ACCURATE ULTRASOUND IMAGING IN HIGH DYNAMIC RANGE SCENARIOS

By

Siegfried Gustav Schlunk

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Biomedical Engineering

Graduation May 12, 2023

Nashville, Tennessee

Approved:

Brett Byram, Ph.D.

Charles Caskey, Ph.D.

Ryan Hsi, M.D.

Bennett Landman, Ph.D.

Michael Miga, Ph.D.

To my family, Herwig, Christine, Sylvia, and Jarl, and all our dogs along the way, Copper, Snoopy, Leela, Kaiser, and Zuko.

ACKNOWLEDGMENTS

Thank you to my advisor, Professor Brett Byram. His flexibility, patience, and understanding helped ensure everything went smoothly, and his knowledge and resourcefulness was invaluable in learning new things and finding new opportunities in our field.

Thank you to the rest of my committee for their time and assistance. To Professor Charles Caskey, whose entire lab has been a constant friend and resource in the exploration of novel ultrasound research, to Doctor Ryan Hsi, who has been an incredible help in understanding the clinical relevance of our work, to Professor Michael Miga, whose lab has had tons of insight and whose seminars have widened my understanding of the general biomedical and medical world, and finally to Professor Bennett Landman, who has provided valuable advice and opportunities in the past.

Thank you to the BEAM lab, for all the questions answered and discussions had. Katie Ozgun was with me from the beginning and was a great friend inside and outside the lab. Kazuyuki Dei was a mentor and guide for all things in the lab. And of course all the others who have been a great help developing ideas and keeping the lab a fun and inviting space.

Thank you to the squad, who kept me up late and helped me maintain my sanity. To my friends from McTyeire who have been there through college and beyond.

Finally, thank you to the National Institutes of Health for funding my research through the grants R01EB020040, S10OD016216-01, R01HL156034, and S10OD023680-01, and additionally to the NSF for award IIS-1750994. Thank you to the Vanderbilt University ACCRE computing cluster for enabling much of my research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Background and Significance

Ultrasound Imaging is a common medical imaging modality due to its low cost, non-invasive, and non-destructive nature (Szabo, 2014; Cobbold, 2007). By emitting sound waves at safe intensities, it can reveal useful medical information and is often used in conjunction with other modalities and treatments since it can be used in real-time. Due to ultrasound's long, excellent safety record and low cost, it is often used for preliminary screening, even in cases where the expected image quality is poor. In particular, ultrasound can struggle in high dynamic range scenarios, where a target and the surrounding reference material may have wildly different intensities, causing artifacts and general image quality concerns.

1.1 Clinical Significance

Judging the quality of an ultrasound beamformer can be subjective, which is why qualitative analysis is often left to experts (clinicians, sonographers, etc.). However, image metrics that quantify image quality have become the standard for evaluation, as they are much easier and convenient to apply. For example, contrast ratio and contrast-to-noise ratio have become the default metrics for evaluating modern beamformers, as they provide estimated measures of the visibility and clarity of the imaging target. As a result, many beamformers are designed explicitly with these metrics in mind, which can have unintended consequences.

Consider the case of imaging a hypoechoic, renal cyst. Renal cysts can be diagnosed with ultrasound, and are split into two categories: simple and complex. Simple cysts are benign, fluid-filled and tend to have a homogeneous appearance to them. On the other hand, complex cysts indicate disease and are heterogeneous in nature (Eknoyan, 2009). Classifying a complex cyst is already considered a difficult task (Warren and McFarlane, 2005; Siegel et al., 1997), but now consider a beamformer that is designed explicitly to enhance the visibility of hypoechoic cysts. In these cases, cysts with intensities lower than -40 dB often look identically, both visually and in terms of image quality metrics, and in cases where there is significant noise or clutter in the image lowering general image quality, this threshold could be as high as -20 dB (Dei et al., 2017; Schlunk et al., 2021). In this case of the renal cyst, this could result in simple and complex cysts looking indistinguishable, further complicating diagnosis.

Alternatively, consider an example of the opposite high dynamic range scenario, the diagnosis of kidney stone disease, where kidney stones are hyperechoic and much brighter than the surrounding tissue. Not only is the detection of kidney stones already a difficult task for ultrasound, but ultrasound is well-known for over-estimating the size of stones, which is necessary for determining intervention. Many clinicians consider 5mm

to be the threshold for when a kidney stone is large enough to require surgical intervention, but ultrasound is generally known to overestimate kidney stone sizes by 2-3 mm (Ulusan et al., 2007; Unal et al., 2003; Ray et al., 2010; Ganesan et al., 2017; Fowler et al., 2002; Dunmire et al., 2015; Sternberg et al., 2016), which would result in more "intervention" diagnoses than are actually necessary. In clinical practice, this results in follow-up computed tomography (CT) imaging, which is the gold standard of sizing stones (Smith and Varanelli, 2000; Smith et al., 1996), but this nonetheless places additional burden on medical and human resources. One study noted that in the absence of CT imaging, 22% of patients diagnosed with ultrasound would have been diagnosed incorrectly (Ganesan et al., 2017).

These situations suggest that there is an additional concern with ultrasound imaging more so than just improving image quality metrics. Specifically, there are imaging cases where imaging *accuracy* is perhaps more important than simply improving contrast or contrast-to-noise ratio. Accurately reporting the contrast ratio and sizing of a target are both critical to certain medical examinations in high dynamic range applications.

1.2 Fundamentals of Ultrasound Imaging

In this section we will introduce relevant fundamental concepts for ultrasound image acquisition and processing.

1.2.1 Linear Acoustic Wave Propagation

Mathematically, acoustic wave propagation can be described in one dimension by the following wave equation,

$$\frac{\partial^2 p}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0, \tag{1.1}$$

where p is pressure, x is the position, c is the speed of sound, and t is time. A classic solution is then

$$p = p_0 \sin(\omega t + kx), \tag{1.2}$$

where p_0 is the amplitude of the wave, ω is the angular frequency, and k is the wave number. Richard Feynman provides a derivation of the wave equation using the one-dimensional continuity equation (conservation of mass), force equation (conservation of momentum), and equation of state (ideal gas law) (Feynman, 2011). He additionally shows that equation 1.1 can be expanded into three dimensions as

$$\nabla^2 p - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0, \tag{1.3}$$

where ∇^2 is the Laplace operator, which simply expands the partial derivatives. There are many alternative forms of the wave equation depending on the situation, such as substituting the sound speed c for compressibility κ with $c^2 = 1/(\rho\kappa)$ (where ρ is the density of the medium) (Anderson and Trahey, 2006).

As the acoustic wave propagates through a medium, it will interact with the inhomogeneities it encounters. These inhomogeneities are caused by the differences in the density or compressibility, which result in acoustic impedance differences between the two mediums. We call scattering the interaction with singular particle inhomogeneities that are smaller than the wavelength, and reflecting as the interaction with larger particles or objects such as organs. In both cases, upon interacting with the boundary, some of the acoustic wave is reflected, which is the key to ultrasound image formation. We generally assume that acoustic wave propagation is a linear process (as we have so far), which enables us to approximate larger surfaces as a collection of scatterers, and calculate the reflections individually and then sum the collection. This provides the foundation for ultrasound simulation programs such as Field II (Jensen and Svendsen, 1992; Jensen, 1996), and additionally allows us to create complex, model-based beamformers (Byram and Jakovljevic, 2014; Byram et al., 2015a; Dei and Byram, 2017), which we discuss in detail in a later section.

Modeling ultrasound acoustic waves as linear plane waves allows us to calculate waves independently and then combine them via the superposition principle. This comes to the foreground with the Huygens-Fresnel principle, alluded to in the prior paragraph. It states that a given wavefront can be decomposed as a collection of point sources, each behaving as the origin of a spherical wave (Anderson and Trahey, 2006). This then leads to the Fraunhofer approximation. Originating in optics, it states that the far-field pressure amplitude pattern of an acoustic wave generated from an aperture is estimated by the 2-D Fourier transform of that transducer aperture amplitude (Anderson and Trahey, 2006). In most cases, the aperture can be modeled as a simple rectangular function, which results in a far-field approximation that is a sinc function. In the case of transmitting a plane wave from the aperture, the Fraunhofer approximation applies for distances greater than $z >> k(e_x^2 + e_y^2)$, for (e_x, e_y) the position of the aperture element farthest from the center. However, conventional ultrasound often makes use of focused waves (further explained in the next section), in which case the approximation only applies at the focus depth.

