Energy Scalable Systems for 2D and 3D Low-Power Ultrasound Beamforming

by

Bonnie Kit Ying Lam

B.A.Sc., University of British Columbia (2008)
S.M., Massachusetts Institute of Technology (2010)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2017

Author ..... Signature redacted .....
Department of Electrical Engineering and Computer Science
May 19, 2017

Certified by..... Signature redacted .....
Anantha P. Chandrakasan
Vannevar Bush Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Certified by.. Signature redacted .....
Gerald J. Sussman
Panasonic Professor of Electrical Engineering
Thesis Supervisor

Accepted by .. Signature redacted .....
Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

Energy Scalable Systems for 2D and 3D Low-Power Ultrasound Beamforming

by

Bonnie Kit Ying Lam

Submitted to the Department of Electrical Engineering and Computer Science
on May 19, 2017, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

In traditional ultrasound imaging systems, bulky and power-intensive mainframes are used to process the high number of waveforms acquired in parallel from a large transducer array. The computational power of these systems scales linearly with transducer count. However, there exist applications where basic functionality in low-power conditions may be favorable to an "all-or-nothing" system that only produces a high resolution image when enough power is supplied. This thesis presents systems designed to support energy-scalability at run-time, enabling the user to make the tradeoff between power and performance.

First, a system-level energy model for a receive-side digital beamforming system is presented. Power-performance tradeoffs for the analog front-end, analog-to-digital converter, and digital beamformer are analyzed individually and then combined to account for the performance dependency between the functional components. These considerations inform a recommendation on design choices for the end-to-end system.

Second, this thesis describes an energy-scalable 2-D beamformer that provides user-controlled run-time tradeoff between image quality and energy consumption. Architectural design choices that enable three operating modes are discussed. A test chip was fabricated in 65-nm low power CMOS technology. It can operate with functional correctness at 0.49 V, with a measured power of 185 μW in real-time operation at 0.52 V.

Finally, a software-based energy-scalable 3-D ultrasound beamformer is implemented on an embedded supercomputer. The energy consumption and corresponding imaging quality are measured and compared.

Thesis Supervisor: Anantha P. Chandrakasan
Title: Vannevar Bush Professor of Electrical Engineering and Computer Science

Thesis Supervisor: Gerald J. Sussman
Title: Panasonic Professor of Electrical Engineering

Acknowledgments

I feel extremely fortunate to have worked under the joint supervision of Professor Anantha Chandrakasan and Professor Jerry Sussman. Since the day that I received that transformative phone call from Anantha notifying me of my acceptance to the PhD program at MIT EECS, he has been a champion for my success. I still remember that first in-person conversation we had during visit weekend, where he assured me that no matter which school I chose to join, I could not make a mistake. I can honestly say now that this was the best decision I could have made. Throughout the many years of ups and downs, he has offered me so much more than the research supervision that I had anticipated. He has gifted me with technical, financial, and even emotional support, pushing me to improve and trusting that I can do it especially when I doubted myself. I will be forever grateful for all the time and energy he has spent to mentor me, and hope this relationship continues well beyond my time at MIT.

I would like to thank Jerry for serving as my co-advisor. Even though I only got to know Jerry near the end of my time at MIT, he has taught me a lot about life and engineering, especially his rule about working on things only if they are fun! Always ready to offer me technical advice and high quality tea, he has truly enriched my graduate career by stepping in to make sure that I got "a round tuit"!

I would like to thank Professor Vivienne Sze for serving in my thesis committee and providing insights about image processing and low-power design tradeoffs. Vivienne was one of the first people I met at MIT and has always been a role model to me. I feel proud to have interacted with her in so many different capacities!

To Ananthagroup past and present, thank you for enriching both the technical and social aspects of my time at MIT. I would like to especially thank Dr. Kailiang Chen and Dr. Michael Price for being awesome collaborators in my research projects, and Dr. Dennis Buss for directly arranging my internships at and fellowship from Texas Instruments. The ISSCC trips we made to San Francisco were so much more than simply educational because of the group I was with, and who can forget the numerous "procrastination circles" that spontaneously formed in the lab. Thank you Joyce and

Vivienne for being awesome examples and giving super helpful life and career advice! To all my Ananthagroup twins over the years (Yildiz, DongNi, and Chu, in order of appearance and subsequent twin-hood), thank you for being wonderful peers and making me feel like I always had a buddy around! Thank you Margaret for being super helpful with reimbursements and other logistics!

To my other friends that I met at MIT, you have all been an irreplaceable part of the MIT experience for me. I was never a popular kid growing up, but somehow managed to stumble into wonderful friendships with all of you. Lei, thank you for inviting me to join Isshinryu Karate with you and for single-handedly awakening the fitness bug in me. You and Yehua have been such good couple friends to me and Mario throughout the years and I hope the distance doesn't weaken our friendship! Ying and Ermin, I really enjoyed rising in the ranks with you both in karate and in graduate school, you are both amazing! Yuanyuan, though we first met in UBC, I feel like you have added so much fun to my social life here in Boston. It has been great seeing you succeed at work, in business school, in your startup, and now as Caelan's mother. Thank you for being there for me! Audrey and Grace, thank you for including me in your wonderful sisterhood and showing me what great girlfriends are for! Karen and Vincent, you provided wonderful company and awesome potlucks when I was first adjusting to life in Boston! Qing, Henna, Shen, thanks for all the fun times! To all of you, I hope that we will stay in touch for years to come!

Last but definitely not least, none of this could have been possible without the support of my parents, Carly and Alfred, my sister Debbie, and my loving husband Mario. Mom, thank you for sacrificing your free time to visit me several times a year so that I would be less homesick. Dad, thank you for believing in my abilities and for bragging about me to all your friends! Debbie, thank you for growing up to become a sort of younger role model to me, and for always being there at the most important moments of my life. Mario, just like it was in our time together at UBC, it has been a pleasure working and fighting towards the finish line next to you. I don't think most people realize how much time and effort you have dedicated towards keeping me focused and on track to graduate!

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Among the numerous medical imaging modalities available, ultrasonography is most commonly used due to its relatively low cost, short imaging time, and low risk to human safety. While most commonly known for its use in obstetric sonography during pregnancy, it can also be used to non-intrusively visualize many other body structures, including tendons, muscles, joints, vessels, and internal organs [9]. Traditional ultrasound imaging systems are expensive - both in terms of cost and power - and bulky, consisting of large processors and cables to accommodate the large amount of data channels on ultrasound probes. The movement to portable ultrasound systems for point-of-care applications has been explored by companies such as General Electric [1, 10] (pictured in Figure 1-1), Clarius [2] (pictured in Figure 1-2 and Philips [11], and the proper tradeoffs between power and performance need to be made when constructing these systems. These advancements are largely enabled by the Moore's Law [12].

1.1 Low Power Diagnostic Applications

Conventional ultrasound systems are often large and expensive due to the intensive computational requirement associated with large transducer arrays and the corresponding downstream electronics. Since the goal is to generate high-resolution, accurate representations of the object to be imaged and mains electricity is usually readily avail-

Figure 1-1: Vscan™: General Electric's handheld, pocket-sized ultrasound tool [1].

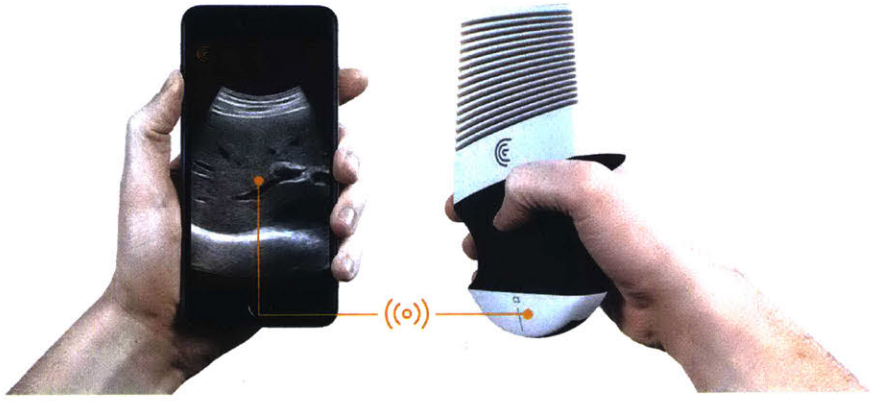


Figure 1-2: Clarius™'s wireless, handheld ultrasound device with mobile app integration [2].

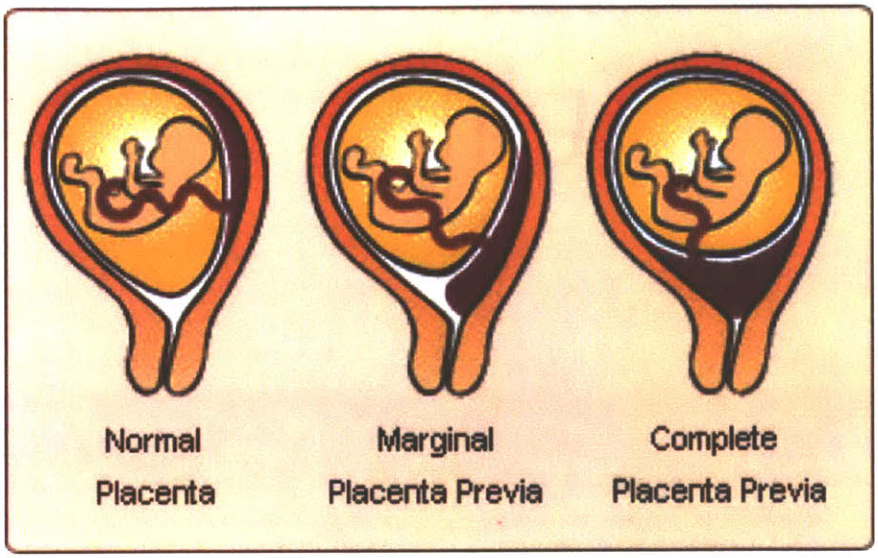


Figure 1-3: Placenta previa [3].

able in the hospital setting, these systems are often built from standalone, off-the-shelf components that independently satisfy the best specifications possible.

In certain circumstances, there are applications where low power operation is preferred or required, even at the expense of image quality. One example is in the case of detecting prenatal risk factors such as placenta previa in rural settings where power is not readily or consistently available. Figure 1-3 shows a side-by-side comparison between the location of a normal placenta and the marginal and complete cases of placenta previa [3]. When the placenta grows at the lowest part of the uterus and partially or complete blocks the opening to the cervix, vaginal delivery is dangerous and can result in maternal and fetal death. After diagnosis, bed rest and cesarean delivery would be recommended in the hospital setting [13]. In rural areas where expectant mothers do not have easy access to hospitals, portable, energy-efficient systems must be used to monitor and diagnose placenta previa instead of conventional, high-performance ultrasound systems. This way, the appropriate arrangements to send the mother to a hospital can be made ahead of expected delivery. In this scenario, the use of at least partially battery-operated systems is preferred so that diagnosis is still possible when mains electricity is intermittent or unavailable.

Because the object that needs to be identified and located in the image is relatively

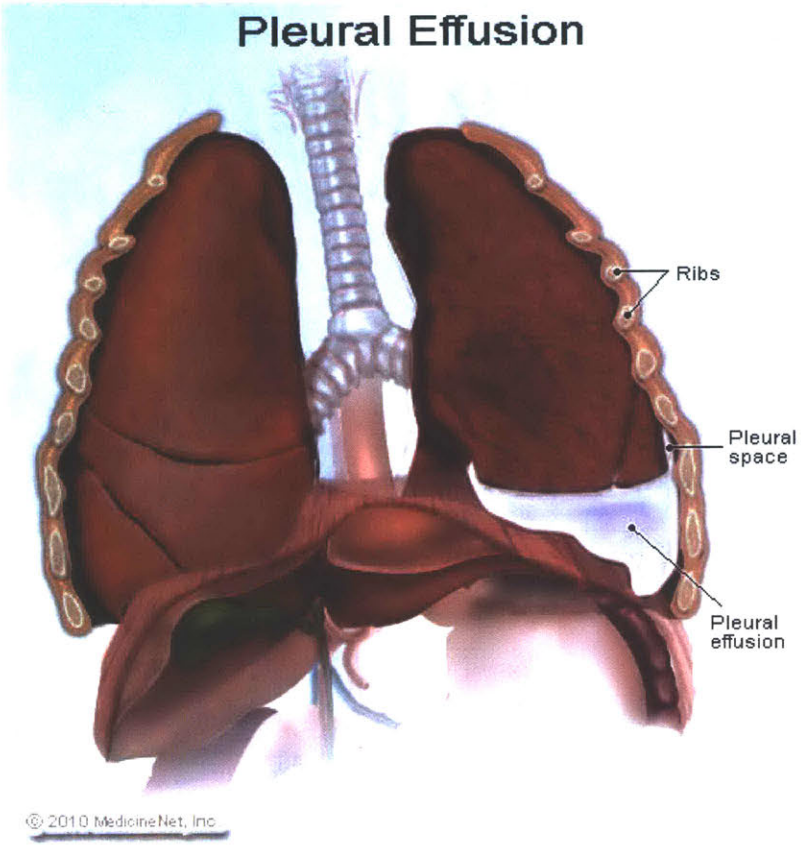


Figure 1-4: Pleural effusion [4].

large, we identified placenta previa as a potential use case for a low-power system where image quality degradations may be acceptable. Another such application is pleural effusion, where the volume of fluid buildup in the chest or lung [14]. A diagram of where this fluid can exist is shown in Figure 1-4. In these use cases, some useful information can still be extracted from a lower quality image when necessary.

1.2 Energy Scalability

As motivated in the last section, there are situations in which generating a degraded image might be preferable to an all-or-nothing strategy under energy-constrained circumstances. However, as high image quality can provide more accurate information to aid in diagnosis, a system that strictly generates degraded images is not as useful as one that is energy-scalable. As in the case of many hardware systems designed to be

energy scalable in a wide range of applications including computational photography [15, 16], sensor signal processing [17], EEG processors [18], discrete cosine transforms [19], and terahertz-wave oscillators [20] the ability to reconfigure the system to operate at various performance and power consumption levels at run time is desirable.

In ultrasound beamforming, a large number of processing channels, each corresponding to and supporting one element in a transducer array, are combined to reconstruct an image. In this thesis, we will analyze a number of parameters specific to ultrasound imaging that can affect the computational complexity (and, as a result, the computational power) and implement systems where these parameters are reconfigurable by the user during run-time. This will allow the systems to operate at a wide range of energy consumption levels, based on the currently available power budget. The resultant tradeoff with performance in terms of frame rate and image quality are quantified and examined.

1.3 Contributions

This thesis explores the use of parallel architectures to enable low-power, scalable 2D and 3D ultrasound imaging while maintaining acceptable performance levels. The energy versus performance tradeoffs are explored and quantified.

1. **System Energy Model.** A system-level model of the receive-side ultrasound beamforming system, including analog front end amplification, analog-to-digital conversion, and digital beamforming, is developed [21].

2. **2D Beamformer Test Chip.** An energy-scalable ASIC for 2D digital beamforming is implemented, fabricated, and tested for power and performance [22].

3. **3D Beamformer Software.** A software-based 3D digital beamformer is executed on an embedded platform and power consumption is measured. The relative energy savings associated with various algorithmic optimizations are quantified along with the corresponding image quality.

1.4 Thesis Organization

This thesis is organized as follows:

In Chapter 2, the background information and nomenclature for ultrasound beamforming are discussed. This includes descriptions of the commonly used ultrasound imaging modes, the coordinate system used for beamforming, the ultrasound imaging operation, beam formation, and delay-and-sum beamforming. Previous implementations of handheld and portable ultrasound systems are also highlighted.

Chapter 3 presents the system-level energy model. The individual power-performance tradeoff of each electronic component is described, and the noise and image quality dependencies examined. The usage model of three spatial decimation modes is discussed at length. Conclusions from this work informs the design of the 2D beamforming ASIC.

In Chapter 4, the architectural design of the 2D beamforming ASIC is detailed. The functional blocks that implement the delay, apodization, and summation stages are presented individually.

Chapter 5 explores the implementation-specific details for the 2D beamformer test chip. Input and output handling are discussed at length. The measurement setup and results are presented.

Chapter 6 extends the concept of energy scalability to 3D ultrasound imaging. The spatial decimation in number of transducer count used in the 2D beamformer design is applied to 3D and the impact on image quality is quantified. The plane wave coherent compounding algorithm and the tradeoff on both computational requirements and performance are examined.

Finally, a summary of thesis contributions and future directions for ultrasound beamforming is provided.

Chapter 2

Background on Ultrasound Imaging

To fully understand the challenges of low-power ultrasound beamforming, several concepts and terminology need to be examined. This chapter develops the framework upon which this thesis builds in developing energy-scalable systems in this application space.

2.1 Overview of Imaging Modes

Depending on the diagnosis required, one of several "imaging modes" is used in medical ultrasound. The amplitude (A), brightness (B), constant depth (C), motion (M), and Doppler (D) modes are most commonly used to visualize tissue anatomy [23–25].

The A-mode describes the use of a single ultrasound tranducer to generate a waveform, where the amplitude at each time t (corresponding to a particular depth z) represents the signal strength of the echo received from that depth. This means that the waveform is a time-varying signal reminiscent of the display on an oscilloscope. The intensity of the ultrasound signal received over time represents changes in reflectivity at various depths in the human body. This mode is mostly obsolete and one of the few uses for this mode is in ophthalmology [26].

The B-mode is the most common and essential modality in diagnostic ultrasound. When using a single transducer to transmit and receive ultrasound pulses, a one-dimensional (1-D) image grayscale is formed, where the brightness at each time

t (again, corresponding to the depth z) corresponds to the amplitude of the echo received from that depth. The operation involved in reconstructing a 1-D B-mode image (or equivalently, an A-mode line when displayed by amplitude rather than brightness) from a single transducer is described in more detail in Section 2.2. More generally, a two-dimensional (2-D) plane in the body is scanned using a 1-D array of transducer elements and a gray-scale image is reconstructed. This image is essentially a combination of multiple 1-D "scan lines" that span the width of the probe, with the added benefit of azimuthal spatial resolution using beamforming, which will be discussed in Section 2.4.

The C-mode refers to a 2-D image formed along the plane perpendicular to the 2-D B-mode image. The equivalent imaging area could be found in a traditional x-ray image at a specified depth.

The M-mode captures the movement of structures toward or away from the probe over time by taking a series of A-mode or B-mode images (essentially generating a video composed of many A- or B-mode frames, and can be used to accurately evaluate rapid movements of internal organs such as the heart.

While each of the four modes described so far provide visualization of tissue anatomy for diagnostic purposes, the Doppler mode uses the Doppler effect [27] to compute velocity information inside the human body (in particular for measuring the speed of blood flow). Three commonly used Doppler modes were developed at various times with different advantages and limitations. The Continuous Wave (CW) Doppler mode detects velocities along a line by continuously transmitting ultrasound pulses and continuously listening for the Doppler frequency shifts from the echo waves [28]. While CW Doppler mode can measure very high velocity flows, this method lacks the ability to determine the location at which those velocities originate. The Pulsed Wave (PW) Doppler mode, on the other hand, sends a pulsed signal for a chosen area or depth, then stops transmission and listens for the reflected frequency shift from that particular depth [29]. The Color Flow Doppler mode enhances the B-mode image by overlaying a color map representing the estimated velocity at each location [30].

2.2 B-Mode Imaging Operation

As introduced in Section 2.1, B-mode imaging refers to the grayscale visual description of an area, with the brightness of each pixel indicating the strength of the echo signal from that particular location in the body. This is achieved by analyzing echoes from acoustic waves traveling through the target medium, which is the human body in medical ultrasonography [31]. Physical boundaries between different materials (e.g. types of tissue) are impedance discontinuities that partially reflect the wavefront created by the transducer array. The length of time it takes for the echo to arrive back at the transducer directly translates into a location of such a boundary, since time and distance are related by the speed of sound propagation in the medium.

This concept is demonstrated through the following sequence, schematically represented in Figure 2-1 for the example of a single transducer forming a 1-D grayscale image. The area being imaged in this example contains two parallel line scatterers. A scatterer refers to a location where there is an acoustic impedance mismatch, typically at the interface of different types of tissues (e.g. between fluid and soft tissue, and between soft tissue and bone). In the case of a prenatal ultrasound, these are the boundaries between the baby, amniotic sac, and the placenta.

The imaging process begins with the transmitter circuit generating an excitation waveform to drive the transducer. This causes the transducer to emit a pulse of a specified amplitude at time = 0, shown in red. Note that this pulse is sent out periodically, and that the period of this signal, labeled T_{PR} and known as the pulse repetition period, limits the frame rate of the system. Once the pulse is sent out from the transducer located at depth $z = 0$ (i.e. the surface of the patient or other phantom being imaged), the transducer toggles into receiver mode, listening for sound pressure waves returning to the transducer. In the meantime, the first red pulse propagates from $z = 0$ away from the transducer (to the right of the figure) at the speed of sound c in human soft tissue. This is generally taken to be $1540m/s$ [32]. It also experiences partial attenuation as it travels through the medium (this phenomenon is further discussed in Section 2.4). After some time $\Delta t = t_1$, the pulse encounters the first

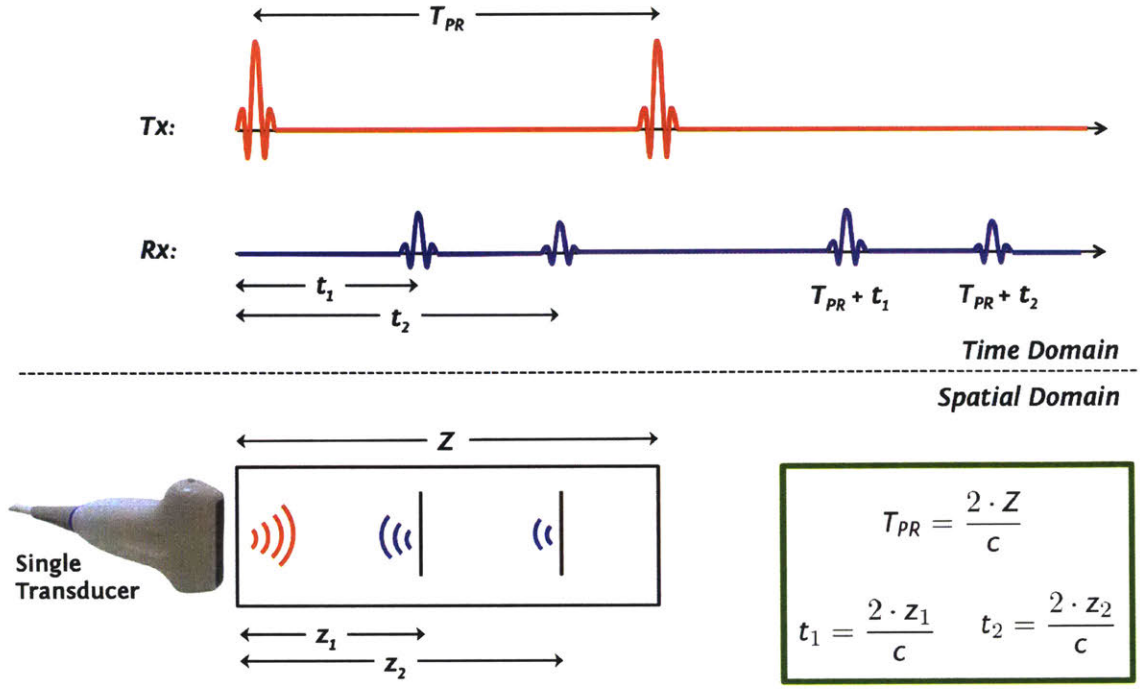


Figure 2-1: Schematic representation of the formation of a 1-D B-mode image.

scatterer at $z = z_1$. The time t_1 and depth z_1 are related through the speed of sound as follows:

$$t_1 = \frac{2 \cdot z_1}{c}, \tag{2.1}$$

The acoustic impedance mismatch at this interface reflects a fraction of the signal back towards the transducer while the remainder of the transmitted signal continues traveling away from the transducer, until it comes into contact with the second scatterer at $z = z_2$ some time later at $\Delta t = t_2$. Similarly, we can write t_2 in terms of z_2 and c, as follows:

$$t_2 = \frac{2 \cdot z_2}{c}, \tag{2.2}$$

At each of those boundaries, the amplitude of the reflected pulse depends on both the extent of impedance mismatch on either side of the boundary. This translates into the brightness of the pixels at those locations in the resulted image. The maximum

depth visible (labeled as Z in Figure 2-1) in the image also directly relates to the pulse repetition period T_{PR} by the speed of sound c:

$$T_PR = \frac{2 \cdot Z}{c}, \tag{2.3}$$

Furthermore, the pulse repetition period is equal to the amount of time it takes to generate one frame of the 1-D "image", which is inversely proportional to frame rate.

2.3 Orientation and Dimensionality

The coordinate system typically assigned to the imaged space for B-mode imaging is shown in Figure 2-2. The direction along the 1-D transducer array is labeled as the x-axis, and this is known as the lateral (or azimuthal, used interchangeably) direction. Since the ultrasound waves are transmitted from the transducers into the imaging medium, the direction along which the waves travel is known as the axial direction, and labeled as the z-axis. This dimension is perpendicular to the surface of the transducer array, and is also known as the depth. Two-dimensional B-mode ultrasound images occupy the $xz-$plane. The third dimension, known as the elevational direction, is labeled as the y-axis.(As described in Section 2.1, a 2-D C-mode image occupies the plane perpendicular to a 2-D B-mode image, or in this coordinate system, the $xy-$plane, at a particular depth $z = D$.) In this system, a 1-D transducer array consists of M array elements located at positions $(x_0, 0, 0)$ to $(x_{M-1}, 0, 0)$.

A 3-D image can be formed using a 1-D transducer array by freehand manipulation of the transducer probe along the elevation direction and combining the resulting B-mode slices into a volume. This method is limited by poor resolution in the direction of freehand motion, and the maximum frame rate attainable is necessarily reduced to accommodate the time it takes for the operator to cover the span of the elevation dimension manually. A single-transmission without elevation focus followed by frequency domain reconstruction along the y-direction has been proposed to improve elevation resolution and frame rate while maintaining axial resolution [5]. A side-by-side comparison of the acoustic field generated in conventional freehand 3-D

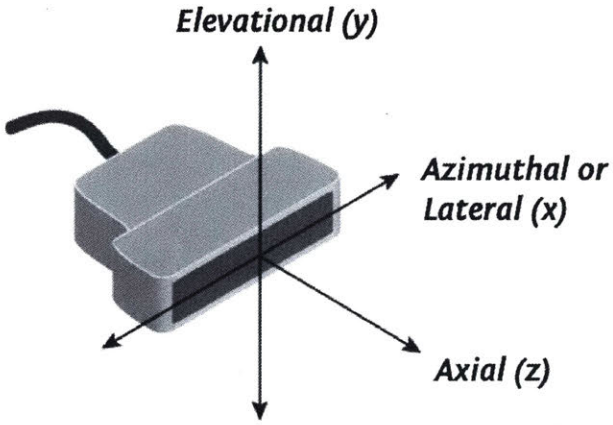


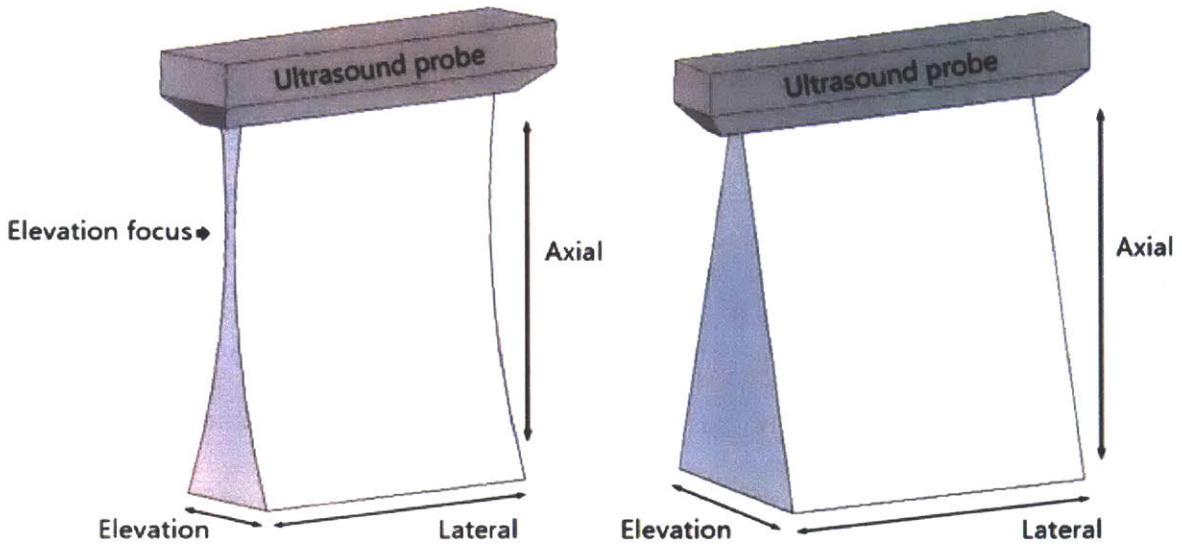
Figure 2-2: Three-dimensional coordinate system for ultrasound imaging.

ultrasound and the proposed system is shown in Figure 2-3.

As advanced process technologies allow for fine-pitched elements, 2-D transducer arrays are used to replace the freehand technique with a 1-D array. Systems using 2-D arrays can support significantly higher elevation resolution and frame rate because no operator motion is required, and the element pitch (distance between neighboring elements) can be made the same in both the x and y directions. In a 2-D transducer array, the N_x by N_y array contains elements located at positions $(x_0, y_0, 0)$ to $(x_{P-1}, y_{Q-1}, 0)$. A pictorial representation of such a system is shown in Figure 2-4.

2.4 Delay-and-Sum Beamforming

A combination of several factors result in the signal received by an individual transducer element having a low signal-to-noise ratio (SNR). Firstly, the maximum acoustic energy transmitted into the human body is limited by FDA safety guidelines [33]. Depending on the combination of the type of transducer element and driver circuit used, the ratio between the input electrical power and the output acoustic power can be as low as 0.66% [34]. For systems where the transmit driver electronics are integrated with the probe, the amount of power dissipated in the transducer is further limited by the transducer surface temperature, the quantity formally regulated by the International Electrotechnical Commission (IEC) [35–37]. Since most of the losses corresponding to the low electro-acoustic efficiency in the transmit phase will convert to heat, these



(a) Acoustic field in conventional freehand. (b) Acoustic field in frequency reconstruction.

Figure 2-3: Comparison between conventional freehand 3-D imaging and a proposed system without elevation focus [5].

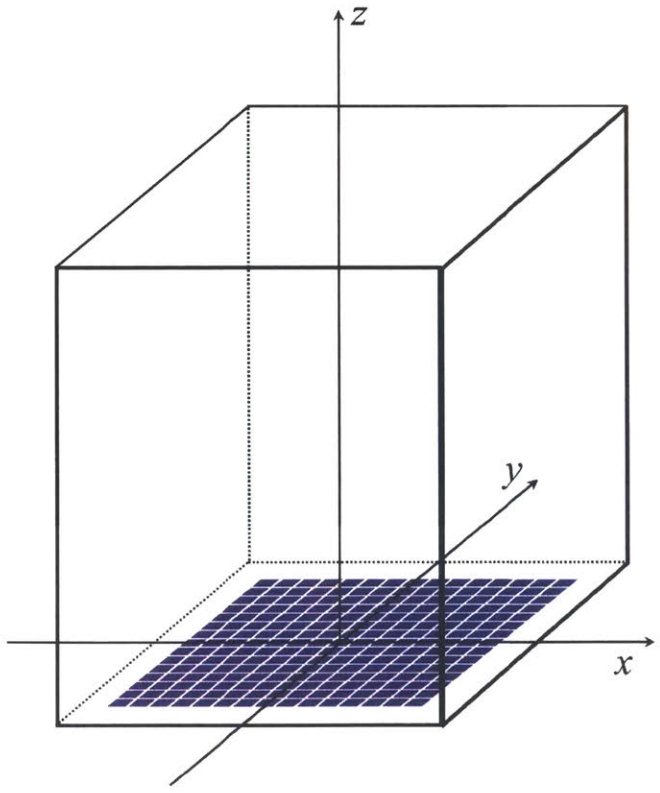


Figure 2-4: An example of a two-dimensional transducer array in the 3-D coordinate system, where $P = Q = 15$.

two constraints on the transmit operation translate to a low energy acoustic wave being sent into the medium. Secondly, as the ideal plane wave propagates through the medium, it is subject to a number of energy loss mechanisms collectively known as attenuation [38–42]. One component of attenuation, known as absorption, refers to the conversion of ultrasound energy into heat, chemical energy, or light [43–46]. The other mechanism, known as scattering, involves the redirection of ultrasound wave energy to a path different from the path of the incident wave [47,48]. Scattering typically accounts for to a small percentage of attenuation in soft tissue, on the order of 10 to 15% [49].

