routed to the FPGA present on the TSW14J56 board. The data is then temporarily stored in memory and then offloaded to the PC via a USB 3.0 connection where the data is interpreted by a piece of proprietary Texas Instruments software called high-speed data converter (HSDC) pro. A picture of the measurement setup can be seen in Figure 5.13.

Measurement Setup

Figure 5.12: Block diagram of the measurement setup used for the testing of the decimation filter parameters.

Using this setup, different filters and decimation factors can be explored to find the optimal parameters that are associated with TCfUS. Apart from the information, these parameters will give on the channel expansion, this information is also important for processing configurations 2 and 3, since these two configurations are also heavily reliant on the decimation feature to provide a near real-time functionality to the system. With an expanded number of channels system as proposed in this section, it is possible to take a closer look at scaling out to multiple systems. Because scaling out to multiple systems is an option with the revised 48 channel AFE PCBs since the clock distribution has become a lot more simple and the efficiency of the number of FireFly modules has been increased. However, first, a closer look has to be taken at the power and heat management of the 48 channel AFE PCBs before any further expansion is made since the amount of heat that has to be dissipated from the front-end subsystem will likely have been tripled. Some heat-exchange solutions have been tried in the past such as a high channel count beamforming probe by Siemens (Munich, Germany) which is water-cooled [24].

5.5.3. Full Custom Design

If a customized PCB is made for an FPGA of choice, this would introduce a new range of possibilities for the back-end subsystem. Because this custom PCB could fully utilize the hardware features of such FPGA, including but not limited to, PCIe gen 4 or 5 with more lanes than the current VCK190 which has 8 PCIe lanes. Additionally, an FPGA could be selected, whose price scales better in proportion to the resources and transceiver count, for example, the Virtex 7 Ultrascale XCVU190. It is difficult to gauge the scalability of the configuration on the FPGA with the current unknowns in parameters and processing, however, with 120 transceivers available, provides enough transceivers for a PCIe 4.0 x16 interface utilizing 16 transceivers. Leaving 104 transceivers for connecting $N_{AFE} = \frac{104}{4} = 26$ AFEs, utilizing 4 lanes per JESD204B Link. This does however imply that the heat and power dissipation issues from the previous section have been solved.

5.6. Discussion

This chapter presented the component selection and design for a TCfUS ultrasound research system. However, because of the lack of real hardware and implementation of the design, it is hard to verify the design and therefore this chapter and thesis do not include these verification steps. The subsequent paragraphs of this section will discuss the limitations encountered in the design and future steps on how to proceed forward, like the measurement plan presented in Section 5.5.2.

The VCK190 was picked as the FPGA of choice for this project, however, due to Covid-19, production limitations, and supply-chain issues, the delivery of the board was postponed by 6 months, making it impossible to do actual experiments with the VCK190 like, for example, testing real-life throughput rates to the workstation or interaction with the AFE58JD48EVM. Besides, the unavailability of the FPGA board, it was hard to find a replacement board that is comparable in specifications, especially with the number of new features introduced on the Versal series of FPGAs. The NoC, for example, provides significant improvement of throughput, incomparable with other FPGA families. Additionally, the documentation and the support for the VCK190 in Vivado, at the time, were not at the point at which simulations of possible designs could be done. Therefore, the focus of this thesis is focused on the hardware aspect of the acquisition system with the FPGA configuration being left for future work.

The AFE58JD48 by Texas Instruments is the top-of-the-line AFE for ultrasound. However, real-life performance using ultrasound signals has not yet been proven. Therefore, no estimation can be made on the actual SNR increase by the system that has been designed in this chapter. The same can be said about the unprecedented oversampling factor that the AFE58JD48 brings in combination with the designed probe and compared to current state-of-the-art systems. However, since the decimation parameters for the AFE and thus processing configuration 2 and 3 are unknown, the increase of SNR with this oversampling and this AFE the specifications of the design are still hypothetical. However, it is known that oversampling and decimation is a proven strategy to increase SNR which is implemented by the Vantage 256, as further explained in Section 3.2.5. Therefore, with 2 bits more resolution in the ADC and the decimation filter in the AFE58JD48 capable of 400 taps at a decimation factor of 25 the specifications of the designed system beat that of the Verasonics, and thus therefore also an increase in SNR is expected. An effort is done into testing the AFE58JD48 using the evaluation module of the AFE, the AFE58JD48EVM, this is presented in Appendix A. The functionality of the decimation filter was verified with the evaluation boards from Texas Instruments and a custom configuration for the TSW14J56. However, because the default filter coefficients cannot be read from the AFE58JD48EVM evaluation board due to the design of the SPI interface, the result cannot be compared to simulated results in Matlab. Additionally, due to problems with the sample clock of the AFE and issues with the configurability of the TSW14J56, both measurements for the SNR with ultrasound signals and decimation have been postponed and are regarded as future work. Aside from the problems, a measurement plan to provide the desired data has been presented in Section 5.5.2.

Another factor that is uncertain in the current design is the alignment of phases of the clocks in between the front-end system, back-end system, and the Verasonics. Because of the difference in wire lengths between the AFE and the other devices, it is unknown if the LMK04826 can compensate for this phase difference. In addition to that, the measurement plan to align the phases in the TCfUS setup still has to be established. Besides, the phase alignment between the different devices or subsystems, another factor is the phase alignment between the reference clock and the JESD204B data going to the FPGA. The reference clock has to be in phase with the JESD204B data but the reference clock is transported over a Cat 7 S/FTP ethernet cable, which is a different media than the JESD204B data which is transported over a fiber optic cable. This can introduce problems when also including conversion time in the FireFly modules. To solve this, multiple features of the AFE can be used to provide debugging patterns and alignment of the signals between individual JESD lanes within the AFE and between AFEs.

Since there is no hardware implementation of the system available, it is hard to get an insight into the real-life power and heat management of the system. Therefore, it is hard to extrapolate the current design even though the interfacing options of the FireFly modules and the VCK190 allow more AFEs to be connected. Therefore scaling the number of channels up or scaling out the design at this moment has no significant value until a first prototype version of the system has been built. It is, however, clear

what the resources are that are needed for the extrapolation of this design and that with the current approach of using off-the-shelf IC components a 1024+ channel design in a portable form factor that is envisioned is not possible without introducing multiplexing or a custom ASIC. However, with both of these factors, the uncertainty of SNR and the development time will both increase which is exactly the advantages that are present in the current design.

It is also hard to predict any scaling of the number of channels, aside from looking at the hardware design of the system since the processing and architecture integrated into the FPGA are still largely unknown and a lot of exploration has to be done in order to provide more data on this subject, therefore this is seen as future work.

In an ideal case, the maximum decimation factor parameters for the probe are known and data from multiple AFEs can be aggregated by utilizing an FPGA on the front-end subsystem without introducing a significant increase in power. With this solution, the number of fiber optic cables to the back-end system can be reduced while the throughput over the fibers can be maximized. Additionally, the throughput that is produced at the front-end subsystem is reduced to a minimum. For this solution to work, more information is required regarding the power and heat envelope of the front-end and exact parameters regarding the decimation filtering on the AFE58JD48 or the next top-of-the-line AFE.

5.7. Conclusion

A hardware design has been presented which incorporates 64 channels and a receive-only capability to increase the SNR by placing the analog-to-digital converters as close to the transducer elements as possible and selecting the AFE58JD48 by Texas Instruments which is the highest performing analog front-end in terms of SNR, sampling frequency and resolution. The VCK190 was selected as the FPGA for the back-end system due to the number of resources available on the VC1902 FPGA and the number of transceivers brought out onto the evaluation board. This allowed custom PCBs to be integrated into the architecture and provide interconnect to the front-end subsystem using the FireFly™ Micro Flyover System™.

The limiting factor in the design, in order to provide near real-time streaming of all the 64 channels raw RF samples to the workstation, is the PCIe 4.0 x8 bus of the VCK190, however, no evaluation board from Xilinx at the moment is available that provides an interface to a workstation with a higher throughput. For this reason, the three processing configurations devised in Chapter 4 provide an outcome. In configuration 1, the raw data samples are temporarily stored on the FPGA and then offloaded to the workstation, therefore, dropping the real-time requirement. Furthermore, in processing configurations 2 and 3 the throughput is reduced by utilizing the decimation filter in the AFE58JD48 to reduce the throughput. A high-level overview of these processing configurations in the form of IP-blocks has been presented, however, due to time limitations, this subject could not be expanded any further.

To scale the system and thus expand the number of channels, the current PCB designs have to be revised, but the VCK190 is not the limiting factor until 192 channels or 12 AFEs. To implement this theoretical up-scaling in the number of channels, the decimation factor in the AFE58JD48 ICs has to be increased to a minimum of 3.5 to provide near real-time offloading of samples to the workstation. This minimum amount of pre-processing is required to reduce the number of JESD204B lanes required per AFE and to reduce the throughput for matching the PCIe 4.0 x8 interface to the workstation. Power dissipation and heat management are, however, not taken into account in this channel expansion feasibility study and only an actual implementation can reveal whether such a high channel count design is possible. It can, however, be concluded that due to these power and heat factors, a 1024+ channel design in a form-factor like this is not possible without introducing multiplexing or a custom ASIC; this would, however, increase development time and/or the number of potential problems to debug.