1.2.2 Conventional Beamforming and Image Formation

Fig. 1.1 shows a simple example of conventional beamforming using a linear transducer array. The transducer both creates the mechanical vibrations (acoustic wave) from a set of input electrical stimuli and also reconverts the received acoustic echoes back into electrical signals for analog or digital processing. In the example depicted, the linear transducer will excite specific subsets of the aperture elements to create a focused wavefront that converges on the focus in the imaging space. By adjusting the delays produced, this focus can

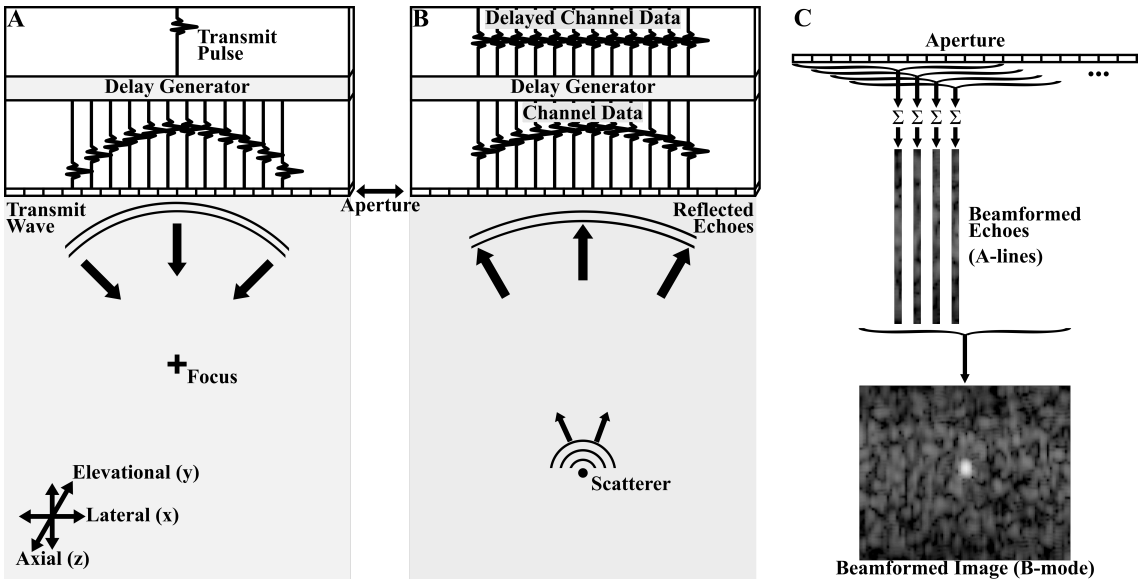

Figure 1.1: Conventional beamforming example from transmit to final beamformed image. (A) shows a transmit pulse being converted to a set of delayed pulses for each element in the transducer aperture. These delays are used to create a transmitted wavefront with a specific shape, in this case a spherical wavefront that will converge at the focus in the imaging space. (B) shows a scatterer in the imaging space creating a reflected echo that travels back to the aperture, which creates the channel data. The transducer will then apply delays (if desired) to create the delayed channel data. (C) in this example, each subset of elements in the aperture create an individual wave and return their own delayed channel data. This delayed channel data is then summed across the channels to create a single vertical line (A-line) in the final image. These individual A-lines then come together to form the final beamformed image. This process is referred to as delay-and-sum (DAS) and represents the most conventional beamformer all other methods are compared against.

be shifted axially closer or farther away from the transducer as needed. Additionally, by changing which set of elements are activated or adjusting the delays, the focus can be shifted or steered laterally. In this linear transducer example, by changing the activated elements, the focus can be moved from left to right across the imaging space, and the resulting data acquired will correspond to the axial band at the focus's lateral location. For conventional beamforming we generally do not consider the elevational dimension.

When the transmitted acoustic wave encounters a scatterer or boundary in the imaging space, a reflected echo will be created that travels back to the transducer and is converted back into electrical signals, which is the raw channel data. This raw channel data is then delayed, either by the transducer or an external beamforming process, and these delayed signals are then summed together to create the beamformed echo corresponding to the lateral location of the focus for that transmit. These beamformed echoes will then be stitched together to create the full beamformed image, which is then enveloped and compressed to produce the final image, referred to as a brightness mode, or B-mode, image. This general beamforming process is referred to as delay-and-sum (DAS) beamforming and is the standard against which all other beamformers are compared.

The acoustic waves (generally) follow linear behavior which allows us to calculate these delays (both transmit and receive) using simple geometry. Consider a specific transducer element e positioned at e_x, e_z, e_y. To simplify, we assume that the transducer elements are positioned at $z = 0$ and $y = 0$ and the center of the aperture is at $x = 0$. Then consider a focus located at f_x, f_z (for most cases we ignore the elevational dimension). Then we can calculate the time delay for that element as the difference between the time it takes for the wave to reach the focus from that element versus an element positioned at f_x:

$$\tau_e = (f_z - \sqrt{(e_x - f_x)^2 + (f_z)^2})/c, \tag{1.4}$$

since f_z is the shortest distance from the transducer to the focus. We can use the same technique to calculate the delays on receive based on the depth at which we want to focus the received data. From this foundation, we can expand to more complicated and robust image formation techniques such as dynamic receive, where the data is delayed to achieve focusing at all depths simultaneously, and synthetic aperture, where we can use the acquired channel data to estimate the beamformed echoes at additional lateral locations.

1.2.3 Image Quality Metrics

When comparing beamforming methods or other processes, we often compare the resulting image quality. After beamforming the channel data, the data (often called radio frequency, or RF, data) is enveloped using a Hilbert transform and then log-compressed to create the B-mode image. However, most image quality metrics

are computed on the enveloped, but not log-compressed data (though some groups do calculate metrics using the log-compressed data). Some common metrics include contrast ratio and contrast-to-noise ratio (CNR), defined as

$$\text{contrast ratio} = 20 * log10(\frac{\mu_{\text{ROI}}}{\mu_{\text{back}}}) \tag{1.5}$$

$$\text{CNR} = \frac{|\mu_{\text{ROI}} - \mu_{\text{back}}|}{\sqrt{\sigma^2_{\text{ROI}} + \sigma^2_{\text{back}}}}, \tag{1.6}$$

where μ is the mean value of the enveloped region defined as the region of interest (ROI) or background (back) and σ^2 is the variance for the same region. Contrast ratio represents a simple ratio of the target region to some background reference, and as defined here will be negative for hypoechoic targets and positive for hyperechoic targets. Contrast-to-noise ratio is a metric of the visibility of the target and accounts for the variance of both the background and the target, so for objects with the same intensity a lower variance leads to an increased CNR.

Additionally, the resolution of a beamformer is often compared in the form of the point spread function (PSF). This is done by imaging a single scatterer and plotting the lateral or axial line from the beamformed image through that point, an example of which is shown in Fig. 1.2. From this plot, we can compute the full width at half maximum (FWHM), or the width of the point target at half of the peak value compared to the background. This represents a measure of resolution for the beamformer, and can be compared against the theoretical axial and lateral resolutions of the hardware, given by

$$\text{res}_{\text{ax}} = \frac{cT}{2} \tag{1.7}$$

$$\text{res}_{\text{lat}} = \frac{\lambda z}{D}. \tag{1.8}$$

Axial resolution depends on T, the temporal length of the transmitted pulse, and the sound speed, c (Cobbold, 2007). Lateral resolution depends on the wavelength, λ, the specific depth, z, and the size of the active aperture, D. These estimates are the theoretical limitations of the hardware, and in practice the resolution observed by the beamformer will depend on the contrast ratio and contrast-to-noise ratio of the target (for example, higher intensity targets often appear physically larger, which may reduce the apparent resolution). In terms of the plot included in Fig. 1.2, the FWHM is located approximately at -6dB, but many groups will additionally report the widths of the PSF at -10 and -20 dB, for added context.

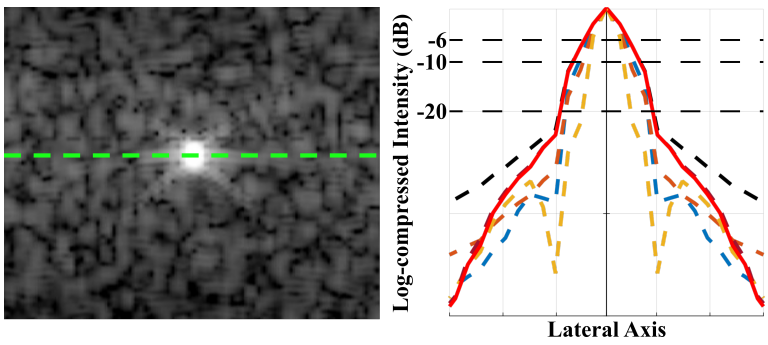

Figure 1.2: (Left) A b-mode image of a simulated point target, where the dashed line is where the lateral profile is measured. (Right) The point spread functions (PSFs) of several different ultrasound beamformers measured on a intense, 100dB bright target simulation. The intensity of the lateral profile is plotted against the lateral axis through the bright target, and the width of the profile can be measured at various amplitudes. Measuring at -6dB approximates the full width at half maximum (FWHM), though measuring at other amplitudes is also common.

1.2.4 Robust Image Quality Metrics

Recent work by several groups indicates that traditional image quality metrics are often insufficient to properly characterize modern adaptive beamformers. Historically, it has been believed that metrics like CNR are correlated with lesion detectability and subjective image quality. However, adaptive beamformers are often able to arbitrarily increase the value of CNR without otherwise improving detectability or "true" image quality. This concept is more specifically covered in Chapter 6, but as an overview this idea has led to the desire for more robust image quality metrics. These are especially important for this work, where image accuracy is our primary goal, since poor quality metrics may misrepresent the quality of our beamformers.

Contrast ratio dynamic range (CRDR) was presented by Dei et al. (Dei et al., 2017), a metric designed to assess the contrast accuracy of a beamformer. By measuring the contrast for a set of lesions with known contrast, we can calculate the range over which a beamformer reports contrast accurately. This metric is intended to combat adaptive beamformers that excessively increase contrast ratio without regard for the content in the image. As suggested in Section 1.1, this could result in complex hypoechoic cysts appearing more homogeneous than intended, which would result in a mis-diagnosis. This metric is the primary focus of our work in Chapter 3.

Generalized contrast-to-noise ratio (gCNR) is designed to be a more robust lesion detectability metric that is transformation resistant compared to traditional CNR (Rodriguez-Molares et al., 2018, 2020). It has quickly gained popularity in recent years due to this robust nature and the ease of implementation. In our work, gCNR is frequently used either to complement or replace CNR, though we do consider some of the challenges and concerns with the method in Chapter 6.