To compensate for the low SNR signals received by individual transducers, large transducer arrays ranging from 128 to 1024 elements are typically used to improve SNR. Beamforming [50] is applied to the received signals to provide spatial selectivity and construct images. Increased array sizes lead to increased beamformer complexity and power consumption.

Figure 2-5 shows a simplified schematic of the ultrasound imaging scenario. The beamformer inserts the appropriate time delay in the received signal to complement the time-of-flight from each transducer element to the desired focal point [51,52]. This coherent summing increases the signal strength to improve image reconstruction.

2.4.1 Two-Transducer Example

Building on this simplified model of the single pulse received by each channel corresponding to a single focal point, we note that there may exist multiple paths through which the sound waves can propagate back towards the transducer array. To analyze this in more detail, the example of an array of two transducer elements (labeled A and B) interacting with two targets (labeled a and b), shown in Figure 2-6, examined. Since there are now two sensors that also acted as the sources of the transmitted waveform in our system, we must consider all paths of propagation. Due to the symmetry in the locations of the sensors and targets, several of these paths have the same length. Section 2.4.1 shows the list of all possible paths taken by sound waves originating from these two transducer elements along with the distance traveled by

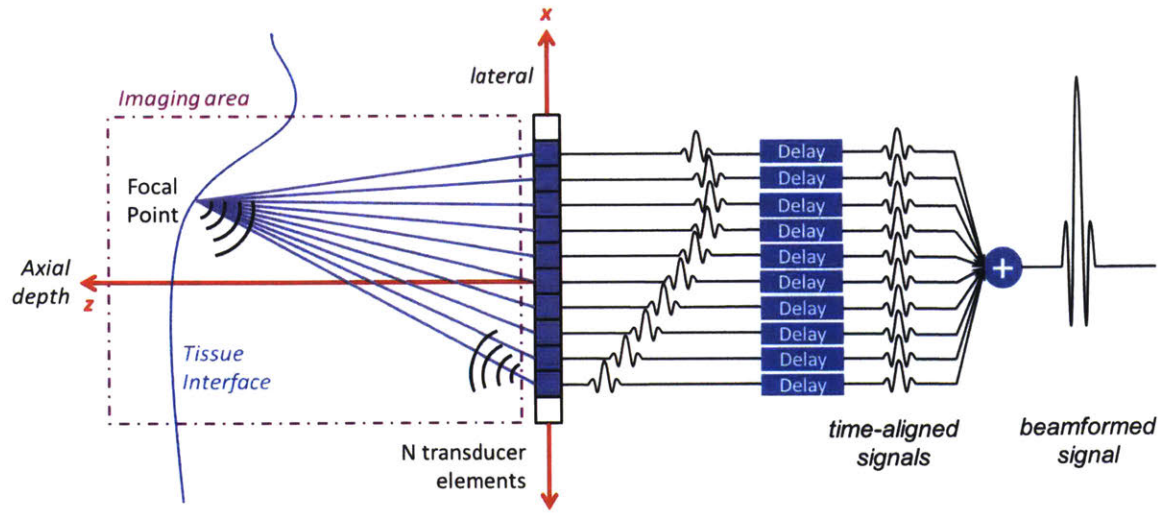


Figure 2-5: Conceptual illustration of beamforming with uniform weights.

the sound wave in each case. The lengths d and L_1 are labeled in the diagram, and the distance between transducer A and target b can be found as:

$$L_2 = \sqrt{d^2 + L_1^2} \tag{2.4}$$

As with the case of a single transducer, explained in Section 2.2, the time of arrival t can be directly related to the distance traveled through the speed of sound c. Assuming $d < L_1$ such that $L_1 < L_2 < 2 \cdot L_1$, we can draw the timing diagram of the signals received by the two transducers in Figure 2-7. Based on this assumption about the relative locations of the sensors and targets, we can also write that $2 \cdot L_1 < L_1 + L_2 < 2 \cdot L_2$. We further assume that the signal strength reduces with the distance traveled in the medium (due to the sources of attenuation described in Section 2.4). According to this assumption, we draw the pulses reflected via different paths to have varying relative amplitudes in Figure 2-7. Note that time gain compensation is usually applied to account for the diminishing magnitude as the wave propagates away from the transducer [53–59]. This is typically normalizing the signal amplitude to compensate for the increasing effect of attenuation [60].

As in the single transducer example illustrated in Figure 2-1, the red waveform

Table 2.1: Enumeration of paths taken by ultrasound waves originating from an array of two transducer elements.

Path	Transmitted from	Reflected at	Received by	Path length
1	A	a	A	$L_1 + L_1$
2	A	a	B	$L_1 + L_2$
3	A	b	A	$L_2 + L_2$
4	A	b	B	$L_2 + L_1$
5	B	a	A	$L_2 + L_1$
6	B	a	B	$L_2 + L_2$
7	B	b	A	$L_1 + L_2$
8	B	b	B	$L_1 + L_1$

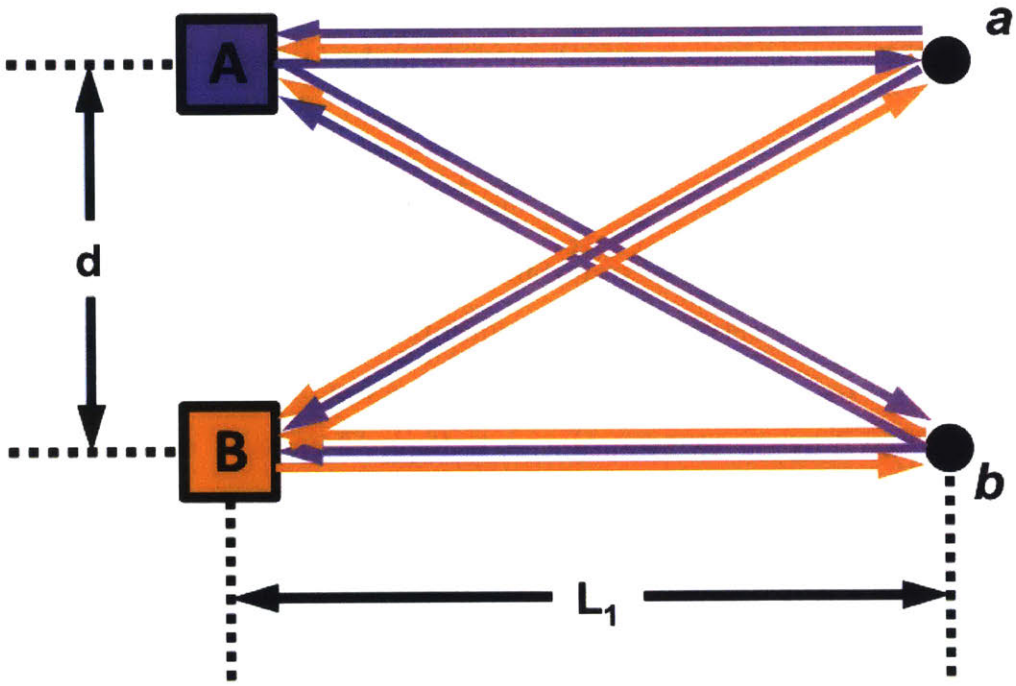


Figure 2-6: Schematic representation of the propagation paths between two neighboring transducer elements (A and B) and two point targets (a and b). Paths of wavefronts originating from transducer A are indicated in purple, whereas those originating from transducer B are colored yellow.

represents the pulse emitted by each transducer. In this example, both transducer A and B send a pulse at the same time $t = 0$. At $t = t_1 = t_8$, transducer A receives the pulse corresponding to path 1 (marked in purple), whereas transducer B receives the pulse corresponding to path 8 (marked in yellow).

At $t = t_5 = t_7 = t_2 = t_4$, transducer A receives the combined effect of the pulses corresponding to paths 5 and 7 (marked in yellow, as these pulses originate from transducer B), whereas transducer B receives the combined effect of the pulses corresponding to paths 2 and 4 (marked in purple, as these pulses come from transducer A).

Note that the amplitude of these pulses are expected to be less than twice that of the first pulse, because of the pulse from path 5 should have a smaller amplitude than that of the first pulse due to attenuation (the longer length of path 5 means more attenuation through the medium). Depending on the degree of attenuation, the combined pulse at $t = t_5$ could have a larger or smaller amplitude than the pulse at $t = t_1$. In the figure, it is assumed that the relative attenuation is less than 50%.

Finally, at $t = t_3 = t_6$, transducer A receives the pulse corresponding to path 3 (marked in purple), whereas transducer B receives the pulse corresponding to path 6 (marked in yellow).

With dynamic focusing where a different delay profile is generated for multiple focal points to equalize resolution across an image (explained in more detail in Section 2.4.3) [61]. In the limit where each pixel in an image is reconstructed using its location as the focal point (also known as software beamforming), only the value of the beamformed signal at the time that corresponds to the location of the pixel is used. Depending on the focal point we are interested in, we delay one of the two received signals relative to the other before summing, the amount of delay being dictated by the difference in propagation paths to that focal point. In our example, when considering target a, we would delay the signal received by transducer A by Δt_a before summing the two waveforms, resulting in the first waveform shown in blue, where

$$\Delta t_a = t_6 - t_1 = \frac{2 \cdot (L_2 - L_1)}{c} \tag{2.5}$$

Similarly, when considering target b, we would delay the signal received by transducer B by Δt_b and sum the two waveforms, resulting in the second waveform shown in blue, where

$$\Delta t_b = t_3 - t_8 = \frac{2 \cdot (L_2 - L_1)}{c} \tag{2.6}$$

The extra pulses at locations other than the location of interest (in this case, t_6 for target a and t_3 for target b) do not come into play because the software beamformer would only consider the value of the combined signal at those times to represent the image content at the corresponding locations.

2.4.2 General Mathematical and Visual Representations

Delay-and-sum beamforming is the most conventional and commonly used beamforming algorithm. Given the ultrasound wave containing echoes from the imaged area is defined as $r(x, y, z, t)$, the signal sensed by a transducer located at location (x_i, y_i, z_i) can be interpreted as $r_i(t) = r(x_i, y_i, z_i, t)$. For an array of M transducer elements, the signal detected by each element depends on its location. If M is odd and the M elements are evenly spaced with the distance between consecutive elements, we can center the array at the origin and define the locations of the elements as follows:

$$\left(\frac{-M}{2} \cdot d, 0, 0\right), \left(\frac{-M+1}{2} \cdot d, 0, 0\right), ...,$$

$$\Big(-d, 0, 0\Big), \Big(0, 0, 0\Big), \Big(d, 0, 0\Big), ...,$$

$$\left(\frac{M-1}{2} \cdot d, 0, 0\right), \left(\frac{M}{2} \cdot d, 0, 0\right)$$

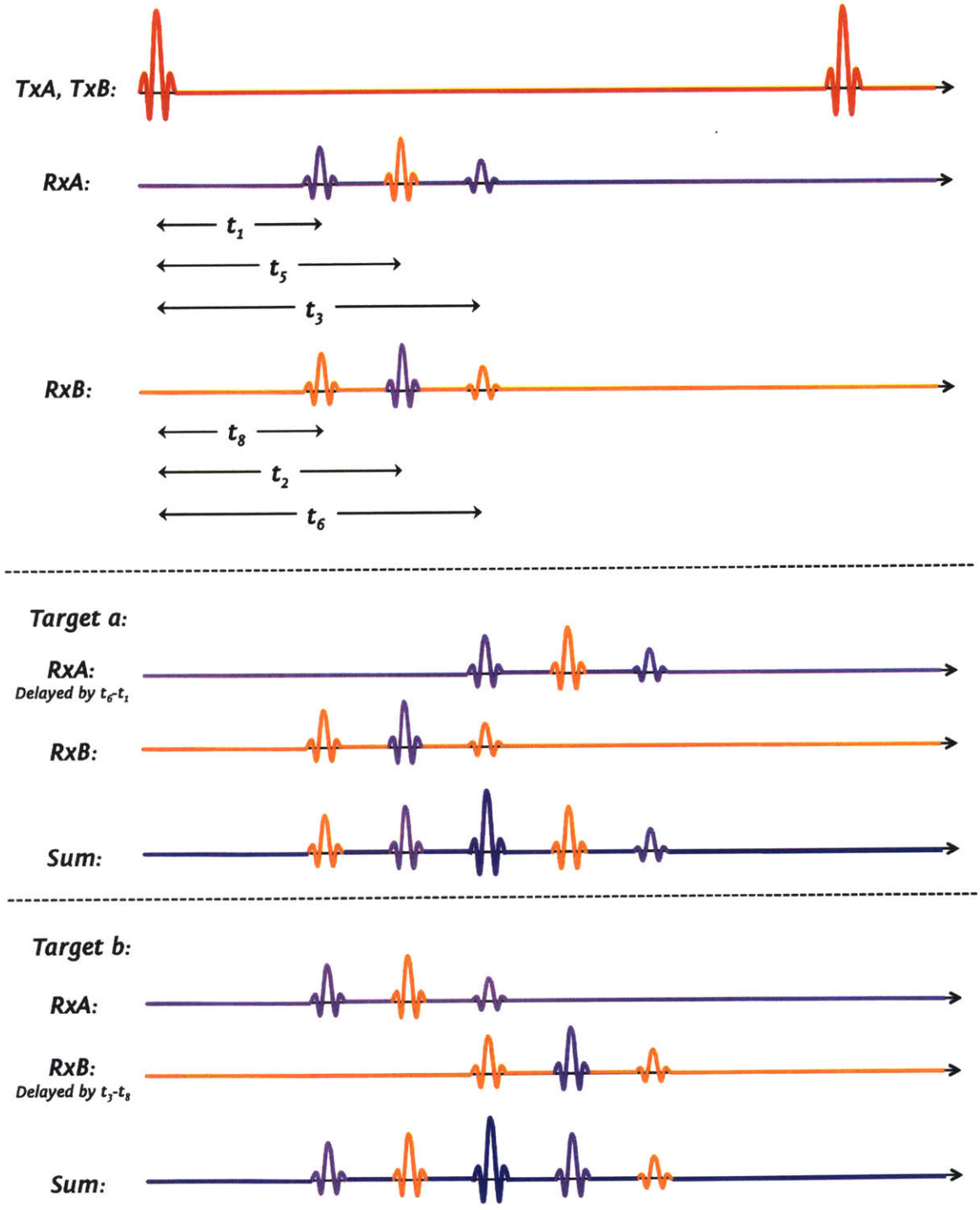


Figure 2-7: Schematic representation of the waveforms received by the two neighboring transducer elements (A and B). Contributions from the waveform transmitted from A are colored purple, whereas those that are from B are colored yellow.

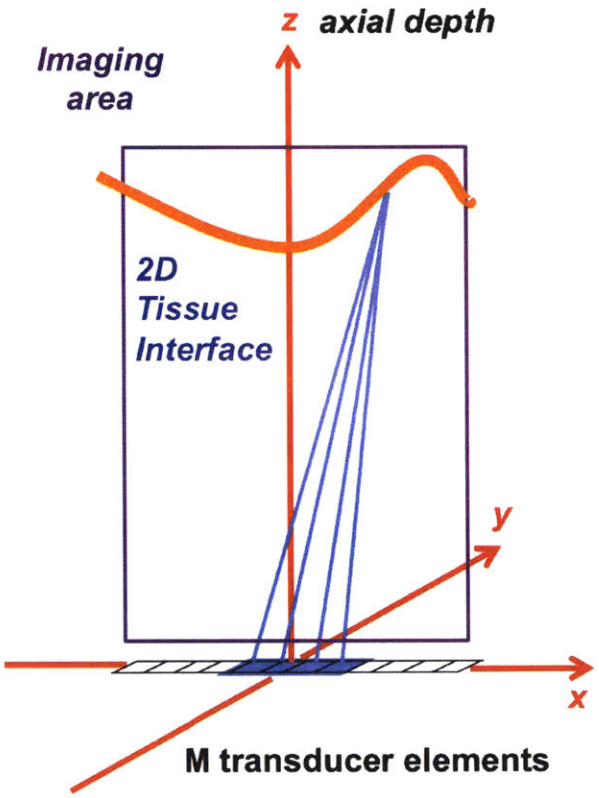


Figure 2-8: Geometry of the transducer and the imaging area for 2D ultrasound imaging relative to the Cartesian coordinate system.

The placement and spacing of the transducer elements effectively correspond to spatially sampling the overall waveform $r(x, y, z, t)$ at these locations.

$$R(t) = \sum_{m=0}^{M-1} A_m \cdot r_m(t - \Delta t_m) \tag{2.7}$$

To better understand the effect of each stage in the delay-and-sum algorithm, we consider the simplified example where each transducer element receives an echo pulse corresponding to the same point in the imaging area. The example further assumes that the point of acoustic impedance discontinuity that reflected each of these waves is located closest to the i-th element, such that it has the same x-coordinate as the i-th element. This also means that this point is located along the line through x_i that is normal to the transducer array, which is aligned with the x-axis.

As shown in Figure 2-9, there is a delay in the time of arrival of that pulse on the off-center channels (indexed $m = i-3, i-2, i-1, i+1, i+2, i+3$) as compared to the signal received by element i. Since the elements are evenly spaced and the line that passes through the i-th element and the focal point bisects the set of seven elements, the delay profile is symmetric around that line, such that:

$$\begin{aligned} \Delta t_{i-1} &= \Delta t_{i+1} \\ \Delta t_{i-2} &= \Delta t_{i+2} \\ \Delta t_{i-3} &= \Delta t_{i+3} \end{aligned} \tag{2.8}$$

These delays are computed based on the difference in propagation paths, p_m, which is the distance that the ultrasound wave travels from the m-th array element at $(x_m, 0, 0)$ to the focal point at $(x_f, 0, z_f)$, gets reflected, and travels back to the m-th element. For the center element (closest to the focal point), $x_i = x_f$. For the off-center elements, $x_{i\pm1} = d$, $x_{i\pm2} = 2 \cdot d$, etc.

$$
\begin{aligned}
p_i &= 2 \cdot \sqrt{0^2 + 0^2 + z_f^2} \\
p_{i\pm1} &= 2 \cdot \sqrt{d^2 + 0^2 + z_f^2} \\
p_{i\pm2} &= 2 \cdot \sqrt{(2 \cdot d)^2 + 0^2 + z_f^2} \\
p_{i\pm3} &= 2 \cdot \sqrt{(3 \cdot d)^2 + 0^2 + z_f^2}
\end{aligned}
\tag{2.9}
$$

$$
\begin{aligned}
\Delta t_{i\pm1} &= \frac{p_{i\pm1} - p_i}{c} = 2 \cdot \left(\frac{\sqrt{d^2 + z_f^2}}{c} - z_f \right) \\
\Delta t_{i\pm2} &= \frac{p_{i\pm2} - p_i}{c} = 2 \cdot \left(\frac{\sqrt{4 \cdot d^2 + z_f^2}}{c} - z_f \right) \\
\Delta t_{i\pm3} &= \frac{p_{i\pm3} - p_i}{c} = 2 \cdot \left(\frac{\sqrt{9 \cdot d^2 + z_f^2}}{c} - z_f \right)
\end{aligned}
\tag{2.10}
$$

When the appropriate delay is applied to each of the signals received by the off-center elements, the times of arrival of the single pulse detected in each channel will be aligned in resultant waveforms, as shown on the right of Figure 2-9.

The collection of weighting coefficients, denoted as A_m in Equation (2.7), is known as the apodization function. Since the use of waveforms received by the array of transducer elements amounts to spatially sampling the four-dimensional overall ultrasound signal $s(x, y, z, t)$, Depending on the relative size of the spacing between consecutive elements, also known as the array pitch, as compared to the wavelength of the signal being sampled, there can be grating lobes within the visible region. Setting a subset of the weights to 0 corresponds to spatial windowing which can help suppress grating lobes but also has the negative effect of broadening the main lobe, thus decreasing azimuthal spatial resolution. A number of typical windowing functions that are used for apodization are summarized in Table 2.2 [62, 63], where center elements are weighed more heavily than the side elements in the given aperture window. Figure 2-10 shows the effect of the apodization stage in the beamformer.

Finally, the delayed and apodized channel signals $A_m \cdot r_m(t - \Delta t_m)$ are summed to

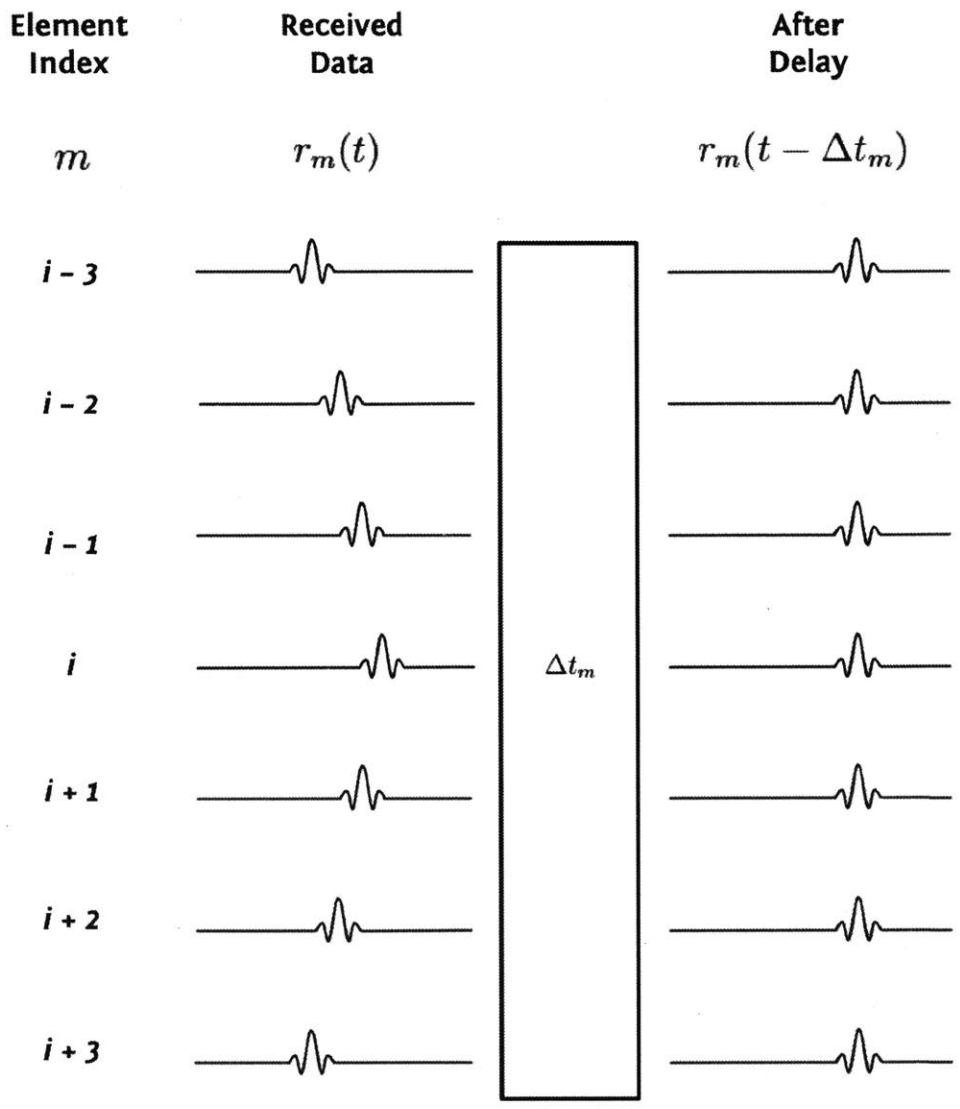


Figure 2-9: Schematic representation of signal progression through the delay block. Applying the appropriate delays aligns the received data from the neighboring transducer elements in time to be added coherently.

Table 2.2: Typical apodization windowing functions and their definitions.

Apodization Window	**Equation** $\left(0 \leq n \leq N\right)$
Hamming	$w(n) = 0.54 - 0.46 \cdot \left(2\pi \cdot n/(N-1)\right)$
Hann	$w(n) = 0.5 \cdot \left(1 - \cos\left(2\pi \cdot n/(N-1)\right)\right)$
Blackman	$w(n) = 0.42 - 0.5 \cdot \cos\left(2\pi \cdot n/(N-1)\right) +$ $0.08 \cdot \cos\left(4\pi \cdot n/(N-1)\right)$
Nuttall	$w(n) = a_0 - a_1 \cdot \cos\left(2\pi \cdot n/(N-1)\right) +$ $a_2 \cdot \cos\left(4\pi \cdot n/(N-1)\right) - a_3 \cdot \cos\left(6\pi \cdot n/(N-1)\right)$
Flat top	$w(n) = a_0 - a_1 \cdot \cos\left(2\pi \cdot n/(N-1)\right) +$ $a_2 \cdot \cos\left(4\pi \cdot n/(N-1)\right) - a_3 \cdot \cos\left(6\pi \cdot n/(N-1)\right) +$ $a_4 \cdot \cos\left(8\pi \cdot n/(N-1)\right)$

increase overall signal strength. The effect of the summation stage in the beamformer is illustrated in Figure 2-11.

2.4.3 Focusing Modes and Delay Profiles

As shown in Figure 2-1, the propagation path taken by the ultrasound wavefront before it arrives at a transducer depends on the relative positioning of a focal point and the transducer. When the focal point coincides exactly with the location where the wave is reflected, the distance between these two points maps to the time of arrival of the reflected pulse. When considering an array of transducers, the location of the focal point affects the length of delays that need to be inserted for each array element. This is schematically shown in Figure 2-12. Considering a point scatterer at z_1, a pulse reflected from the scatterer travels back to a group of five elements, with the pulse being detected first by the center element and later by the outer elements. The amount of delay between the transducers is indicated by the blue dashed curve (referred to as

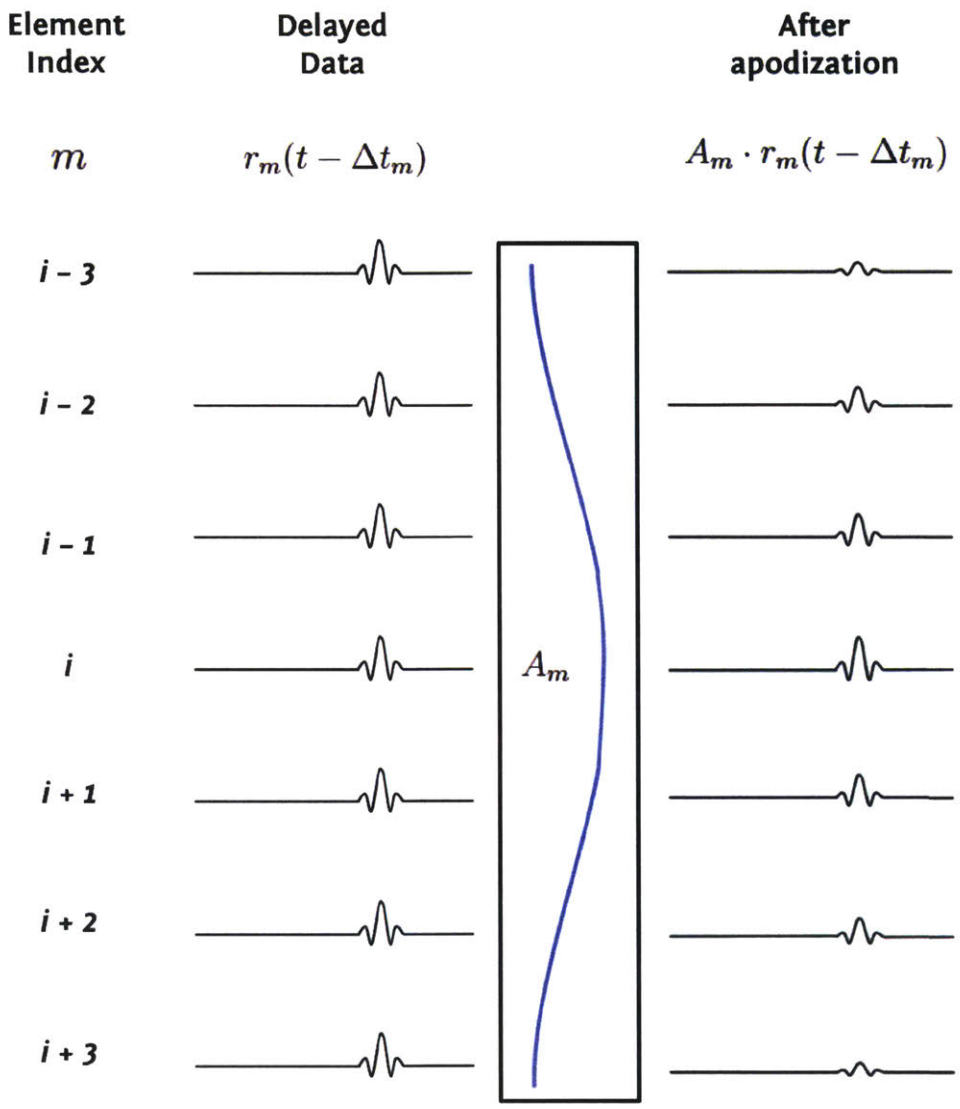


Figure 2-10: Schematic representation of signal progression through the apodization stage. This example assumes the apodization weights, A_m, are chosen such that the waveforms received by the outermost elements (farthest away from the i-th element in the center) contribute to the waveform after summation to a lesser degree than those received by the center elements.

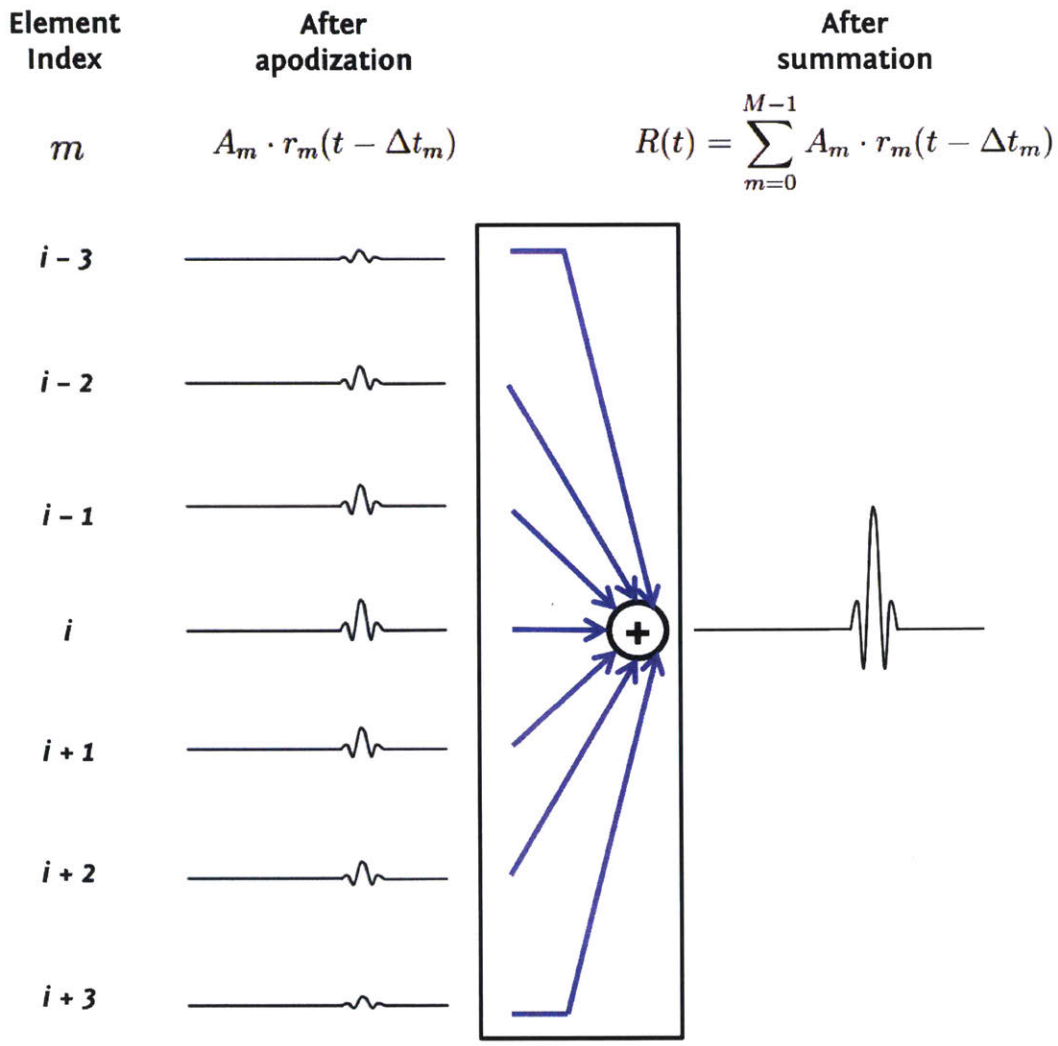


Figure 2-11: Schematic representation of signal progression through the summation stage. The time-aligned signals add coherently to produce a high SNR pulse at the time corresponding to the location of discontinuity in acoustic impedance.

the delay profile), which translates to the amount of delay that needs to be inserted to realign the pulses in the time domain. Now consider another point scatterer at z_2, where $z_2 > z_1$ (i.e. z_2 is farther away from the transducer array). Because the propagation path to all five elements have increased, the difference between them becomes less pronounced, resulting in the shallower delay profile denoted by the green dashed curve.