Additionally, the design of the high-SNR TCfUS ultrasound system presented in this chapter comes with a few shortcomings. At first, the system is reduced to only an acquisition system. This helps to rule out any sources of noise or distortion, however, comes at a cost. The complexity of the system's hardware is increased to also incorporate features to synchronize with external systems such as

the Verasonics. In addition to that, the synchronization in software is also still at a conceptual level. Second, the number of channels in the system is significantly smaller than current research systems such as the Verasonics [79] and ULA-OP [47], but this comes with the added benefit of full control of the system. The raw RF data can be offloaded to the workstation to be processed offline or the data can be filtered and decimated in the front-end of the system to provide a near real-time dataflow for processing on the workstation or the FPGA. Third, the current design is not as portable as originally conceived in Chapter 4 because compromises had to be made in order to make this first version of the design debuggable and producible in a small time frame. The system can still be regarded as portable since the front-end subsystem can be used without any external machinery and contains a minimal number of cables. However, the front-end system is not as small in dimensions as a regular commercial ultrasound transducer. Fourth, a major drawback in the current design is the limitation of the PCIe interface available on the VCK190. With the real-life bandwidth performance of the PCIe interface unknown, it is hard to judge how much of a throughput deficit there is compared to the data being generated by the AFEs, but it is certain that with the PCIe 4.0 x8 the throughput will by definition not be enough since the throughput from the front-end exceeds the throughput of the PCIe 4.0 x8 standard. Therefore, near real-time raw RF data transfer for all of the 64 channels to the workstation is impossible with an off-the-shelf Xilinx evaluation board without any concessions, be that either on the sample frequency, resolution, number of channels, or mandatory pre-processing. Fifth, besides the interface to the workstation being known and PCIe 4.0 x8, the implementation for the storage on the workstation is still known except for writing the data to system memory. Future work would include performance metrics and sizes of back-end storage solutions and what kind of throughput can be achieved there. Also, the incorporation of DMA and CCIX/CXL can be researched for data transfers from the FPGA directly to a GPU or solid-state drive(s).
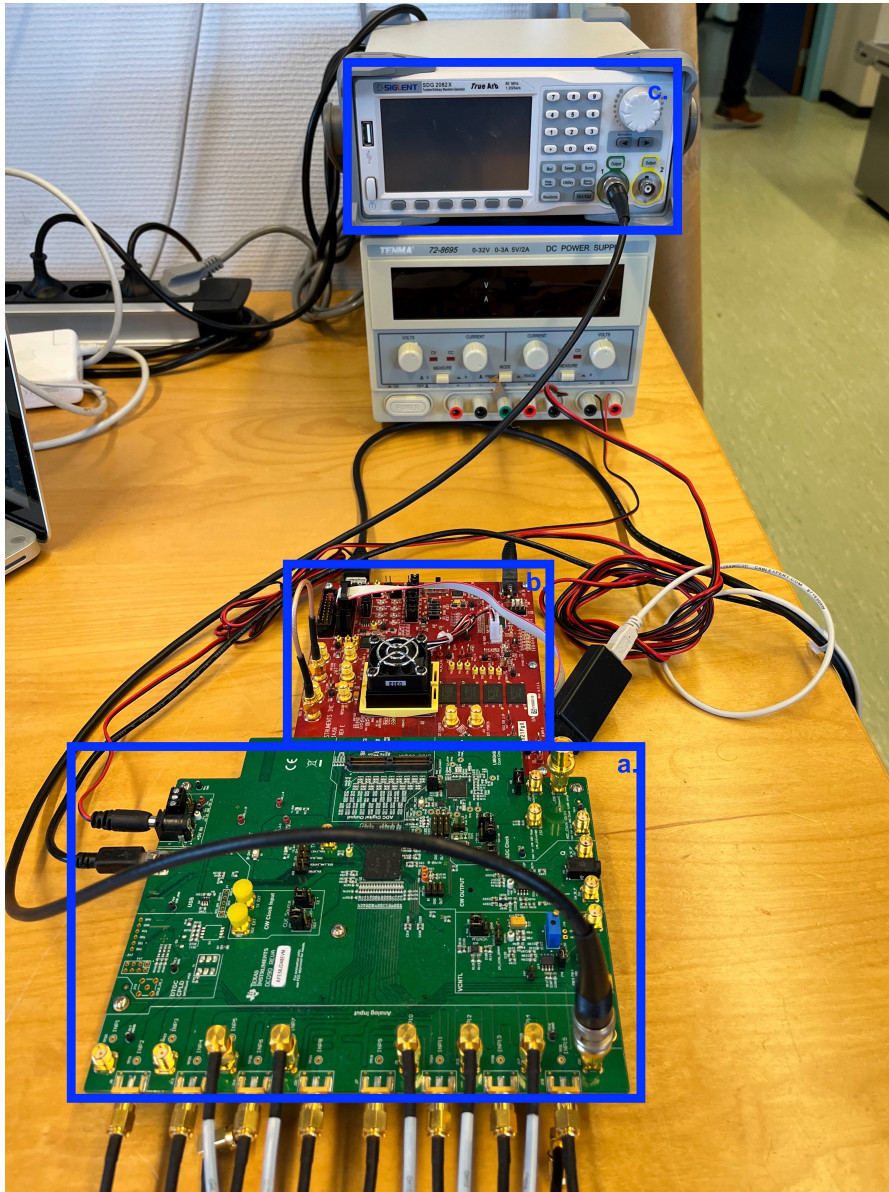
Figure 5.13: A picture of the measurement setup used for the testing of the decimation filter parameters. a) the AFE58JD48EVM by Texas Instruments with channel 16 connected to an arbitrary waveform generator b) the TSW14J56 by Texas Instruments, connecting the AFE58JD48EVM to the workstation c) the SDG2082X arbitrary waveform generator by siglent.

6

Conclusions

In this chapter, a concise summary of Chapters 2 to 5 of the thesis is given in Section 6.1. Hereafter, Section 6.2 will provide an overview of the main contributions of this work. Finally, Section 6.3 will provide an overview of possible future work following from the work that is presented in the thesis.

6.1. Summary

Ultrasound is a technique for non-invasive imaging of the human body and brain by using acoustic waves and the reflection of these waves on tissue boundaries which causes reflections. Doppler ultrasound imaging uses the Doppler effect to determine the blood velocity and volume and the cerebral blood volume in its turn can be correlated with brain activity using the mechanism of neurovascular coupling and this is called functional ultrasound. In small animals like mice and rats, it is possible to do transcranial functional ultrasound. However, the human skull reflects, aberrates, and absorbs the majority of the transmitted signal and therefore the signal reflected by the red blood cells is very weak compared to the transmitted signal. Therefore, it is currently only possible to do transcranial ultrasound using the acoustic windows within the human skull. In order to do transcranial imaging of the human brain, it is therefore critical to focus on the signal-to-noise ratio of the power of the signal of interest versus the power of the noise that is introduced. Since the thesis focus is to design an ultrasound system that is focused on reaching the highest possible SNR several techniques are discussed in order to improve the SNR in ultrasound imaging. 1) Ultrafast Doppler uses multiple plane waves at different angles to increase the sensitivity of the Doppler signal while also increasing/maintaining the SNR at framerates up to 10 kHz. This technique is the foundation enabling the required sensitivity for functional ultrasound. 2) Contrast agents are injected into the blood and increase the number of reflective objects besides the red blood cells already present therefore increasing the reflected signal amplitude. The major drawbacks of contrast agents are the short presence in the blood and the extra invasiveness introduced. 3) Oversampling and decimation are a combination of techniques in order to reduce the quantization noise that is introduced when the received signal is converted to the digital domain. This quantization noise can be spread over a broader spectrum by sampling the signal and bandwidth of interest with a significantly larger sample frequency. This however causes a significant data burden and by filtering and decimation, the amount of data can be reduced to only the signal/bandwidth of interest while reducing the quantization noise that is introduced. An important step in the decimation process is the filter that precedes it; without this filter, the decimation process would introduce aliasing artifacts into the bandwidth of interest and the quantization noise of the whole spectrum would still be present. FPGAs are a tool commonly used in ultrasound research machines due to their versatility and reconfigurability while performing application-specific tasks.

In Chapter 3, a general description of the receiving end of an ultrasound system is given and depicted in Figure 3.2. Recent work on the design and use of ultrasound systems for functional and transcranial functional ultrasound are discussed. Several key factors influence the design that is envisioned. The physical location where digitization and processing take place influences the requirements on the interconnects between the AFE and the processing and between the processing and storage. Additionally,

for a portable system, it can be concluded that the processing is inversely related to the degree of portability and thus power/heat dissipation. From the design of the system by Pietrangelo [55] and the Lightprobe [26] it can be seen that noise can be reduced if the proximity of the transducer to the analog-to-digital converter is reduced. Additionally, Xu et al. [86] advise not to introduce multiplexing to reduce additional noise in the receive path of the ultrasound signal. Therefore, if portability is a key design goal, only the number of channels and physical location of processing can be parameterized. Because of the computational load, part of the processing has to be physically separated from the digitization and transducer part due to power and heat limitations. Processing can be software-based, where the data is directly transported to a workstation, which provides all of the processing capabilities, or a system can be hardware-based, where the data is processed in the system using FPGAs. In a hardware-based processing approach the signals can be processed more efficiently, however, this comes at the cost of ease of configurability. Some ultrasound systems use the concept of oversampling and filtering and decimation to increase the SNR of the ultrasound signal. However, without any pre-processing like filtering and decimation the oversampling causes the interconnect to reach the throughput limits of the current-day interconnect technology. Therefore, the processing stage can be split into two parts – pre-processing and processing – and several trade-offs are introduced: 1) a fixed amount of pre-processing can achieve a higher channel count at the cost of information loss. 2) A design that is flexible for the type and parameters of the processing pipeline and thus research but has a limited number of channels because of the throughput that is reached if no pre-processing is applied and the user requires access to the raw RF samples. 3) if memory is added to a system, near real-time constraints cannot easily be satisfied with the addition that the temporary storage also needs to be placed close to the processing stage to retain signal integrity and timing.

Requirements are set up from which design decisions have been drawn. The design requirements of the system are limited to a receive-only system with 64 channels and split into 1) a front-end subsystem containing the AFEs and pre-processing and 2) a back-end system containing the processing and storage. This split is made to relieve power and heat constraints off the front-end system. The interconnect between these subsystems is realized with a fiber optic link due to length and throughput specifications. The system can be configured into three modes of operation also called processing configurations to provide different functionality to the user. Ranging from non-real-time raw RF data storage for exploration of processing parameters to a near real-time solution for direct feedback to the operator. Depending on the processing configuration described in Section 4.3, the configuration inside the FPGA will change, as does the throughput to the workstation. However, the hardware used for all the processing configurations has to remain the same. Processing plays a large role in the design of the system to balance throughput and accuracy, and therefore the envisioned processing pipeline is split up into a pre-processing part containing filtering and decimation, and a processing part that contains beamforming followed by subsequent processing steps for Doppler processing. Depending on the configuration, the secondary processing stage in the back-end subsystem is executed either on the workstation or the FPGA that handles the signals coming from the front-end subsystem. In either case, the incoming samples from the FPGA are transferred to the workstation and buffered in the system memory, to be later stored in a storage raid of solid-state drives. It is critical that the system is designed in the same way the data flows when receiving an ultrasound pulse since the specification and interface of every component influences the next component.

A hardware design has been presented which incorporates 64 channels and a receive-only capability to increase the SNR by placing the analog-to-digital converters as close to the transducer elements as possible and selecting the AFE58JD48 by Texas Instruments which is the highest performing analog front-end in terms of SNR, sampling frequency and resolution. The VCK190 was selected as the FPGA for the back-end system due to the number of resources available on the VC1902 FPGA and the number of transceivers brought out onto the evaluation board. This allowed custom PCBs to be integrated into the architecture and provide interconnect to the front-end subsystem using the FireFly™ Micro Flyover System™.