Finally, histogram matching is a newer method introduced by Bottenus et al. (Bottenus et al., 2020a). Ar-

guably, histogram matching is the most powerful of the three methods, as it works by calculating a transformation between two target images. This allows us to apply traditional image quality metrics to the "matched" images. This is valuable, since if a beamformer is merely a transform of another method (i.e. is not making any content-based improvements), the matched images and traditional metrics applied to these images will be similar. In theory, using histogram matching and CNR would be similar to just applying gCNR, as both methods would detect transformations. However, histogram matching is more complex to apply compared to gCNR, and there remain challenges and concerns with its use, which we detail in Chapter 7. As a result, while we have used histogram matching at various times in this work, the results are generally not included, as interpretation is ambiguous, as we will demonstrate.

Overall, the ethos of these metrics and this work is the same. We desire methods that produce real, clinical improvements to ultrasound images, and we desire metrics that can measure those improvements accurately.

1.3 Image Degradation in High Dynamic Range Applications

A high dynamic range scenario refers to a imaging scenario where the target is at a much higher or lower intensity compared to other features in the image. For example, blood or a cyst may be much lower intensity whereas a kidney stone may be much higher. These situations can be prone to certain artifacts or concerns due to these significant intensity differences, as we will explain here.

1.3.1 Sidelobe or Off-axis Clutter Artifacts

The Fraunhofer approximation earlier stated that the far-field approximation of an transducer pattern will be the 2-D Fourier transform (in the lateral and elevational dimensions). Since a transducer element can be modeled as a rectangular function, this means that the far-field pattern will be a sinc function. The primary function has a main lobe with smaller sidelobes, and for large element sizes, grating lobes will appear farther off-axis, as seen in Fig. 1.3. These grating lobes can be mostly eliminated by using an element spacing, or pitch, of $\lambda/2$ (Anderson and Trahey, 2006). The sidelobes, however, can create sidelobe artifacts in the presence of relatively strong acoustic scatterers.

When focusing at a location that is adjacent to a stronger acoustic target, the sidelobes will reflect off of that target and will falsely appear to originate from the focused location instead. In an image, this can appear as high intensity regions to the sides of the bright objects, such as in the bottom of Fig. 1.3. As the point target increases in intensity the sidelobe artifacts become increasingly visible, causing significant distortion in the final image. In hyperehcoic cases, where a target is much brighter than the surrounding area, this will cause distortions in the areas around the target, and may also result in the target appearing larger than it actually is. Hypoechoic cases suffer similarly, where the surrounding area is much brighter than the target, causing the

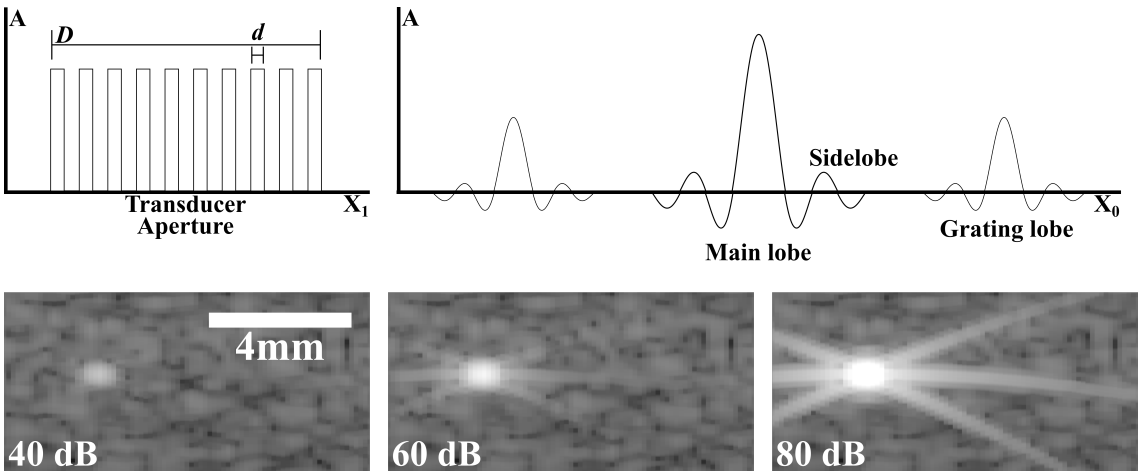

Figure 1.3: (Top left) a transducer aperture can be modeled in 2-D as a collection of rectangular functions with width d and aperture size D. (Top right) the Fourier transform in the far-field is a sinc function, with main lobes, sidelobes, and grating lobes. Image adapted from Anderson and Trahey (Anderson and Trahey, 2006). (Bottom) delay-and-sum beamformed images of single point targets at intensities of 40, 60, and 80 dB compared to the background. The sidelobe artifacts become increasingly visible as the scatterer's intensity increases.

target to appear smaller and sometimes brighter than it actually is. This is demonstrated in Fig. 1.4, where the two hypoechoic and hyperechoic cysts are indistinguishable and also appear incorrectly sized.

1.3.2 Dark Region Artifacts

Adaptive beamformers are a popular area of research in ultrasound, as they will adapt weights or other features of the beamforming process based on the content of the image. However, the dark region artifact described by Rindal et al. (Rindal et al., 2017) can occur as a result of an adaptive beamformer encountering sidelobe clutter. In this situation, some beamformers will remove the sidelobe clutter, but in doing so also remove the underlying signal, resulting in a dark region. Similar to sidelobe clutter, this will occur in the presence of strong acoustic targets. Fig. 1.5 shows examples of both a hypoechoic (left) and hyperechoic (right) case producing a dark region artifact when the adaptive beamformer removes the sidelobe clutter present in the delay-and-sum (DAS) image. In the hypoechoic blood vessel case this results in the complete loss of the blood signal in that region.

1.3.3 Noise due to Reverberation Clutter or Multipath Scattering

Reverberation clutter, or noise caused by multipath scattering, occurs when an acoustic wave encounters multiple scattering sources before returning to the transducer. This will result in echoes from different locations returning to the transducer at the same time, such as shown in Fig. 1.6. Normally, we assume that these near-field scatterers are low amplitude, and thus the echoes from additional reflections are weak so that we need

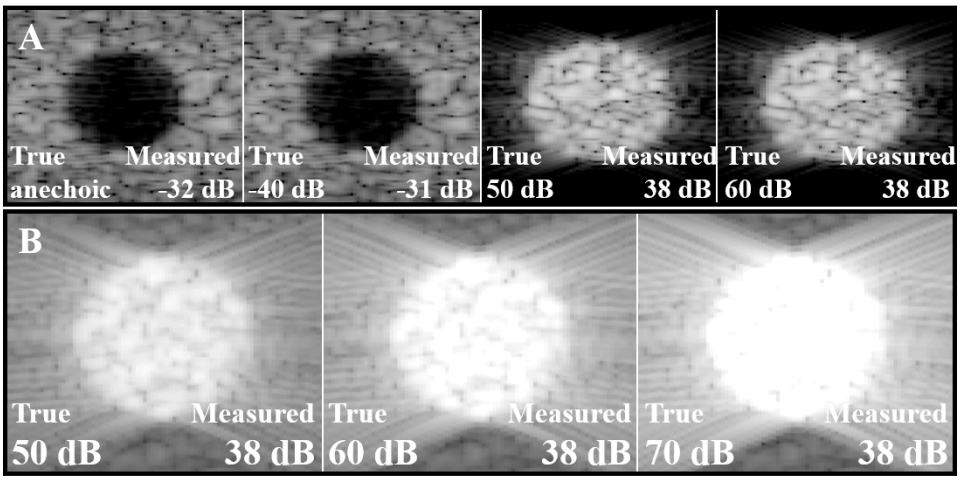

Figure 1.4: 5mm cysts of varying intensities displayed with delay-and-sum. (A) hypoechoic and hyperechoic cysts that have the same measured contrast ratio and are visibly identical. Note that the sidelobe clutter makes the hypoechoic cysts appear slightly smaller and the hyperechoic cysts appear slightly larger than reality. (B) Three hyperechoic cysts that have the same measured contrast ratio due to the increasingly strong sidelobes being produced.

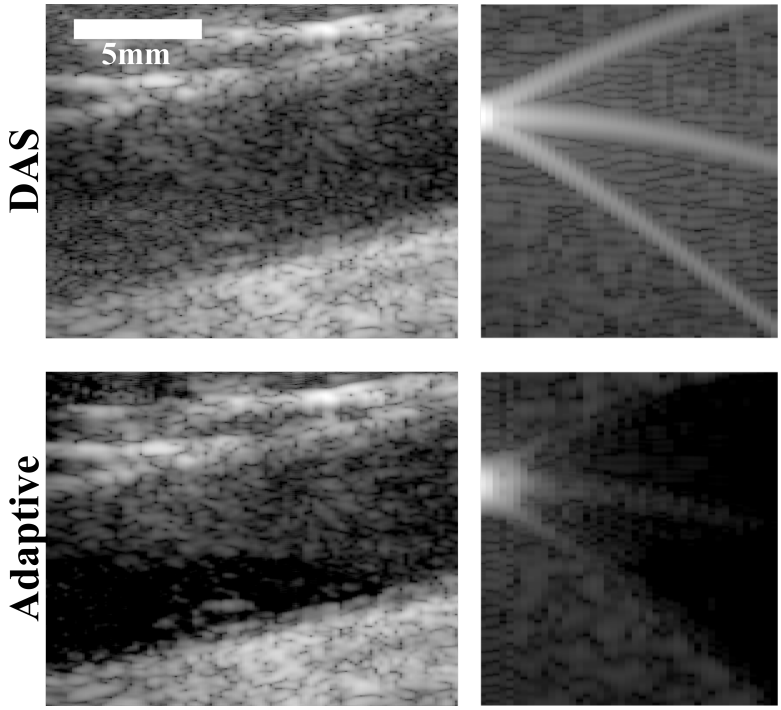

Figure 1.5: Delay-and-sum (DAS) and an adaptive beamformer (ADMIRE) images of a blood vessel (left) and a single 100 dB point target (right). In both cases, the adaptive beamformer removes the sidelobe clutter present, but in doing so removes the important underlying signal as well.