Systems that employ a single delay profile to image all points in a given area effectively has a single focal point, i.e. have fixed focusing. Dynamic focusing can be achieved by updating the delay profile for imaging different points in the imaging area, thus changing the focal length at different imaging depths.

2.4.4 Increasing SNR

As previously explained, one of the motivations of array processing and beam formation for ultrasound imaging is the need to increase the SNR of the reconstructed waveform [64–66]. Deconstructing the signal received by a single transducer as follows:

$$r(t) = s(t) + n(t) \tag{2.11}$$

where $s(t)$ is the signal component and $n(t)$ is the noise (including noise from the imaging medium and thermal noise in the electronics), and assuming that the noise and signal components of the waveform are uncorrelated, the SNR is given as:

$$SNR_{sensor} = \frac{E\{|s(t)^2|\}}{E\{|n(t)^2|\}} = \frac{\sigma_s^2}{\sigma_n^2} \tag{2.12}$$

When we perform the analogous analysis on an array of M transducer elements, whose total waveform after beamforming is described in Equation (2.7), we get:

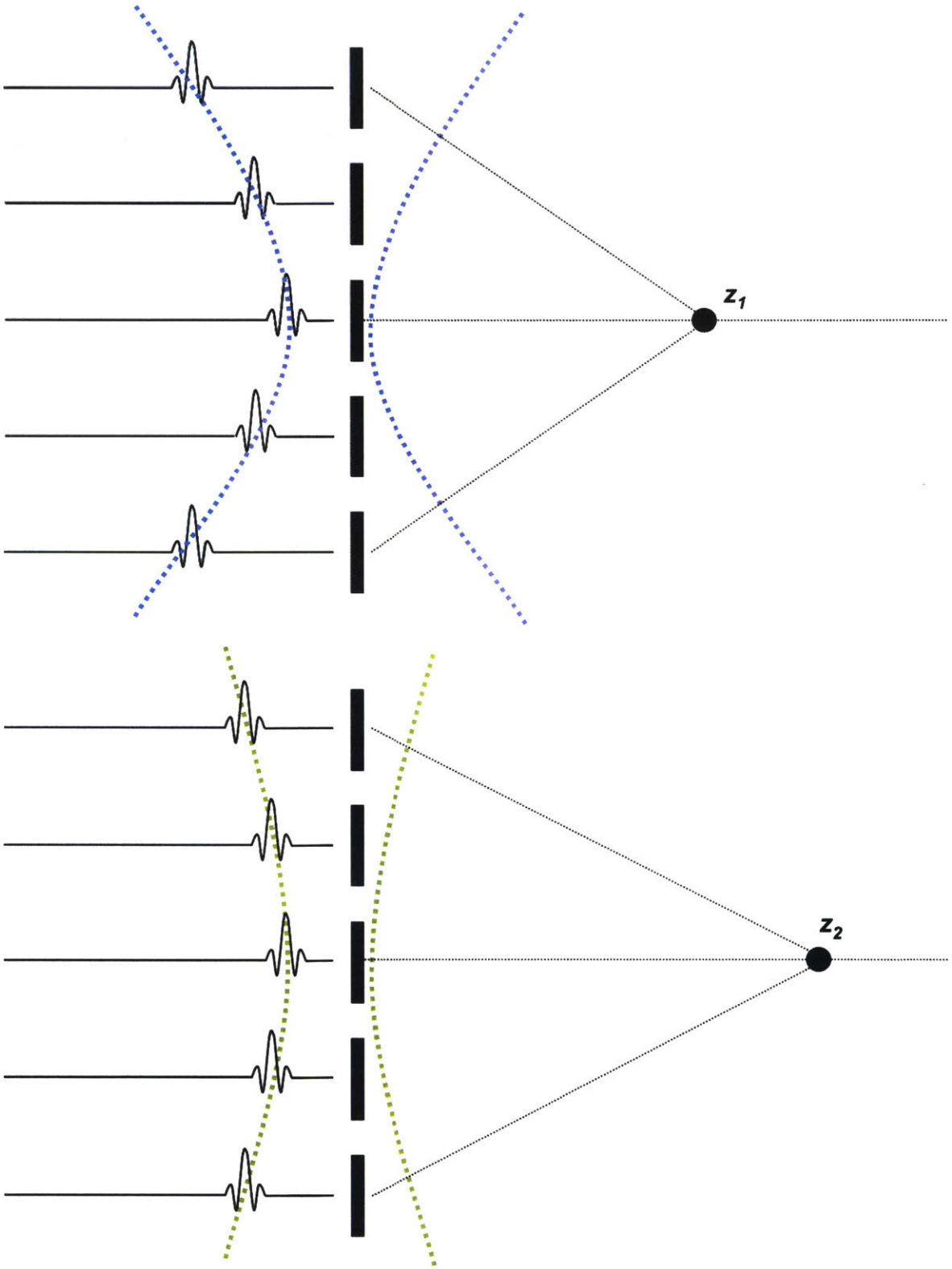


Figure 2-12: Dynamic focusing example for a five-element transducer array and scatterers at locations z_1 and z_2.

$$
\begin{aligned}
R(t) &= \sum_{m=0}^{M-1} A_m \cdot r_m(t - \Delta t_m) \\
&= \sum_{m=0}^{M-1} A_m \cdot \Big(s_m(t - \Delta t_m) - n_m(t - \Delta t_m) \Big) \\
&= R_s(t) + R_n(t)
\end{aligned}
\tag{2.13}
$$

Assuming the delays Δt_m are chosen to match the direction of wave propagation, so that the channel waveforms add constructively, Equation (2.7) reduces to:

$$
\begin{aligned}
R(t) &= \sum_{m=0}^{M-1} A_m \cdot s(t) \\
&= s(t) \cdot \left[\sum_{m=0}^{M-1} A_m \right]
\end{aligned}
\tag{2.14}
$$

We can write the following expression for signal power of the array, with the above assumption:

$$
\begin{aligned}
E\Big\{ |R_s(t)^2| \Big\} &= E\left\{ \Big| \sum_{m=0}^{M-1} A_m \cdot s_m(t) \Big|^2 \right\} \\
&= E\left\{ \left(\sum_{m=0}^{M-1} A_m \cdot s_m(t) \right) \cdot \left(\sum_{l=0}^{M-1} A_l^* \cdot s_l^*(t) \right) \right\} \\
&= E\left\{ \Big| \sum_{m=0}^{M-1} A_m \Big|^2 \right\} \cdot E\Big\{ \big| s(t) \big| \Big\} \\
&= \left| \sum_{m=0}^{M-1} A_m \right|^2 \cdot \sigma_s^2
\end{aligned}
\tag{2.15}
$$

We can also write the expression for noise power of the array using the same approach:

$$
\begin{aligned}
E\Big\{|R_n(t)^2|\Big\} &= E\left\{\left|\sum_{m=0}^{M-1} A_m \cdot n_m(t)\right|^2\right\} \\
&= E\left\{\left(\sum_{m=0}^{M-1} A_m \cdot n_m(t)\right) \cdot \left(\sum_{l=0}^{M-1} A_l^* \cdot n_l^*(t)\right)\right\} \\
&= E\left\{\sum_{m=0}^{M-1}\sum_{l=0}^{M-1} A_m \cdot A_l^* \cdot R_{noise}(m-l)\right\}
\end{aligned}
\tag{2.16}
$$

where R_{noise} is the autocorrelation of the noise. With the assumption that noise detected by one sensor is spatially uncorrelated to noise detected by any other sensor in the array, we can write:

$$
R_{noise}(m-l) = \begin{cases} \sigma_n^2, & \text{if } m = l \\ 0, & \text{otherwise} \end{cases}
\tag{2.17}
$$

Therefore, we can find the total noise power of the array by substituting Equation (2.17) into Equation (2.16):

$$
E\Big\{|R_n(t)^2|\Big\} = \sum_{m=0}^{M-1} |A_m|^2 \cdot \omega_n^2
\tag{2.18}
$$

The SNR of the overall array is the ratio between Equation (2.14) and Equation (2.18):

$$
SNR_{array} = \frac{\left|\sum_{m=0}^{M-1} A_m\right|^2 \cdot \sigma_s^2}{\sum_{m=0}^{M-1} |A_m|^2 \cdot \sigma_n^2}
\tag{2.19}
$$

Putting Equation (2.12) and Equation (2.19) together, we can quantify the increase in SNR provided by an array of M transducer elements:

$$
\begin{aligned}
SNRgain = \frac{SNR_{array}}{SNR_{sensor}} &= \frac{\left|\sum_{m=0}^{M-1} A_m\right|^2 \cdot \sigma_s^2}{\sum_{m=0}^{M-1} |A_m|^2 \cdot \sigma_n^2} \cdot \frac{\sigma_n^2}{\sigma_s^2} \\
&= \frac{\left|\sum_{m=0}^{M-1} A_m\right|^2}{\sum_{m=0}^{M-1} |A_m|^2}
\end{aligned}
\tag{2.20}
$$

Thus, the gain in SNR is maximized when the apodization weights are equal (i.e. $A_m = A$:

$$SNRgain = \frac{\left|\sum_{m=0}^{M-1} A\right|^2}{\sum_{m=0}^{M-1} |A|^2} = \frac{M^2 \cdot |A|^2}{M \cdot |A|} = M \tag{2.21}$$

We can conclude that the larger the transducer count, the more significant the increase in SNR that beamforming provides.

2.5 Previous Hardware Implementations

Researchers are pursuing diverse strategies for the size and power reductions required of portable ultrasound systems. It is now possible to integrate a complete ultrasound system in a smartphone-attached device, but further power reductions are necessary to extend its battery life [67]. A single-channel beamformer for handheld ultrasound that dissipates 25 mW at 40 MHz was proposed in [68]. Another general technique aimed at power reduction is beamforming of delta-sigma modulated probe signals [69], eliminating the need for Nyquist-rate ADCs.

As systems are scaled to many channels, we witness tradeoffs and synergies between analog and digital beamforming approaches in an effort to reduce overall power consumption. A switched-current architecture allows analog beamformers to scale efficiently [70], but the analog delay lines incur a large area overhead. Digital beamformers are more flexible; architectures accommodating 32 40-MHz channels have been demonstrated on an FPGA [71]. However, care must be taken to mitigate the increased demands on ADCs and interfaces. Multichannel systems can exploit multiplexing to reduce wire count [72], or perform subarray beamforming with analog ICs closely coupled to the transducer [73]. Zhou [74] applied compressed sensing techniques to reduce ADC power, with minor tradeoffs in image quality. [75] demonstrated a hybrid concept of several analog beamformers feeding a digital beamformer on a single CMOS chip.

Similar to these system approaches, [21] first proposed the idea of scaling image

quality to reduce power consumption. It also demonstrated this tradeoff using a three-part system model: an FPGA beamformer, a custom analog front end (AFE), and a stand-alone ADC. This energy model will be discussed at length in Chapter 3.

Chapter 3

System Energy Model for Digital Beamformer

For applications such as point-of-care diagnostics in rural areas, the movement to a portable and low-power ultrasound imaging system is warranted. Since system energy consumption becomes a priority, it is worthwhile to examine the interaction between its various components and optimize for best achievable image quality at a given energy budget for the overall system. This chapter explores the impact of the overall energy budget on the performance of each functional component and how the performance of one component affects that of the others. The goal of this work is to provide a recommendation for design choices in the end-to-end system based on this understanding.

3.1 Observations on Beamforming Operation

The basic operation of an ultrasound imaging system can be divided into two stages: transmit and receive. During the transmit stage, a pulse-based digital waveform is generated and converted into the analog domain, then again converted into high frequency sound waves (on the order of 2 to $18MHz$) via the transducer element. The transmit-side electronics corresponding to a single transducer element, including a digital-to-analog converter, pulser, and a T/R switch, are schematically shown in

Figure 3-1.

As the ultrasound waves come into contact with the tissue interfaces where there are changes in acoustic impedance, the waves are partially reflected back towards the transducer array. Some examples of such interfaces include the boundaries between fluid and soft tissue (as is the case of fetal ultrasound imaging), and between soft tissue and bone. The system then switches to the receive stage, where the transducer elements listen for the reflected ultrasound waves and convert them back into the electrical domain. The signals undergo low-noise amplification and analog-to-digital conversion, and are then processed in the digital domain. The receive-side electronics, again for a single transducer element, are represented in Figure 3-2.

Bearing in mind the large number of transducer array elements in a typical ultrasound imaging system, and the transmit and receive side electronics that correspond to each element, we note that reduction in the number of elements has a significant impact on the overall system power. Based on this observation, we propose a receive-side ultrasound beamforming system that allows the user to select the number of transducer elements used to form an image at run-time [21]. With a digital "channel on/off" signal feeding back to the front end electronics, this run-time control can be used to reduce the system power consumption by eliminating a subset of the parallel analog front end (AFE) and analog-to-digital converter (ADC) channels that supply digital channel data to the scalable digital beamformer.

3.2 Receive-side Digital Beamformer

The proposed ultrasound imaging system is schematically represented in Figure 3-3. At run-time, the user can choose to form an image under one of three beamforming modes: quarter, half, and full resolution. Selecting the quarter and half resolution modes reduces the effective transducer array size while maintaining uniform element spacing (increasing the effective element pitch by 4x and 2x, respectively). These modes provide the direct benefit of system-level power reduction (by reducing the number of active AFE and ADC channels), but also directly impacts the image quality.

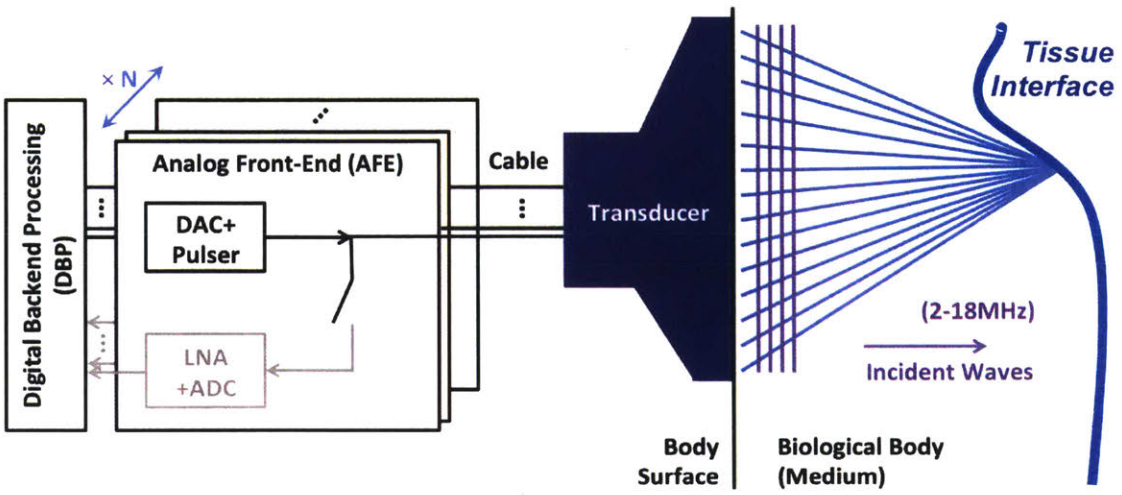


Figure 3-1: Schematic representation of transmit-side electronics.

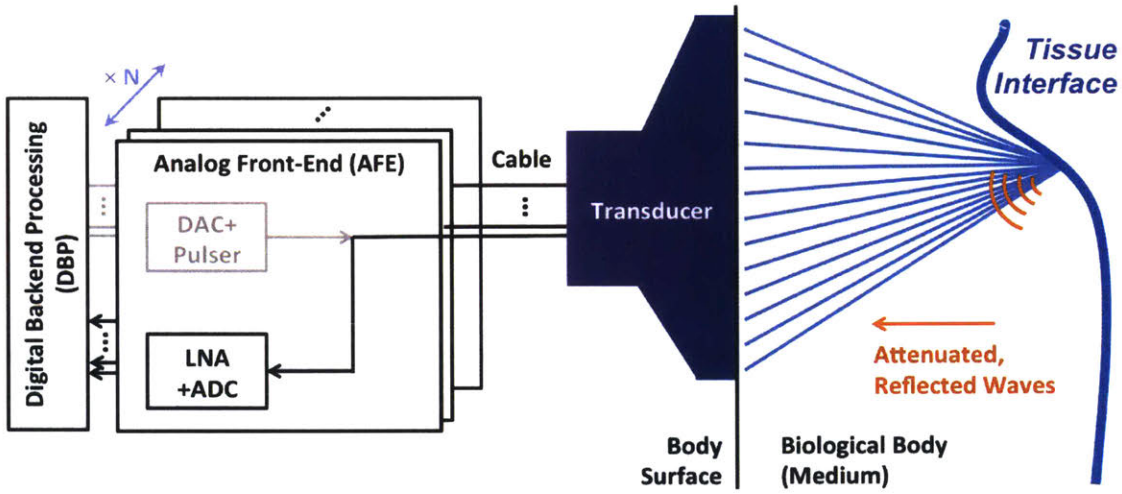


Figure 3-2: Schematic representation of receive-side electronics.

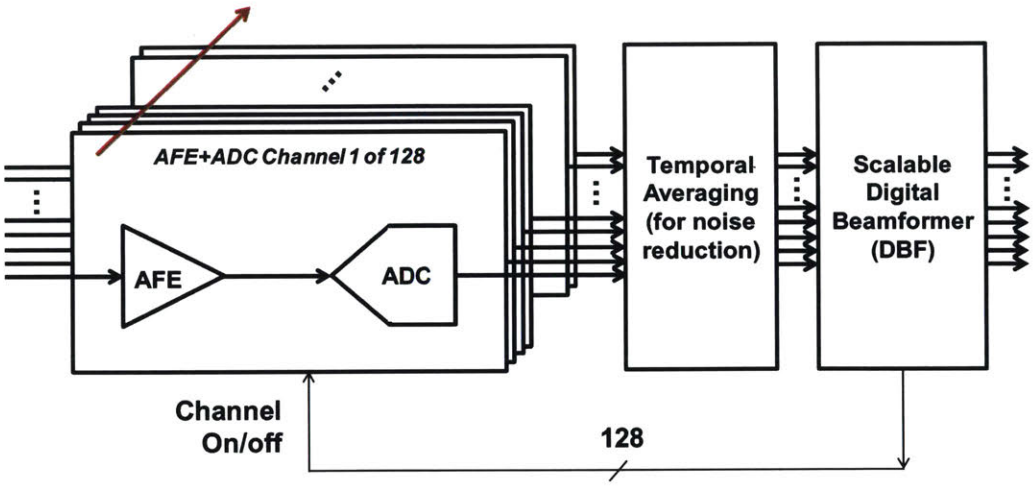


Figure 3-3: Ultrasound system block diagram.

The envisioned usage of the lower quality modes are either for direct use in specific imaging applications where the image quality is deemed "good enough" (such as prenatal diagnosis for placenta previa [3]), or for the user to specify a smaller area of interest to be imaged at full resolution in a second step.

This section discusses the functional blocks in the receive-side digital beamformer and the individual tradeoffs between performance and power consumption.

3.2.1 Analog Front End

Depending on the topology and impedance characteristics of the transducer device used in the probe, the AFE needs to be designed differently. The AFE in this work is designed to support the micro-electrical-mechanical systems (MEMS) ultrasonic device known as the Capacitive Micromachined Ultrasonic Transducer (CMUT). A transimpedance amplifier (TIA) topology is used in the circuit schematic shown in Figure 3-4. The noise contributions from the passive components and the operational amplifier (op-amp) are modeled for noise optimization. Analyzing the TIA circuit schematic, we can write the following equivalent impedances for the input and feedback branches as:

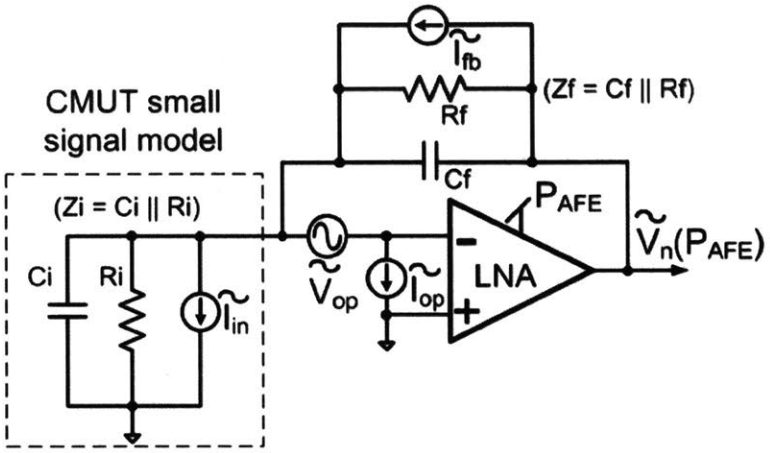


Figure 3-4: Small-signal model and noise sources of the CMUT transducer element and the low-noise amplifier.

$$\begin{aligned} Z_i = Z_{Ci} \parallel Z_{Ri} &= \frac{1}{sC_i} \parallel R_i \\ &= \frac{\frac{1}{sC+f} \cdot R_f}{\frac{1}{sC_i} + R_i} = \frac{R_i}{1 + sR_iC_i} \end{aligned} \tag{3.1}$$

$$\begin{aligned} Z_f = Z_{Cf} \parallel Z_{Rf} &= \frac{1}{sC + f} \parallel R_f \\ &= \frac{\frac{1}{sC+f} \cdot R_f}{\frac{1}{sC+f} + R_f} = \frac{R_f}{1 + sR_fC_f} \end{aligned} \tag{3.2}$$

Therefore, we can express the noise figure (NF) [34, 76] and the output noise spectral density (V_n) are expressed as:

$$NF = 1 + \frac{R_i}{R_f} + \frac{V_{op}^2}{I_{in}^2 \cdot |Z_i \parallel Z_f|} + \frac{I_{op}^2}{I_{in}^2} + \frac{2 \cdot |V_{op} \cdot I_{op}|}{I_{in}^2 \cdot |Z_i \parallel Z_f|} \tag{3.3}$$

$$V_n = |Z_f| \cdot \sqrt{\frac{4kT}{R_i} + \frac{4kT}{R_f} + \left| \frac{V_{op}}{Z_i || Z_f} + I_{op} \right|^2} \tag{3.4}$$

In Equation (3.3) and Equation (3.4), the noise sources from the op-amp, V_{op} and I_{op}, are affected by the op-amp power consumption (P_{AFE}). Assuming a two-stage op-

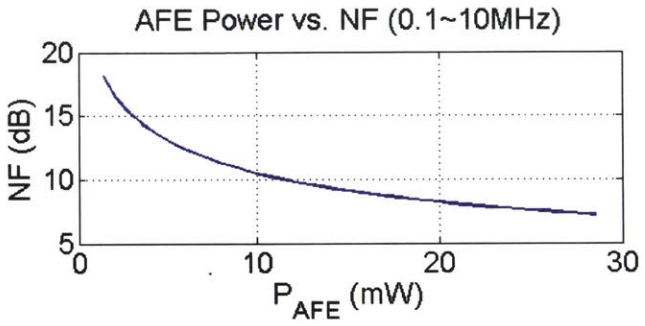


Figure 3-5: AFE noise figure as a function of P_{AFE}

amp structure, the relationship between the noise sources and the power consumption is modeled as follows [77]:

$$V_{op} = \sqrt{2 \cdot \alpha \cdot (4kT) \frac{1}{g_{m1}} \left(1 + \frac{g_{m3}}{g_{m1}}\right)}, \tag{3.5}$$

$$I_{op} = (jwC_{in}) \cdot V_{op}, \tag{3.6}$$

where g_{m1}, g_{m3} are transistor transconductance for op-amp's input stage and the current mirror; C_{in} is op-amp's input capacitance. g_{m1} is related to P_{AFE} as follows:

$$g_{m1} \propto P_{AFE} = g_{m1(\text{nominal})} \cdot \frac{P_{AFE}}{P_{AFE(\text{nominal})}} \tag{3.7}$$

Combining Equations (3.3) to (3.7), NF and V_n are expressed as functions of P_{AFE}. Based on the TIA design in [76], the modeling parameters are: $C_{in} = 1.5pF$, $\alpha = 0.5$, $g_{m1(\text{nominal})} = 41mS$, $\frac{g_{m3}}{g_{m1}} = 0.35$, $P_{AFE(\text{nominal})} = 14.3mW$. Figure 3-5 shows the AFE noise performance affected by the power consumption.

3.2.2 Analog to Digital Converter

The ADC converts the amplified continuous-time analog signal at the output of the AFE into a digital discrete-time signal, before feeding it to the DBF block. The ADC performance metrics include speed (sampling frequency) and accuracy (quantization noise), which are directly related to the block power consumption. The most commonly used figure-of-merit (FOM) for an ADC captures this tradeoff between

power consumption and performance:

$$FOM = \frac{P_{ADC}}{2^{ENOB} \cdot f_s} \tag{3.8}$$

where P_{ADC} is the power consumption of the ADC and ENOB is its effective number of bits (which is a measure of ADC accuracy). With the units fJ/step, a smaller FOM corresponds to a more energy-efficient and accurate ADC.

The ENOB of the ADC relates to its quantization noise (V_q) by:

$$V_q = \frac{ADC\ Full\ Scale}{2^{ENOB}} \tag{3.9}$$

where ENOB is assumed to be 1 bit less than the ADC's bit width (M):

$$ENOB \approx M - 1 \tag{3.10}$$

Assuming the ADC's FOM is relatively constant over a range of operating conditions, the sampling frequency is thus linearly proportional to the power consumption:

$$P_{ADC} = FOM \cdot 2^{ENOB} \cdot f_s \tag{3.11}$$

Equations (3.8) to (3.11) contribute to the ADC's design-time optimization, so that the ADC specs are compatible to other building blocks in the ultrasound system. A pipeline ADC design is used for modeling [78], with an FOM of $41fJ$/step. We can use this FOM to model the relationship between the ADC's power consumption and accuracy based on the assumptions made above. Figure 3-6 shows the relationship between ENOB and ADC power given the sampling frequency of 40 MHz.

3.2.3 Energy-Scalable Digital Beamformer

Since the system-level AFE and ADC power scales linearly with the number of channels in receive-side beamforming, the energy model attempts to offer run-time energy scalability by reducing the number of channels being used. The energy-scalable digital beamformer provides this flexibility by adjusting its internal parameters according to

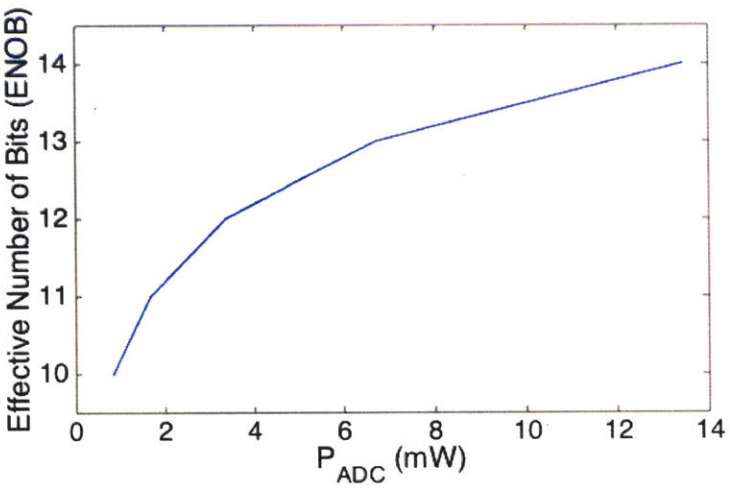


Figure 3-6: ADC ENOB as a function of ADC power, for FOM of $41fJ$/step and sampling frequency of 40 MHz.

Table 3.1: Specifications of L14-5/38 Linear Transducer used for 2D data acquisition.

Center frequency	7.5 MHz
Bandwidth	14-5 MHz
Number of elements	128
Element pitch	300 μm
Physical footprint	4 mm $\times$ 39 mm
Depth range	2-9 cm
Geometric Focus	16 mm
Steered Angle	17.5 $^\circ$
Elevation aperture	4 mm

whether 50% or 75% of the AFE and ADC channels are turned off. This supports the usage model described in Figure 3-7.

3.3 Dataset Description

The 2D imaging dataset used to demonstrate correct beamforming functionality was provided by Ultrasonix Medical Corporation (now part of Analogic Corporation), an ultrasound transducer probe and system supplier based in Vancouver, Canada. The Sonix L14-5/38 Linear transducer probe was used to acquire the receive-side, discrete-time digitized data [79]. The specifications of the probe are summarized in Table 3.1.

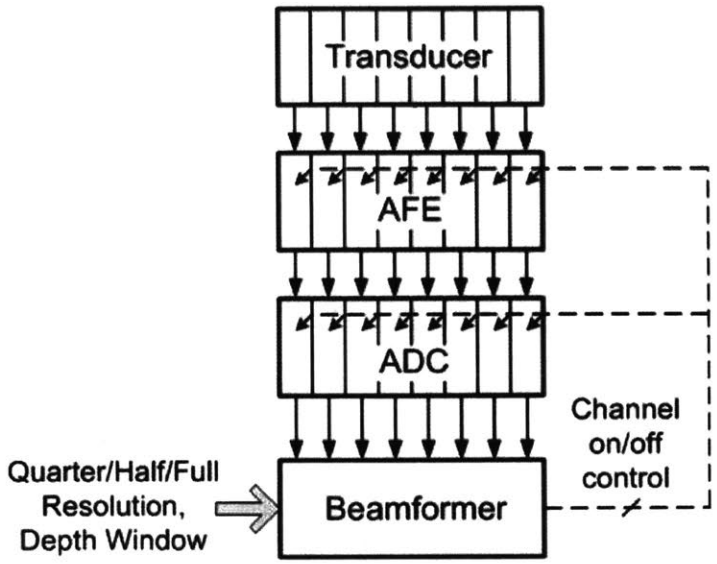


Figure 3-7: Usage model for energy-scalable beamformer with feedback to turn off unused AFE and ADC channels.

The medium being imaged in the dataset is a commercial multipurpose ultrasound phantom, shown in Figure 3-8. Urethane rubber is used to mimic the behavior of soft tissue [6]. The phantom provides four surfaces on which a transducer can be applied, each providing a different set of imaging targets made of monofilament nylon. In this work, the scanning surface #1 was used to collect image data corresponding to the inset area shown in Figure 3-9. This area has the azimuthal (x, the direction along which the probe is aligned) dimension of 38.4 mm, corresponding to the width of the 128-element transducer array, and the axial (z, or the direction of ultrasonic wave propagation away from the array) dimension of 57.5 mm. Within this area, there are a set of five off-center, uniformly spaced and cascaded point targets (5.0 mm apart in the azimuthal direction) and four more point targets at different axial depths.

3.4 Design-Time Tradeoffs

At the interface between the AFE and ADC blocks, coordinating the noise and signal range is a major design consideration. Because the highly capacitive CMUT source leads to a relatively high noise figure performance of the AFE, the dynamic range at the AFE output is limited. The time gain compensation (TGC) block (introduced in

Figure 3-8: ATS Laboratories Model 539 Multipurpose Ultrasound Phantom [6].