The limiting factor in the design, in order to provide near real-time streaming of all the 64 channels raw RF samples to the workstation, is the PCIe 4.0 x8 bus of the VCK190, however, no evaluation board from Xilinx at the moment is available that provides an interface to a workstation with a higher

throughput. For this reason, the three processing configurations devised in Chapter 4 provide an outcome. In configuration 1, the raw data samples are temporarily stored on the FPGA and then offloaded to the workstation, therefore, dropping the real-time requirement. Furthermore, in processing configurations 2 and 3 the throughput is reduced by utilizing the decimation filter in the AFE58JD48 to reduce the throughput. A high-level overview of these processing configurations in the form of IP-blocks has been presented, however, due to time limitations, this subject could not be expanded any further.

To scale the system and thus expand the number of channels, the current PCB designs have to be revised, but the VCK190 is not the limiting factor until 192 channels or 12 AFEs. To implement this theoretical up-scaling in the number of channels, the decimation factor in the AFE58JD48 ICs has to be increased to a minimum of 3.5 to provide near real-time offloading of samples to the workstation. This minimum amount of pre-processing is required to reduce the number of JESD204B lanes required per AFE and to reduce the throughput for matching the PCIe 4.0 x8 interface to the workstation. Power dissipation and heat management are, however, not taken into account in this channel expansion feasibility study and only an actual implementation can reveal whether such a high channel count design is possible. It can, however, be concluded that due to these power and heat factors, a 1024+ channel design in a form-factor like this is not possible without introducing multiplexing or a custom ASIC; this would, however, increase development time and/or the number of potential problems to debug.

Additionally, the design of the high-SNR TCfUS ultrasound system presented in this chapter comes with a few shortcomings. At first, the system is reduced to only an acquisition system. This helps to rule out any sources of noise or distortion, however, comes at a cost. The complexity of the system's hardware is increased to also incorporate features to synchronize with external systems such as the Verasonics. In addition to that, the synchronization in software is also still at a conceptual level. Second, the number of channels in the system is significantly smaller than current research systems such as the Verasonics [79] and ULA-OP [47], but this comes with the added benefit of full control of the system. The raw RF data can be offloaded to the workstation to be processed offline or the data can be filtered and decimated in the front-end of the system to provide a near real-time dataflow for processing on the workstation or the FPGA. Third, the current design is not as portable as originally conceived in Chapter 4 because compromises had to be made in order to make this first version of the design debuggable and producible in a small time frame. The system can still be regarded as portable since the front-end subsystem can be used without any external machinery and contains a minimal number of cables. However, the front-end system is not as small in dimensions as a regular commercial ultrasound transducer. Fourth, a major drawback in the current design is the limitation of the PCIe interface available on the VCK190. With the real-life bandwidth performance of the PCIe interface unknown, it is hard to judge how much of a throughput deficit there is compared to the data being generated by the AFEs, but it is certain that with the PCIe 4.0 x8 the throughput will by definition not be enough since the throughput from the front-end exceeds the throughput of the PCIe 4.0 x8 standard. Therefore, near real-time raw RF data transfer for all of the 64 channels to the workstation is impossible with an off-the-shelf Xilinx evaluation board without any concessions, be that either on the sample frequency, resolution, number of channels, or mandatory pre-processing. Fifth, besides the interface to the workstation being known and PCIe 4.0 x8, the implementation for the storage on the workstation is still known except for writing the data to system memory. Future work would include performance metrics and sizes of back-end storage solutions and what kind of throughput can be achieved there. Also, the incorporation of DMA and CCIX/CXL can be researched for data transfers from the FPGA directly to a GPU or solid-state drive(s).

6.2. Main Contributions

This section points out the main contributions of the work presented in this thesis. The main contributions are highlighted by answering the research questions presented in Chapter 1. However, only part of the main research question could be answered in this thesis due to the fact that the scope of the main research question is much larger than can be fit in a master's graduation project. Therefore, this thesis only covers the part of the main question that regards the specifications, system organization, hardware selection and architecture of a prototype design, and a channel count exploration. However, the sub-questions that are also listed in Section 1.3, can be answered. We will now state the main

contributions this work introduced per sub-question.

What is the status of related works with respect to transcranial functional ultrasound systems?
A table with the specifications of commonly used systems in work regarding fUS or TCfUS is listed
in Table 3.1. These research systems provide a large degree of freedom with regard to 1) access to
the raw RF data and 2) flexibility and configurability of the processing of the signals into a PDI. The
Aixplorer by Supersonic image and the Verasonics Vantage 256 are the leading systems for fUS and
TCfUS and integrate a software-based processing approach of the signals on a workstation separate
from the ultrasound system, utilizing multiple GPUs. This software-based approach allows for quick
changes of parameters in the processing as well as full flexibility of the processing pipeline. However,
all of the systems, with the exception of the Lightprobe, cannot stream the raw RF samples from all
the channels to the back-end and workstation in near real-time without introducing any form of 1) pre-
processing using filtering and decimation or 2) temporary storage due to limiting the throughput of the
interface implemented in the systems. The mandatory pre-processing can be a major disadvantage
since it is unknown if any of the information that is decimated will be needed later on during the pro-
cessing of the signals. The disadvantage of the temporary storage is that the system can only record
for a limited time before the data has to be offloaded and the system is therefore thus also not real-time.
The system realized by Pietrangelo (Section 3.2.7) provides a small portable system for transcranial
imagining that can be fixated to the head. The imaging is only possible through the acoustic windows
and is limited by 1) the number of channels and 2) on-board processing power due to the portability
aspect of the design. It is therefore hard to explore new processing techniques and parameters on
this platform and therefore highlights the two limiting factors of an ultra-portable design. The research
systems from Table 3.1 also provide more in-depth information about their system organization. It can
be seen that most designs are split between multiple front-end boards connected to a back-end board
using a high-speed connector such as a backplane, or via a copper/fiber optic cable interconnect. The
split of these components has been made to 1) create an easily expandable design in the number of
channels by reusing the same front-end design multiple times and 2) to split the processing to reduce
the data throughput to the back-end board. For this reason, front-end boards contain at least one
FPGA and/or DSP IC which can be configured by the user. Some research systems from Table 3.1
also provide the ability to scale out by connecting multiple systems together and therefore expanding
the number of total channels. However, this presents a challenge because the sample clock of all the
systems has to be synchronized for Doppler imaging, requiring a complex system for clock distribution.

What are the design trade-offs involved in this system when optimizing for SNR?
Several trade-offs were found while studying related works and during the design of the prototype,
these are:

1. Channel count comes at the cost of portability if digitization takes place as close to the source of
   the signal as possible due to the fact that the digitization and processing require adequate heat
   dissipation and space, which both form a problem for a portable design.

2. The system organization can provide offloading of the processing to a secondary location at the
   cost of channel count depending on the throughput that state-of-the-art high-speed interconnect
   can provide.

3. A research system providing all-round configurability and flexibility of the processing pipeline,
   while using oversampling, comes a the price of a high channel count. Any pre-processing located
   near the digitization can alleviate this problem at the price of information that is lost and cannot be
   used for offline processing, however, the pre-processing can introduce the opportunity to provide
   near real-time processing (<0.5 s).

4. Processing can be software-based or hardware-based. The software-based processing provides
   quick parameter changes and is convenient for exploring parameters of a processing pipeline
   but requires a workstation with a lot of processing capabilities. The hardware-based processing
   approach can often provide a more efficient solution in terms of execution time compared to the
   software-based approach, however, this comes at the cost of a more cumbersome configuration
   process.

5. Channel expansion comes at the cost of design complexity. A one-off design can be optimized for a specific purpose like optimizing for SNR, however, it is hard to integrate the features necessary to expand such a design like exposing high-speed interfaces and synchronization of clock lines.

What is the system organization that best serves these trade-offs?

A system organization that incorporates the SNR improvement techniques of oversampling, increasing the ADC resolution, and placing the ADCs close to the ultrasound probe, requires a partitioning of the ultrasound system. If we want to design a system that is also portable, either processing power or data throughput will become infeasible to implement. With a split of a system into a portable front-end system that handles: 1) the transducer, 2) digitization, and 3) pre-processing of the digitized data. Thereafter the data is transferred via a high throughput interface to a back-end system that is not necessarily portable and can handle the processing power and storage speed/volume required to generate and store the power Doppler images required for functional ultrasound imaging or raw RF data samples for future parameter exploration. By designing such a research system, the number of channels needs to be fixed depending on the maximum throughput that the front-end system can generate and the interfacing options available. Since a research system requires both access to raw RF data samples and pre-processed data, this thesis fixes the number of channels to 64, in order to keep a good balance between the data rate produced by these 64 channels and the spatial resolution achieved.

Because it is impossible to adhere to all the requirements of the system that are listed in this thesis, three processing configurations are introduced that each implements a specific subset of the requirements that are listed in Section 4.1. In configuration 1 the system is able to transfer the raw RF samples from the front-end system without any pre-processing to the back-end system and to a workstation for software-based processing. Since this process requires the maximum bandwidth of the interface between the front-end, back-end, and the workstation, is not feasible to execute this process in near real-time. In configuration 2 minimal pre-processing, in the form of filtering and decimation, is executed on the front-end subsystem and data is reduced to the exact point that it can be processed, in software, by the workstation, in near real-time. In configuration 3 the samples are pre-processed in the front-end system in the same way as in configuration 2, however, the decimation factor M can be increased depending on the processing in the back-end. The processing in the back-end is completely done using dedicated hardware, therefore providing near-real-time processing with techniques not possible using software-based processing on a workstation.

What are the minimal technical specifications for a TCfUS system?

The most important specifications for the TCfUS acquisition system are: 1) the transducer needs to have a center frequency of 1.5 MHz in order to penetrate the skull. 2) for the analog front-end the SNR specifications have to be minimally 78 dB for transcranial ultrasound [48]. To compete with the state-of-the-art a minimal sample frequency of 62.5 MHz and 14 bits of resolution are required. 3) the interconnect between the front-end and the back-end has to be implemented using fiber optic cables since the throughput reached while using 64 channels in combination with the specifications of the AFEs is larger than can be transported using copper interconnect longer than 3 meters. 4) the pre-processing has to be capable of decimation and filtering with a selectable bypass and a minimum range of $M = 1 - 25$. 5) the interconnect between the front-end and the processing has to be implemented using an FPGA since this is the only device in the price range that can adapt the interfaces that are encountered. 6) the processing can be implemented on the workstation or the FPGA, on the workstation resources for a minimum of 50 TFlop [73] is needed, and for delay-and-sum beamforming, on the FPGA the resources can be found in Table 4.3. 7) the workstation will need a PCIe 4.0 x16 slot for the FPGA, 2 TB of memory, and a RAID array consisting of solid-state storage with a minimum write speed equal to the maximum throughput reached by the interconnect between the FPGA and the workstation.