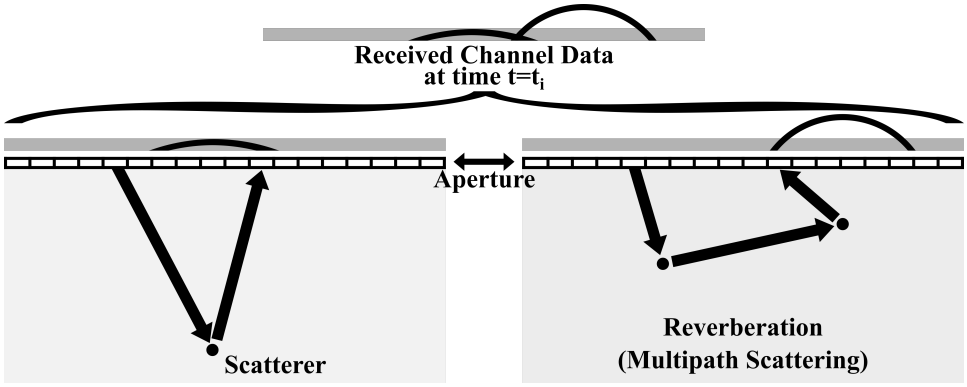

Figure 1.6: Echoes returning from a single reflection (left) and multiple reflections (right) at the same time, causing distortions in the received channel data.

only model the first reflections as described by the first-order Born approximation (Jensen, 1991). However, in patients with thick fat layers the effect of multipath scattering increases, causing significant distortions to the received channel data (Dahl et al., 2017).

1.3.4 Accurate Image Representation

These artifact and noise sources can cause significant degradation to image quality, especially for imaging in high dynamic situations or otherwise "difficult-to-image" patients. Additionally, these artifacts can make it difficult to produce images that accurately reflect the imaging target. As we showed earlier, sidelobe clutter can cause targets to appear larger or smaller than reality. The dark region artifact has the chance to completely remove important signals. Reverberation will generally degrade image quality across the entire image. Finally, beamformers will sometimes have difficulty reporting accurate levels of intensity, such as an anechoic or -40 dB cyst being measured as -30 dB. For example, in this case, sidelobe clutter with an intensity of -30 dB will mask lower intensity structures or objects, resulting in an incorrect appearance.

1.4 Beamforming Algorithms for Improved Image Quality

Parts of the descriptions of the beamforming algorithms in this section are adapted from work originally published in (Schlunk et al., 2021), ©2021 IEEE. Reprinted, with permission, from Schlunk, S., Dei, K., & Byram, B. (2021). Iterative Model-Based Beamforming for High Dynamic Range Applications. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 68(3), 482–493.

There are a variety of imaging techniques and beamforming algorithms that can be used to combat some of the concerns described in the previous section. We include one example of an imaging acquisition technique, but mainly focus on alternative beamformers that will be featured throughout the rest of this work.

1.4.1 Tissue Harmonic Imaging

Though we often only consider the linear aspects of ultrasound imaging, there are non-linear effects that we can take advantage of. As a wave propagates through a material with frequency f, it slowly begins to distort slightly, and the frequency content begins to show the increased presence of harmonics at $2f, 3f, 4f$, and so on. This means that more of these harmonics are generated as depth increases, but conversely means there are very few harmonics at shallow depths. Tissue harmonic imaging is designed to take advantage of this fact, focusing primarily on the second harmonics that are generated, as the later harmonics tend to be too low amplitude (potentially low signal-to-noise ratio) or high frequency (beyond the limits of the sampling rate) to be useful. By separating the second harmonics from the fundamental frequency, we can generate a significantly improved image (Choudhry et al., 2000). Two common methods for separating the two signals are frequency filtering and pulse inversion (Szabo, 2014).

Since the frequencies of interest occur at $2f$, the simplest solution is to use a bandpass filter centered at $2f$ that will remove those frequencies around the fundamental frequency. This requires no additional effort on the part of transmission and will work as long as the transducer is able to receive the second harmonic frequencies. However, we can improve the quality of the second harmonics by narrowing the transmitted bandwidth to prevent it from overlapping with our desired second harmonics. This can be accomplished by lengthening the transmitted pulse echo, but will result in worse axial resolution by equation 1.7 (Szabo, 2014).

The alternative method makes use of pulse inversion. By transmitting first a normal pulse and then following that with an inverted version of that pulse, the linear fundamental frequencies will be canceled while the nonlinear second harmonics will be enhanced (Szabo, 2014). Pulse inversion relies on a capable ultrasound machine, but otherwise does not rely on the input frequency or bandwidth in the same ways as filtering does.

As long as the ultrasound unit is able to pick up the low amplitude harmonic frequencies, tissue harmonic imaging can be a powerful tool for improving image quality without a computationally expensive or complex beamformer. Since the harmonic frequencies are generated from inside the body, they are less prone to artifacts that occur near the skin or imaging surface (Choudhry et al., 2000), such as reverberation clutter.

1.4.2 Delay-and-Sum (DAS)

Though not a beamformer designed for improving image quality, we include a formal definition of delay-and-sum for reference with the other beamformers in this section. We define a given image pixel as

$$S_{\mathrm{DAS}}(x,z) = \sum_{i=1}^{M} \mathrm{w}_i(x,z) s_i(x,z). \tag{1.9}$$

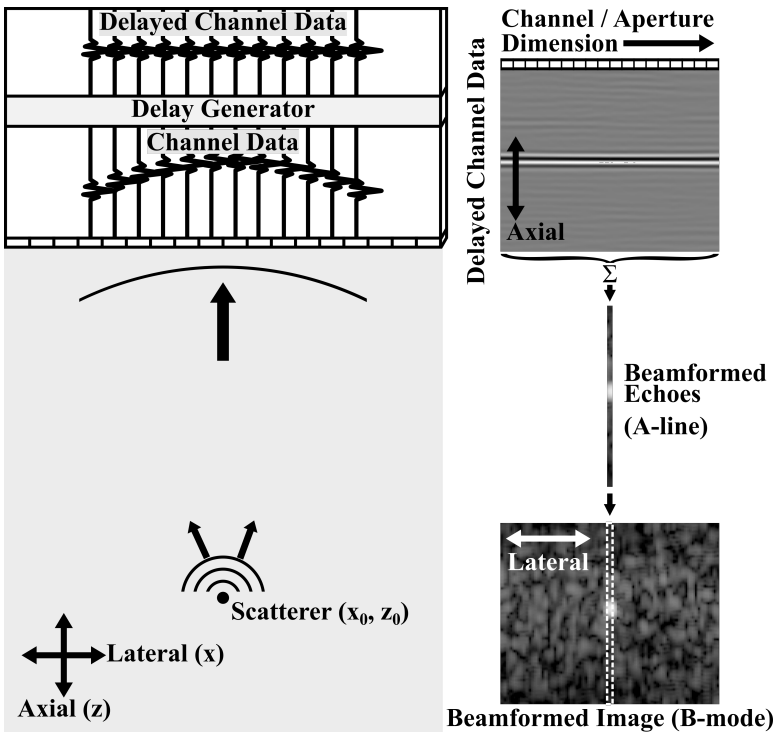

Figure 1.7: Visual reference for delay-and-sum (DAS) beamforming. A final b-mode image is composed of many separate acquisitions focused at different lateral locations (x axis). In this example, we are focused at x_0, the location of the point target. The returning echo is delayed to create the delayed channel data shown. This data corresponds to that single lateral location, and as such is composed of the same axial dimension and a different channel or aperture dimension. DAS sums along the channel dimension, which results in the single a-line, a small part of the overall beamformed image. Because of this, for a single picture our data is often in the form (z,i,x) referring to the (axial, channel, lateral) dimensions.

where x is the lateral location of the a-line of the final image, z is a discrete time index (related to axial depth), M is the total number of channels (active aperture elements), $\mathrm{w}_i(z)$ is the weighting factor for channel i, and $s_i(x,z)$ is the delayed channel data across the aperture for a given channel i. The weighting factor is a weighting that is applied to the channels after being received and delayed, the most simple of which is a rectangular weighting (i.e. all weights are equal). We include a quick visual representation to clarify the dimensions here and for the rest of this section in Fig. 1.7.

While DAS is the most basic beamformer, sidelobe artifacts can be somewhat mitigated by changing the weighting function used during summation. It has been demonstrated that Hamming weighting schemes can be used to reduce sidelobes (Matrone et al., 2015), though this is at the cost of broadening the mainlobe, which results in lower lateral resolution.

DAS is often considered to be the gold standard for beamforming due to the sheer simplicity, and therefore computational efficiency, of the algorithm. No other beamforming method can compete with the speed of DAS, and in clinical machines DAS can be computed at the hardware level, further improving efficiency

by reducing the amount of data an ultrasound machine needs to handle. Though there is no inherent noise reduction performed by DAS, combining it with tissue harmonic imaging and clever filtering can result in impressive clinical images. As a result, DAS is always considered for comparison when evaluating alternative beamformers.