Section 2.4.1) is therefore not always necessary. In this work as shown in Figure 3-10, the AFE gain is adjusted such that the AFE output signal range is within the ADC input signal range without the use of a TGC. A noise constraint is also introduced at the AFE-ADC boundary:

$$V_q(M) \approx \frac{1}{2} \cdot V_n(P_{AFE}) \tag{3.12}$$

With this design-time constraint, the ADC bit width is adjusted to work compatibly with a given AFE noise performance. Figure 3-11 shows the design-time tradeoffs. As the AFE block burns more power, i.e. P_{AFE} increases, the AFE output noise is reduced (blue curve in Figure 3-11(a)). Accordingly, the ADC is required to have a higher accuracy, i.e. more bits, resulting in lower quantization noise (green curve in Figure 3-11(a)). Because V_q satisfies the constraint described in Equation (3.12), noise contribution from the ADC is small, as can be seen from the fact that $V_{tot} \approx V_n$ in Figure 3-11(a). The block-level power consumptions are plotted in Figure 3-11(b).

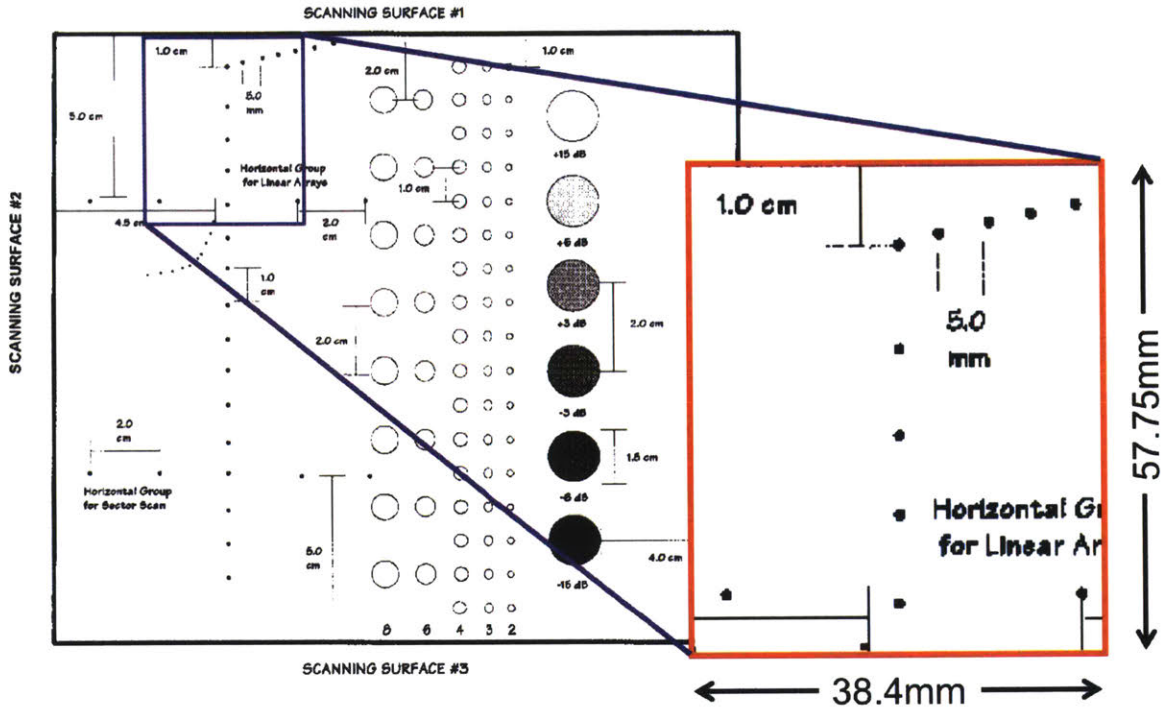


Figure 3-9: Specifications of the locations of scatterers to be imaged in the ATS Multipurpose Ultrasound Phantom.

3.5 Performance Evaluation

3.5.1 Image Quality

To assess the quality of images generated by the system, the peak signal-to-noise ratio (PSNR) metric is introduced [80]. Other image quality metrics such as the structural similarity index measure (SSIM) have been defined and compared with PSNR [81], but we selected PSNR for its easy of use and direct comparison between two images of equal resolution. The mean square error (MSE) between a noise-less image, $I(i,j)$, and a noise-corrupted one, $K(i,j)$, is first computed as:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2 \tag{3.13}$$

The PSNR is then defined as the ratio between the maximum signal range (MAX_I) and the MSE:

$$PSNR = 10 \cdot log_{10}(\frac{MAX_I^2}{MSE}) \tag{3.14}$$

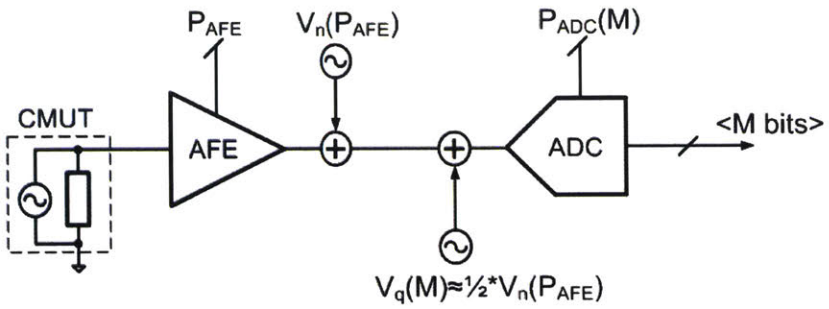


Figure 3-10: Block diagram of the interface between the AFE and the ADC

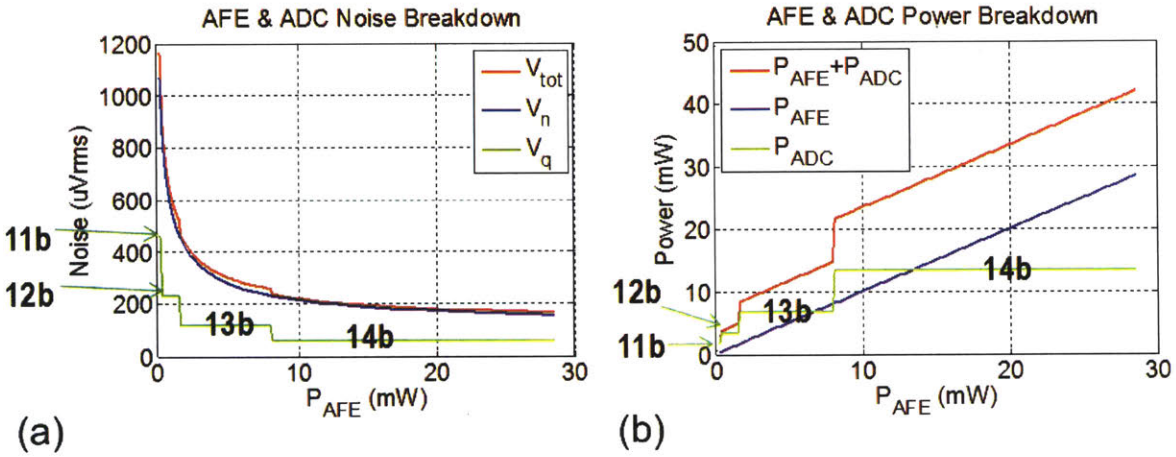


Figure 3-11: (a) Noise amplitude of ADC is maintained less than half of that from AFE. (b) The power consumption breakdown between AFE and ADC.

3.5.2 PSNR vs. Front-End Power

To quantify the relationship between the AFE&ADC noise performance and the image quality in terms of PSNR, the AFE is varied while the same ADC sampling rate (40MHz) and DBF operation are used to generate the same image. The corresponding ADC bit width is determined using the quantization noise constraint inEquation (3.12)) Moreover, noise averaging is modeled, which helps improve image quality at the cost of reducing imaging frame rate. As shown in Figure 3-12, under the same amount of noise averaging, more power consumed in the AFE and ADC leads to a higher PSNR for images; more noise averaging also considerably improves PSNR. However, at a given noise averaging number, the marginal gain of PSNR improvement with increasing power consumption diminishes roughly beyond $20mW$. It follows that the AFE and ADC power should not be pushed beyond 20mW, but rather the number of

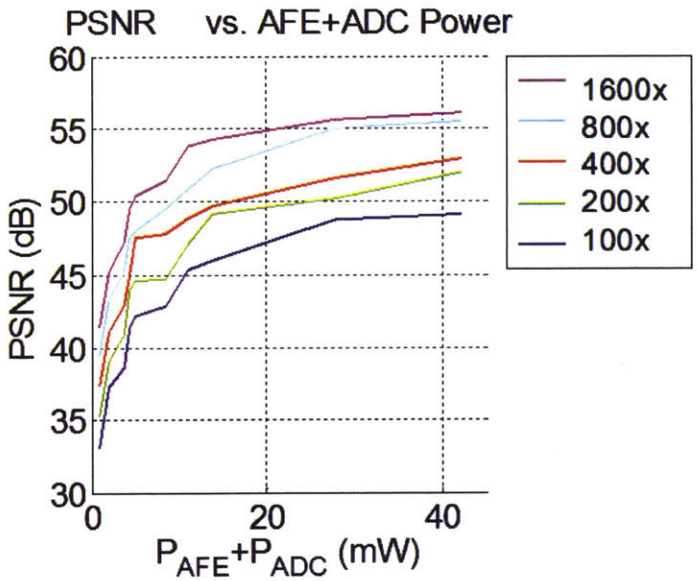


Figure 3-12: The relationship between the PSNR performance and the power consumption in AFE&ADC, under different amount of noise averaging.

noise averaging should be increased for further PSNR improvement.

3.5.3 Image Resolution vs. Beamformer Power

The scalable DBF architecture described in Section 3.2.3 imposes a limitation on the maximum frame rate attainable, since a fixed number of pixels is formed per processing cycle. Given a target frame rate, the required throughput of the DBF to match that frame rate can be computed and the DBF blocks can be powered down for the duration between the completion of one frame and the beginning of the next. By implementing the DBF block on a Virtex-4 FPGA and performing power analysis using the Xilinx ISE Power Analyzer (XPA), the power consumption of the DBF block assuming 100% duty-cycling (i.e. DBF turned on during all processing cycles) can be obtained.

Based on the target frame rate and assuming minimal leakage when the DBF blocks are placed in sleep mode, the power consumption obtained from XPA is scaled according to the duty cycling used. The linear scaling is likewise applied to the quarter and half resolution modes (power reduction due to lower number of pixels to be generated). An example image is shown for each of the three modes in Figure 3-13

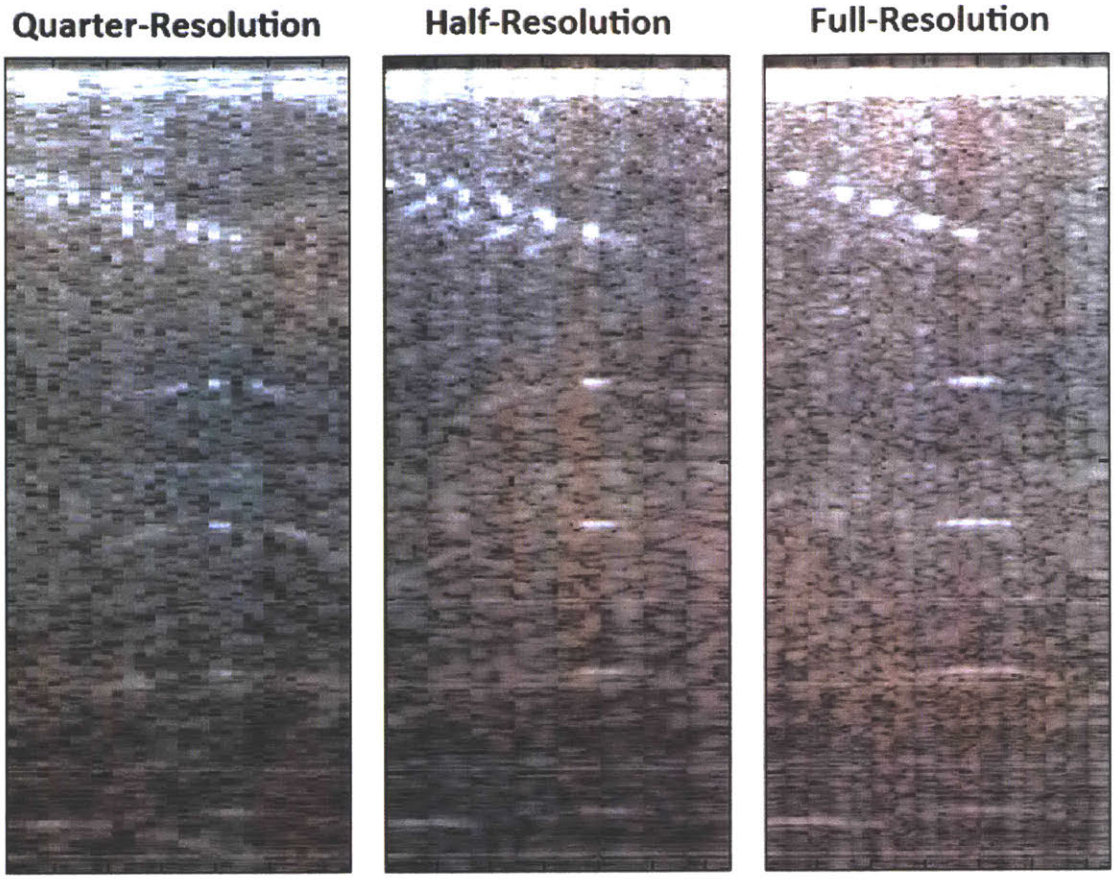


Figure 3-13: Image quality comparison for the quarter, half, and full resolution settings at run-time.

to demonstrate image quality degradation corresponding to this power reduction.

3.5.4 System Performance

System-level performance can be related to system power consumption by modeling the AFE, ADC, and DBF blocks together, as shown in Figure 3-3. The system-level performance metrics include the image PSNR and the frame rate. The frame rate is limited by two factors: the amount of noise averaging and the DBF throughput. In other words, for a fixed frame rate, the maximum number of noise averaging and the DBF operating frequency to achieve that frame rate are determined. Moreover, the AFE&ADC power consumption corresponding to a target PSNR and frame rate can be obtained from Figure 3-12, while the DBF power is determined from operating

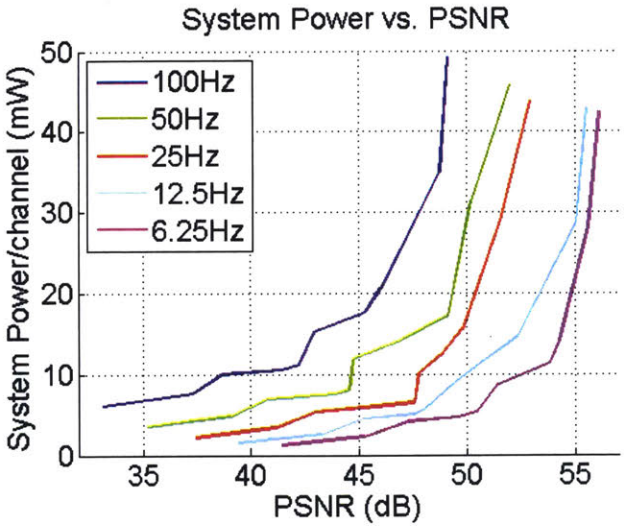


Figure 3-14: System power dissipation vs. image PSNR at different frame rates.

frequency. Combining the block-level power at each fixed frame rate, the relationship between the PSNR and the system power is derived, which is plotted in Figure 3-14 for a range of conventional imaging frame rates, assuming a Pulse Repetition Frequency (PRF) of 10KHz.

To visually represent the distribution of the system power that is dissipated by each block, the power consumption breakdown is summarized in Figure 3-15 at a frame rate of 100Hz. The ADC bit width used is also labeled in the figure. It is noted that the DBF consumes relatively constant amount of power with respect to PSNR, while the AFE power drastically increases as the image PSNR requirement raises. The ADC power also increases with the bit width to meet the noise requirements.

3.6 Summary and Conclusions

Overall, the system-level tradeoffs are clearly shown, and design-time optimizations can be made based on these relationships. Once design decisions such as the AFE noise requirement and the ADC bit width are made, the run-time tradeoff can be realized with selection of the fine/coarse beamforming modes implemented in the DBF, which trades the image PSNR performance for the frame rate.

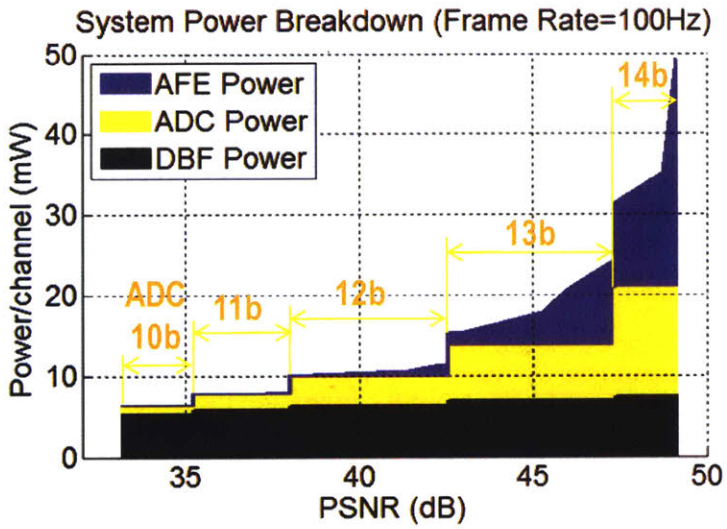


Figure 3-15: System power distribution at a frame rate of 100Hz, DBF at the fine imaging mode.

Chapter 4

Architecture of 2D Energy-Scalable Beamformer

This chapter details the architecture of the 2D energy-scalable beamformer. Where image quality is sacrificed for energy efficiency, the expected degradation is evaluated in simulation.

In this work, we extend this adaptive-resolution scheme and demonstrate an energy-scalable ASIC that replaces the digital beamformer on FPGA in [21]. To support a number of operating points to be selected under a given power budget and image reconstruction frame rate requirement, several design parameters are made reconfigurable at run-time. An overview of the processing units in the beamformer is presented, followed by a discussion on each of these parameters and its impact on those processing steps. The overall quality degradation of the reconstructed image is evaluated.

4.1 Processing Units

Figure 4-1 summarizes the signal flowchart involved in the delay-and-sum beamforming algorithm described in Section 2.4. They can be categorized into the delay insertion and summation components. The functional units detailed in this section are designed to perform these steps.

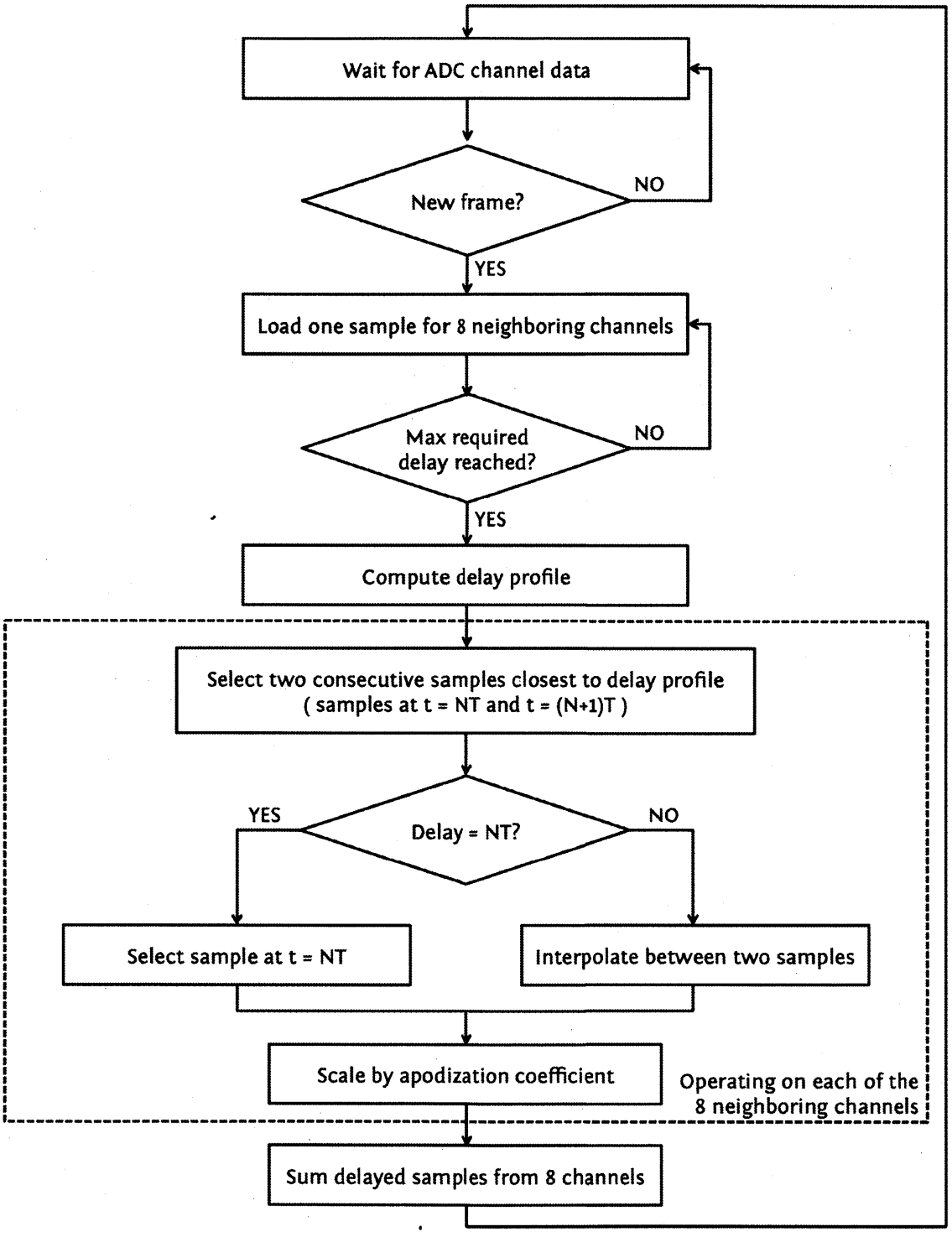


Figure 4-1: Control flowchart of the 2D delay-and-sum beamformer.

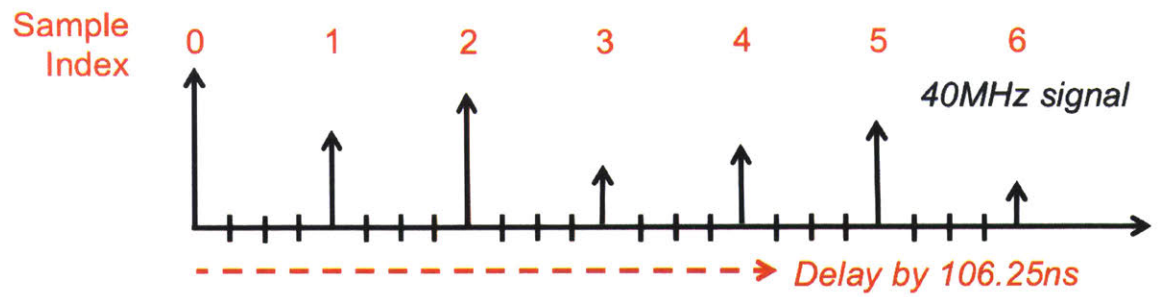


Figure 4-2: Schematic representation of integer and fractional components to a given delay. For a sampling frequency of 40 MHz, data samples are spaced apart by 25 ns, so a delay of 106.25 ns can be decomposed into an integer delay of 4 and a fractional delay of 1/4.

4.1.1 Precomputed Delays

In Section 2.4.3, we discussed the dependence of the delay profile on the depth at which we are focusing in order to achieve dynamic focusing. This means that a slightly different delay profile needs to be computed for every depth location imaged. In an ultrasound image consisting of 128 pixels in the lateral (x) direction and 2000 pixels in the axial (y) direction, this translates to 2000 distinct delay profiles that must be calculated.

Since our beamformer operates in the digital domain (on the outputs of ADCs), the received waveforms are discrete-time signals, sampled at a fixed frequency. This means that the exact values of the signal acquired are only available at those fixed intervals, spaced apart by the sampling period, T_s. In our system, the received waveform for each transducer element is sampled at 40 MHz and quantized to 12 bits by an ADC:

$$r_m[n] \approx r_m(nT_s), \tag{4.1}$$

where $T_s = 25$ ns is the sampling interval of the ADC.

The sampling period limits the spatial resolution of the delay line in a system that only supports integer delays, where the received waveform can only be delayed by integer multiples of T_s. However, the rule of thumb for delay resolution is 1/16 to 1/32 of the wavelength of the transmitted pulse [82, 83]. For our transducer, the center frequency is 7.5 MHz (see Table 3.1, the wavelength is approximately 133 ns.

Table 4.1: Maximum delay required under each azimuthal array decimation mode.

Array decimation mode	Maximum delay
Full	8
Half	16
Quarter	64

This means the desired delay resolution is approximately 4.17 ns.

To improve accuracy, we allow channel delays Δt_m to be quantized to $\frac{1}{8}$ of the sampling interval (3.125 ns):

$$s(t) = \sum_{m=1}^{M} r_m(t - \Delta t_m) \tag{4.2}$$

$$\Delta t_m \approx \Delta \hat{t}_m = \left(\tau_m + \frac{k_m}{8}\right) \cdot T_s \tag{4.3}$$

where $\tau_m = \lfloor f_s \cdot \Delta t_m \rfloor$ is the integer component of the delay (up to 64 cycles), and $k_m = \lfloor 8 \cdot (\Delta t_m - \tau_m \cdot T_s) \rfloor$ is the fractional component of the delay.

In order to reduce computational complexity, τ_m and k_m are precomputed and stored in static lookup tables. According to Equation (2.9) and Equation (2.10) (also illustrated in Figure 2-12), the delay is calculated from d (the array pitch, which is static for a given transducer array) and the focal depth z_f, and the delay profile is symmetric around the center element (see Equation (2.8)). This means that only four sets of lookup tables are necessary to support summation of up to 9 channels (one center channel plus four off-center channels on each side). We also noted in Section 2.4.3 that z_f changes as different pixels are reconstructed using dynamic focusing. To minimize on-chip storage of channel data, we only buffer enough data samples for each channel to match the steepest delay profile. Section 4.1.1 summarizes the maximum delay computed for the three azimuthal array decimation modes.

This translates to each lookup table being addressed by a 6-bit depth value to select the appropriately delayed sample up to $2^6 = 64$ samples ago.

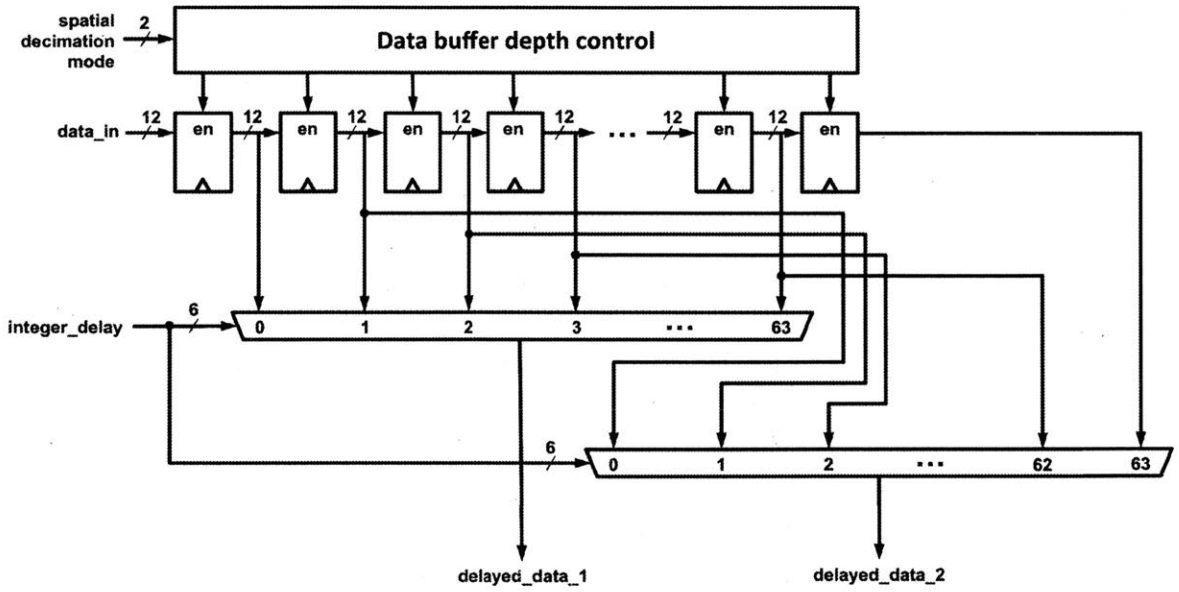


Figure 4-3: Shift register and multiplexer implementation of sample selection; two consecutive samples from each channel are chosen based on the integer delay values.

4.1.2 Sample Selection

The integer delay line (shown as the "Sample Selection" block in Figure 5-1) is implemented using a shift register and multiplexer for each channel, shown in Figure 4-3. The length of this "sliding window" for data storage depends on the effective array pitch (which increases as we decimate the array, as explained in Section 4.2) and is adjusted using the enable signals to reduce unnecessary switching.

4.1.3 Interpolation

The fractional delay line is implemented via two-tap linear interpolation, depicted in Figure 4-4; interpolation is bypassed if two consecutive samples for channel m are equal.

$$r_m(t - \Delta t_m) \approx \frac{8 - k_m}{8} \cdot r_m[n - \tau_m] + \frac{k_m}{8} \cdot r_m[n - \tau_m - 1] \tag{4.4}$$

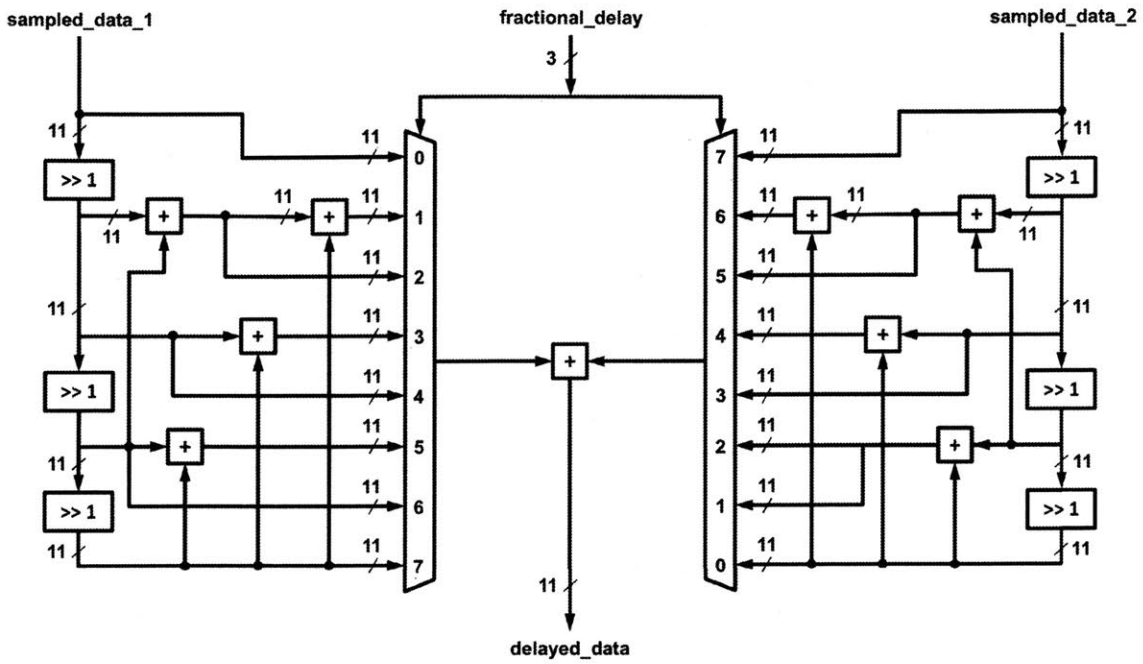


Figure 4-4: Shift-based two-tap interpolator implementation.