How do we guarantee that our system organization is possible with off-the-shelf components?

In order to design a system that can be implemented using off-the-shelf components, the specifications of the state-of-the-art components have to be compared against the listed specifications. Once confirmed that the components meet these specifications, the inter-compatibility of the components has to be checked. For example, the analog front-end IC can provide multiple interfaces to communicate to the back-end, however, the interface also has to be compatible with the components that are selected for the high-speed interconnect to the back-end. When the compatibility between the components is

verified, a design for the implementation can be created, taking into account the system organization with the split between the front-end and the back-end. Depending on the selection of the components, the physical design of the front-end and back-end will have limitations. If a design can be created that meets the specifications, a prototype can be fabricated, otherwise, the design process has to be reiterated or alternative components have to be sourced.

How does a prototype of a system with these specifications look like and how can it be validated?

Figure 5.3 depicts an overview of the hardware architecture of the designed system. This system consists of four AFE PCBs which connect to the 64 channels of the transducer, where each AFE PCB contains an AFE58JD48 IC and a FireFly™ module. An overview of this PCB is depicted in Figure 5.5. To transport the data to the back-end subsystem the FireFly™ modules convert the data into light and the data is transported via fiber optic cables. In the back-end subsystem, the data is converted back to the electrical domain by the FireFly FMC PCB and the JESD204B encoded data is decoded by the VCK190 FPGA board and offloaded to a workstation. Depending on the processing configuration that is applied this process is executed in near real-time for processing configurations 2 and 3, or acquiring a set number of samples and offloading the data afterwards in processing configuration 1. The configuration, timing, and synchronization of the sample clock of the AFE PCBs are all managed by the clocking & management PCB.

In Section 5.5.2 a measurement plan is presented to validate the hardware design. Figure 5.12 depicts the measurement setup of a preliminary set of evaluation boards but the same measurement plan can be applied to the designed system. An SDG2082X arbitrary waveform generator is used to generate and mimic the signals coming from the transducer. Due to the fact that the system has not yet been built and because of sample clock issues with the evaluation boards that were available the measurement plan has not yet been executed.

How does the designed prototype scale with regard to the number of channels?

The design that is presented in Chapter 5 is hard to scale in terms of the number of channels because of the design of the geometry of the probe, the design of the PCBs, and the clocking distribution circuit. The system that is presented, however, is also designed in the scope of a prototype and focused on SNR, a dedicated design based on scalability could produce a more scalable design. The VCK190 that is selected to connect the front-end subsystem to the workstation does provide options to connect more analog front-end chips. However, the current implementation of the front-end subsystem would need a significant redesign. With the introduction of decimation and filtering in the processing pipeline, as seen in processing configurations 2 and 3, the number of analog front-end chips, and thus the number of channels, can be increased significantly because the throughput decreases significantly, which means the number of JESD204B lanes can be reduced to two lanes per AFE. The VCK190, which is selected for the high number of transceivers that are brought out to the FMC+ connectors, can accommodate a maximum of 12 connections to analog front-ends before the maximum number of transceivers is reached. This results in 192 ultrasound channels when using the AFE58JD48 as the implementation for the analog front-end.

6.3. Future work

Since no hardware implementation of the design from Chapter 5 is available and because of the issues with the evaluation boards and hardware that was available, seen in Appendix A, measurements and results in the thesis are minimal. Therefore, future work should be focused on acquiring preliminary data that will further back up the design choices made in the design that is presented. Moreover, the results of these measurements will provide the parameters that are critical in the implementation phase of the presented design. Because future work would also include the implementation and debugging of the assembled PCBs, testing the FireFly interconnect to the VCK190, and the implementation of the three processing configurations on the FPGA. With the help of the processing configurations, it will also become clear what the limitations are of throughput to the workstation. Additionally, a storage solution such as a RAID array will need to be found that can handle the throughput that the system produces. Once such a basis is provided, High Level Synthesis (HLS) can be utilized to design and

integrate different processing pipelines on the FPGA in a short amount of time to explore the different processing strategies and parameters.

From the hardware implementation of the design, valuable lessons can be learned, which can be adapted to the next revision of the system. Power consumption and heat management are one of them for the front-end subsystem. This would also mean that more data will be available for the estimation of the expandability of the system and what the theoretical limit would be regarding the number of channels in the system. Future work on the revision of the hardware in the system should include the integration of multiple AFEs on one PCB with a single clock distribution chip, and a single FireFly module. This would imply a more efficient use of the space, more effective use of the number of fiber optic cables, as well as the synchronization of the clocking between the AFEs. By also including transmit hardware, such as the TX7516 by Texas Instruments, the system can become independent of a secondary system for transmission of ultrasound waves and the system can then show its full potential as a portable transcranial functional ultrasound modality. In the second iteration of the design, more complex and elaborate designs can be considered, for example, custom FPGA boards for the back-end subsystem, or the implementation of a digital signal processing (DSP) chip or midrange FPGA in the front-end subsystem to provide more pre-processing capabilities.

Besides, the future already provides standards that can alleviate some of the throughput problems in the current design. With JESD204B capped at a maximum throughput of 12.8 Gbps per lane, the JESD204C standard introduces throughput rates of 32 Gbps per lane [88], which means the number of lanes required per connection can be shrunk drastically. With the introduction of version C of the JESD standard, throughput limitations for future analog front-end are significantly reduced, and therefore their maximum sample frequency, resolution, and channel count can increase. Additionally, with the adoption of PCIe generation 5 in hardware, the interconnection to the workstation also provides more possibilities for real-time imaging with an increased number of channels.

A
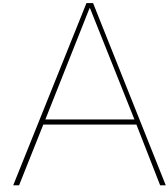
Vermont Probe Adapter PCB

In this chapter, a description will be given of a custom printed circuit (PCB) board that has been designed and produced during the thesis. The purpose of this PCB is to connect a commonly used probe to an evaluation board of the selected analog front-end for the design in Chapter 5. The reason this adapter PCB is made is twofold. First, with this PCB we are able to verify the parameters for the decimation filter of the selected AFE. This exploration of parameters is done to verify a preliminary implementation of processing configuration 2 and 3 and the designed architecture proposed in Chapter 5. Second, this PCB will also serve as a testbed for the application of a technique called compressed sensing, as mentioned in Section 4.4.2 and in [40], using a distributed array of transducer elements. How these elements are distributed will become clear in Section A.1.

In Section A.1 the design for the PCB will be presented after which Section A.2 will discuss the implementation of the design and the test setup. Thereafter, Section A.3 will present the results following from the measurements taken with the setup described in the implementation section. Following Section A.4 and A.5 will discuss the results and conclude on the work of this appendix and provide some leads for future work.

A.1. Design

The design of the board begins with the interfaces available on the devices to be connected. The evaluation board of the selected AFE58JD48 called the AFE58JD48EVM [74] is also produced by Texas Instruments and contains, aside from the AFE, multiple debug points, SMA connectors for the individual channel inputs, a clocking IC, the LMK04826 [75] to produce a sample clock for the AFE, and an FMC connector to connect the evaluation board to an FPGA for data acquisition. The probe that is used is produced by Vermont (Tours, France), contains 128 channels, and has a center frequency of 18 MHz. This probe has been selected because it is regularly and successfully used for functional ultrasound experiments in the lab [39]. The probe has two Hirose (Kanagawa, Japan) FX8 120-pin connectors, with one connected to the odd half and the other connector connected to the even half of the 128 total transducer elements. Because the AFE58JD48EVM board does not have any capabilities to send a high-voltage ultrasound signal and can only receive, the connector which is connected to the 64 odd channels is connected to a Verasonics Vantage 256 [79]. The connector which includes the 64 even channels is then connected to the custom PCB which is connected to the AFE58JD48EVM. Therefore, the Verasonics controls the transmit part of the probe using the odd transducer elements and the AFE58JD48EVM provides the receive capabilities using the even elements of the probe.

In order to achieve a distributed array as an input to the AFE58JD48EVM, and in order to reduce the 64 channels of the Vermont probe to the 16 channels available on the AFE58JD48EVM, the signal output of 4 channels of the Vermont probe are summed to become one channel that is presented at the input of the AFE58JD48EVM. This summation, or 4:1 mapping, is achieved by taking a Gaussian distributed subset of four of the 64 channels and summing these channels. One extra restriction on this randomly distributed subset is however that the elements are separated by a minimum of four chan-

Output channel	ch1	ch2	ch3	ch4	ch5	ch6	ch7	ch8	ch9	ch10	ch11	ch12	ch13	ch14	ch15	ch16
	50	114	82	48	38	120	100	62	44	26	122	76	42	68	110	10
Vermont probe	128	96	64	32	88	90	124	94	2	30	8	72	104	54	40	52
channel number	36	46	102	24	4	56	28	14	18	116	34	6	16	112	22	84
	78	108	74	66	70	126	12	106	86	98	80	20	60	118	92	58

Table A.1: This table contains the mapping of the channels from the Vermont probe to the channels of the AFE58JD48EVM, each column contains the four channel numbers of the Vermont probe and the channel of the AFE58JD48EVM on which they are mapped on. Only the even channels are used due to the connector layout of the probe, the channels of the probe are mapped with a minimum separation of 4 channels and a Gaussian distribution.

nels that are part of another subset. This separation gives both the spatial information and distribution needed for sparse imaging. A Matlab script was written that provided the mapping listed in Table A.1. Additionally, this summation of channels also increases the total area per channel connected to the AFE and thus more energy can be received per AFE channel.

The adapter PCB implements the counterpart of the Hirose FX8 120 pin board-to-board connector implemented on the Vermont probe. The even input channels of the Vermont probe are then summed by connecting the traces of the mapping provided in Table A.1. The summed 16 channels then each go through a Pi-type impedance matching network in order to provide any matching necessary originated from the summation or for the cabling to the AFE58JD48EVM. The matching network also contains room for a clamping diode in order to protect the analog front-end. Hereafter the channels are routed to a grid of SMA connectors which each correspond to one channel of the AFE58JD48EVM.