1.4.3 Filtered Delay-Multiply-and-Sum (F-DMAS)

Filtered delay-multiply-and-sum (F-DMAS) is a simple adaptive beamformer originally used in RADAR and adapted for ultrasound by Matrone et al. (Matrone et al., 2015). It works by combinatorially coupling and multiplying the delayed channel data before summing across the channels. Bold characters here and for the rest of the text indicate vectors. The unfiltered DMAS signal can be calculated as

$$\mathbf{S}^{*}_{\mathrm{DMAS}}(z) = \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \hat{\mathbf{s}}_{ij}(z), \tag{1.10}$$

and $\hat{\mathbf{s}}_{ij}(z)$ is defined as

$$\hat{\mathbf{s}}_{ij}(z) = \mathrm{sign}(\mathbf{s}_i(z)\mathbf{s}_j(z))\sqrt{|\mathbf{s}_i(z)\mathbf{s}_j(z)|}. \tag{1.11}$$

The filtered signal can then be calculated from $\mathbf{S}^{*}_{\mathrm{DMAS}}$ by removing the DC and high frequency components using a band-pass filter centered around $2f_c$.

Since the F-DMAS beamformer introduces a measure of correlation into the calculation, this can reduce some clutter and noise sources that are uncorrelated compared to tissue. Additionally, the calculation artificially doubles the frequency, which halves the wavelength leading to an improved lateral resolution compared to DAS.

1.4.4 Generalized Coherence Factor (GCF)

Generalized coherence factor (GCF) is designed to reduce focusing errors caused by phase-aberrations (wavefront distortions caused by sound speed inhomogeneities in a medium). It is a weighting of the DAS image by the ratio of the energy in some chosen low-frequency region to the total energy (Li and Li, 2003). This ratio for a pixel at (x,z) is defined as

$$\mathrm{GCF}(x,z) = \frac{\sum_{k=0}^{M_0} |P_k(x,z)|^2}{\sum_{k=0}^{M} |P_k(x,z)|^2}, \tag{1.12}$$

where M is the number of channels, M_0 is the chosen cutoff frequency index, and $P(k)$ is the M-point discrete Fourier transform across the aperture as defined by

$$P_k(x,z) = e^{j\pi k} \sum_{i=0}^{M} s_i(x,z) e^{-j2\pi(ik/M)}, \tag{1.13}$$

where $s_i(x,z)$ is the delayed channel data for channel i. GCF allows us to vary between pure coherence factor ($M_0 = 0$) and DAS ($M_0 = M$). We chose $M_0 = 5$ as done in (Rindal et al., 2017). The final GCF image is then calculated simply as

$$S_{\text{GCF}}(x,z) = \text{GCF}(x,z) S_{\text{DAS}}(x,z). \tag{1.14}$$

The original GCF work showed that the low-frequency cutoff M_0 impacts the performance of the beamformer. The smaller the value of the cutoff, the stronger the effect. However, at $M_0 = 0$ and sufficiently small values, artifacts are introduced and the authors suggest it is only suitable for imaging point targets. On the other hand, increasing the cutoff too high approaches DAS, which defeats the purpose of the beamformer. Generally, increasing the cutoff lowers the variation while sacrificing contrast ratio improvements.

1.4.5 Minimum Variance (MV)

Similar to GCF, minimum variance (MV) beamforming did not originate with ultrasound, and has been used prior in RADAR and seismic activity monitoring applications (Capon, 1969). MV can be thought of as a way to calculate an optimized set of apodization weights for DAS beamforming (Holfort et al., 2009), (Synnevåg et al., 2007). These weights are calculated as

$$\mathbf{w} = \frac{R^{-1}\mathbf{e}}{\mathbf{e}^H R^{-1} \mathbf{e}}, \tag{1.15}$$

where $\mathbf{e}$ is the steering vector, H is the conjugate transpose, and R is the covariance matrix defined as

$$R(z) = \text{E}[\mathbf{s}(z)\mathbf{s}(z)^H], \tag{1.16}$$

where $\text{E}[\cdot]$ denotes the expectation and $\mathbf{s}(z)$ is the delayed aperture signal at depth z.

To properly adapt the original method to ultrasound, many groups have proposed various modifications. Most commonly are methods to ensure the covariance matrix is invertible, specifically averaging subarrays to estimate the matrix and using diagonal loading. The standard recommendation is to use subarray lengths of L, equal to 50% of the total aperture length, and diagonal loading defined as $\varepsilon = \Delta \cdot \text{tr}(\hat{R})$, where $\Delta = 1/(10L)$

(Synnevåg et al., 2007). The MV signal estimate is then defined as

$$\hat{\mathbf{S}}_{\mathrm{MV}}(z) = \frac{1}{M-L+1} \sum_{l=0}^{M-L} \mathbf{w}(z)^H \bar{\mathbf{s}}(z), \tag{1.17}$$

where M is the total aperture length and $\bar{\mathbf{s}}(z)$ is the delayed aperture signal for the corresponding subarray.

MV was originally designed to help distinguish between seismic waves arriving at a receiver, which can be thought of as a resolution problem. In ultrasound, MV similarly helps to isolate wavefronts arriving at the transducer by suppressing the interference that other wavefronts cause off-axis (off-axis clutter). This helps to narrow the mainlobe, which improves lateral resolution. MV is a very popular beamformer due to its rigorous mathematical basis, however, computational complexity and relatively poor results *in vivo* often hold it back.

1.4.6 Short-Lag Spatial Coherence (SLSC)

Short-lag spatial coherence (SLSC) is different from many other beamformers in that it only uses the spatial coherence of received echoes to form images, resulting in improved contrast-to-noise ratio compared to DAS. Introduced by Lediju et al. (Lediju et al., 2011), the foundation of SLSC is the van Citter-Zernike (VCZ) theorem, which demonstrates a prediction for the spatial coherence, or covariance, for the backscattered echoes (Mallart and Fink, 1991). SLSC is computed per pixel using a correlation kernel as

$$S_{\mathrm{SLSC}}(x,z) = \sum_{l=1}^{L} \frac{1}{M-l} \sum_{i=1}^{M-l} \frac{\sum_{n=z}^{z+n_k} s_i(x,n) s_{i+l}(x,n)}{\sqrt{\sum_{n=z}^{z+n_k} s_i^2(x,n) \sum_{n=z}^{z+n_k} s_{i+l}^2(x,n)}}, \tag{1.18}$$

where l is the lag, L is the number of lags to sum, and n_k is the size of the correlation kernel. This produces an SLSC image calculated per a-line x and each depth z. The choice of number of lags is recommended as $<$30% of the total number of channels (Bell et al., 2015). In our works, we chose to use $L = 20$ lags (17% of our total number of channels), with a correlation kernel equivalent to 1 wavelength.

Unlike most other beamformers, and all those presented here, SLSC is not usually log-compressed when displayed and instead uses a linear scale. This is chosen since the normalization used in the calculation results in a much smaller dynamic range than traditional beamformers. This can make it difficult to accurately assess comparisons between SLSC and other traditional beamformers. Similar to F-DMAS and GCF which employ correlation, SLSC achieves clutter reduction and improved CNR due to the low coherence of many types of clutter.

CHAPTER 2

ADMIRE: Description and Preliminary Work

2.1 Aperture Domain Model Image Reconstruction (ADMIRE)

This section describing the ADMIRE method is adapted from work originally published in (Schlunk and Byram, 2022b) ©2022 IEEE and (Schlunk et al., 2022), ©2022 Ultrasonics. Reprinted, with permission, from Schlunk, S., & Byram, B. (2022). Combining ADMIRE and MV to Improve Image Quality. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 69(9), 2651–2662. Reprinted, with permission, from Schlunk, S., Hsi, R., & Byram, B. (n.d.). Enhancing sizing accuracy in ultrasound images with an alternative ADMIRE model and dynamic range considerations. Ultrasonics, *In Review*.

Aperture domain model image reconstruction (ADMIRE) is a pre-processing method designed for the removal and suppression of reverberation clutter, off-axis interference, and wavefront aberration (sound speed inhomogeneities). The original details of the work can be found in (Byram et al., 2015a) with more specifics available elsewhere (Dei and Byram, 2017; Byram and Jakovljevic, 2014), but we will reproduce most of the concepts and specifics here as it forms the foundation for much of this work.

ADMIRE is somewhat unique among beamformers as it exists as a pre-processing step compared to other beamformers that are strictly post-processors. When any single data acquisition is made, the initial data is (usually) three dimensional: depth (or time), lateral, and along the aperture. The depth and lateral axes are the actual imaging space, while the aperture dimension is where summation occurs for DAS. Most adaptive beamformers perform some sort of calculations along this aperture dimension to modify the summation step, which results in a two-dimensional final image. As a result, these beamformers can be considered post-processing, as we are not able to apply multiple post-processors due to the loss of the aperture dimension. However, ADMIRE's processing is performed without the need for the summation to occur and maintaining the aperture dimension, allowing other beamformers to be applied afterwords.

As with other beamformers, we start with the properly time-delayed channel data. The actual ADMIRE processing, however, occurs in the frequency domain. This is done by subdividing the delayed channel data axially into many overlapping windows, and applying a short-time Fourier transform (STFT) to break the broadband signal into many narrowband components. We generally use 90% overlap between windows to maintain good speckle texture. This produces an aperture signal for each lateral index, depth index, and primary frequency component of the windowed broadband signal. Each aperture domain signal is analyzed against a physics-based model that contains the predicted aperture domain signal responses for scatterers

throughout the imaging space. These responses can be calculated due to the linear nature of sound wave propagation, and in theory any received signal can be represented as the linear combination of these responses with the equation

$$p_s(x;t,\omega) = \sum_{n=0}^{N-1} A(x;x_n,z_n,\tau_n,\omega)e^{jk\tau(x;x_n,z_n,\tau_n)}. \tag{2.1}$$

Here, x is indexing the aperture location and ω defines the post-STFT frequency of the signal. For the signal arriving at the aperture at time t, there are a total of N responses from scatterers in the medium arriving at that time. The wavefront delay $\tau(x;x_n,z_n,\tau_n)$ is for a signal originating from (x_n,z_n) at time τ_n, and k is the wavefront number. Finally, a measure of amplitude modulation across the aperture $A(x;x_n,z_n,\tau_n,\omega)$ is included, which is based on effects of the STFT windows and element sensitivity.