4.1.4 Apodization and Summation

A beamformer that sums the time-aligned received channels with uniform weights (as in Eq. 4.2) will be sensitive to echoes from the desired focal point as well as other locations (side lobes). The unwanted side lobes can be suppressed by spatial filtering (apodization), or summing the received channels with non-uniform weights A_m:

$$s(t) = \sum_{m=1}^{M} A_m \cdot r_m(t - \Delta t_m) \tag{4.5}$$

The apodization filter is designed to weigh the elements closest to the focal point more heavily than those farthest away. Filter performance depends on the number of taps M, which we call the aperture size. To simplify the computation, we allow the smallest coefficients to be zeroed out, neglecting the more distant transducer elements. Figure 4-5 illustrates this concept by showing the normalized spatial responses of a beamformer with and without apodization. Apodization suppresses the side lobes of the spatial response from -12.8 dB to -28.3 dB (in exchange for reduced gain and increased beam width). Using a smaller aperture further reduces the gain and increases

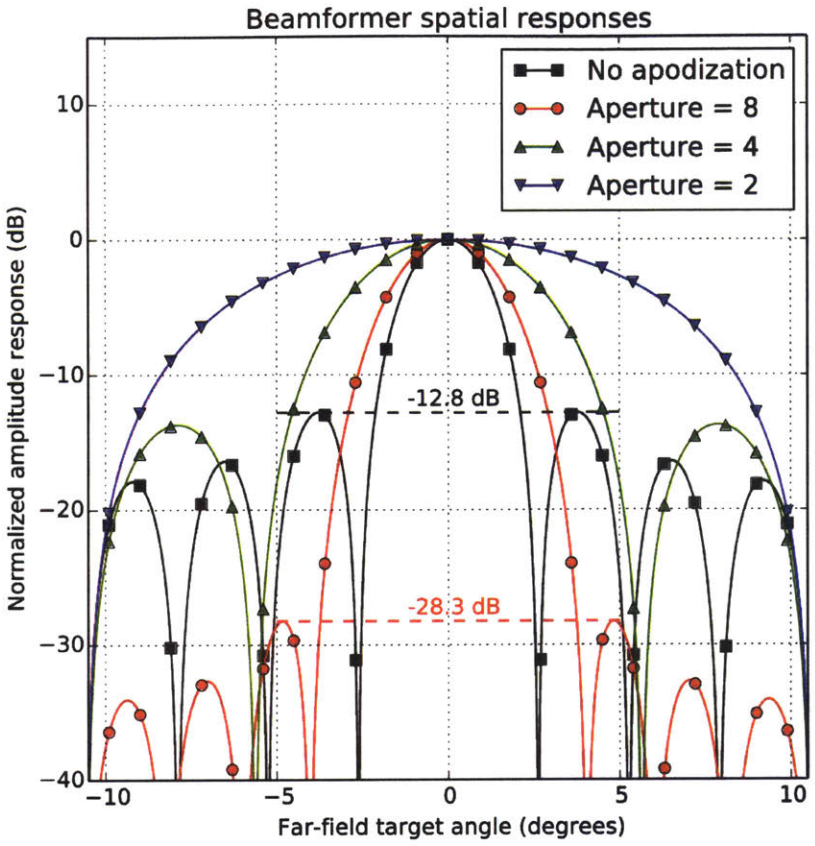


Figure 4-5: Spatial filter responses of different apodization schemes.

beam width.

4.1.5 Image Depth FSM

Most of the intermediate variables, including the delay profile and apodization coefficients, depend on the z-coordinate of the image pixel being reconstructed (i.e. the image depth in the axial direction). The finite state machine keeps track of this depth of the current pixel.

4.2 Variations for Energy-Scalability

In this section we present variations on the 2D beamforming algorithm allowing trade-offs between energy consumption and image quality, which we quantify in Section 5.5.2.

These variations pertain to individual components of the beamformation process, which are summarized in Figure 4-1.

4.2.1 Azimuthal Array Spatial Sampling

As discussed in Chapter 3, the system energy can be significantly reduced with the number of parallel AFE and ADC channels being used, due to the high per-channel power consumption of the analog portions of the system (including transmitters). When the user opts to turn off a subset of AFE and ADC channels, we expect the reconstructed image to be an approximation to the true image that can be reconstructed using the full system. The beamformer should therefore adapt to the change in the element pitch and total number of channels of data it can use for reconstruct the image. This concept is equivalent to spatially sampling the received data in the lateral or $x-$direction. This design supports two additional modes where every second and every fourth consecutive transducer channel is sampled, catered to use cases with tight power constraints where a lower resolution image is acceptable. The three modes are referred to as the Full, Half, and Quarter spatial sampling modes, and are schematically represented in Figure 4-6, Figure 4-7, and Figure 4-8.

The half and quarter spatial sampling modes effectively double and quadruple the transducer array pitch, respectively, so a different delay profile needs to be computed for each of the three modes. Figure 4-9 shows a direct comparison of images beamformed using these three modes.

4.2.2 Aperture Selection

After beamforming, we constructed images (using Matlab) with a dynamic range of 60 dB. Figure 4-10 shows the impact of varying aperture sizes on horizontal cross-sections of a beamformed image. For values of $M = \{2, 4, 8, 16\}$, we show an overlay of normalized intensity at two different depths. The peaks correspond to bright spots in the beamformed image, as shown in Figure 4-9. Decreasing M introduces additional image artifacts caused by reflections originating away from the focal point. Our VLSI

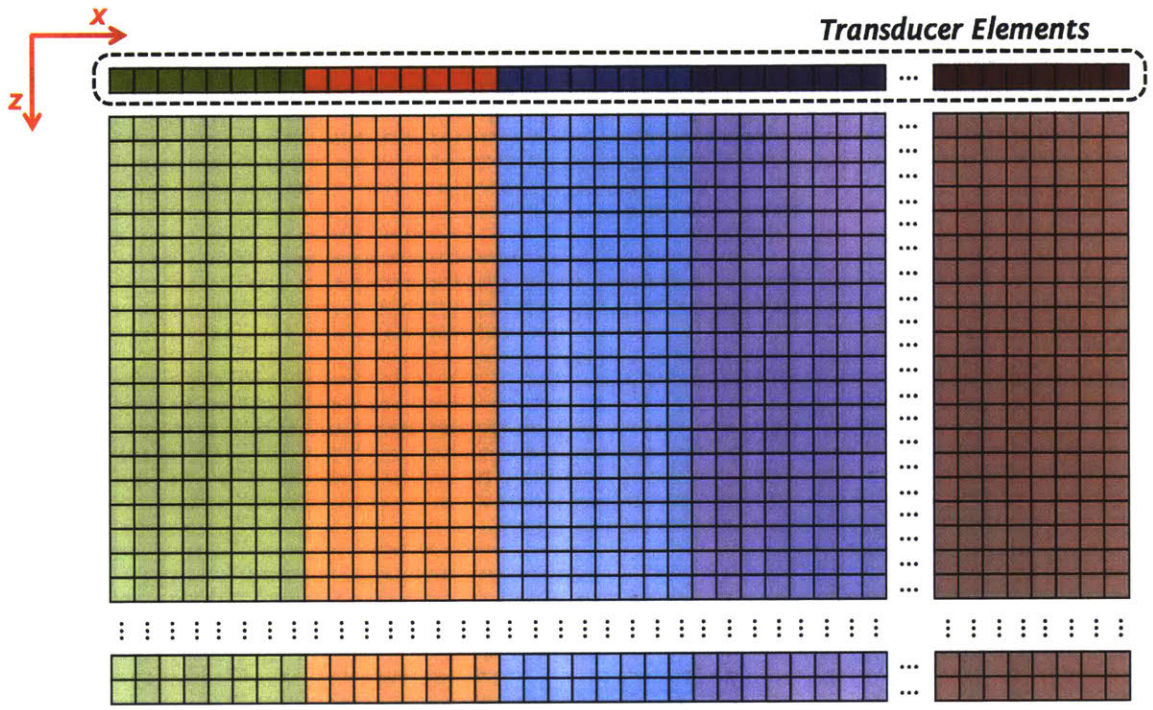


Figure 4-6: Imaging mode where transducer array is not decimated (full array of 128 channels used for imaging a 128 pixel-wide image).

implementation allows the user to select between aperture sizes of $M = \{2, 4, 8\}$.

4.3 Dataset Description

As with the dataset used in the system-level energy model presented in Chapter 3, the 2D imaging dataset used to demonstrate correct beamforming functionality was provided by Ultrasonix Medical Corporation. The same linear probe was used to acquire this data, and its specifications are summarized in Table 3.1.

4.4 Post-Synthesis Area Breakdown

The 2D energy-scalable beamformer has a gate count of 160 kgates after synthesis. This includes the input/output interfacing logic that will be discussed in Section 5.2. Table 4.2 summarizes the area breakdown within the beamformer design. The eight 12-bit shift registers store a sliding window "snapshot" of eight input waveforms, which

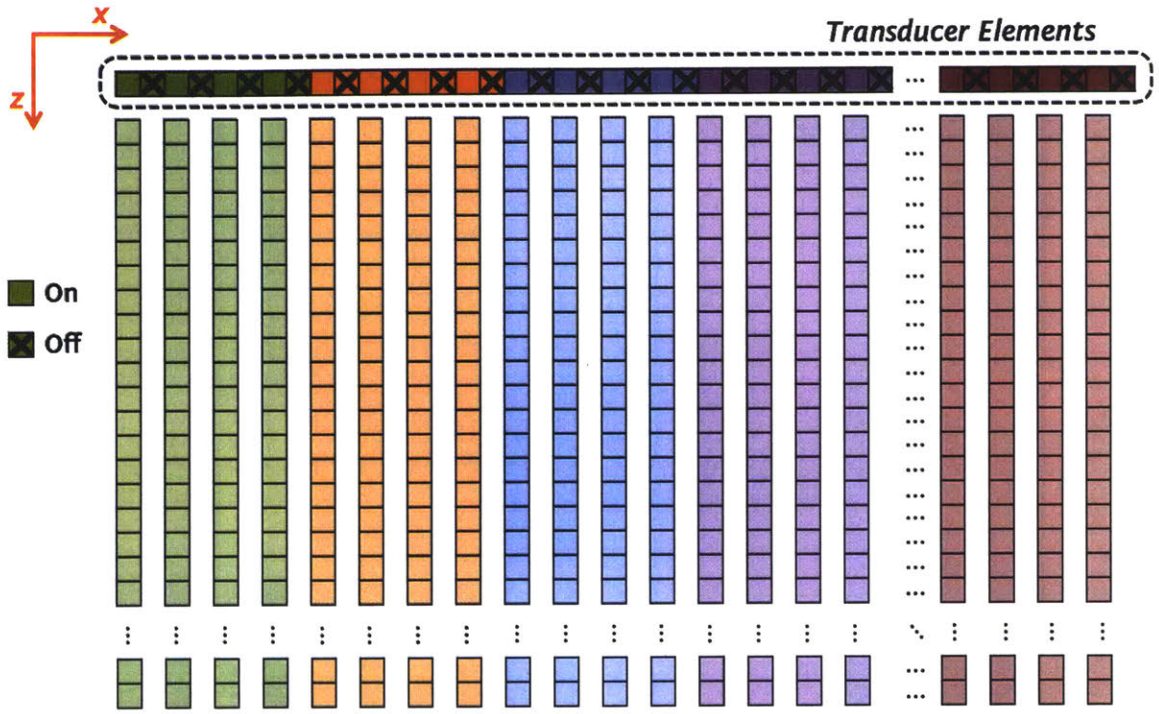


Figure 4-7: Imaging mode where transducer array is decimated by a factor of 2 (every other channel is turned off, so that an equivalent array of 64 channels is used for imaging a 64 pixel-wide image).

account for 37.69% of the total core area. The static lookup tables for integer and fractional delays (τ_m and k_m, respectively) account for less than 2%. Multipliers and adders account for around 55%, and the rest is made up of scanline circuitry that generates the four 16-bit beamformed outputs.

Table 4.2: Post-layout ASIC area breakdown of the beamformer test chip.

Functional Block	**Quantity**	**% Total Area**
Channel Data Storage	8	37.69
Adder trees	4	29.79
Apodization multipliers	22	25.32
Interpolation	7	5.42
Lookup tables for τ_m	9	0.85
Lookup tables for k_m	11	0.92

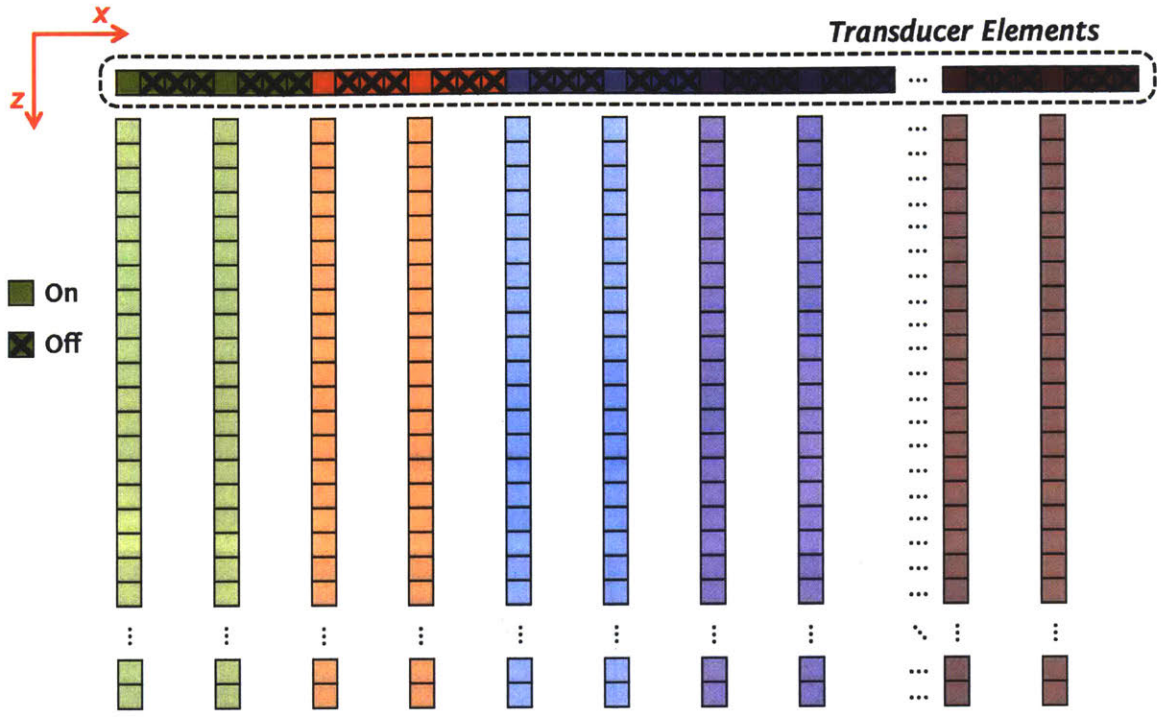


Figure 4-8: Imaging mode where transducer array is decimated by a factor of 4 (three out of four adjacent channels are turned off, so that an equivalent array of 32 channels is used for imaging a 32 pixel-wide image).

4.5 Summary and Conclusions

Building on the idea that system power reduction is achievable through an array decimation scheme, the 2D beamformer with variable-length sample buffering and three supported delay profiles for dynamic focusing is designed. This system is able to support multiple operation modes to demonstrate a wide range of power-performance combinations.

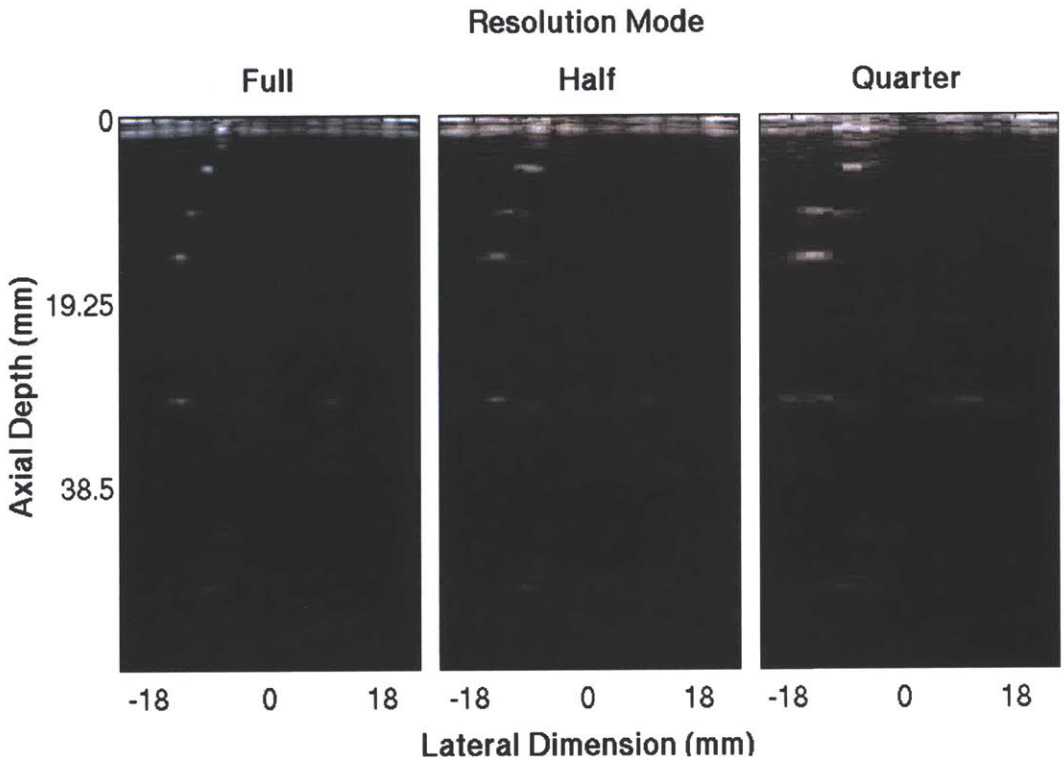


Figure 4-9: Side-by-side comparison of beamformed images for three resolution modes.

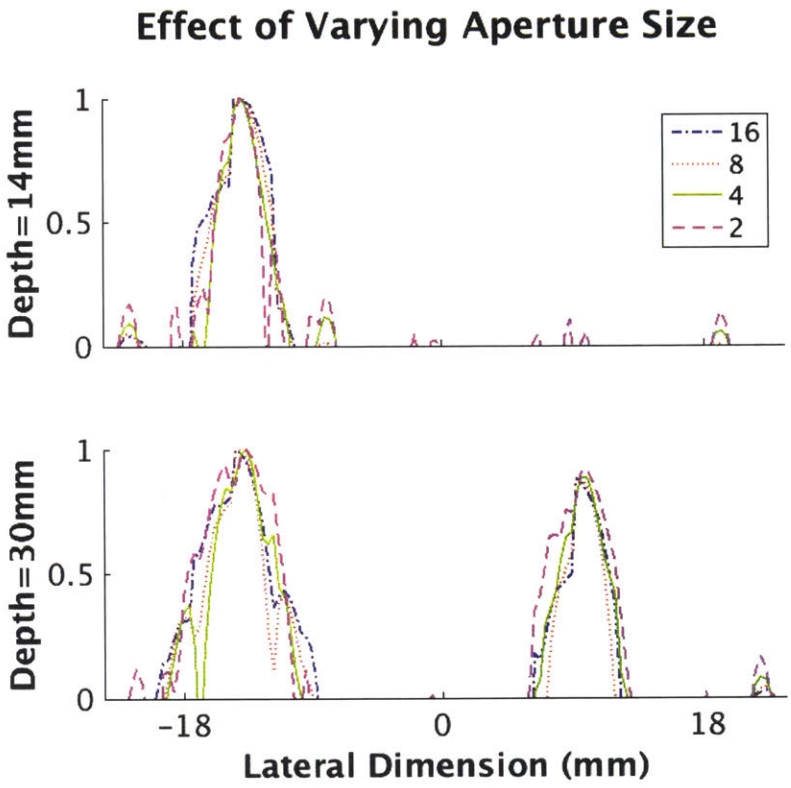


Figure 4-10: Analysis of point spread function for aperture selection.

Chapter 5

2D Energy-Scalable Beamformer Test Chip

This chapter describes the implementation of the test chip used to demonstrate the power and performance tradeoffs of the 2D energy-scalable beamformer. The challenges... The test setup used for the measured results will be presented and

5.1 Functionality

The top level architecture of the 2D energy-scalable beamformer test chip is shown in Figure 5-1.

5.2 Input/Output Interface

Our chip has a relatively small core area and is I/O-limited. To reduce the number of data pins, we supply the 12-bit input channel data and expect the 16-bit output line data to appear serially over four clock cycles. This means that the core clock frequency is one-fourth that of the I/O clock frequency. The timing diagram in Fig. 5-2 illustrates the 1-to-4-bit deserialization applied to the input channel data and the 4-to-1-bit serialization applied to the output line data.

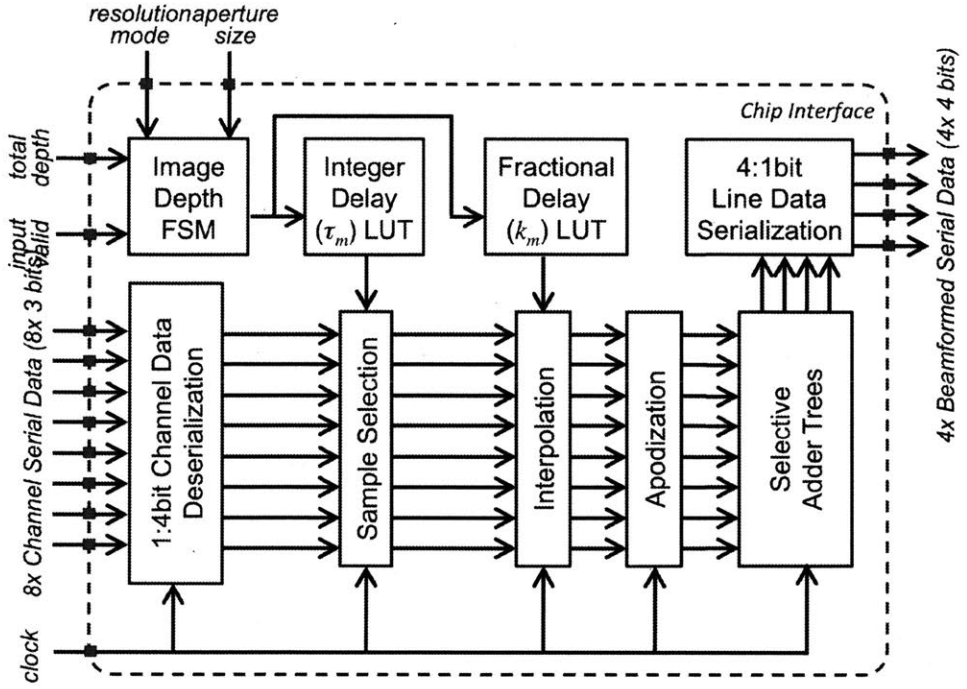


Figure 5-1: Top-level interfaces and architecture of beamformer chip.

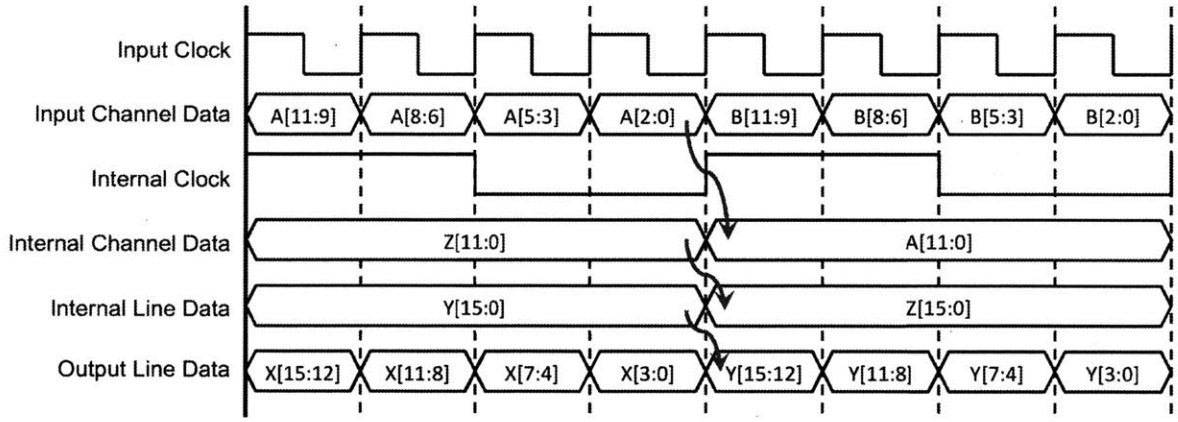


Figure 5-2: Timing diagram depicting I/O deserialization and serialization.

5.3 Configurability

ASIC design and testing allowed us to evaluate the tradeoffs described above. Our beamforming chip is designed to fit into a low-power system that performs imaging with a 128-element transducer array. The image is 128 $\times$ 2000 pixels, and is split into 8-pixel-wide strips which are independently constructed from an 8-element subarray of the transducer. In synchronous mode, the chip operates at the ADC sampling rate

of 40 MHz and produces one strip of image data from each stimulus pulse in 100 μs. External multiplexers cycle through the 16 subarrays to construct the complete image. Buffering the samples from a single stimulus pulse allows asynchronous operation, where the chip's clock frequency is determined by the video frame rate (4 pixels per clock, or 1.92 MHz at 30 fps). In addition to adjustable aperture and resolution, we used sliding window data storage to reduce area and selective adder trees to reduce power:

- **Sliding window data storage:** Our beamformer stores only the samples that are needed for constructing the image as a stream of pixels. This eliminates the need for a large on-chip SRAM to hold all transducer samples across a frame. The length of this sliding window depends on the delay value computed for the outermost transducer element in the summation, which depends on the aperture size chosen and the z-coordinate of the focal point being processed. We defined the length of our window by accounting for the worst case aperture as $M = 8$ and the worst case depth (with the steepest delay profile) as the near field region where z is small.

- **Selective adder trees:** Pixels near the edge of the strip have only 4 nonzero aperture coefficients, compared with 8 for the central pixels. Our architecture provides multiple adder trees and selects a tree of the appropriate size for each pixel, in order to reduce dynamic power.

5.4 Test Setup

The end-to-end system usage model is illustrated in Figure 5-3. Three commercially available development boards from Texas Instruments that each support 8 parallel channels are used for the transmit analog circuitry, receive analog front end and ADC, and digital data buffering on an FPGA. The 128-wide Ultrasonix transducer probe is to be connected to a time-multiplexing breakout board that will reuse those 8 channels to cover all 128 channels in the probe. A daughter card that is pitch-matched to

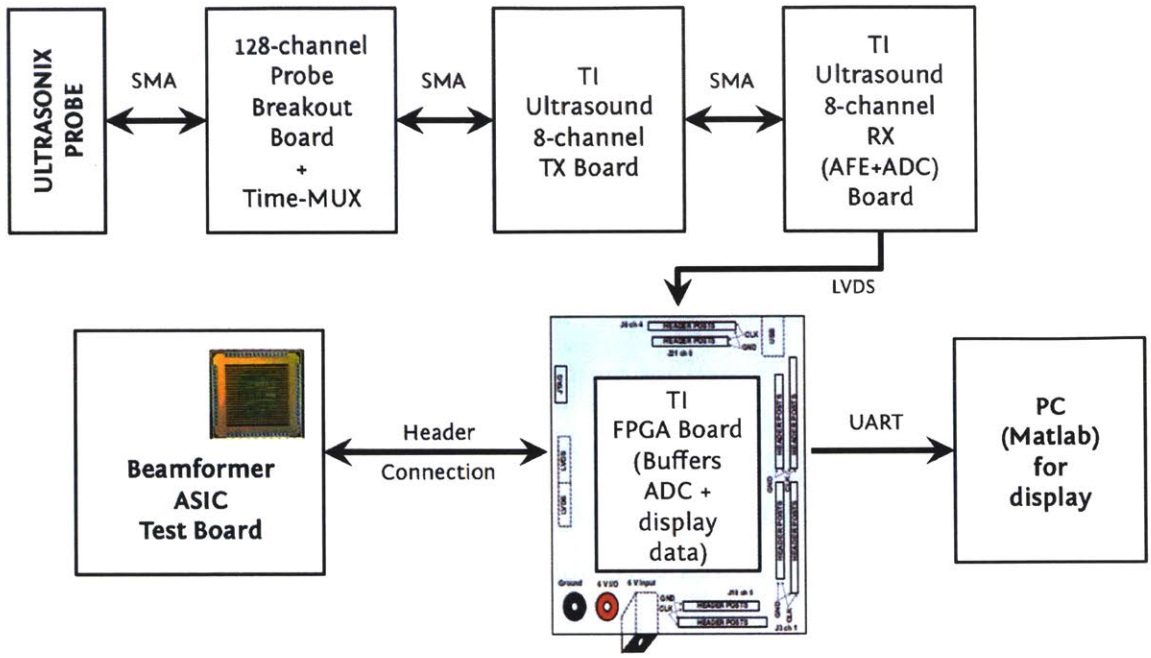


Figure 5-3: Transmit/receive end-to-end usage model for 2D Beamformer ASIC.

the Xilinx FPGA board connects the beamformer chip's I/O pins with the buffering FPGA, which then relays the reconstructed image to a PC via a UART connection.

To simplify the testing process, the daughter card were used in isolation for testing (as shown in Figure 5-4). Instead of real-time data being captured using the probe and analog components, previously acquired digital data is supplied from a pattern generator (such that it replaces the FPGA data buffer described above). This also allows us to easily adjust the clock frequency of the chip as we scaled its voltage to demonstrate power reduction via Dynamic Voltage and Frequency Scaling (DVFS).

5.5 Measurement Results

The test chip shown in Figure 5-5 was designed using a standard cell methodology and fabricated in a 65 nm low-power CMOS process. The total area is fixed by the 64 I/O pads at 1.435 $\times$ 1.434 mm. Of the 59 pads carrying I/O signals, 24 carry 4:1 serialized input waveforms for eight parallel transducer channels and 16 carry serialized beamformed output pixels. (The interfaces run at four times the internal

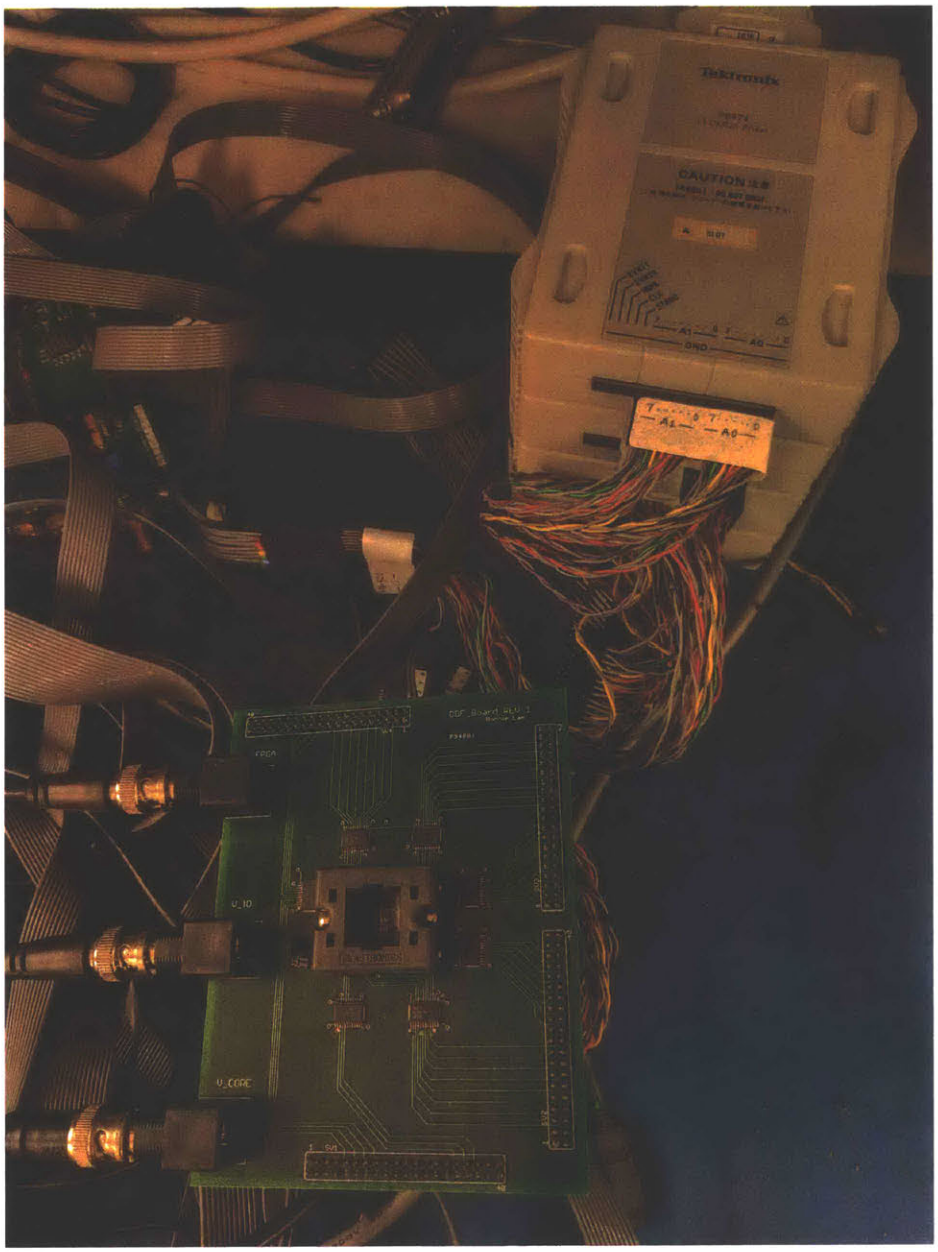


Figure 5-4: Test setup for 2D Beamformer ASIC.