Since the signal velocity of a microstrip line on an FR4 board is equal $V_s = 1.5 \cdot 10^8$ m/s, the wavelength of the ultrasound signal is thus equal to $\lambda = \frac{V_s}{f_c} = \frac{1.5 \cdot 10^8}{18 \cdot 10^6} = 8.33$ m. With f_c equal to the center frequency of the probe. Because the wavelength is much longer than the distances to be routed on the PCB the length of the individual traces relative to each other is neglected, and the same is true for the length of the traces from the Hirose FX8 connector to the point of summation or between channels. Because of this disparity in length between traces the characteristic impedance of the traces also has not been taken into account. The PCB consists of 4 layers, where the outer layers provide signal routing and the middle two contain a ground plain to improve the ground return path and therefore improving signal integrity. The PCB layout incorporating the 4:1 mapping of the 64 even Vermont channels is depicted in Figure A.1.
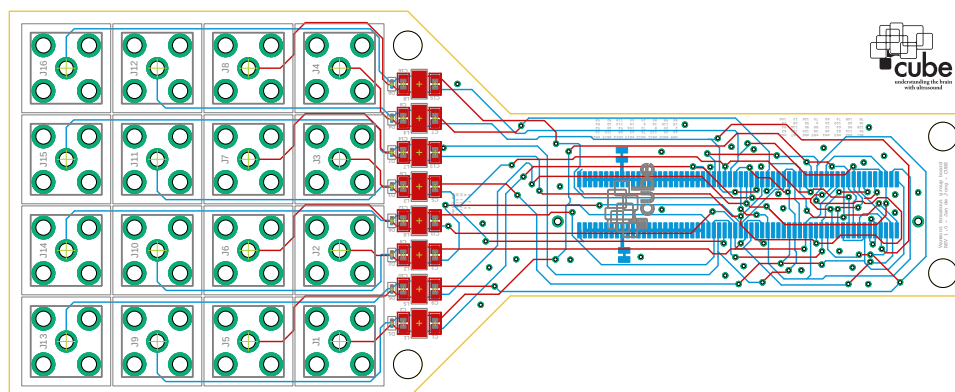

Figure A.1: PCB layout of the adapter PCB, including a Pi-type impedance matching circuit, clamping diodes, and a custom 4:1 mapping of the even channels of a 128 channel, 18 MHz Vermont probe. Signals are brought out to the AFE58JD48EVM using 16 SMA connectors.

A.2. Implementation

The adapter PCB was produced at Eurocircuits Belgium and assembled at the lab at the Erasmus MC. During assembly some mistakes came up, one of which was that the tracks and vias underneath the

FX8 board-to-board connector which were touching the side, but also below, mounted pins of the FX8 connector. This was solved using Kapton tape as an isolation material and no issues were experienced after. The Pi-type matching network was not populated and only the 50ω resistors matching for the coax cabling were populated. An attempt was made to measure the impedance for any corrections using a vector network analyzer (VNA), however, a lack of time and experience with the specific device prevented any improvement to the current implementation. No clamping diodes were implemented on the PCB since these were installed on the AFE58JD48EVM and thus not necessary.

The adapter PCB could then be connected to the probe, the 128 channel Vermont probe was disassembled and the PCB for the even channels carrying the microcoaxial cabling to the Verasonics was disconnected and replaced by the adapter PCB. The PCB for the odd channels was left installed in order to provide the transmission of the acoustic waves. The assembly of these boards was then isolated using isolation and copper tape to provide protection and reduce noise and interference from outside sources. The adapter PCB contains M3 mounting holes to connect to a clamp using a custom 3D printed mount.

The adapter PCB is then connected to the AFE58JD48EVM using 500 mm RG174 coaxial cabling. The testbed setup has been depicted in Figure A.2. The AFE58JD48EVM connects to the TSW14J56, using an FMC connector. TSW14J56 board contains an Intel Arria® V GZ FPGA, 4 GB of DDR3 SDRAM, and an FX3 IC from Infineon implementing USB 3.0 to offload the data to a workstation. The workstation to which the TSW14J56 connects could in theory be the same workstation that connects to the Verasonics Vantage 256, however, it was more convenient to keep the systems separated due to the installation of the proprietary data acquisition software of the TSW14J56 called HSDC pro. The HSDC pro software handles the configuration parameters of the TSW14J56 and the AFE58JD48EVM and the raw data samples from HSDC pro can then be exported to any other software for post-processing such as Matlab. The parameters of the AFE58JD48EVM are set to use eight JESD lanes, 18 dB of gain (medium), and the anti-aliasing filter was set to cut off at 30 MHz. The acquisition of the TSW14J56 was triggered using the Verasonics Vantage 256 and then capturing 16384 samples. 16 plane waves were sent, each under a different angle, and this was repeated 200 times. Thereafter the 838MB of data was offloaded from the RAM of the TSW14J56 to PC 2. The targets that are insonified and imaged in the test setup are 1) a static phantom of a wire and 2) a flow phantom submerged in water mimicking blood vessels and containing a blood-like fluid. The flow in the plastic tubing mimicking the vessels is done by using a syringe pump. Pictures of the test setup can be seen in Figures A.4 and A.5.

A.3. Results

The data that was collected by the setup from the flow phantom and a static phantom looked promising from a single channel perspective of the AFE. In Figure A.3 transmission and a reflection can be seen, afterwards, a new transmission can be seen starting at ~0.7 seconds. It is assumed that the transmissions can be seen on the receive channels because of the crosstalk of the transducer elements. From a conventional ultrasound perspective, the results of these reflections would be enough to create a B-mode image, however, it is hard to interpret the results of a single channel without taking into account the compressed sensing perspective and processing that still has to be done. A post-doc who is currently with us at the Erasmus MC and is specialized in ultrasound processing did the post-processing for these experiments and found that there is a significant difference between the defined sample frequency of $f_s = 125$ MHz and the actual sample frequency. This is detrimental to the beamforming that is required for the reconstruction of the distributed array, be this either in the frequency domain or in the time domain. Because of this reason, any further experiments including the checking of decimation with ultrasound signals have been halted.

It was further investigated what the cause of the sample clock problem was and why the sampling frequency of the AFE did not match that of the specification of the AFE58JD48EVM datasheet. The sample clock of the AFE58JD48 is distributed by the LMK04826 and is generated by a 125 MHz crystal with a frequency stability of 50 ppm. To check the disparity between the specified sample clock frequency and the actual sample clock frequency two methods were used. First, a source with a known frequency was presented at the input of the AFE58JD48, and the fast Fourier transform (FFT) was cal-
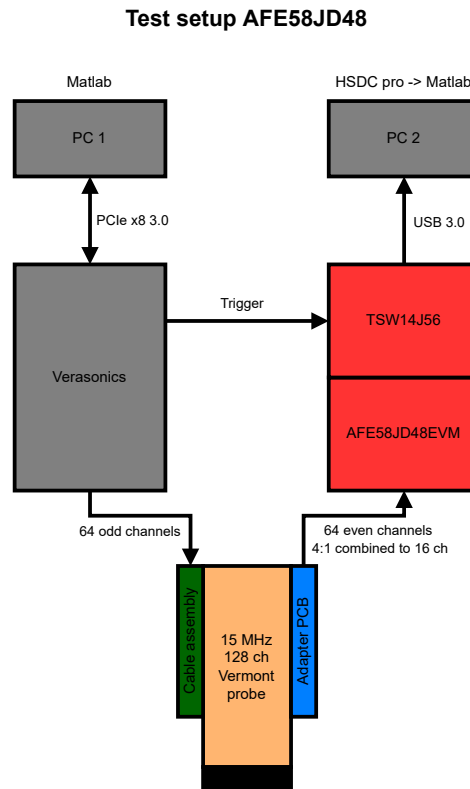
Test setup AFE58JD48

Figure A.2: The test setup of the adapter PCB in combination with the AFE58JD48, providing the receive functionality and the TSW14J56, providing the interface between the AFE and the workstation. Triggering is done from the Verasonics Vantage 256 using a coaxial RG58 cable. And a Matlab script on PC1 manages the beginning and ending of a measuring sequence.

culated to obtain the frequency spectrum. Since the FFT is dependent on a known sample frequency the results of the known input source did not match the results of the FFT spectrogram. By changing the sampling frequency for the FFT the actual sample frequency could be found by comparing the peak in the spectrogram with the actual input frequency. This was done for multiple frequencies as an input source to the EVM to check the deviation over the spectrum. The actual sample frequency was found to be around 118.38 MHz. In order to verify this, we used a second approach. This second approach included connecting an oscilloscope to one of the extra outputs of the LMK04826. An oscilloscope is not the ideal tool here for verifying the frequency of the sample clock, because of the limited bandwidth available on the DS1054Z oscilloscope that is used. However, at the moment no spectrum analyzer was available. The LMK04826 was programmed to divide the input clock with a factor of 2 and the frequency was measured at 59.5 MHz which is to be expected as this is half of the frequency derived from the FFT spectrogram. It is however not clear what the exact problem is, if the crystal is not accurate or if the settings of the LMK04826 are incorrect. Unfortunately, due to time limitations, no further investigation was possible and this is left for future work.

A.4. Discussion

It has been proven that the signals from multiple individual transducer elements can be summed up with a custom designed PCB and can be offered to the analog front-end evaluation board. However, it seems like important matching specifications and characteristics may have been oversimplified in the current design by looking at work like Cummins [17], that do take specifications like impedance matching and equal trace length into account. Therefore, the complexity of this adapter PCB has been underestimated. However, a first prototype has been presented that can lay the groundwork for a future revision of the PCB.