The true value for N, the number of scatterers that make up an observed signal, is unknowable. However, by oversampling the imaging space, we can combine all of the individually modeled signals into a model matrix, X, which gives the model the flexibility to represent even a complex observed aperture domain signal, y, by its component sources as

$$y = \mathrm{X}\beta, \tag{2.2}$$

where every received signal y corresponds to some location (x_n,z_n) and frequency ω, and we are solving for β, the unknown set of model coefficients that will reveal the origins of the scatterer responses that combine to make the received signal. Fig. 2.1 shows a visual example of how signals in the model $\mathrm{X} = [\mathrm{X}_{\text{clutter}}\ \mathrm{X}_{\text{ROI}}]$ could correspond to specific physical locations in the field of view. By design, the matrix consists of both clutter signals $\mathrm{X}_{\text{clutter}}$, which includes image degradation sources such as reverberation clutter and off-axis interference, and region of interest (ROI) signals X_{ROI}, which is all the signals originating near the target (x_n,z_n). The model coefficients β then represent the specific signals that linearly combine to form y. By solving for β and zeroing those coefficients corresponding to signals in $\mathrm{X}_{\text{clutter}}$, the decluttered aperture signal can be reconstructed with

$$y_{\text{decluttered}} = \mathrm{X}\beta_{\text{ROI}}. \tag{2.3}$$

We can specify which signals we want to keep by how we define the ROI, an ellipse, around the target (x_n,z_n). The major and minor radii of this ellipse are calculated as $c_{axial}\text{res}_{axial}$ and $c_{lateral}\text{res}_{lateral}$, with c_{axial} and $c_{lateral}$ scalars based on how large we desire the ROI to be in the axial and lateral dimensions. The axial sampling is approximated with $\text{res}_{axial} \approx 2\text{res}_{lateral}$, and the lateral sampling is estimated as $\text{res}_{lateral} \approx \lambda_w z_n \mathscr{F}\{|p_s(x_n,z_n,0)|\}_{BW}$, where λ_w is the wavelength and $\mathscr{F}\{|\cdot|\}_{BW}$ is the lateral bandwidth at (x_n,z_n). The specific choices for this application are included in Table 2.1, with the scalars chosen empirically from previous testing.

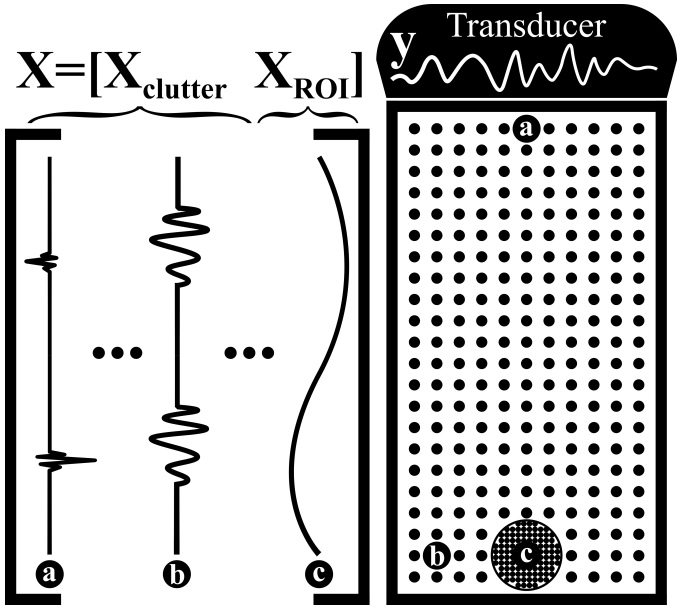

Figure 2.1: An example ADMIRE model X, which consists of signals from clutter locations (sparsely sampled, e.g a and b), and signals from region of interest locations (highly sampled, e.g. c). By solving the matrix equation, we can estimate which physical sources in the modeled imaging space make up the received echo. Signals a and b represent reverberation clutter signals and off-axis signals, respectively, while c is an example of a signal from the ROI. Figure adapted from Schlunk et al. (Schlunk et al., 2021) ©2021 IEEE

Table 2.1: Default ADMIRE Model Space Parameters. Table ©2022 IEEE.

Parameter	Value
α	0.9
λ	$(0.00189)y_{\text{RMS}}$
$\text{X}_{\text{clutter}}$ Sampling	$3.577\text{res}_{lateral}, 7.154\text{res}_{axial}$
$\text{X}_{\text{clutter}}$ Size	Full Imaging Space
X_{ROI} Sampling	$0.179\text{res}_{lateral}, 0.715\text{res}_{axial}$
X_{ROI} Ellipse Radii	$(0.5\text{res}_{lateral}, \text{res}_{axial})$

However, the inverse matrix calculation that would be used to solve (2.2) is ill-posed due to the much smaller size of the aperture signal y compared to the much larger size of the oversampled model X. Instead, the elastic-net regularization technique (Zou and Hastie, 2005) is used to solve for the model coefficients. The optimization equation is

$$\hat{\beta} = \arg\min_{\beta}(||y - \text{X}\beta||^2 + \lambda(\alpha||\beta||_1 + (1-\alpha)||\beta||_2^2/2)), \tag{2.4}$$

where $||\beta||_1$ and $||\beta||_2$ are the L1 and L2 norm, and α can be set between 0 and 1 to control the weighting between them. The degrees of freedom of the solution can be controlled by a separate regularization parameter, λ (Tibshirani and Taylor, 2012). The chosen values for these parameters are included in Table 2.1, based

on values from previous work (Byram et al., 2015a; Dei and Byram, 2017). These values have been observed to be robust across different imaging applications.

With the model coefficients estimated by the elastic-net, the decluttered aperture domain signal can be calculated with (2.3) as described. The decluttered signal can then be returned to the time domain via the inverse STFT (Yang, 2008). Since this entire process is performed along the channel dimension, the result is decluttered channel data.

Since ADMIRE returns decluttered channel data, we can, in theory, apply any beamforming method to data that has already been processed with ADMIRE. Most commonly, we simply sum the decluttered channel data like in DAS beamforming, but rather than write ADMIRE+DAS we simply refer to this as ADMIRE. At various points, we have additionally considered combinations of ADMIRE and MV, GCF, and others where the results will be written as ADMIRE+MV, ADMIRE+GCF, etc... This does lead to additional questions about how ADMIRE parameters and other methods may be tuned to further enhance the combined method, which will be addressed as these combined methods are introduced later.

Computation time is variable since it increases at deeper depths (larger models), for larger apertures, and for increased overlap between STFT windows. For example, for the simulations in Chapter 3, we use 90% overlap between windows and process a 10 mm band centered at a depth of 3 cm. This was split across 50 computational cores, with each core requiring 30-40 minutes. Though this timing is acceptable for research, a variety of ADMIRE development efforts are focused on architectural changes in order to achieve near real-time processing (Dei et al., 2019; Khan et al., 2019, 2020, 2021).

Overall, the power of ADMIRE is being able to separate any received ultrasound echo into the components that form the linear combination, as described by the superposition principle. By determining the origin location in the imaging space of these components, components that are related to reverberation clutter (originating from the near-field, close to the transducer) or off-axis clutter (similar depth, but laterally separated) can be removed easily. Additionally, the process of fitting the signals using the elastic-net can help to reduce errors caused by phase aberration, unless the degrees of freedom are set too high to cause overfitting of the signal (Dei and Byram, 2017).

2.2 Preliminary Work

In this section, we briefly consider how we choose the ADMIRE parameters and further analyze how ADMIRE behaves in high dynamic range scenarios, and how it behaves in the presence of the discussed imaging artifacts and noise sources.

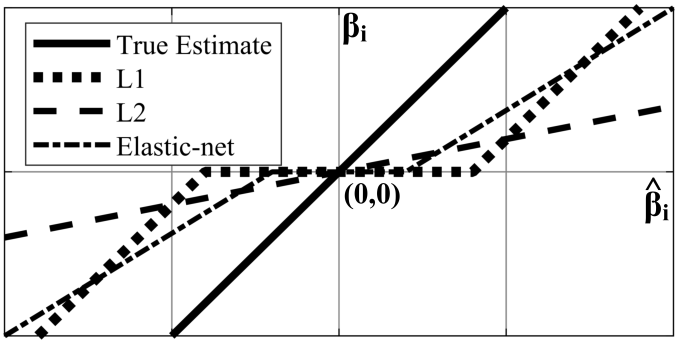

Figure 2.2: Qualitative example of the L1, L2, and elastic-net estimates for some specific model coefficient, β_i. In the L1 case, coefficients that are sufficiently small are zeroed in favor of larger ones. No coefficients are zeroed in L2, but instead there is significant coefficient shrinkage (loss of amplitude) for stronger components. ADMIRE uses the elastic-net, which balances the two extremes but includes features of both. Figure adapted from Schlunk et al. (Schlunk et al., 2021) ©2021 IEEE

2.2.1 ADMIRE Parameters

As described in the previous section, there are several user-chosen parameters in ADMIRE, and though the recommended values are generally robust, understanding the role of the parameters in a resulting image is important. The main two parameters adjusted are the values of α and λ, which control the balance of L1 and L2 regularization and the degrees of freedom, respectively.