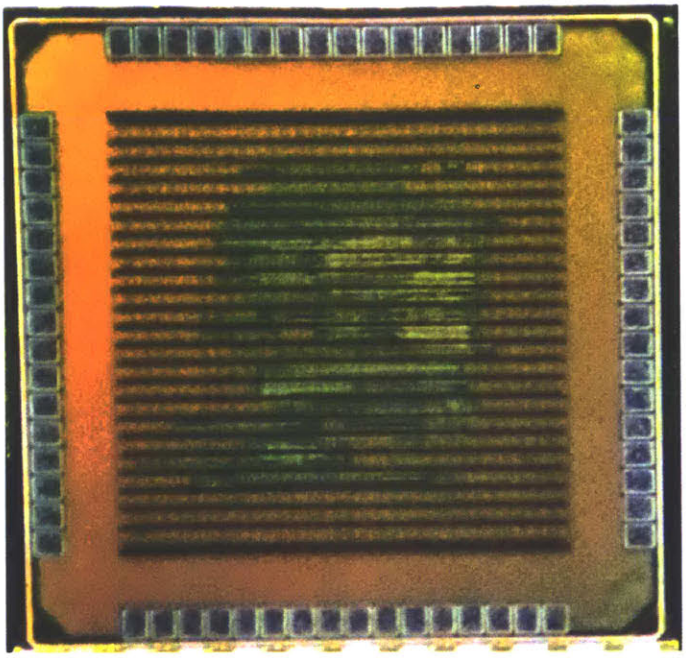


Figure 5-5: Die photo of 2D energy-scalable beamformer test chip.

clock frequency.) The active area is 0.911 mm^2.

To demonstrate functional correctness and evaluate the tradeoff between image quality and energy consumption, the test chip power was measured when beamforming a set of previously captured channel data under various resolution and aperture modes. The digital channel data, provided by Ultrasonix Medical Corporation, was collected from a single transmit firing and sampled at 40 MHz, using a linear array probe with a center frequency of 14 MHz and a bandwidth of 5 MHz. A Tektronix pattern generator was used to provide the clock and logic inputs to the test chip. Four lines of 4-bit serialized beamforming output were captured using a Tektronicx logic analyzer for further analysis. A Keithley 2400 series Sourcemeter was used to regulate the chip's power supply and measure the current drawn by the core. Core power was measured instead of total chip power because the beamformer would be integrated with AFE and ADC components in our target application.

Table 5.1: Summary of 2D beamformer chip implementation.

Channel width	8
Technology	65-nm
Core Area	0.955 mm × 0.954 mm
Logic Gate Count (NAND-2)	160k
I/O Pads (package)	64-pin TQFN
Supply Voltage (core)	0.49 to 1.0 V
Supply Voltage (I/O)	1.8 V
Operating Frequency	24.1kHz to 66.7 MHz
Core Power Consumption	3.38 μW to 4.048 mW

5.5.1 Power

The chip is operational at supply voltages from 0.49 V to 1.0 V. As the core supply voltage is decreased, power consumption decreases along with the maximum clock frequency. The power measurements and equivalent energy per frame are plotted in Figure 5-6. Core power consumption when operating in full-resolution, $M = 8$ mode at 0.49 V is measured to be 3.38 μW. It is well established that there exists a minimum energy point, beyond which energy per operation increases as leakage power becomes significant compared to active power [84]. This occurs at 0.54 V, where power consumption is measured to be 235 μW and clock frequency is 4.0 MHz. The energy dissipated to construct a 2000-by-128-pixel frame is 3.765 μJ.

5.5.2 Performance Tradeoffs

The maximum asynchronous processing speed (in frames per second) for all three resolution modes at various core voltages is shown in Figure 5-7. For core voltages above 0.8 V, frame rate plateaus at 1000 fps due to the frequency limitations of our test setup. Frame rate rapidly decreases below 0.52 V. An equivalent frame rate of at least 30 fps (a typical frame rate for video) is achieved for core voltages above 0.52 V in the full resolution mode. Since the number of pixels computed (and the required clock frequency) decreases by 50% and 75% for the half and quarter resolution modes, respectively, the core voltage can be decreased slightly in those modes while maintaining 30 fps. The peak signal-to-noise ratio (PSNR), a image

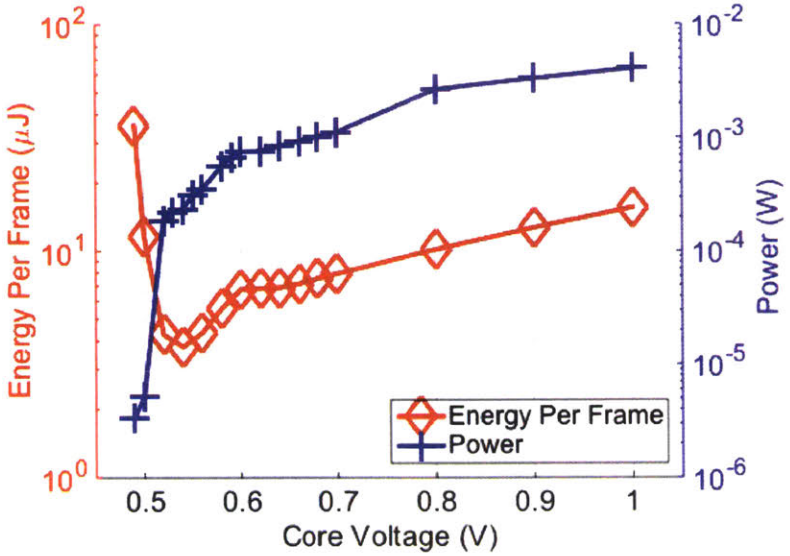


Figure 5-6: Energy per frame is minimized at 3.765 μJ when core voltage is set to 0.54 V in full resolution mode.

quality metric defined and analyzed in [21], is plotted with energy per frame for a fixed supply voltage of 0.52 V and equivalent frame rate of 30 fps in Figure 5-8. We note that the degradation in PSNR is much more pronounced in the red curve where the maximum number of array elements are summed (as we spatially decimate the array). This could be attributed to the fact that in the decimated modes, the aperture size effectively doubles and quadruples for the same number of elements summed. The gain in azimuthal resolution from the increased aperture size may have partially compensated for the loss of information due to decimation. The energy consumed per frame reconstructed under the "Quarter" array spatial decimation mode increases by very little as the number of elements summed per pixel increases. This is most likely due to the reduced number of computations in a frame formed under this mode - the total number of pixels to be reconstructed in this mode is one quarter of that of the "Full" case (where the waveforms received by every element in the array is used).

As we proposed in [21], the aperture size and spatial decimation mode can be selected to trade between energy consumption and PSNR.

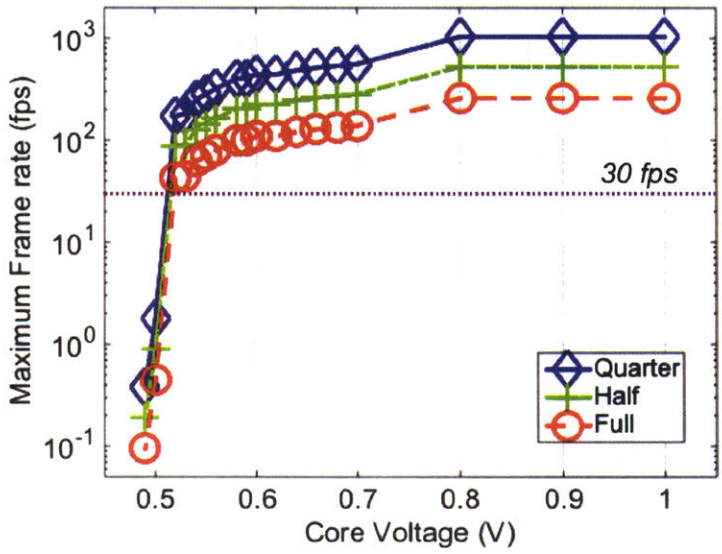


Figure 5-7: A frame rate of at least 30 fps is attainable for core voltages above 0.52 V.

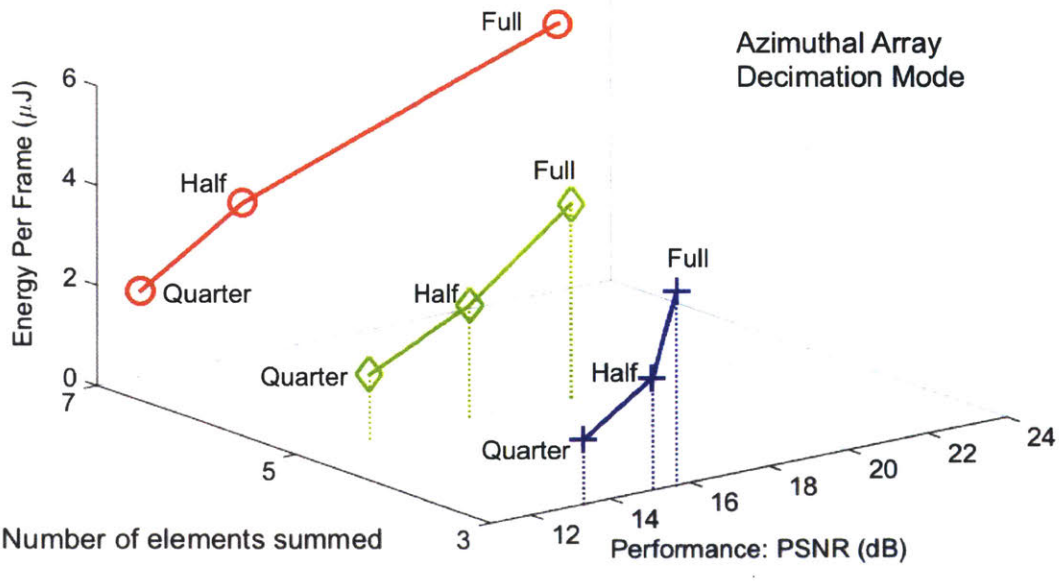


Figure 5-8: Energy-performance tradeoff at core voltage of 0.52 V and 30 fps.

5.6 Summary and Conclusions

In this work, we implemented, fabricated, and tested an 8-channel beamforming ASIC At least 30 fps image reconstruction has been demonstrated at 0.52 V with 185 μW power consumption. Configuration settings allow for user-selected azimuthal decimation modes (full, half, and quarter) and aperture sizes ($M = 3, 5, 7$) as needed to meet the appropriate power constraints for portable applications. Our parallel architecture helped to increase performance and counter the effects of voltage scaling. Sharing delay lookup tables between the eight parallel processing units helped minimize chip area.

Chapter 6

Software Demonstration of Energy-Scalable 3D Ultrasound

The use of two-dimensional transducer arrays to generate 3-D ultrasound images is gaining importance as the volumetric view provides improved ease of use and accuracy in diagnosis. However, these benefits come with the cost of increased complexity, which in turn limits both the maximum attainable frame rate and the battery life in portable or hand-held systems. The resolution in each of the lateral, axial, and elevational directions, labeled in Figure 6-1, contributes to the overall accuracy of the 3-D image constructed.

Assuming a transducer array containing P-by-Q elements and an imaging volume consisting of X-by-Y-by-Z voxels, the 3D equivalent of pixels in 2D imaging (as shown in Figure 6-2), the intensity of each voxel is computed as the sum of $P \times Q$ time-delayed and weighted time series (each captured by one of the transducer elements) in conventional delay and sum beamforming. Even with a reduced 2-D aperture size of $M \times N$, the number of computations is only slightly decreased, and is still on the order of $M \times N \times X \times Y \times Z$ to generate one volumetric frame. The computational complexity further increases if frame averaging is to be used to improve image quality, as is the case in 3-D plane-wave coherent compounding (PWCC3D) that we introduce in this chapter. On the other hand, extending the array decimation scheme presented in Section 4.2.1 to 3D imaging, we can expect to reduce energy consumption at the

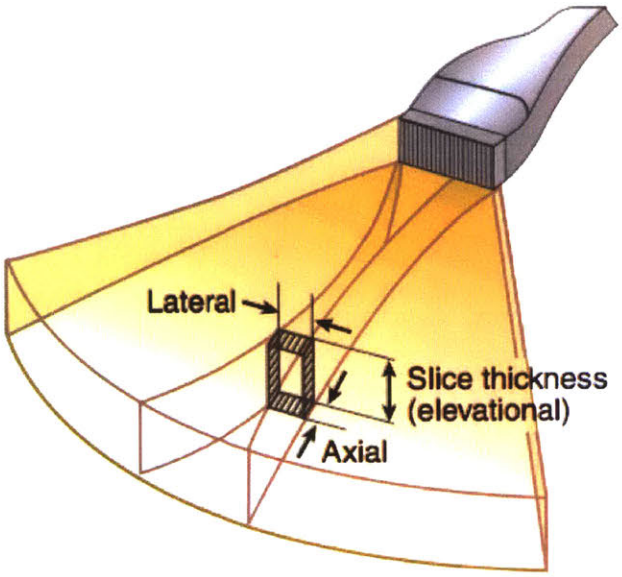


Figure 6-1: Geometry of the lateral (x), elevational (y), and axial (z) axes relative to the ultrasound probe [7].

expense of image quality. This work aims to quantify the tradeoff between energy consumption and performance by varying parameters under these two frameworks.

6.1 Dataset Description

The imaging data used in this work is the result of experimental work by Dr. Kailiang Chen [8]. A custom-built oil tank, shown in Figure 6-3, was integrated with a 16×16 CMUT and column-row-parallel front-end ASIC through flip-chip bonding. Taking into account that a number of the edge elements were either short-circuited or otherwise non-functional due to non-idealities in the CMUT process, only 15×15 of the array elements are used for beamforming. The 2D transducer array specifications are summarized in Table 6.1.

A metal wire shaped into a circular ring (with a known discontinuity where the two ends of the wire meet) with a diameter of $0.48mm$. The ring phantom is secured to an XYZ translation stage so that it can be placed at an exact location relative to the CMUT transducer array at the bottom of the oil tank. In our case, the axial

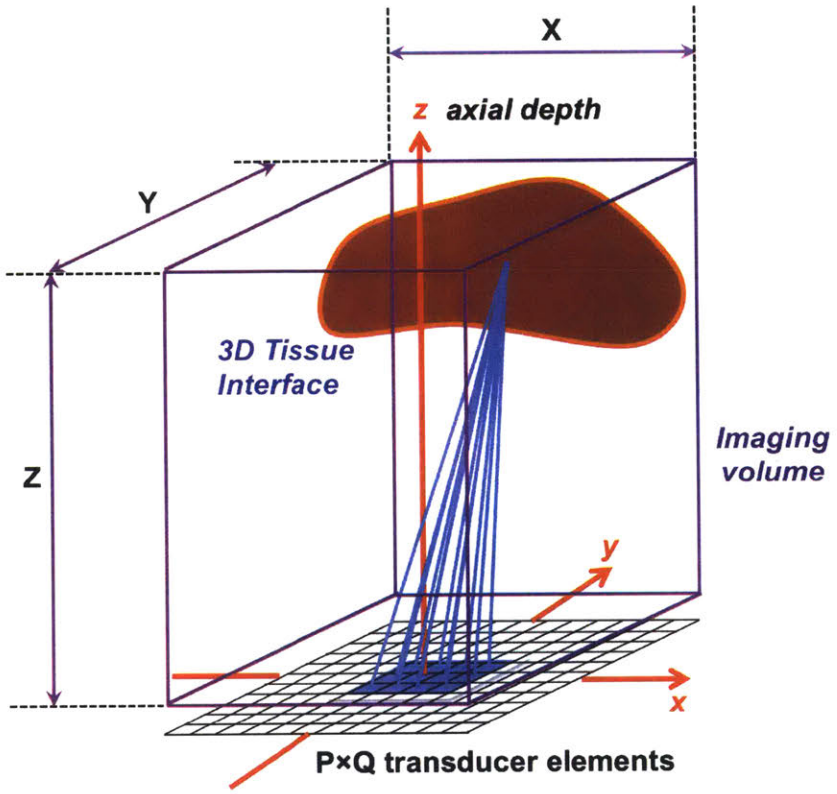


Figure 6-2: Geometry of the transducer and the imaging volume for 3D ultrasound imaging relative to the Cartesian coordinate system.

Table 6.1: Specifications of 2D CMUT Transuducer array used for 3D data acquisition.

Number of elements in x-direction (P)	15
Number of elements in y-direction (Q)	15
Element pitch	250 μm
Transmit pulses	2 bursts, 8.3 MHz
Depth range	6.8-8.8 mm

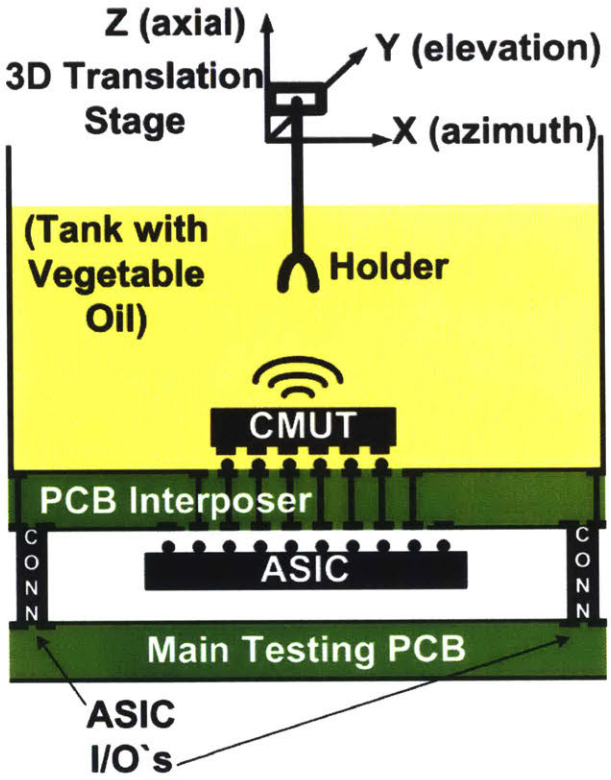


Figure 6-3: Diagram of data acquisition system used to obtain 3D image data [8].

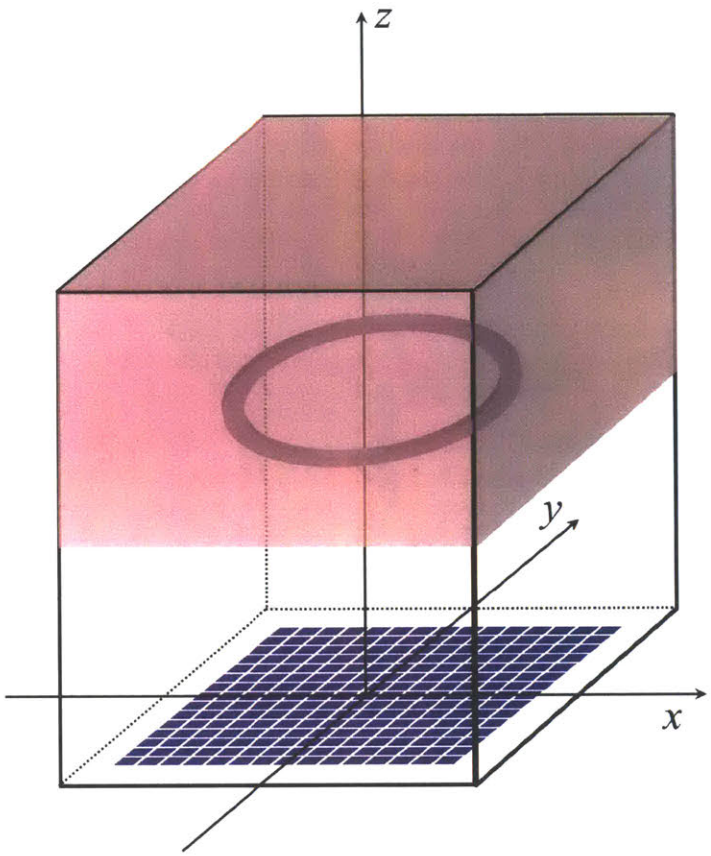


Figure 6-4: Diagram of wire ring position relative to 2D transducer array.

distance between the transducer element and the ring is $7.5mm$.

We divide the imaging volume into a matrix containing 31 points in each of the $x-$, $y-$, and $z-$ directions, the beamforming operations are performed on $31 \times 31 \times 31 = 29,791$ voxels. When reconstructing a voxel located at (x, y, z), a squared aperture centered around $(x, y, 0)$ is used for beamforming, with the exception of the pixels near the edge, which have an asymmetric aperture window.

The naive algorithm, where the volume consisting of 31 by 31 by 31 pixels, is imaged by traversing the x, y, and z dimensions, computing its distance from each transducer element, and applying the corresponding delay before summing the delayed signals, is implemented as the baseline. Using this baseline, we will quantify the amount of power reduction using several other techniques. We will also compare the image quality of all techniques with this baseline, giving the tradeoff between image

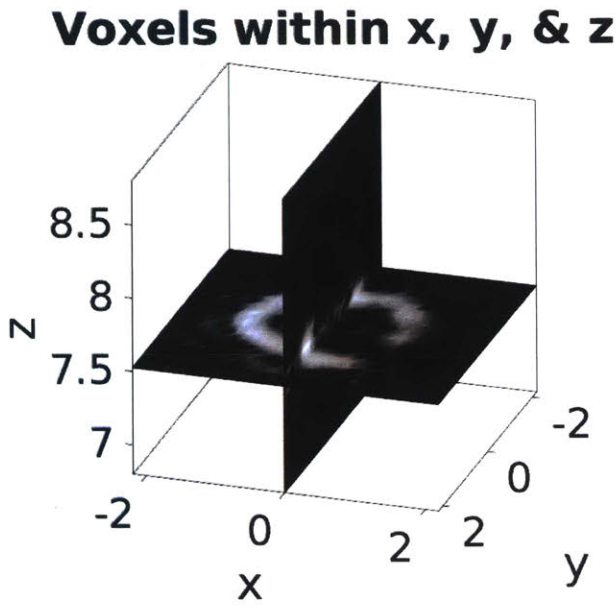


Figure 6-5: Visualization of 3D beamformed image of the circular ring phantom, generated using the baseline algorithm.

quality and power consumption. The visualization of the 3D beamformed image of a circular ring, generated using this baseline algorithm, is shown in Figure 6-5.

6.2 Variations for Energy Scalability

In this section we present variations on the 3D beamforming algorithm that allow tradeoffs between energy consumption and image quality, which we measure and quantify in Section 6.4.

6.2.1 Plane Wave Coherent Compounding

Plane wave coherent compounding for 2D imaging involves transmitting plane waves towards the imaging area by exciting the entire array at once [85–88]. This reduces the number of excitations required to form an image when compared to element-by-element excitation with the goal of increasing the frame rate. To compensate for the loss in image quality, this scheme proposes the coherent combining of multiple frames,

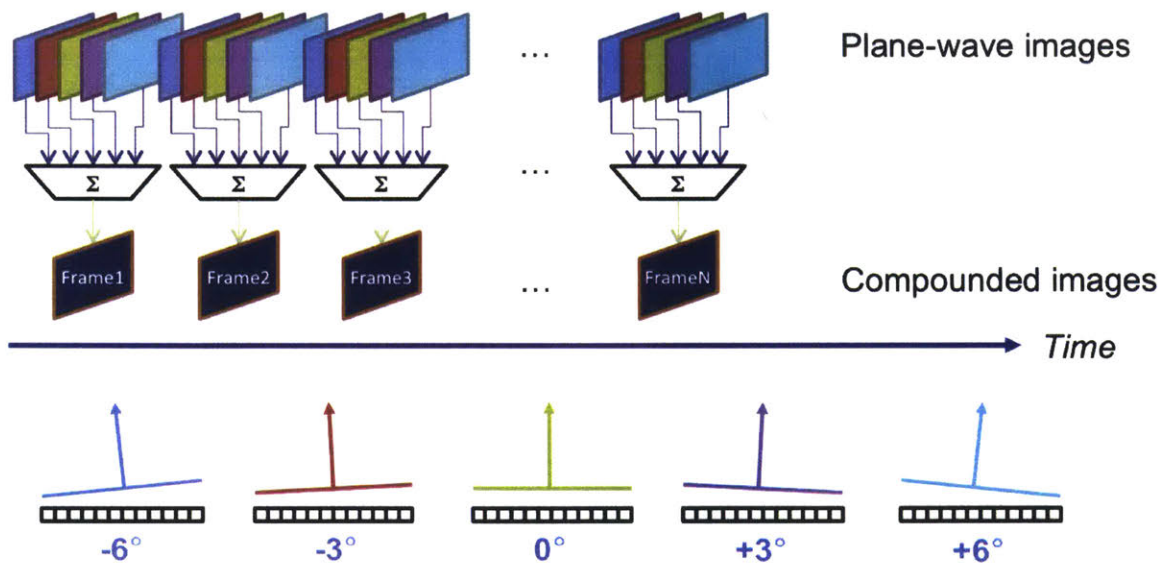


Figure 6-6: Pictorial description of 2D plane wave coherent compounding using 5 angles.

each acquired by transmitting a different incident angle towards the area of interest. Figure 6-6 shows a schematic of involving 5 angles: the 0° angle indicates that the plane wave travels in the path normal to the surface of the transducer array, and two pairs of off-center angles that direct the waves in opposite sides of that normal path.

The extension of this scheme to 3D imaging was proposed in [8], where plane waves are transmitted from a 2D array into the volume of interest. By varying the number of azimuthal and elevational angles along which the plane waves are transmitted, we effectively increase the number of volumes (and thus, the associated computations) used to reconstruct a frame with higher image quality. Figure 6-7 shows the conceptual diagram of varying the incident angle independently of the transmitted waves in the x- and y- directions.

Depending on the incident angle used in the transmit waveform, a different transmit path delay is computed and included in the delay calculations for beamforming that volume. Figure 6-8 labels an example of an incident angle α in the azimuthal (x-)

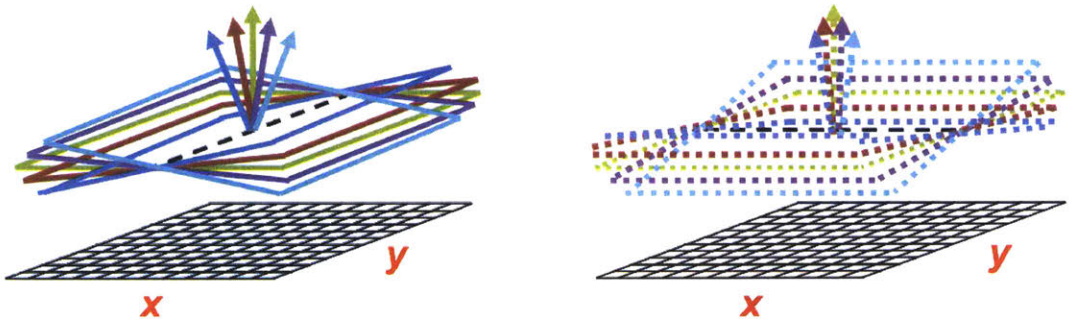


Figure 6-7: Pictorial description of 3D plane wave coherent compounding using 5 angles in x-direction and 5 angles in y-direction.

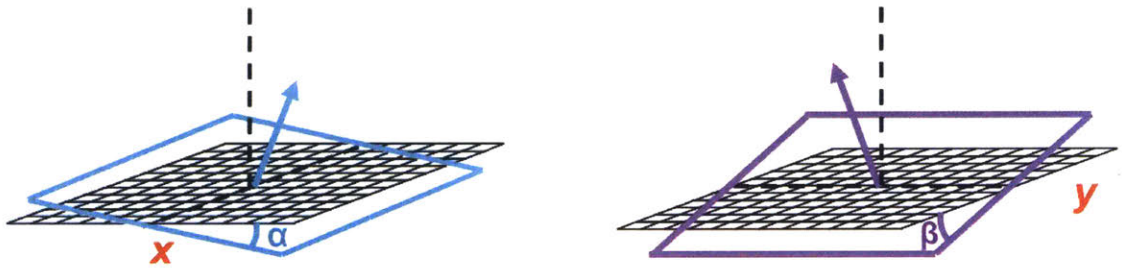


Figure 6-8: Diagrams of example incident angles (α in the azimuthal direction and β in the elevational direction).

direction and an incident angle β in the elevational (y-) direction. Equations (6.1) and (6.2) compute the transmit path delays to a focal point (x, y, z) of a plane wave traveling in these two directions. The power-performance tradeoff is examined by reconstructing an image under varying levels of plane-wave coherent compounding in the azimuthal and elevational directions, measuring the power consumption, and calculating the PSNR for the image.

$$\tau_{TX\_azimuthal}(\alpha, x, y, z) = \frac{z \cdot \cos(\alpha) + x \cdot \sin(\alpha)}{c} \tag{6.1}$$

$$\tau_{TX\_elevational}(\beta, x, y, z) = \frac{z \cdot \cos(\beta) + y \cdot \sin(\beta)}{c} \tag{6.2}$$

Table 6.2: List of possible configurations for 3D plane wave coherent compounding.

Number of angles in		Total number of volumes compounded per frame
X	Y	
1	1	1
1	3	3
1	5	5
3	1	3
3	3	5
3	5	7
5	1	5
5	3	7
5	5	9

Adding this transmit path delays to the receive path delays computed based on the distance between each transducer element and the location of each voxel, conventional delay and sum beamforming is applied to the received waveforms to reconstruct one intermediate volume. Finally, the intermediate volumes are compounded by summing the values of each voxel across as many volumes as are compounded, and the result is mapped to intensity values in a fixed dynamic range, effectively averaging the volumes to create one frame.

The exhaustive list of all possible combinations of azimuthal and elevational angles and the resultant number of volumes compounded is summarized in Table 6.2. The number of computations changes more or less linearly as we increase the number of volumes being compounded to reconstruct each frame, since the same 3D beamforming operation takes place for each intermediate frame, and the compounding algorithm simply repeats this process for more than one transmission. Therefore we expect the energy consumption to increase linearly as compute time increases. The portion of energy consumed that appears to be fixed across the sets of three configurations as we increase the number of angles compounded are likely external to our algorithm and can be attributed to the power dissipated as the power monitoring program is run as a parallel thread to the main beamforming program. As the number of volumes compounded increases, we would expect the image quality to increase as well. In this work, we quantify this tradeoff between power and performance.