During the testing of the AFE58JD48, the TSW14J56 has been used as an acquisition platform and
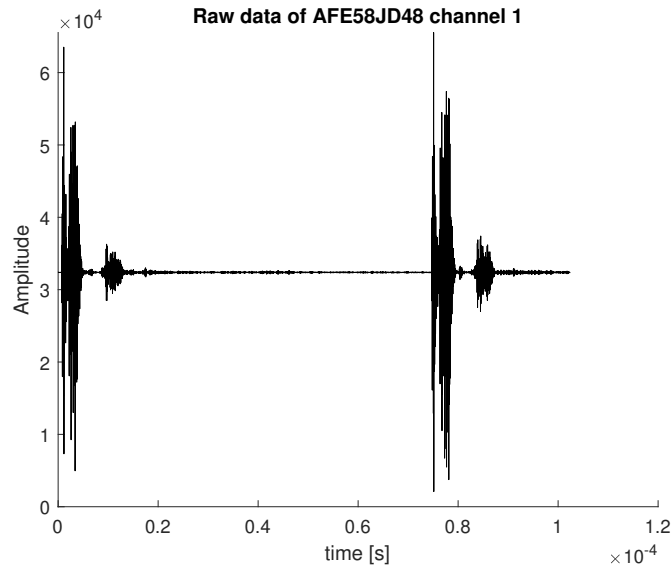
Figure A.3: Transmission and reflection of a static phantom captured using the 4:1 PCB adapter and the AFE58JD48 in combination with the TSW14J56.

interface between the AFE and the workstation. The TSW14J56 has been designed as a universal JESD204B board to connect to a multitude of Texas Instruments devices. However, from an FPGA development perspective, it is interesting to see if any changes can be made to the configuration of the TSW14J56 to also support some preliminary signal processing. For example to compare the decimation filtering on the AFE58JD48 versus an FPGA implementation on the Arria® V GZ FPGA on the TSW14J56. The source for a Quartus project for the TSW14J56 was available, however, the right version for the AFE58JD48, documentation, and support for any changes to changes of the basic configuration was very limited.

A.5. Conclusion and Future Work

A working adapter PCB has been designed for a 128 channel 18 MHz Vermont probe. The PCB successfully combines four of the even channels of the probe into one channel connected to an evaluation board of the highest specification analog front-end chip from Texas Instruments (AFE58JD48EVM). The summation of the four channels is Gaussian distributed and the minimum separation is four channels. A test setup with a flow phantom and a static phantom is made and samples from this setup have been successfully acquired while being triggered by the Verasonics Vantage 256. No images have been recreated from these acquired samples because of a hardware issue where the sample clock of the AFE58JD48 was not exactly 125 MHz and the exact sample clock could not be established with the tools available at the time. A test-bed has been produced which allows for future measurements and testing of hardware in combination with the AFE58JD48, which is a potent tool for the ultrasound industry.

Future work includes further investigation of the sample clock problem of the AFE58JD48EVM board since this is one of the most important limiting factors at the moment to conduct any further experiments and recordings. Therefore, the configuration of the LMK04826 will be further investigated. Additionally, connecting an external clock to the LMK04826 could potentially solve the problem. One of the main advantages of connecting an external clock to the AFE EVM board is that this would allow for the synchronization of the AFE EVM board and the Verasonics machine. The AFE EVM board would then connect to the core clock of the Verasonics Vantage 256. Since the core clock of the Verasonics runs at 250 MHz it can be seen that this is an exact multiple of two of the sample clock in the AFE EVM board. Because the AFE EVM board contains the LMK04826 IC, this clocking chip can from there on divide the input clock from the Verasonics to the AFE58JD48 which could synchronize both systems. To implement such a solution would however require the design of another custom PCB, since the

clock output of the Verasonics is run via an HDMI connector and is differential, and the input to the AFE58JD48EVM board uses an SMA connector. This requires impedance matching and conversion, hence another custom PCB.

Future work also includes further investigation of the impedance matching on the current PCB by using a VNA. Results of these measurements and calculations would give more insight into the design of the next revision of the adapter PCB.
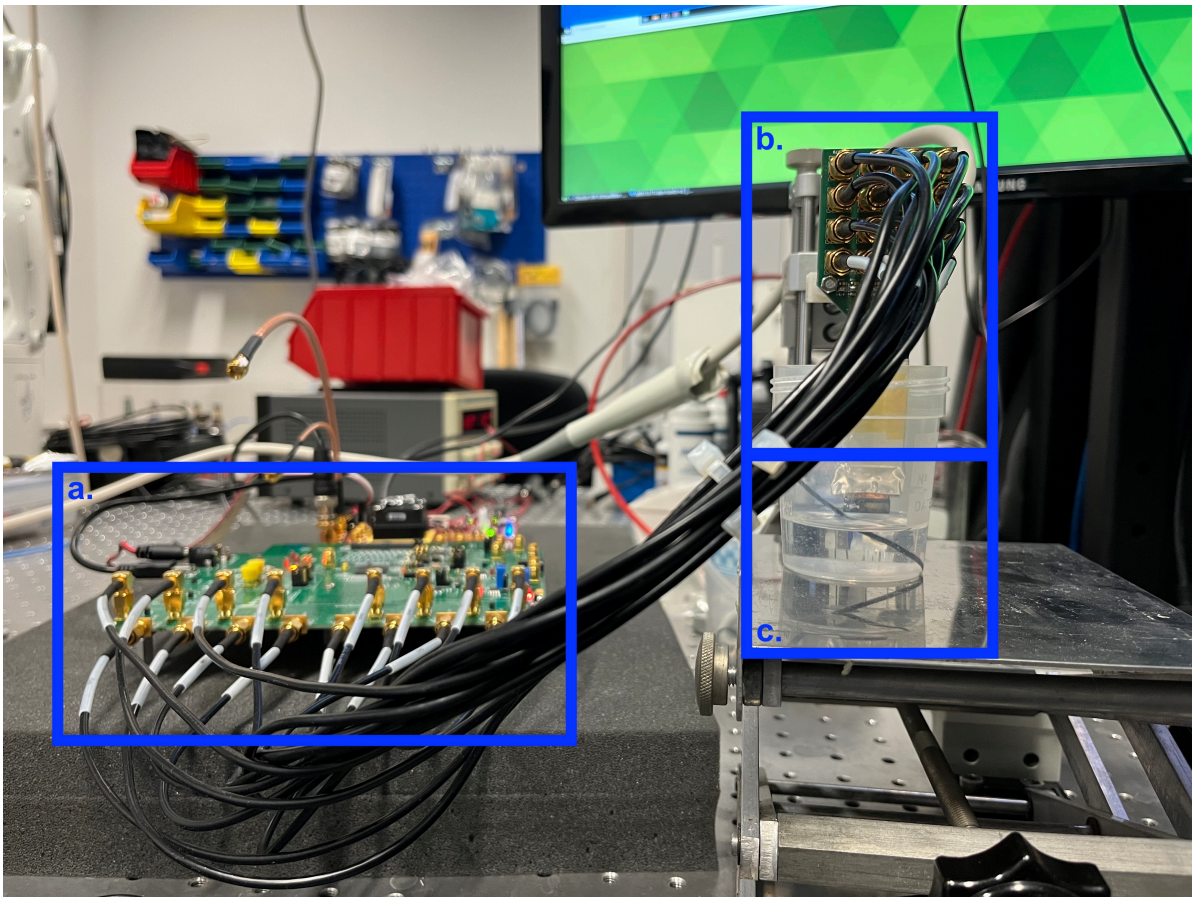
Figure A.4: Picture of the test setup of the static phantom in the form of a piece of wire (c) showing the probe (b) connected to adapter PCB connected to AFE58JD48EVM connected to the TSW14J56 (a).
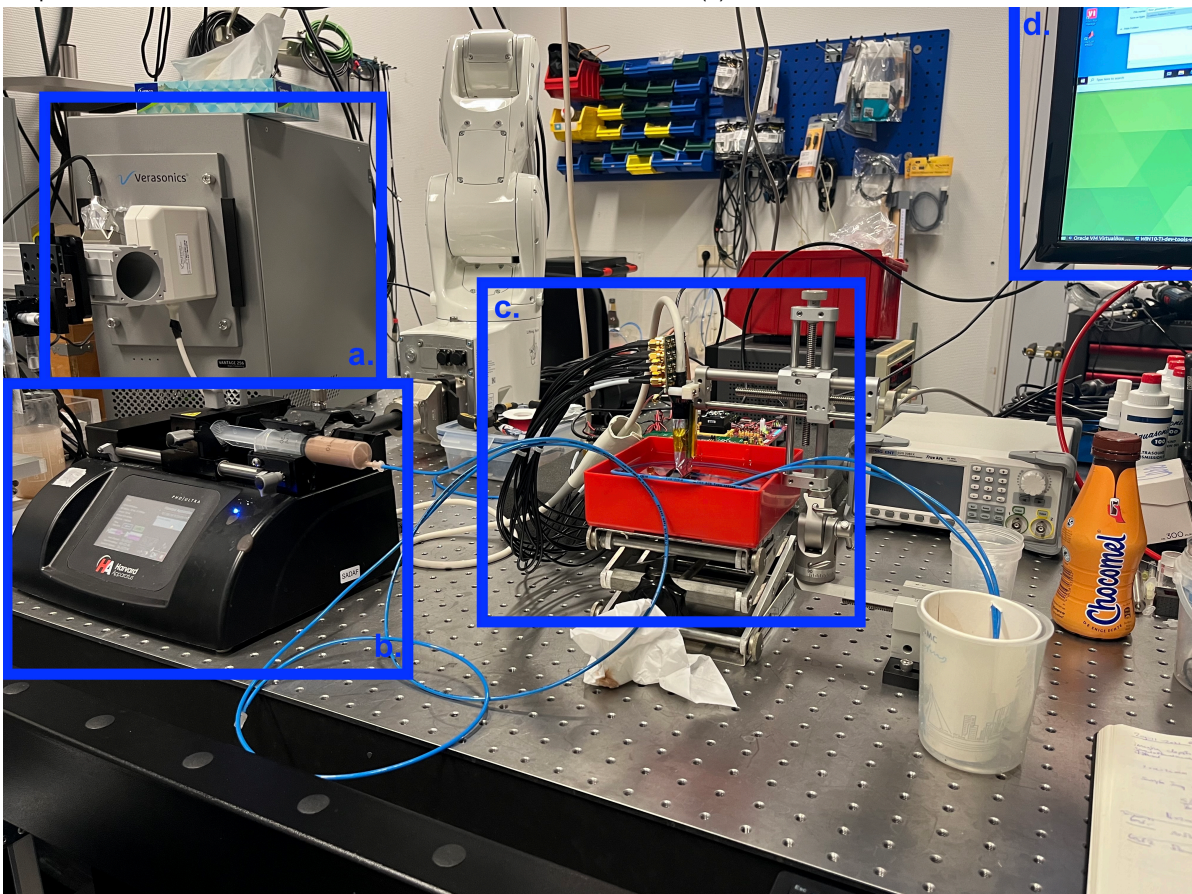

Figure A.5: Picture of the test setup of the flow phantom using the syringe pump to create a flow. The fluid is pushed at a constant flow from the syringe pump (b) to the measurement area (c) in which the fluid flows through a narrow channel with the same characteristics as a blood vessel. In (c) it can also be seen that the custom PCB has been fitted to one part of the probe and the Verasonics Vantage 256 (a) to the other side of the probe. In (d) the data is visualized on the workstation. The setup is placed on a vibration-isolated table.