Fig. 2.2 shows a graphical example of L1, L2, and the elastic-net solve for a coefficient slightly differently. L1 regularization, or Lasso (least absolute shrinkage and selection operator) regression, penalizes the absolute value of non-zero coefficients in a solution, and will therefore try to shrink sufficiently small coefficients to zero. This results in a more sparse solution compared to L2, where few, and generally uncorrelated, coefficients are selected, hence it is often referred to as feature selection. Any feature that is sufficiently larger or strong will be selected, but there is a potential for weak signal components to be lost. In comparison, L2 regularization, or Ridge regression, does not remove any coefficients and instead penalizes larger coefficients as it uses the square of the coefficients, resulting in coefficient shrinkage. L1 producing sparse solutions can seem preferable, but many signals in ultrasound can be correlated, especially those with spatial proximity to each other, and losing those correlated signals can create high variance images, and occasionally loss of some signals entirely. On the other hand, L2 can lead to situations where a strong signal may be underfit due to the coefficient shrinkage.

The large size of the model matrix in ADMIRE suggests that using L1 to create sparser solutions would be preferable, but the high variance and potential loss of signal means that a balanced approach is optimal. Hence, the elastic-net, which provides a balance of the two regression methods is chosen, and we tune α to balance between them. At $\alpha = 0$, the result would be entirely L2, and $\alpha = 1$ would correspond to all L1.

Initial testing by Byram et al. (Byram et al., 2015b) suggests that L1 minimizes error of the overall observed ultrasound echo, while L2 minimizes the error of specifically the components from the region of interest. An L1 solution would generally produce poorer final images, while L2 would frequently fail to reject certain sources of clutter. As a result, the balanced approach of the elastic-net helps to mitigate the shortcomings of each method, and the use of $\alpha = 0.90$ is recommended.

The tuning parameter for the degrees of freedom, λ, can be controlled separately. Similar to the choice of regularization method, increasing or decreasing the degrees of freedom will influence the sparsity of the final solution. Smaller values of λ produce higher degree of freedom solutions, which can result in more precise fits of the aperture signal, while lower degrees of freedom favor sparsity. This has similar drawbacks to the consideration of L1 and L2, and as a result we sometimes consider both high and low degree of freedom solutions. Generally, the default recommendation used is $\lambda_{\text{HDF}} = (0.0189/10)y_{\text{RMS}}$, which is the value included in Table 2.1, which corresponds to a high degree of freedom (HDF) case. In situations where a low degree of freedom (LDF) case is considered, we instead would use $\lambda_{\text{LDF}} = (0.0189/2)y_{\text{RMS}}$. Unless otherwise stated, the high degree of freedom case is used.

Finally, we can additionally consider controlling the specifics of the model generation. Both the sampling of the model space and the physical size can be considered. Increasing the sampling of the model can result, in theory, in more accurate fits of the aperture signal, but also can greatly increase run time without necessarily improving image quality. Considerations have been made in the past to reduce the size of the model to improve computation time, such as removing the estimates for near-field scatterers in cases without reverberation clutter, using only the ROI part of the model and relying and not overfitting the signals (Schlunk and Byram, 2021a), and using dimensionality reduction methods to reduce the size of the models (Dei et al., 2019). These have been met with mixed results, and generally for more complex problems and applications we continue to use "full" models that consider the entire scope of the imaging space with a reasonably small, but well-sampled, ROI.

2.2.2 Sidelobe and Dark Region Artifact Mitigation

2.2.2.1 Simulated Bright Target Phantom

There are many scenarios where sidelobe and dark region artifacts can manifest, but perhaps the easiest way for testing is to create a single target of sufficiently great intensity that it casts a sidelobe artifact into the surrounding area. To this end, we used an ultrasound simulation package called Field II to simulate bright point targets at the focal depth with contrasts ranging from 20 dB to 120 dB relative to the background speckle. We created 6 independent speckle realizations for each amplitude using the parameters found in Table 2.2. These parameters are the same for all simulations performed in this work.

Table 2.2: Field II Simulation Parameters for Phantoms

Parameter	Value
Number of elements	117
Number of mathematical elements (lateral)	7
Number of mathematical elements (elevation)	11
Element height	4 mm
Element width	0.254 mm
Kerf	0.003 mm
Lateral pitch	0.257 mm
Center frequency (f_c)	3 MHz
Sampling frequency (simulation) (f_s)	640 Mhz
Sampling frequency (downsampled) (f_s)	40 MHz
Bandwidth	60%
Transmit focal depth	3 cm
Transmit/receive f-number	1

2.2.2.2 Visual Artifact Confirmation

In early work by Dei et al., combinations of DAS, ADMIRE, and coherence factor (CF) were applied to these simulations demonstrating the dark region artifact in all cases except for base DAS (Dei and Byram, 2018). Since the DAS method is a linear process, a dark region artifact will never present itself. However, in testing DAS+CF began showing signs of the artifact as low as 20 dB. ADMIRE using a low degree of freedom showed the dark region artifact earlier than the high degree of freedom variant, showing that ADMIRE (HDF) is somewhat robust to both the sidelobe and dark region artifacts up to an intensity of around 60 dB. Furthermore, the ADMIRE+CF (HDF) variant showed resilience to introducing the dark region artifact compared to DAS+CF, meaning that the sidelobe reduction of ADMIRE (HDF) can help other adaptive beamformers when used in combination from suffering the artifact. A similar set of example results can be visually confirmed in Fig. 2.3, where GCF, SLSC, and ADMIRE (HDF) all present the dark region artifact and all methods (except SLSC) present sidelobe artifacts.

2.2.2.3 Measuring Speckle Accuracy

Removing sidelobe artifacts without introducing dark region artifacts means we must be able to verify that the missing signal being restored is the correct signal. We can do this by calculating the 2-D correlation coefficient r by comparing each simulated phantom to its corresponding clean speckle realization in DAS (with no cyst or bright scatterer present) to determine the degree of speckle corruption from the sidelobe and

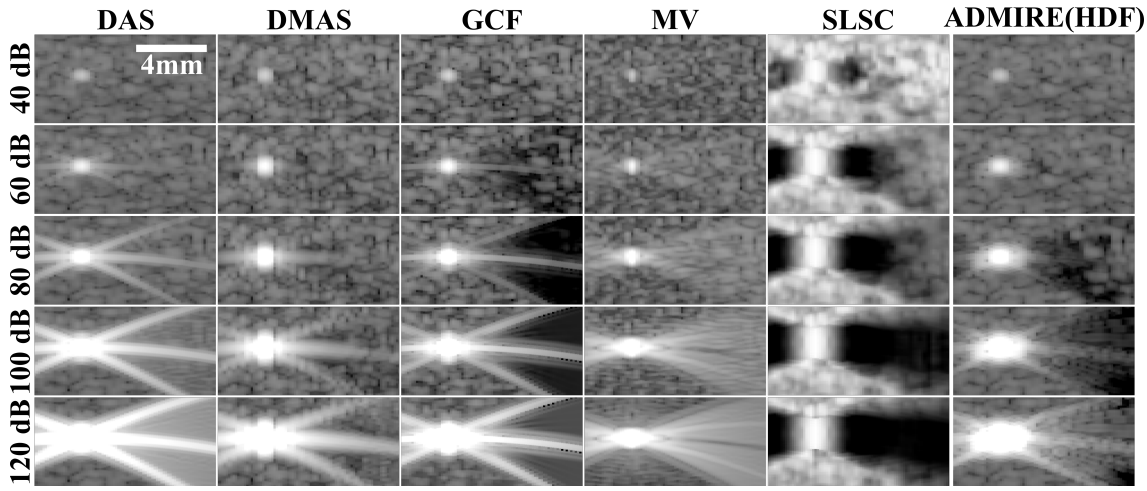

Figure 2.3: Delay-and-sum (DAS), delay-mulitply-and-sum (DMAS), generalized coherence factor (GCF), minimum variance (MV), short-lag spatial coherence (SLSC), and ADMIRE (high DoF) applied to bright target phantoms of varying intensities. At the strongest intensity, all beamformers except DAS and DMAS exhibit the dark region artifact. Figure adapted from Schlunk et al. (Schlunk et al., 2021). Figure ©2021 IEEE.

dark region artifacts. The correlation was calculated as

$$r = \frac{\sum_x \sum_z (S(x,z) - \mu_S)(D(x,z) - \mu_D)}{\sqrt{(\sum_x \sum_z (S(x,z) - \mu_S)^2)(\sum_x \sum_z (D(x,z) - \mu_D)^2)}}, \tag{2.5}$$

where S is the enveloped, uncompressed region of the data for the beamformer of interest, and D is short-hand for the reference DAS enveloped, uncompressed region of data. Values of r closer to 1 indicate more accurate speckle in the target region. A DAS image with only speckle would represent the "true" background signal, having a correlation, or accuracy, of 1. As the sidelobe artifacts increasingly corrupt that background, the accuracy decreases. Table 2.3 shows preliminary results of the beamformers described in Section 1.3 applied to the bright target phantoms. At low target intensities, DAS maintains near perfect speckle representation, but MV performs better for greater intensities. However, all beamformers have relatively poor speckle accuracy (r<0.80) at 100 dB and greater.