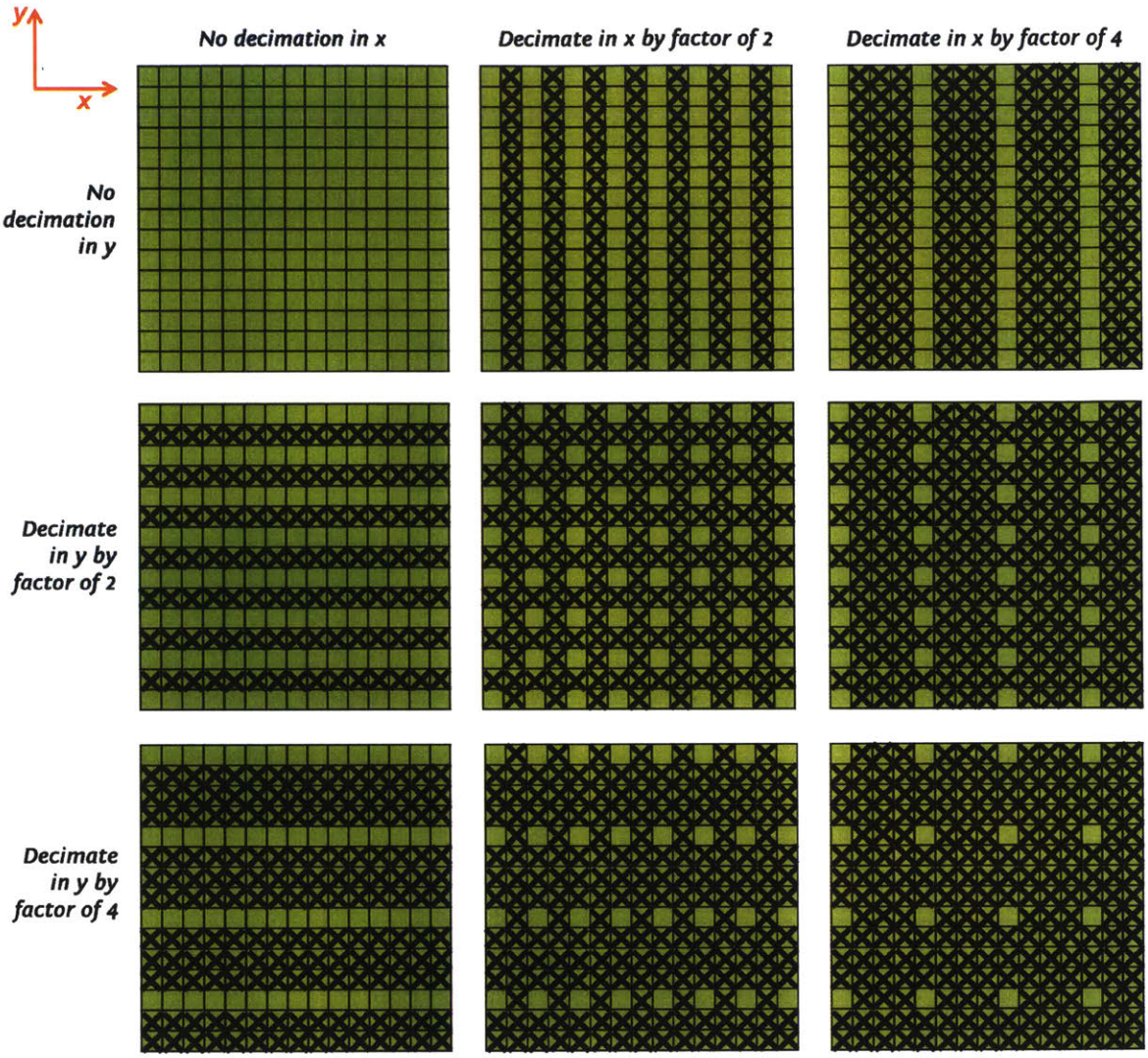


Figure 6-9: Visual representation of 2D array of 15 $\times$ 15 transducer elements, with the elements that are not in use marked in black.

6.2.2 2D Array Spatial Decimation

Just as with the 2D imaging case, we propose spatially decimating the array elements that we use to reconstruct the image. The motivation behind this is again to reduce the number of parallel front-end electronics required in the receive signal chain. Since the array is two-dimensional, we can vary the decimation factor in the azimuthal (or x) direction and the elevational (or y) directions. Figure 6-9 illustrates this scheme by indicating a subset of array elements as turned off, and Table 6.3 summarizes these 9 different configurations.

Table 6.3: List of possible configurations for varying decimation factors in azimuthal (x-) and elevational (y-) directions.

Decimation factor X	**Decimation factor Y**	**Effective array size**
No decimation	No decimation	15 × 15
No decimation	2	15 × 7
No decimation	4	15 × 4
2	No decimation	7 × 15
2	2	7 × 7
2	4	7 × 4
4	No decimation	4 × 15
4	2	4 × 7
4	4	4 × 4

When decimation is applied to the array, the algorithm simply skips those elements that occurs that are turned off (shown in Figure 6-9 as black cross marks). Since the overall number of voxels formed remains constant, and only the number of delayed waveforms being summed for each of these voxels is reduced, we would expect the energy reduction as the array is decimated in either the x or y direction to be much less drastic than that of reducing the number of volumes compounded (described in Section 6.2.1

6.3 Implementation

The beamformer design is implemented in C++ and compiled to be run on an NVIDIA Jetsen™ TX1 embedded platform to demonstrate 3-D beamforming capabilities. The power monitor communicates with an Arduino UNO board via an I2C interface, which executes a measurement program to log the real-time power consumption of the GPU, CPU, and the total power (including DRAM and peripheral controllers). The logged power data is then sent via UART back to the Jetson (which doubles as the host PC) for plotting. This power measurement setup is schematically represented in Figure 6-10, and the physical setup is photographed in Figure 6-11.

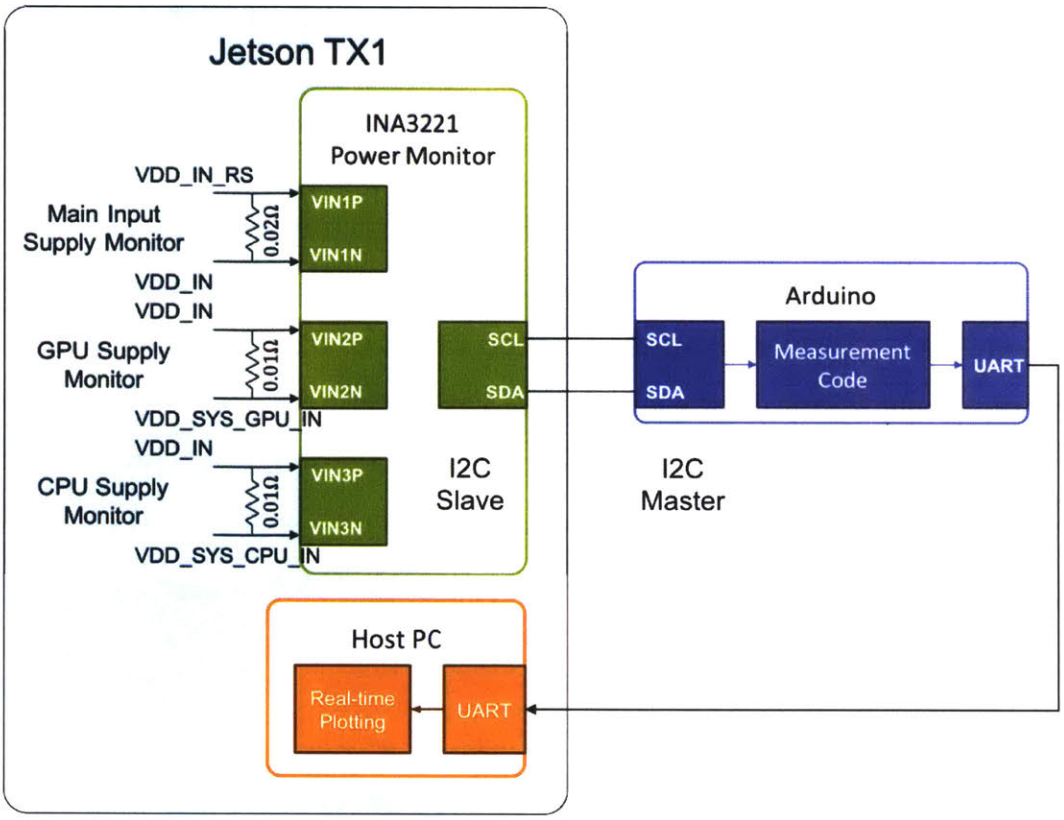


Figure 6-10: Schematic of the Jetson power monitor setup.

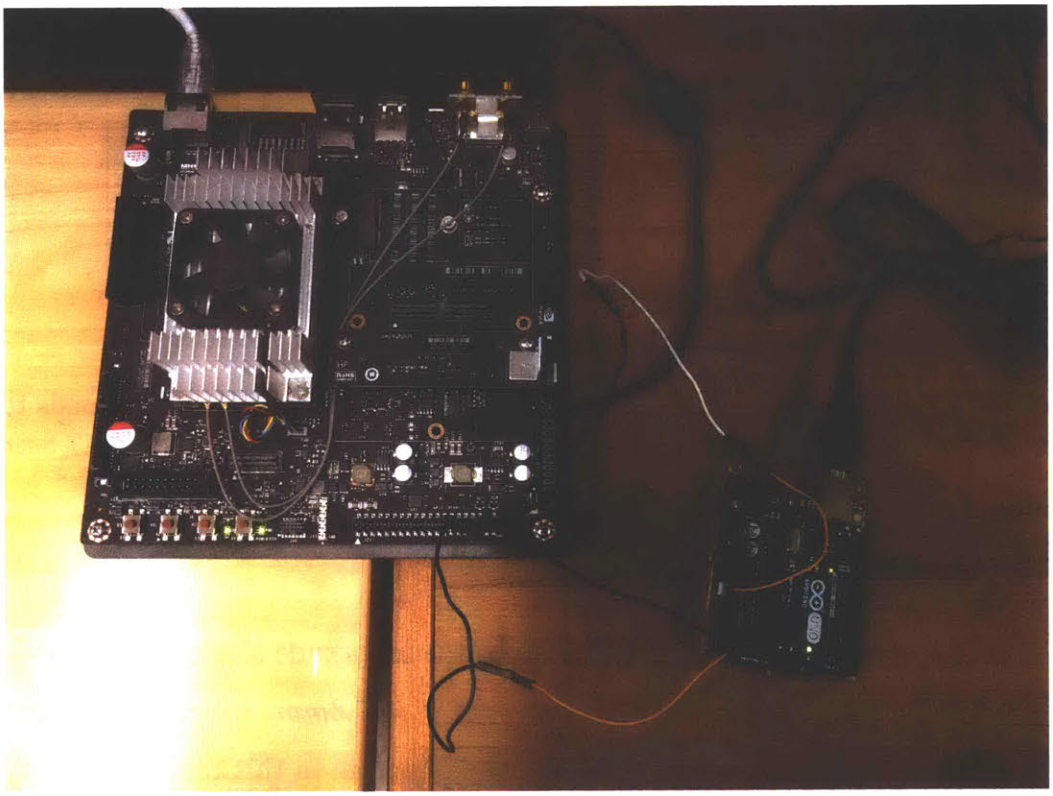


Figure 6-11: Measurement setup of Jetson embedded platform with Arduino Duo board for voltage and current monitoring.

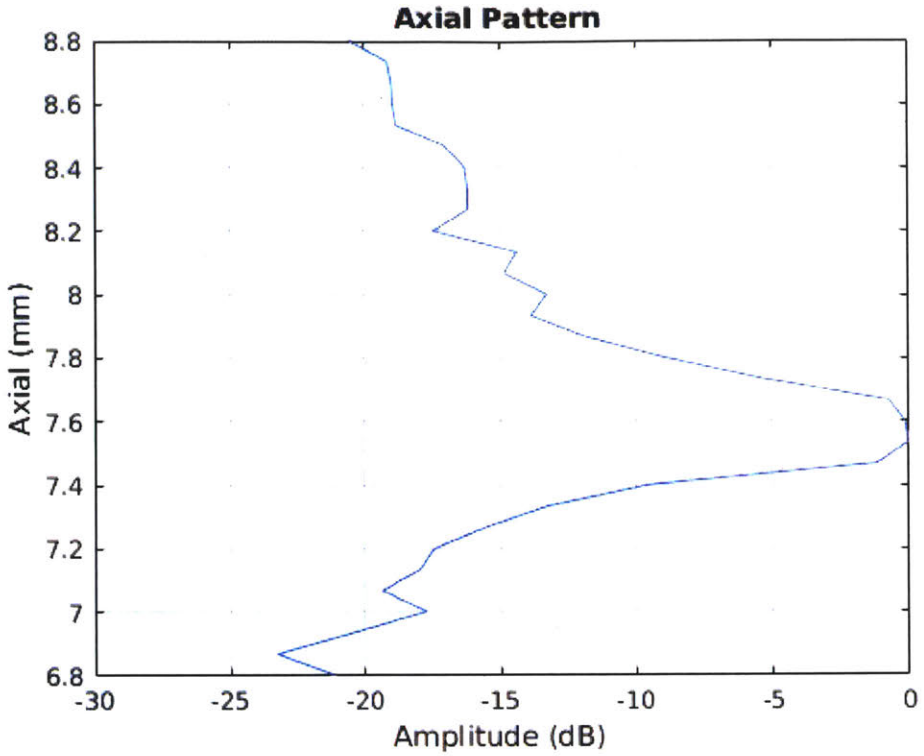


Figure 6-12: Axial pattern of ring phantom. The peak in amplitude corresponds to the depth at which the ring phantom is located (between $z = 7.4mm$ and $z = 7.8mm$).

6.4 Measurement Results

The axial pattern plotted in Figure 6-12 plots the amplitude along the z-axis. There is a clear maximum between $z = 7.4mm$ and $z = 7.8mm$, which can be used to automatically identify the axial location of the phantom in this case.

Figure 6-13 shows a cross-sectional 2D image of the reconstructed volume taken at $z = 7.533mm$. While the ring structure is clearly visible, there is also some blurring in the center of the ring that should be occupied by the oil.

To better visualize the relative magnitude between the desired image of the ring phantom and the unwanted bright spot at its center, we plot the lateral (or azimuthal) pattern at $z = 7.533mm$. We note that the unwanted bright spot is likely an artifact of the center edges of the two main lobes that correspond to the ring phantom adding together.

Finally, this same experiment is carried out for each of the cases specified in

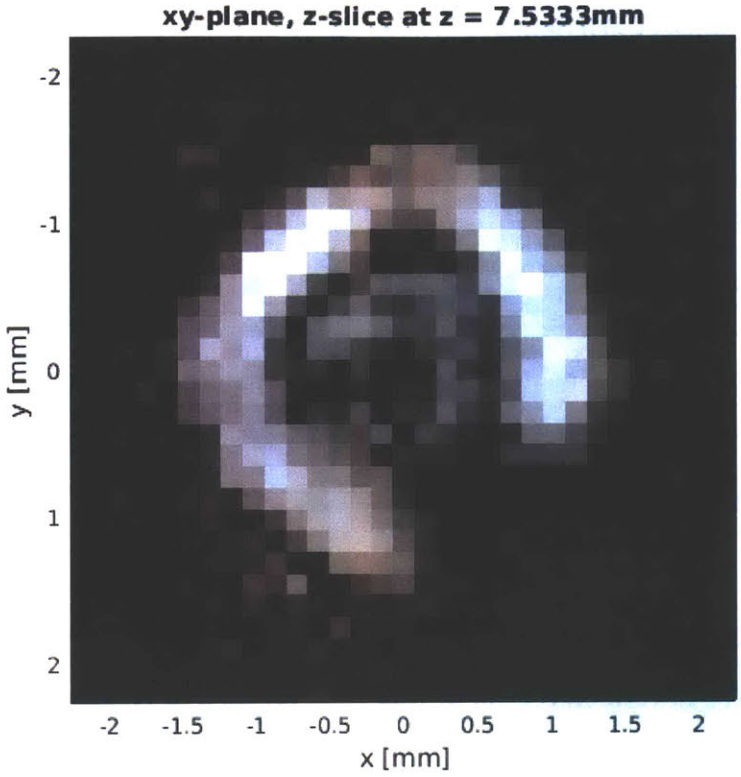


Figure 6-13: 2D slice at $z = 7.533mm$.

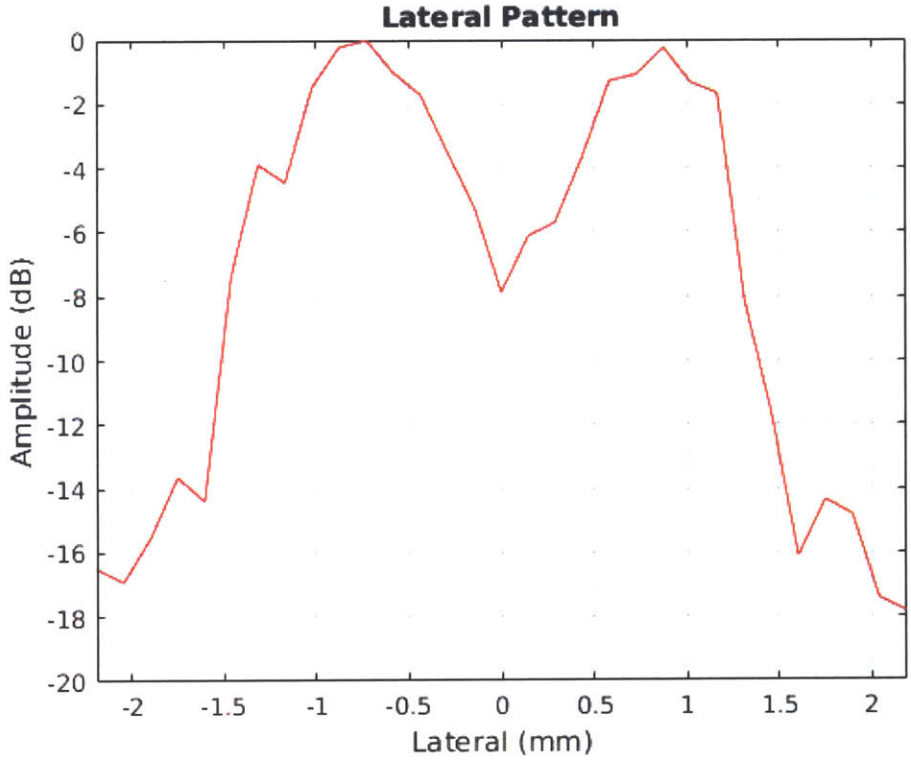


Figure 6-14: Lateral pattern of ring phantom at $z = 7.533mm$.

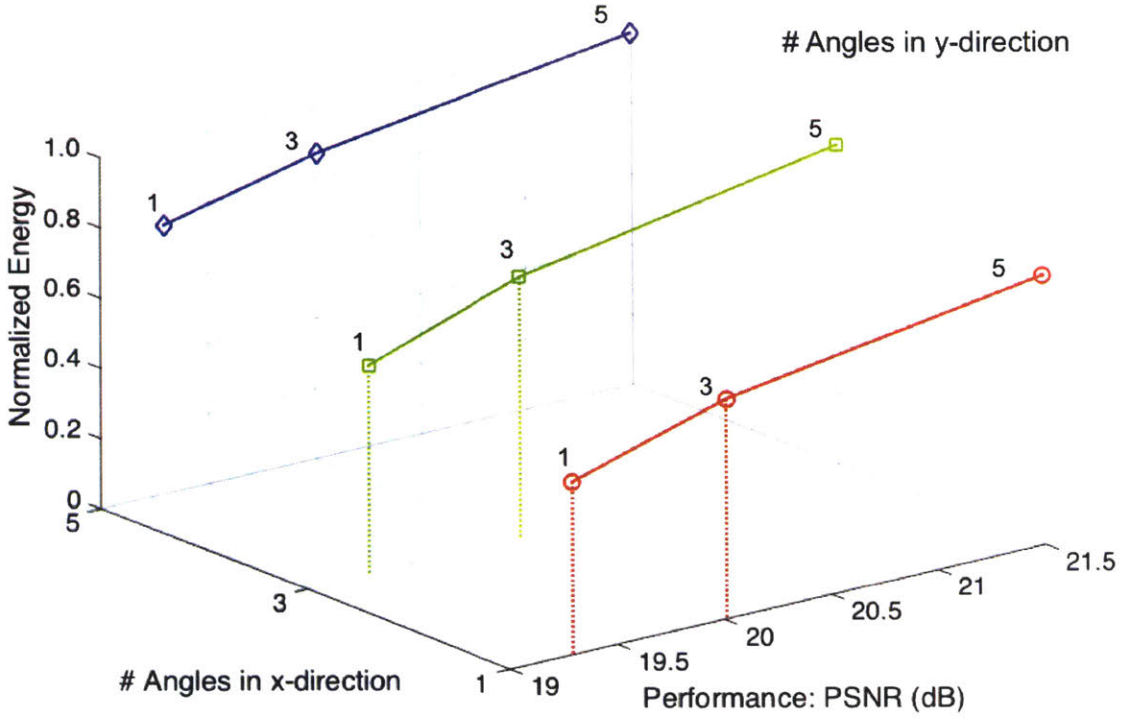


Figure 6-15: Tradeoff between normalized energy consumption (per volume reconstructed) and PSNR while varying number of volumes averaged in 3D plane wave continuous compounding.

Table 6.2, and the power versus time plots are included in Appendix A. The energy consumption is computed by integrating power consumption over the run time for one volume formed, and summarized against PSNR in a 3D plot in Figure 6-15.

Similarly, power measurements were made for all scenarios in Table 6.3, energy consumption per frame computed and plotted against PSNR in Figure 6-16.

6.5 Summary and Conclusions

This chapter presents the software-based analysis of an energy-scalable 3D ultrasound beamforming system. The tradeoffs between energy consumption and volume image quality are quantified and can be used to inform the design choices for a hardware implementation.

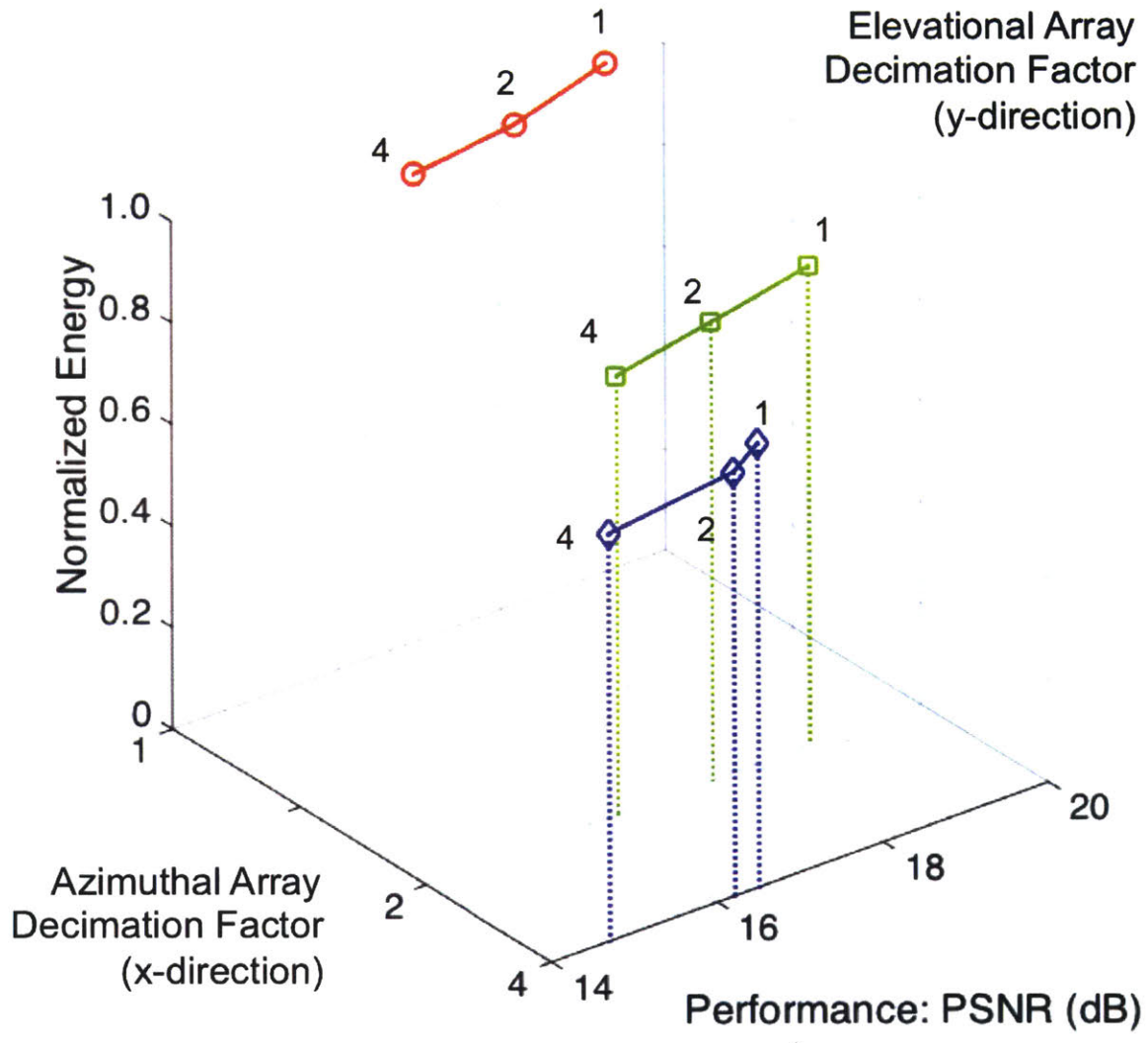


Figure 6-16: Tradeoff between normalized energy consumption (per volume reconstructed) and PSNR for various factors of spatial decimation in the x and y directions.

Chapter 7

Conclusions and Future Work

This thesis presented a number of key ideas that can be used to enable energy scalability in portable and wearable ultrasound imaging systems. Energy scalability is especially important for applications where diagnosis does not require high image quality. Two potential use cases that were identified are placenta previa and pleural effusion, where the general location and approximate size of a relatively large mass may be identifiable in a low-resolution image. The ability for the user to switch between multiple operation modes that correspond to different energy budgets is particularly useful in scenarios where access to mains power is limited or intermittent, and a battery must be used instead.

To understand the system-level optimizations and noise dependencies between the various stages of a receive-side ultrasound beamforming signal-chain, an energy model that includes the ADC, AFE, and digital beamformer was constructed. Based on the findings that system power is dominated by the analog electronics which are dedicated per transducer element, we developed energy scalable scheme that spatially decimates the array such that the number of such channels are reduced in low energy modes. An ASIC for a beamformer that supports three such decimation modes was designed, fabricated and tested. Finally, these ideas were extended to a 3D imaging system and software-based power measurements were carried out to quantify power-performance tradeoffs.

7.1 Summary of Contributions

In this section, we summarize the key ideas explored, implemented, and demonstrated in this thesis.

7.1.1 2D Receive Beamforming System-level Energy Modeling

An energy model for an ultrasound receive beamforming system (including AFE, ADC, and digital beamformer) was constructed to explore tradeoffs between system power and performance while accounting for interdependencies between these functional blocks. Given a desired frame rate and image quality (in terms of PSNR), the appropriate system topology and individual component specifications can be assigned based on these tradeoffs. Targeting power-limited applications where lower quality images may be desirable, we further proposed run-time selection from three resolution modes using the same transducer array, allowing additional energy savings beyond the design-time optimizations

7.1.2 2D Beamforming Test Chip Implementation

A low power energy-scalable ultrasound beamformer was implemented in 65-nm CMOS. This design targets a receive-side beamforming system which trades off image quality with power consumption. The reduced number of transducer channels used result in significant system-level power savings by turning off analog front end and analog-to-digital conversion electronics for those channels that are not in use. The test chip can generate ultrasound images at 30 fps with a measured power of $185\mu W$ at $0.52V$.

7.1.3 Power-Performance Tradeoff for 3D Beamforming

A reconfigurable 3D beamformer that averages a varying number of frames using plane wave coherent compounding and uses a varying number of transducer elements of receive beamforming was implemented. The software program was run on an embedded platform that supports voltage and current measurements. Using these measurements,

the tradeoff between energy consumption per frame and image quality in terms of PSNR is quantified. This platform can be used to further explore algorithmic tradeoffs prior to hardware design.

7.2 Future Work

There are a number of exciting new directions that can build on the work presented in this thesis:

- *Transmit and Receive Beamforming Co-design for Synthetic Aperture.* Our beamformer designs assumed plane wave transmitted waveforms (i.e. no transmit beamforming). An implementation that takes advantage of synthetic aperture schemes that can be generated using various combinations of transmit and receive apertures could be used to improve image quality without increasing transducer count.

- *Reconfigurable System Energy Model.* Our model assumed a specific AFE, ADC, and digital beamformer design. It could be instructive to implement a model that supports a wide portfolio of commercial, off-the-shelf AFEs and ADCs and a more parametrized beamformer to support more flexible energy modeling. Time gain compensation should also be included in the new model since the dynamic range requirement on the ADC can be relaxed with the normalization of signal amplitude over depth. This could inform future co-design of end-to-end systems.

- *Extending 2D Beamformer ASIC to Larger Apertures.* Our 2D beamformer test chip operates on groups of waveforms from eight consecutive transducer elements. This limits the subaperture sizes that can be used. Increasing the maximum number of channels while maintaining area and power efficiency using parallelism could improve image quality beyond what is possible with the current test chip. Automated mode selection may also be desirable in a system that is able to identify the image quality threshold above which it should operate.

- *Energy Scalable 3D Beamformer ASIC.* The findings and optimizations demonstrated using the 3D beamformer embedded platform can be used to inform an energy-efficient and reconfigurable ASIC for 3D beamforming and image reconstruction. An additional low-energy mode that can be added would be for an automatic measurement of width or volume to be provided in lieu of or alongside an image when measuring the amount of fluid in pleural effusion.

Appendix A

Power Measurements on Plane Wave Coherent Compounding

Using the Jetson platform, the power consumption data was acquired and plotted over the duration of forming one frame using the baseline algorithm (see Figure A-1). The CPU power measurements for various combinations of azimuthal (x) and elevational (y) angles in 3D plane wave coherent compounding are summarized as follows:

- For no averaging in the azimuthal direction (only one angle in x): Figure A-2 shows no averaging in the elevational direction (only one angle in y), Figure A-3 shows averaging three angles in the elevational direction, and Figure A-4 shows averaging five angles in the elevational direction.
- For averaging in three angles in the azimuthal direction: Figure A-5 shows no averaging in the elevational direction (only one angle in y), Figure A-6 shows averaging three angles in the elevational direction, and Figure A-7 shows averaging five angles in the elevational direction.
- For averaging in five angles in the azimuthal direction: Figure A-8 shows no averaging in the elevational direction (only one angle in y), Figure A-9 shows averaging three angles in the elevational direction, and Figure A-10 shows averaging five angles in the elevational direction.

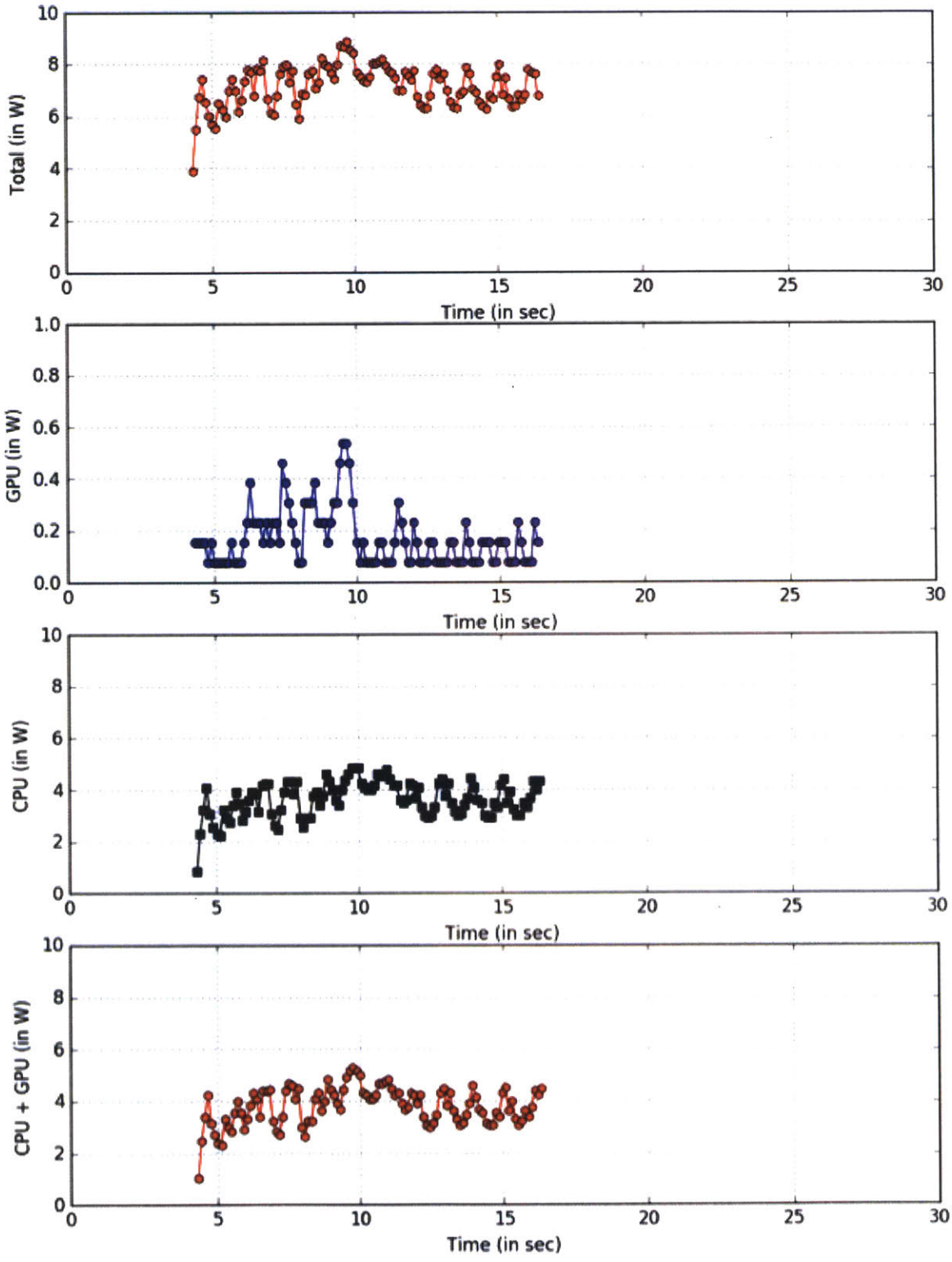


Figure A-1: Power consumption breakdown of Jetson system over time for the baseline algorithm (no averaging in both azimuthal and elevational directions.