Bibliography

[1] Analog Devices, Inc. *AFE58JD48 Datasheet*. Last accessed 24 March 2022. URL: `https://www.ti.com/lit/ds/symlink/afe58jd48.pdf?ts=1648134213122&ref_url=https%253A%252F%252Fwww.ti.com%252Fproduct%252FAFE58JD48`.

[2] Analog Devices, Inc. *JESD204B Octal Ultrasound AFE with Digital Demodulator Datasheet*. Last accessed 24 March 2022. URL: `https://www.analog.com/media/en/technical-documentation/data-sheets/AD9671.pdf`.

[3] Analog Devices, Inc. *JESD204B Survival Guide*. Last accessed 13 January 2022. 2011. URL: `https://www.analog.com/media/en/technical-documentation/technical-articles/JESD204B-Survival-Guide.pdf`.

[4] F. Angiolini et al. "1024-Channel 3D ultrasound digital beamformer in a single 5W FPGA". In: *Design, Automation Test in Europe Conference Exhibition (DATE), 2017*. 2017, pp. 1225–1228. DOI: `10.23919/DATE.2017.7927175`.

[5] Arm, Ltd. *Introduction to AMBA AXI*. Last accessed 11 April 2022. 2021. URL: `https://developer.arm.com/documentation/102202/0200/AXI-protocol-overview`.

[6] David Attwell et al. "Glial and neuronal control of brain blood flow". In: *Nature* 468.7321 (2010), pp. 232–243.

[7] Erol Başar. "Brain oscillations in neuropsychiatric disease". In: *Dialogues in clinical neuroscience* 15.3 (2013), p. 291.

[8] Jeremy Bercoff et al. "Ultrafast compound Doppler imaging: Providing full blood flow characterization". In: *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 58.1 (2011), pp. 134–147.

[9] Enrico Boni et al. "Architecture of an Ultrasound System for Continuous Real-Time High Frame Rate Imaging". In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 64.9 (2017), pp. 1276–1284. ISSN: 15258955. DOI: `10.1109/TUFFC.2017.2727980`.

[10] Enrico Boni et al. "Multi-channel raw-data acquisition for ultrasound research". In: *Proceedings - 2014 17th Euromicro Conference on Digital System Design, DSD 2014* (2014), pp. 647–650. DOI: `10.1109/DSD.2014.41`.

[11] Enrico Boni et al. "ULA-OP 256: A 256-Channel Open Scanner for Development and Real-Time Implementation of New Ultrasound Methods". In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 63.10 (2016), pp. 1488–1495. ISSN: 08853010. DOI: `10.1109/TUFFC.2016.2566920`.

[12] Enrico Boni et al. "Ultrasound open platforms for next-generation imaging technique development". In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 65.7 (2018), pp. 1078–1092. ISSN: 08853010. DOI: `10.1109/TUFFC.2018.2844560`.

[13] Richard S.C. Cobbold. *Foundations of biomedical ultrasound*. Oxford university press, 2006.

[14] Giulio Corradi and Jørgen Arendt Jensen. "Real Time Synthetic Aperture and Plane Wave Ultrasound Imaging with the Xilinx VERSAL™ SIMD-VLIW Architecture". In: *2020 IEEE International Ultrasonics Symposium (IUS)*. 2020, pp. 1–4. DOI: `10.1109/IUS46767.2020.9251749`.

[15] R.E. Crochiere and L.R. Rabiner. "Interpolation and decimation of digital signals—A tutorial review". In: *Proceedings of the IEEE* 69.3 (1981), pp. 300–331. DOI: `10.1109/PROC.1981.11969`.

[16] Crucial by Micron Technology, Inc. *DDR Memory Speeds and Compatibility*. Last accessed 30 March 2022. URL: `https://www.crucial.com/support/memory-speeds-compatability`.

[17]   Thomas Cummins, Payam Eliahoo, and K. Kirk Shung. "High-frequency ultrasound array designed for ultrasound-guided breast biopsy". In: *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 63.6 (2016), pp. 817–827.

[18]   Antonello D'Andrea et al. "Transcranial Doppler ultrasonography: From methodology to major clinical applications". In: *World journal of cardiology* 8.7 (2016), p. 383.

[19]   Thomas Deffieux, Charlie Demené, and Mickael Tanter. "Functional Ultrasound Imaging: A New Imaging Modality for Neuroscience". In: *Neuroscience* 474 (2021), pp. 110–121. ISSN: 18737544. DOI: `10.1016/j.neuroscience.2021.03.005`.

[20]   Charlie Demené et al. "Transcranial ultrafast ultrasound localization microscopy of brain vasculature in patients". In: *Nature Biomedical Engineering* 5.3 (2021), pp. 219–228. ISSN: 2157846X. DOI: `10.1038/s41551-021-00697-x`. URL: `http://dx.doi.org/10.1038/s41551-021-00697-x`.

[21]   George Diniz. *JESD204B vs. Serial LVDS Interface Considerations for Wideband Data Converter Applications*. Last accessed 13 January 2022. URL: `https://www.analog.com/en/technical-articles/jesd204b-vs-serial-lvds-interface-considerations-for-wideband-data-converter-applications.html`.

[22]   Valery L. Feigin et al. "Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016". In: *The Lancet Neurology* 18.5 (2019), pp. 459–480.

[23]   Fraunhofer Institute for Biomedical Engineering. *ULTRASOUND BEAMFORMER PLATFORM - DiPhAS*. Last accessed 13 January 2022. URL: `https://www.ibmt.fraunhofer.de/content/dam/ibmt/de/Dokumente/PDFs/ibmt-produktblaetter/ibmt-ultraschall-2018/_ULTRASOUND_BEAMFORMER_PLATFORM.pdf`.

[24]   Gregg Frey and Richard Chiao. *4Z1c Real-Time Volume Imaging Transducer*. last accessed 14 april 2022. URL: `https://www.siemens-healthineers.com/ultrasound/information-gallery/whitepapers/4z1c-real-time-volume-imaging-transducer.html`.

[25]   Gary H. Glover. "Overview of functional magnetic resonance imaging". In: *Neurosurgery Clinics* 22.2 (2011), pp. 133–139.

[26]   Pascal Alexander Hager. "Design of Fully-Digital Medical Ultrasound Imaging Systems". In: 25812 (2019). URL: `https://www.research-collection.ethz.ch:443/handle/20.500.11850/344325`.

[27]   Pascal Alexander Hager and Luca Benini. "LightProbe: A Digital Ultrasound Probe for Software-Defined Ultrafast Imaging". In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 66.4 (2019), pp. 747–760. ISSN: 15258955. DOI: `10.1109/TUFFC.2019.2898007`.

[28]   Sevan Harput et al. "3-D super-resolution ultrasound imaging with a 2-D sparse array". In: *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 67.2 (2019), pp. 269–277.

[29]   Max W. Hauser. "Principles of Oversampling A/D Conversion*". In: *J.AudioEng.Soc.* 39.1/2 (1991).

[30]   Baptiste Heiles et al. "Ultrafast 3D Ultrasound Localization Microscopy Using a 32 × 32 Matrix Array". In: *IEEE transactions on medical imaging* 38.9 (2019), pp. 2005–2015. ISSN: 1558254X. DOI: `10.1109/TMI.2018.2890358`.

[31]   Baptiste Heiles et al. "Volumetric ultrasound localization microscopy of the whole brain microvasculature". In: *bioRxiv* (2021). DOI: `10.1101/2021.09.17.460797`.

[32]   Suzana Herculano-Houzel. "The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost". In: *Proceedings of the National Academy of Sciences* 109.Supplement 1 (2012), pp. 10661–10668.

[33]   Holger Hewener et al. "Integrated 1024 channel ultrasound beamformer for ultrasound research". In: *IEEE International Ultrasonics Symposium, IUS* 2020-September (2020), pp. 2020–2023. ISSN: 19485727. DOI: `10.1109/IUS46767.2020.9251700`.

[34]   Aya Ibrahim et al. "Single-FPGA complete 3D and 2D medical ultrasound imager". In: *2017 Conference on Design and Architectures for Signal and Image Processing (DASIP)*. IEEE. 2017, pp. 1–6.

[35] Jørgen Arendt Jensen. *Estimation of blood velocities using ultrasound: a signal processing approach*. Cambridge university press, 1996.

[36] Jørgen Arendt Jensen et al. "SARUS: A synthetic aperture real-time ultrasound system". In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 60.9 (2013), pp. 1838–1852. ISSN: 08853010. DOI: 10.1109/TUFFC.2013.2770.

[37] K. Kato. "On the mechanism of generation of detected sound in ultrasonic flow meter". In: *Memoirs of the Institute of Scientific and Industrial Research, Osaka University* 19 (1961), pp. 51–57.

[38] Kioxia Holdings Corporation. *Enterprise SSDs*. Last accessed 11 May 2022. URL: https://business.kioxia.com/content/dam/kioxia/shared/business/ssd/doc/EnterpriseSSD_DataSheet_E.pdf.

[39] S.K.E. Koekkoek et al. "High Frequency Functional Ultrasound in Mice". In: *2018 IEEE International Ultrasonics Symposium (IUS)*. 2018, pp. 1–4. DOI: 10.1109/ULTSYM.2018.8579865.

[40] Pieter Kruizinga et al. "Compressive 3D ultrasound imaging using a single sensor". In: *Science advances* 3.12 (2017), e1701423.

[41] Lattice Semiconductor Corporation. *8b/10b Encoder/Decoder*. Last accessed 11 May 2022. 2015. URL: https://www.latticesemi.com/-/media/LatticeSemi/Documents/ReferenceDesigns/1D/8b10bEncoderDecoder-Documentation.ashx?la=en.

[42] John P. Lawrence. "Physics and instrumentation of ultrasound". In: *Critical Care Medicine* 35.8 SUPPL. (2007). ISSN: 00903493. DOI: 10.1097/01.CCM.0000270241.33075.60.

[43] Emilie Macé et al. "Functional ultrasound imaging of the brain". In: *Nature Methods* 8.8 (2011), pp. 662–664. ISSN: 15487091. DOI: 10.1038/nmeth.1641.

[44] Emilie Macé et al. "Functional ultrasound imaging of the brain: Theory and basic principles". In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 60.3 (2013), pp. 492–506. ISSN: 08853010. DOI: 10.1109/TUFFC.2013.2592.

[45] Chris Martin. "Contributions and complexities from the use of in vivo animal models to improve understanding of human neuroimaging signals". In: *Frontiers in neuroscience* 8 (2014), p. 211.