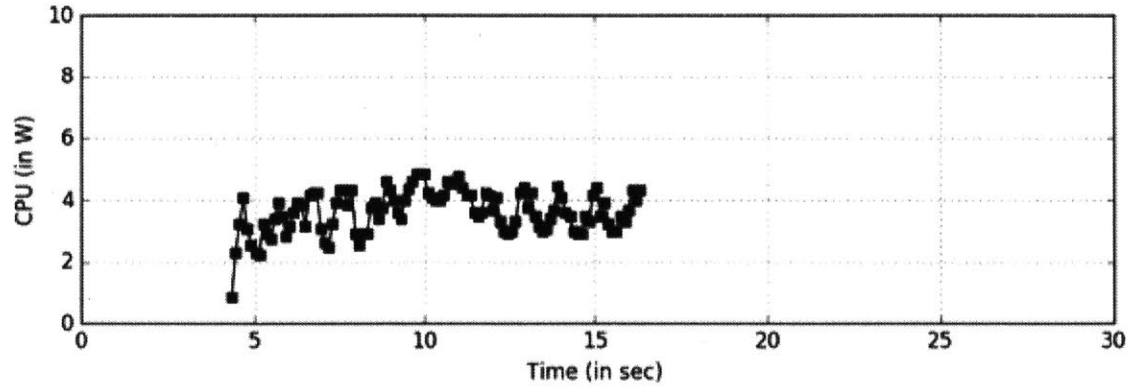


Figure A-2: Measured CPU power versus time for no averaging in both the azimuthal and elevational directions.

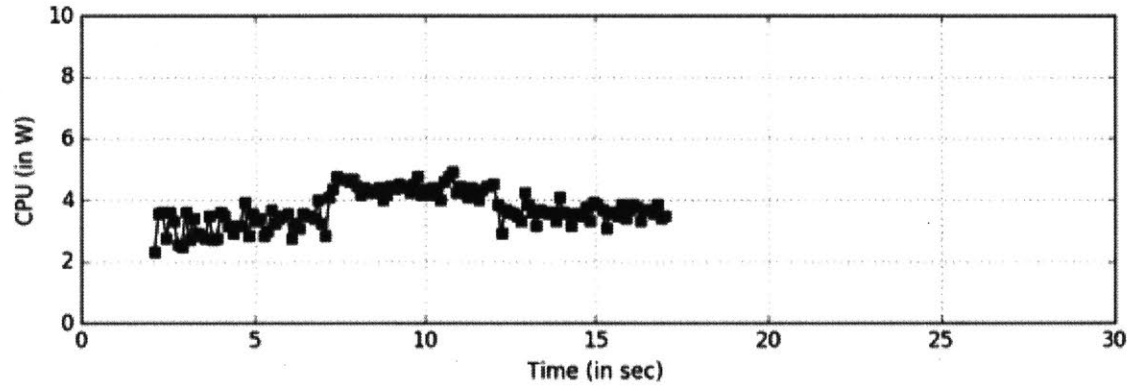


Figure A-3: Measured CPU power versus time for averaging three volumes (no averaging in x, three angles in y).

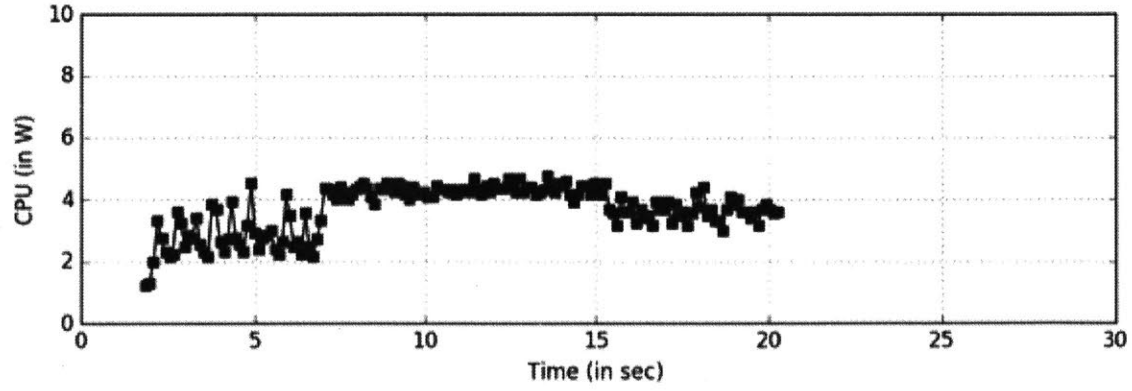


Figure A-4: Measured CPU power versus time for averaging five volumes (no averaging in x, five angles in y).

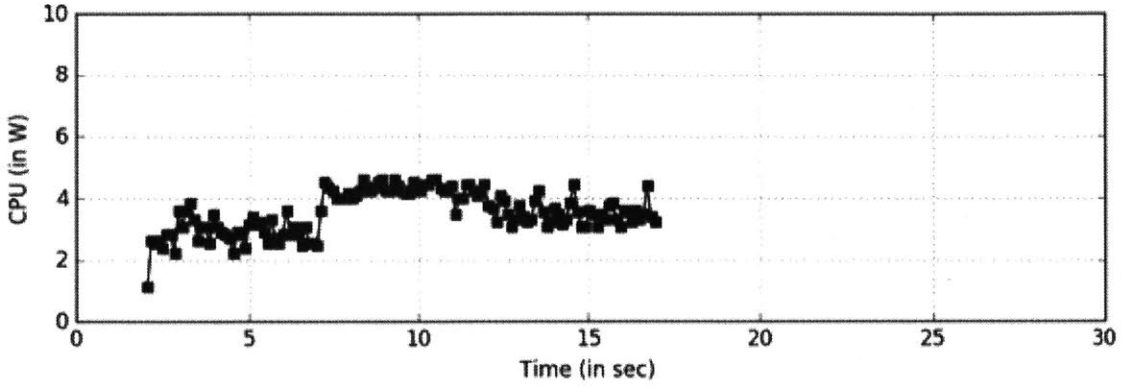


Figure A-5: Measured CPU power versus time for averaging three volumes (three angles in x, no averaging in y).

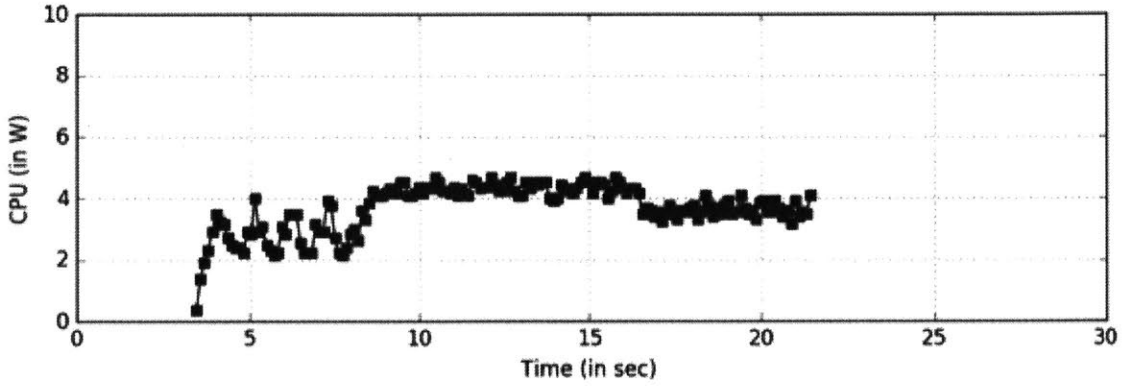


Figure A-6: Measured CPU power versus time for averaging five volumes (three angles in x, three angles in y).

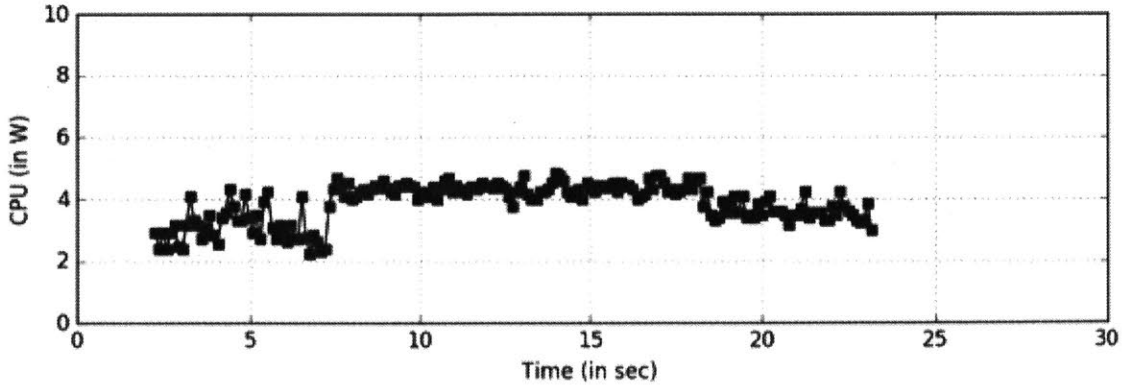


Figure A-7: Measured CPU power versus time for averaging seven volumes (three angles in x, five angles in y).

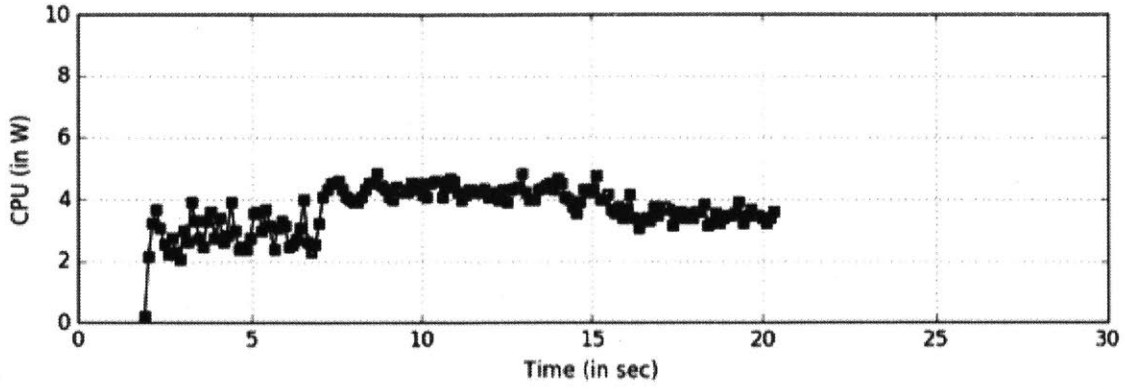


Figure A-8: Measured CPU power versus time for averaging five volumes (five angles in x, no averaging in y).

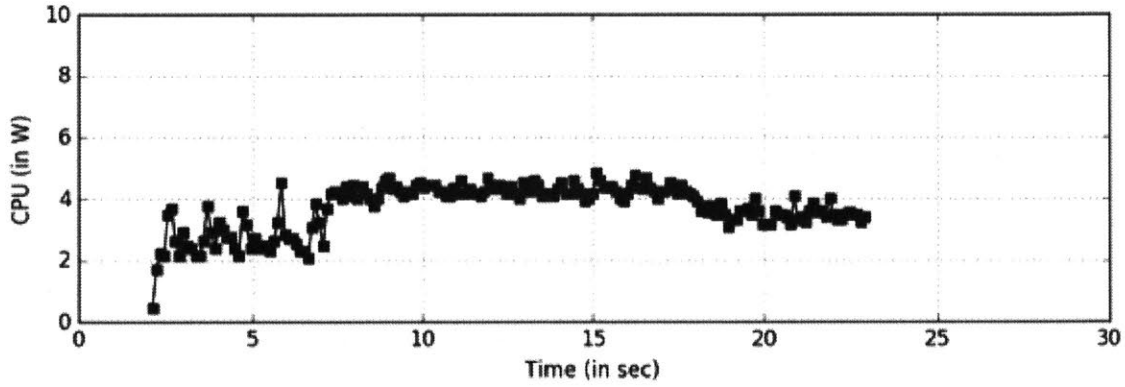


Figure A-9: Measured CPU power versus time for averaging seven volumes (five angles in x, three angles in y).

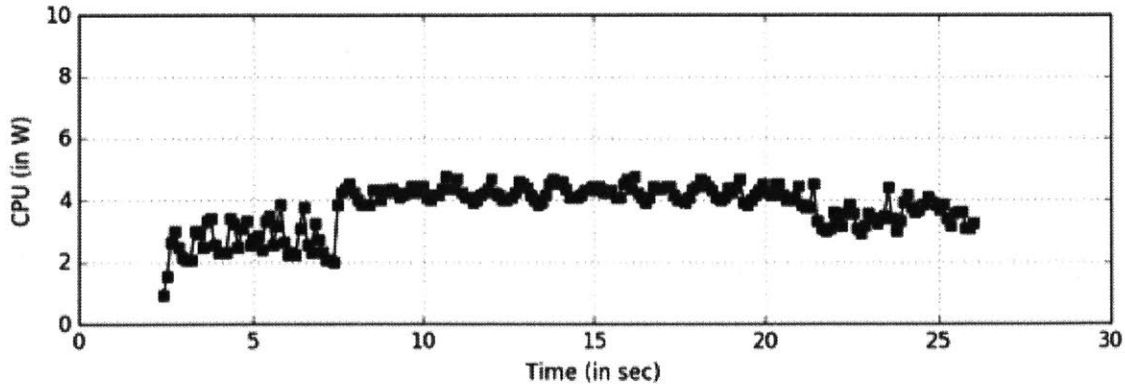


Figure A-10: Measured CPU power versus time for averaging nine volumes (five angles in x, five angles in y).

Bibliography

[1] General electric, vscan official website.

[2] clarius ultrasound anywhere - product.

[3] Lee Y. H. Chou M. M., Ho E. S. Prenatal diagnosis of placenta previa accreta by transabdominal color doppler ultrasound. *Ultrasound Obstet Gynecol*, 15(1):28 –35, jan 2000.

[4] C. P. Davis. Pleural effusion (fluid in the chest or on the lung), Aug. 9 2016.

[5] Zhihui Han, Hu Peng, and Putian Lu. 3d ultrasound imaging in frequency domain with 1d array transducer. *Ultrasonics*, 76:28 – 34, 2017.

[6] ATS Laboratories. Model 539 multipurpose phantom, 2016.

[7] C.M. Otto. *Textbook of Clinical Echocardiography.* MD Consult. Saunders, 2004.

[8] Kailiang Chen. *A Column-Row-Parallel ASIC Architecture for 3D Wearable / Portable Medical Ultrasonic Imaging.* PhD thesis, Massachusetts Institute of Technology, 2014.

[9] John M. Reid. Ultrasound: its applications in medicine and biology. *Ultrasound in Medicine and Biology*, 4(3):307–308, 1978.

[10] C. Prinz and J.U. Voigt. Diagnostic accuracy of a hand-held ultrasound scanner in routine patients referred for echocardiography. *Journal of the American Society of Echocardiography*, 2010.

[11] Philips, lumify official website.

[12] G.E. Moore. Cramming more components onto integrated circuits. *Proceedings of the IEEE*, 86(1):82 –85, jan 1998.

[13] Placenta previa: Symtoms, causes, and treatment, May 3 2017.

[14] Lung ultrasound - pleural effusion, 2011.

[15] R. Rithe, P. Raina, N. Ickes, S. V. Tenneti, and A. P. Chandrakasan. Reconfigurable processor for energy-scalable computational photography. In *2013 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pages 164–165, Feb 2013.

[16] P. Raina, M. Tikekar, and A. P. Chandrakasan. An energy-scalable accelerator for blind image deblurring. In *ESSCIRC Conference 2016: 42nd European Solid-State Circuits Conference*, pages 113–116, Sept 2016.

[17] L. Guo, M. Scott, and R. Amirtharajah. An energy scalable functional unit for sensor signal processing. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, volume 2, pages II–73–II–76, April 2007.

[18] M. Shoaib, K. H. Lee, N. K. Jha, and N. Verma. A 0.6-107 uw energy-scalable processor for directly analyzing compressively-sensed eeg. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 61(4):1105–1118, April 2014.

[19] B. Garg, N. K. Bharadwaj, and G. K. Sharma. Energy scalable approximate dct architecture trading quality via boundary error-resiliency. In *2014 27th IEEE International System-on-Chip Conference (SOCC)*, pages 306–311, Sept 2014.

[20] T. Ikari, H. Minamide, and H. Ito. Energy-scalable thz-wave parametric oscillator and its application to scanning-beam terahertz-wave reflection imaging. In *2007 Conference on Lasers and Electro-Optics (CLEO)*, pages 1–2, May 2007.

[21] K. Chen, B. Lam, C. G. Sodini, and A. P. Chandrakasan. System energy model for a digital ultrasound beamformer with image quality control. In *Ultrasonics Symposium (IUS), 2012 IEEE International*, pages 615–618, Oct 2012.

[22] B. Lam, M. Price, and A. P. Chandrakasan. An asic for energy-scalable, low-power digital ultrasound beamforming. In *2016 IEEE International Workshop on Signal Processing Systems (SiPS)*, pages 57–62, Oct 2016.

[23] D. Olendorf, C. Jeryan, and K. Boyden. *The Gale encyclopedia of medicine.* Gale Research (Detroit, MI), 1999.

[24] Richard S. C. Cobbold. *Foundations of Biomedical Ultrasound.* Oxford University Press, 2006.

[25] Gordon S. Kino. *Acoustic Waves: Devices, Imaging, and Analog Signal Processing.* Prentice Hall, 1987.

[26] F. I. Lizzi and D. J. Coleman. History of ophthalmic ultrasound. *J. Ultrasound Med.*, 23(10):1255–1266, 2004.

[27] D. H. Evans and W. N. McDicken. *Doppler Ultrasound (Second ed.).* John Wiley and Sons, New York, 2000.

[28] P. Satomura. Study of the flow patterns in peripheral arteries by ultrasonics. *J. Acoust. Soc. Japan*, 15:151–158, 1959.

[29] D.W. Baker. Pulsed ultrasonic doppler blood-flow sensing. *Sonics and Ultrasonics, IEEE Transactions on*, 17(3):170 – 184, jul 1970.

[30] C. Kasai, K. Namekawa, A. Koyano, and R. Omoto. Real-time two-dimensional blood flow imaging using an autocorrelation technique. *Sonics and Ultrasonics, IEEE Transactions on*, 32(3):458 – 464, may 1985.

[31] G. Elko. Principles of acoustic devices. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(3):766–767, June 1985.

[32] ME Anderson, MS McKeag, and GE Trahey. The impact of sound speed errors on medical ultrasound imaging. *The Journal of the Acoustical Society of America*, 107:3540, 2000.

[33] U.S. Food & Drug Administration. Radiation-emitting products and procedures: Ultrasound imaging, 2016.

[34] K. Chen, H. S. Lee, A. P. Chandrakasan, and C. G. Sodini. Ultrasonic imaging transceiver design for cmut: A three-level 30-vpp pulse-shaping pulser with improved efficiency and a noise-optimized receiver. *IEEE Journal of Solid-State Circuits*, 48(11):2734–2745, Nov 2013.

[35] Iec standard 60601-2-37 : 2001/a1 : 2005: Medical electrical equipment – particular requirements for the safety of ultrasonic medical diagnostic and monitoring equipment. Technical report, International Electrotechnical Commission, Geneva, 2005.

[36] J. Calvert, F. Duck, S Clift, and H. Azaime. Surface heating by transvaginal transducers. *Ultrasound Obstet Gynecol*, 29:427–432, 2007.

[37] G.F.S. The physical and biological effects of high-frequency sound-waves of great intensity. *Journal of the Franklin Institute*, 205(1):151 – 153, 1928.

[38] C. M. Langton and C. F. Njeh. The measurement of broadband ultrasonic attenuation in cancellous bone-a review of the science and technology. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 55(7):1546–1554, July 2008.

[39] K. R. Waters, B. K. Hoffmeister, and J. A. Javarone. Application of the kramers-kronig relations to measurements of attenuation and dispersion in cancellous bone. In *IEEE Ultrasonics Symposium, 2004*, volume 1, pages 561–564 Vol.1, Aug 2004.

[40] R. C. Chivers and C. R. Hill. Ultrasonic attenuation in human tissue. *Ultrasound in Medicine and Biology*, 2(1):25–29, 10 1975.

[41] Bradley E. Treeby, Edward Z. Zhang, Alison S. Thomas, and Ben T. Cox. Measurement of the ultrasound attenuation and dispersion in whole human blood and its components from 0-70 mhz. *Ultrasound in Medicine and Biology*, 37(2):289–300, 2 2011.

[42] J. C. Bamber. *Attenuation and Absorption*, pages 93–166. John Wiley & Sons, Ltd, 2005.

[43] A. B. Temsamani, S. Vandenplas, and L. van Biesen. Ultrasonic waves propagation in fluid-saturated materials. In *Proceedings of the 2000 International Symposium on Underwater Technology (Cat. No.00EX418)*, pages 361–366, 2000.

[44] A. P. Berkhoff, J. M. Thijssen, and R. J. F. Homan. Simulation of ultrasonic imaging with linear arrays in causal absorptive media. *Ultrasound in Medicine and Biology*, 22(2):245–259, 1996.

[45] P. N. T. Wells. Absorption and dispersion of ultrasound in biological tissue. *Ultrasound in Medicine and Biology*, 1(4):369–376, 3 1975.

[46] Kevin J. Parker. Ultrasonic attenuation and absorption in liver tissue. *Ultrasound in Medicine and Biology*, 9(4):363–369, 7 1983.

[47] Pascal Laugier and Guillaume Haïat. *Introduction to the Physics of Ultrasound*, pages 29–45. Springer Netherlands, Dordrecht, 2011.

[48] Shyh-Hau Wang and K. K. Shung. An approach for measuring ultrasonic backscattering from biological tissues with focused transducers. *IEEE Transactions on Biomedical Engineering*, 44(7):549–554, July 1997.

[49] JeffreyC Bamber. chapter Ultrasonic Properties of Tissues. Series in Medical Physics and Biomedical Engineering. Taylor & Francis, Jan 1998. 0.

[50] B. D. Van Veen and K. M. Buckley. Beamforming: a versatile approach to spatial filtering. *ASSP Magazine, IEEE*, 5(2):4–24, April 1988.

[51] A. Macovski. Ultrasonic imaging using arrays. *Proceedings of the IEEE*, 67(4):484–495, April 1979.

[52] O. T. Von Ramm and S. W. Smith. Beam steering with linear arrays. *IEEE Transactions on Biomedical Engineering*, BME-30(8):438–452, Aug 1983.

[53] M. Tang, F. Luo, and D. Liu. Automatic time gain compensation in ultrasound imaging system. In *2009 3rd International Conference on Bioinformatics and Biomedical Engineering*, pages 1–4, June 2009.

[54] A. Vasjanov and V. Barzdenas. Design of a time-gain-compensation amplifier for ultrasonic echo signal processing. In *2015 Open Conference of Electrical, Electronic and Information Sciences (eStream)*, pages 1–6, April 2015.

[55] J. Yao, Z. Yu, M. A. P. Pertijs, G. C. M. Meijer, C. T. Lancée, J. G. Bosch, and N. de Jong. Design of a low power time-gain-compensation amplifier for a 2d piezoelectric ultrasound transducer. In *2010 IEEE International Ultrasonics Symposium*, pages 841–844, Oct 2010.

[56] D. Shi, Z. Fan, H. Yin, and D. C. Liu. Fast gpu-based automatic time gain compensation for ultrasound imaging. In *2010 4th International Conference on Bioinformatics and Biomedical Engineering*, pages 1–4, June 2010.

[57] J. M. Mari, K. Hibbs, E. Stride, R. J. Eckersley, and M. X. Tang. An approximate nonlinear model for time gain compensation of amplitude modulated images of ultrasound contrast agent perfusion. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 57(4):818–829, April 2010.

[58] Zili Yu and G. C. M. Meijer. A programmable time-gain-compensation (tgc) amplifier for medical ultrasonic echo signal processing. In *2008 9th International Conference on Solid-State and Integrated-Circuit Technology*, pages 1725–1728, Oct 2008.

[59] S. R. Snare and H. Torp. Estimating frequency dependent attenuation to improve autmatic time gain compensation in b-mode imaging. In *2008 IEEE Ultrasonics Symposium*, pages 1322–1325, Nov 2008.

[60] Rahul Prakash Sanjay Pithadia. Time gain control (compensation) in ultrasound applications. Technical report, Texas Instruments Incorporated, Dallas, TX, 12 2016.

[61] K.E. Thomenius. Evolution of ultrasound beamformers. In *Ultrasonics Symposium, 1996. Proceedings., 1996 IEEE*, volume 2, pages 1615 –1622 vol.2, nov 1996.

[62] S. A. Mohamed, E. D. Mohamed, M. F. Elshikh, and M. A. Hassan. Design of digital apodization technique for medical ultrasound imaging. In *2013 INTERNATIONAL CONFERENCE ON COMPUTING, ELECTRICAL AND ELECTRONIC ENGINEERING (ICCEEE)*, pages 541–544, Aug 2013.

[63] Lee Y. H. Chou M. M., Ho E. S. *American Journal of Biomedical Science and Engineering*, 1(1):1–8, jan 2015.

[64] Don H. Johnson and Dan E. Dudgeon. *Array Signal Processing: Concepts and Techniques*. Simon & Schuster, 1992.

[65] J. Grythe. Array gain and reduction of self-noise. Technical report, Norsonic AS, Oslo, Norway.

[66] Jie Liu, Kang-Sik Kim, and Michael F. Insana. Snr comparisons of beamforming strategies. *IEEE Trans Ultrason Ferroelectr Freq Control*, 54(5):1010–1017, May 2007. 17523565[pmid].

[67] Sewoong Ahn, Jeeun Kang, Pilsu Kim, Gunho Lee, Eunji Jeong, Woojin Jung, Minsuk Park, and Tai-Kyong Song. Smartphone-based portable ultrasound imaging system: Prototype implementation and evaluation. In *Ultrasonics Symposium (IUS), 2015 IEEE International*, pages 1–4, Oct 2015.

[68] V. S. Gierenz, R. Schwann, and T. G. Noll. A low power digital beamformer for handheld ultrasound systems. In *Solid-State Circuits Conference, 2001. ESSCIRC 2001. Proceedings of the 27th European*, pages 261–264, Sept 2001.

[69] P. Song, K.-T. Tiew, Y. Lam, and K. L. Mong. A CMOS 3.4 mW 200 MHz continuous-time delta-sigma modulator with 61.5 dB dynamic range and 5 MHz bandwidth for ultrasound application. In *Circuits and Systems, 2007. MWSCAS 2007. 50th Midwest Symposium on*, pages 152–155, Aug 2007.

[70] S. Sharma and T. Ytterdal. In-Probe Ultrasound Beamformer Utilizing Switched-Current Analog RAM. *Circuits and Systems II: Express Briefs, IEEE Transactions on*, 62(6):517–521, June 2015.

[71] G. D. Kim, C. Yoon, S. B. Kye, Y. Lee, J. Kang, Y. Yoo, and T. k. Song. A single FPGA-based portable ultrasound imaging system for point-of-care applications. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 59(7):1386–1394, July 2012.

[72] Qilong Liu, Chao Chen, Zu yao Chang, C. Prins, and M. A. P. Pertijs. A mixed-signal multiplexing system for cable-count reduction in ultrasound probes. In *Ultrasonics Symposium (IUS), 2015 IEEE International*, pages 1–4, Oct 2015.

[73] M. Pertijs, C. Chen, S. Raghunathan, Z. Yu, M. ShabaniMotlagh, Z. Chen, Z. y. Chang, E. Noothout, S. Blaak, J. Ponte, C. Prins, H. Bosch, M. Verweij, and N. de Jong. Low-power receive electronics for a miniature real-time 3D ultrasound probe. In *Advances in Sensors and Interfaces (IWASI), 2015 6th IEEE International Workshop on*, pages 235–238, June 2015.

[74] J. Zhou, S. Hoyos, and B. M. Sadler. Asynchronous compressed beamformer for portable diagnostic ultrasound systems. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 61(11):1791–1801, November 2014.

[75] J. Y. Um, Y. J. Kim, S. E. Cho, M. K. Chae, J. Song, B. Kim, S. Lee, J. Bang, Y. Kim, K. Cho, B. Kim, J. Y. Sim, and H. J. Park. An Analog-Digital Hybrid RX Beamformer Chip With Non-Uniform Sampling for Ultrasound Medical Imaging With 2D CMUT Array. *IEEE Transactions on Biomedical Circuits and Systems*, 8(6):799–809, Dec 2014.

[76] K. Chen, A. P. Chandrakasan, and C. G. Sodini. Ultrasonic imaging front-end design for cmut: A 3-level 30vpp pulse-shaping pulser with improved efficiency and a noise-optimized receiver. In *2012 IEEE Asian Solid State Circuits Conference (A-SSCC)*, pages 173–176, Nov 2012.

[77] R. Behzad. Design of analog cmos integrated circuits. *McGraw-Hill*, 2001.

[78] S. Lee, A.P. Chandrakasan, and H.S. Lee. A 12 b 5-to-50 ms/s 0.5-to-1 v voltage scalable zero-crossing based pipelined adc. *Solid-State Circuits, IEEE Journal of*, 47(7):1603–1614, 2012.

[79] L14-5/38 linear, Jul. 31 2014.

[80] Q. Huynh-Thu and M. Ghanbari. Scope of validity of psnr in image/video quality assessment. *Electronics letters*, 44(13):800–801, 2008.

[81] A. Hore and D. Ziou. Image quality metrics: Psnr vs. ssim. In *2010 20th International Conference on Pattern Recognition*, pages 2366–2369, Aug 2010.

[82] J.T. Bushberg. *The essential physics of medical imaging.* Williams & Wilkins, 2002.

[83] H.A.D. PETTERSSON. The encyclopaedia of medical imaging: V. 1. physics, techniques and procedures. 1998.

[84] R Dreslinski, Michael Wieckowski, D Sylvester Blaauw, and T Mudge. Near threshold computing: Overcoming performance degradation from aggressive voltage scaling. In *Proc. Workshop Energy-Efficient Design*, pages 44–49, 2009.

[85] Jeremy Bercoff. Ultrafast ultrasound imaging. In Igor V. Minin and Oleg V. Minin, editors, *Ultrasound Imaging - Medical Applications*, chapter 01. InTech, Rijeka, 2011.

[86] O. Couture, M. Fink, and M. Tanter. Ultrasound contrast plane wave imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 59(12), Dec 2012.

[87] G. Montaldo, M. Tanter, J. Bercoff, N. Benech, and M. Fink. Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 56(3):489–506, March 2009.

[88] J. Bercoff, M. Tanter, and M. Fink. Supersonic shear imaging: a new technique for soft tissue elasticity mapping. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 51(4):396–409, April 2004.