[46] Maxim Integrated Products, Inc. *MAX2088, 16-Channel, High-Performance, Low-Power, Fully-Integrated Ultrasound Receiver*. Last accessed 24 March 2022. URL: https://www.maximintegrated.com/en/products/analog/data-converters/analog-front-end-ics/MAX2088.html/product-details/tabs-3.

[47] Daniele Mazierli et al. "Architecture for an ultrasound advanced open platform with an arbitrary number of independent channels". In: *IEEE Transactions on Biomedical Circuits and Systems* (2021), pp. 1–11. ISSN: 19409990. DOI: 10.1109/TBCAS.2021.3077664.

[48] M. A. Moehring and M. P. Spencer. "Power M-mode Doppler (PMD) for observing cerebral blood flow and tracking emboli". In: *Ultrasound in Medicine and Biology* 28.1 (2002), pp. 49–57. ISSN: 03015629. DOI: 10.1016/S0301-5629(01)00486-0.

[49] Asraf Mohamed Moubark et al. "Selection of Excitation Signals and Acoustic Pressure Measurement for In-Vivo Studies with Ultrasound Array Research Platform II". In: *IOP Conference Series: Materials Science and Engineering* 1070.1 (2021), p. 012095. ISSN: 1757-8981. DOI: 10.1088/1757-899x/1070/1/012095.

[50] Gabriel Montaldo et al. "Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography". In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 56.3 (2009), pp. 489–506. ISSN: 08853010. DOI: 10.1109/TUFFC.2009.1067.

[51] Nvidia Corporation. *NVIDIA RTX A6000 datasheet*. Last accessed 17 March 2022. URL: https://www.nvidia.com/content/dam/en-zz/Solutions/design-visualization/quadro-product-literature/proviz-print-nvidia-rtx-a6000-datasheet-us-nvidia-1454980-r9-web%5C%20(1).pdf.

[52] Peripheral Component Interconnect Special Interest Group. *PCI Express Base Specification Revision 6.0, Version 1.0*. Last accessed 10 May 2022. URL: https://pcisig.com/specifications/pciexpress/technical_library/pciexpress_whitepaper.pdf.

[53] Lorena Petrusca et al. "Fast volumetric ultrasound B-mode and Doppler imaging with a new high-channels density platform for advanced 4D cardiac imaging/therapy". In: *Applied Sciences* 8.2 (2018), p. 200.

[54] Sabino Joseph Pietrangelo. "An Electronically Steered , Wearable Transcranial Doppler Ultrasound System". PhD thesis. 2013.

[55] Sabino Joseph Pietrangelo, Hae Seung Lee, and Charles G. Sodini. "A wearable transcranial doppler ultrasound phased array system". In: *Acta Neurochirurgica, Supplementum* 126.February (2018), pp. 111–114. ISSN: 21978395. DOI: 10.1007/978-3-319-65798-1_24.

[56] PJRC. *Teensy® USB Development Board*. last accessed 17 april 2022. URL: https://www.pjrc.com/teensy/.

[57] John G. Proakis and Dimitris G. Manolakis. *Digital Signal Processing: Pearson New International Edition*. 4th ed. Pearson Education Limited, 2014.

[58] Christoph Risser et al. "High channel count ultrasound beamformer system with external multiplexer support for ultrafast 3D/4D ultrasound". In: *IEEE International Ultrasonics Symposium, IUS* 2016-November (2016), pp. 13–16. ISSN: 19485727. DOI: 10.1109/ULTSYM.2016.7728714.

[59] Christoph Risser et al. "Real-Time Volumetric Ultrasound Research Platform with 1024 Parallel Transmit and Receive Channels". In: *Applied Sciences (Switzerland)* 11.13 (2021). ISSN: 20763417. DOI: 10.3390/app11135795.

[60] Tony J. Rouphael. "RF and Digital Signal Processing for Software-Defined Radio". In: ed. by Tony J. Rouphael. Burlington: Newnes, 2009. Chap. 10, pp. 319–376. ISBN: 978-0-7506-8210-7. DOI: https://doi.org/10.1016/B978-0-7506-8210-7.00010-2.

[61] Jonathan M. Rubin et al. "Fractional moving blood volume: estimation with power Doppler US." In: *Radiology* 197.1 (1995), pp. 183–190.

[62] Jonathan M. Rubin et al. "Power Doppler US: a potentially useful alternative to mean frequency-based color Doppler US." In: *Radiology* 190.3 (1994), pp. 853–856.

[63] Dario Russo and Stefano Ricci. "FPGA Implementation of a Synchronization Circuit for Arbitrary Trigger Sequences". In: *IEEE Transactions on Instrumentation and Measurement* 69.7 (2020), pp. 5251–5259. ISSN: 15579662. DOI: 10.1109/TIM.2019.2952478.

[64] Samtec, Inc. *14 GBPS FIREFLY™ FMC DEVELOPMENT KIT*. last accessed 16 april 2022. URL: https://www.samtec.com/kits/optics-fpga/14g-firefly-fmc.

[65] Samtec, Inc. *FIREFLY™ MICRO FLYOVER SYSTEM™*. last accessed 16 april 2022. URL: https://www.samtec.com/optics/optical-cable/mid-board/firefly.

[66] Laurent Sandrin et al. "Shear modulus imaging with 2-D transient elastography". In: *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 49.4 (2002), pp. 426–435.

[67] Sadaf Soloukey et al. "Functional Ultrasound (fUS) During Awake Brain Surgery: The Clinical Potential of Intra-Operative Functional and Vascular Brain Mapping". In: *Frontiers in Neuroscience* 13 (2020). ISSN: 1662-453X. DOI: 10.3389/fnins.2019.01384. URL: https://www.frontiersin.org/article/10.3389/fnins.2019.01384.

[68] SuperSonic Imagine. *Aixplorer®– Innovative UltraFast Ultrasound Imaging*. Last accessed 31 March 2022. URL: https://www.supersonicimagine.com/Aixplorer-MACH2/Aixplorer.

[69] Thomas L. Szabo. "Acoustic Wave Propagation". In: *Diagnostic Ultrasound Imaging: Inside Out* (2014), pp. 55–80. DOI: 10.1016/b978-0-12-396487-8.00003-3.

[70] Thomas L. Szabo. "Doppler Modes". In: *Diagnostic Ultrasound Imaging: Inside Out*. 2014, pp. 431–500. DOI: 10.1016/b978-0-12-396487-8.00011-2.

[71] Katarzyna M. Szostak, Laszlo Grand, and Timothy G. Constandinou. "Neural interfaces for intracortical recording: Requirements, fabrication methods, and characteristics". In: *Frontiers in Neuroscience* 11 (2017), p. 665.

[72] Li Tan. *Digital Signal Processing Fundamentals and Applications*. Elsevier/Academic Press, 2007.

[73] Mickael Tanter and Mathias Fink. "Ultrafast imaging in biomedical ultrasound". In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 61.1 (2014), pp. 102–119. ISSN: 08853010. DOI: `10.1109/TUFFC.2014.2882`.

[74] Texas Instruments, Inc. *AFE58JD48 12.8-GB JESD204B ultrasound AFE with 16-bit 125-MSPS ADC evaluation module*. Last accessed 29 April 2022. URL: `https://www.ti.com/tool/AFE58JD48EVM`.

[75] Texas Instruments, Inc. *Ultra low-noise JESD204B compliant clock jitter cleaner with integrated 1840 to 1970-MHz VCO0*. Last accessed 29 April 2022. URL: `https://www.ti.com/product/LMK04826`.

[76] Elodie Tiran et al. "Transcranial functional ultrasound imaging in freely moving awake mice and anesthetized young rats without contrast agent". In: *Ultrasound in medicine & biology* 43.8 (2017), pp. 1679–1689.

[77] Ultrasonics and Instrumentation - School of Electronic and Electrical Engineering - University of Leeds. *Instrumentation*. Last accessed 13 January 2022. URL: `https://institutes.engineering.leeds.ac.uk/ultrasound/facilities_instrumentation.html`.

[78] Verasonics, Inc. *Signal Conditioning in the Vantage Research Ultrasound System Receive Path Adds Flexibility and Efficiency to Research Projects*. Last accessed 13 January 2022. 2020. URL: `https://verasonics.com/signal-conditioning-in-the-vantage-research-ultrasound-system/`.

[79] Verasonics, Inc. *The Vantage ™ Research Ultrasound Systems*. Last accessed 13 January 2022. URL: `https://verasonics.com/vantage-systems/`.

[80] Rijksinstituut voor Volksgezondheid en Milieu (RIVM). *Een op vier Nederlanders heeft hersenaandoening*. Nov. 2017. URL: `https://www.rivm.nl/nieuws/op-vier-nederlanders-heeft-hersenaandoening`.

[81] Yonghao Wang and Joshua D. Reiss. "Time domain performance of decimation filter architectures for high resolution sigma delta analogue to digital conversion". In: *Audio Engineering Society Convention 132*. Audio Engineering Society. 2012.

[82] Xilinx, Inc. *VCK190 Evaluation Board - User Guide*. 2021. URL: `https://docs.xilinx.com/r/en-US/ug1366-vck190-eval-bd`.

[83] Xilinx, Inc. *Versal ACAP DSP Engine Architecture Manual (AM004)*. Last accessed 11 April 2022. URL: `https://docs.xilinx.com/r/en-US/am004-versal-dsp-engine`.

[84] Xilinx, Inc. *What is an FPGA?* Last accessed 10 May 2022. URL: `https://www.xilinx.com/products/silicon-devices/fpga/what-is-an-fpga.html`.

[85] Xilinx, Inc. *Xilinx AI Engine Technology*. Last accessed 11 April 2022. URL: `https://www.xilinx.com/products/technology/ai-engine.html`.

[86] Xiaochen Xu et al. "Open platform for accelerating smart ultrasound transducer probe development". In: *IEEE International Ultrasonics Symposium, IUS* 2020-Septe.c (2020). ISSN: 19485727. DOI: `10.1109/IUS46767.2020.9251594`.

[87] Jaesok Yu et al. "Design of a volumetric imaging sequence using a Vantage-256 ultrasound research platform multiplexed with a 1024-element fully sampled matrix array". In: *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 67.2 (2019), pp. 248–257.

[88] Richard Zarr. *What to Know About the Differences Between JESD204B and JESD204C*. Last accessed 17 January 2022. 2020. URL: `https://www.ti.com/lit/wp/sbaa517/sbaa517.pdf?ts=1642379186363`.