2.2.2.4 Measuring Contrast Ratio in the Artifact Region

A final simple metric for testing the artifacts is to measure the contrast ratio in the area where the sidelobe and dark region artifacts manifest compared to a region of the image with no artifacts. This means that positive ratios indicate sidelobes, while negative ratios would indicate dark region artifacts. Table 2.4 shows these preliminary results. Results close to 0 dB indicate that the background contrast ratio is close to the expected

Table 2.3: 2-D Speckle Correlation of Bright Target Phantoms

	2-D Speckle Correlation (close to 1 is better)				
Intensity (dB)	40	60	80	100	120
DAS	**1.00±0.00**	**0.99±0.00**	0.54±0.05	0.05±0.04	0.02±0.03
DMAS	0.80±0.02	0.81±0.02	0.76±0.02	0.40±0.06	0.07±0.05
GCF	0.96±0.00	0.87±0.01	0.10±0.06	0.03±0.03	0.03±0.03
MV	0.71±0.02	0.71±0.03	**0.90±0.01**	**0.54±0.03**	0.02±0.03
SLSC	0.49±0.04	0.38±0.05	0.29±0.04	0.18±0.09	**0.13±0.05**
ADMIRE (HDF)	0.97±0.01	0.96±0.01	0.82±0.05	0.04±0.05	0.04±0.04

Table 2.4: Contrast Ratio of Artifacts in Bright Target Phantoms

	Contrast Ratio (dB) (close to 0 is better)				
Intensity (dB)	40	60	80	100	120
DAS	1.9±0.9	**1.9±0.9**	2.0±0.9	**4.6±0.6**	20.0±0.6
DMAS	5.3±1.1	2.9±1.0	**0.1±1.0**	-5.0±1.3	**0.6±1.1**
GCF	4.9±1.2	-6.3±1.8	-34.8±1.4	-20.7±0.7	-1.8±0.6
MV	1.2±1.3	5.1±1.1	10.6±0.8	15.5±0.7	21.2±0.3
SLSC	**1.1±0.3**	-2.9±0.3	-5.8±0.4	-16.1±0.7	-12.6±3.9
ADMIRE (HDF)	3.0±0.8	2.3±0.9	-7.0±2.1	-24.4±0.6	-4.8±0.8

value, however, this only indicates that the background appears to have the correct intensity, not that the background is free of artifacts. Ideally this metric would be combined with speckle accuracy for a more complete analysis.

2.2.3 Contrast Ratio Dynamic Range (CRDR)

In an effort to better characterize true contrast in beamformers, Dei et al. proposed the contrast ratio dynamic range (CRDR) method (Dei et al., 2017). By applying a given beamformer to simulated cysts at varying intensities, in this case ranging from -50 to 70 dB plus an anechoic cyst, the measured contrast can be compared to the true contrast over a wide range. For a given realization, the measured contrast is then linearly interpolated to get a well sampled curve. The CRDR measurement can then be estimated from that measured versus true contrast curve (such as in Fig. 2.4) by determining the range on the curve for which the measured contrast does not deviate from the true contrast (two-tailed t-test, $\alpha = 0.05$).

2.2.3.1 Simulated Contrast Phantoms

Field II (Jensen, 1996; Jensen and Svendsen, 1992) was used to simulate 5 mm diameter cysts of known contrasts ranging from -50 dB to 70 dB relative to the background, plus an anechoic case. Each level of contrast was simulated with 6 independent realizations of speckle. Simulation parameters used in Field II are the same as detailed in Table 2.2.

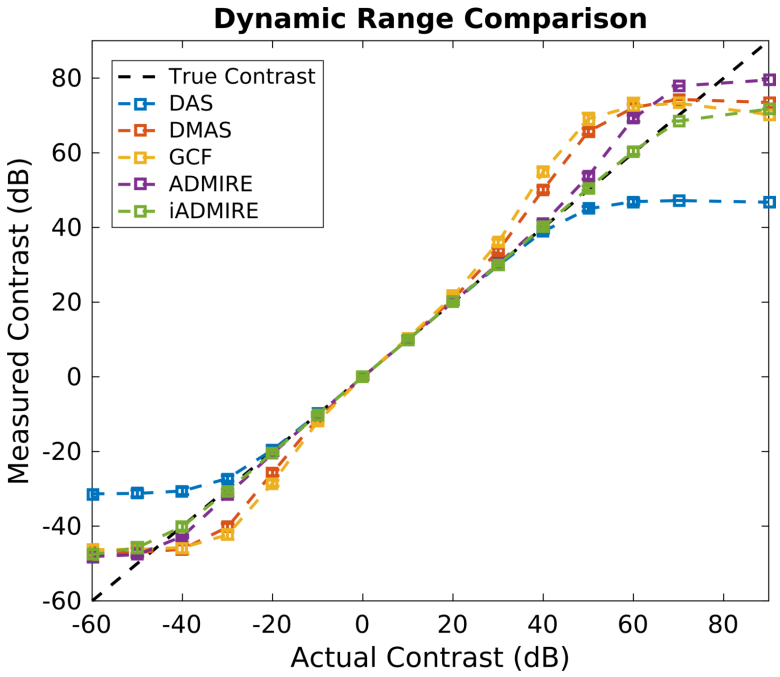

Figure 2.4: Measured contrast versus actual (true) contrast for various beamformers in cases with no added clutter. Figure adapted from Schlunk et al. (Schlunk et al., 2018). Figure ©2018 IEEE.

2.2.3.2 Simulated Reverberation Clutter

For each of the contrast phantoms, additional realizations were created that included simulated reverberation clutter. This reverberation clutter was simulated using the technique detailed by Byram and Shu (Byram and Shu, 2016b), (Byram and Shu, 2016a). The clutter was added directly to the channel data such that it satisfied the signal-to-clutter ratio (SCR) calculated by

$$\text{SCR} = 10\log_{10}\left(\frac{P_{\text{SOI}}}{\alpha^2 P_{\text{clutter}}}\right), \tag{2.6}$$

where P_{SOI} is the power of the channel data of the original phantom, P_{clutter} is the power of the channel data of the reverberation clutter, and α is the scalar for the reverberation clutter in order to achieve the desired SCR. A region of background was used to calculate the power for each phantom so that the amount of clutter added was consistent across the different magnitudes of cysts. Equation (2.6) can then be rewritten to solve for α as

$$\alpha = \sqrt{\frac{P_{\text{SOI}}}{P_{\text{clutter}} 10^{\text{SCR}/10}}}. \tag{2.7}$$

The scalar α is then calculated and the clutter channel data is scaled and combined with the phantom channel data. In our works, the reverberation clutter is often scaled against the background speckle to achieve realizations with 20, 10, 0, -10, and -20 dB SCR compared to the channel data. We believe these values

cover a realistic range of *in vivo* possibilities, though we expect most clinical cases probably fall between 20 dB (easy-to-image patients) and 0 dB (difficult-to-image patients). However, there is a significant lack of research into the quantification of reverberation, making it difficult to know the level of reverberation clutter for any given scenario. Some groups have investigated reverberation clutter, but mainly do so in the context of harmonic imaging (Pinton et al., 2011; Fatemi et al., 2019). Studies involving the addition of simulated reverberation clutter fall within our simulated range (Bell et al., 2015; Fatemi et al., 2019), and one *in vivo* study found that bladder images had bladder wall to clutter ratios between 30 and 0 dB for all sources of clutter including reverberation (Lediju et al., 2008), further supporting our choices.

2.2.3.3 CRDR for Common Beamformers

Fig. 2.4 shows a summary of example results from Schlunk et al. (Schlunk et al., 2021). In the original work by Dei et al., they found that DAS and MV reported good contrast accuracy over a range of 55dB when there is no additional noise present. However, this drops dramatically as strong reverberation clutter is introduced such that the accurate range is only 27dB in the presence of 0dB SCR reverberation clutter. ADMIRE (HDF) is shown to produce similar results as DAS in the noise-free cases, and improves upon DAS in the presence of reverberation clutter, reporting an accurate range of 36dB (Dei et al., 2017).

2.2.4 Lateral Point Spread Function and Resolution

The lateral point spread function (PSF) can be a valuable tool not only to get a measure of the resolution of a beamformer (which can be related to the sizing accuracy of said beamformer), but also as an indicator of the presence of sidelobe artifacts and dark region artifacts. Though the full width at half maximum (FWHM) would not change with the intensity of the target, the appearance of the PSF will have higher sidelobes when the sidelobe artifact is present, and likewise lower sidelobes when a dark region artifact is present. Though the PSF is frequently measured for beamformers to assess lateral resolution performance, the sizing accuracy is less often considered or evaluated.

CHAPTER 3

Addressing Contrast Accuracy: Iterative ADMIRE

This chapter is adapted from work originally published in (Schlunk et al., 2021), ©2021 IEEE. Reprinted, with permission, from Schlunk, S., Dei, K., & Byram, B. (2021). Iterative Model-Based Beamforming for High Dynamic Range Applications. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 68(3), 482–493. https://doi.org/10.1109/TUFFC.2020.3012165.

3.1 Introduction

Adaptive beamforming methods have become increasingly popular in recent years due to the reportedly great improvements to contrast ratio or contrast-to-noise ratio (CNR) that they can produce (Matrone et al., 2015; Li and Li, 2003; Asl and Mahloojifar, 2009). However, there is growing concern that not all of these methods produce "true" improvements to contrast ratio and CNR. Because of this concern, some groups have begun expressing the importance of developing new robust measurements to test performance (Rindal et al., 2016; Dei et al., 2017; Rodriguez-Molares et al., 2018, 2020; Rindal et al., 2019). These methods include contrast ratio dynamic range (CRDR) (Dei et al., 2017), which seeks to quantify measured versus true contrast, and generalized contrast-to-noise ratio (gCNR) (Rodriguez-Molares et al., 2018, 2020), which seeks to generalize CNR by providing a detectability metric that is less susceptible to artificial improvements.

Many adaptive beamformers characterize their performance based on the contrast ratio of anechoic cysts, where the average target intensity is compared to some nearby background region. However, sidelobe clutter as well as the dark region artifact described by Rindal et al. (Rindal et al., 2017) can interfere with this assessment. The dark region artifact can arise in the presence of strong acoustic targets, where sidelobe clutter will obscure nearby, lower strength acoustic targets. Some adaptive beamformers will remove this clutter, but in doing so also remove the underlying signal of interest, resulting in the dark region artifact. This lowers the intensity in that region leading to a reported contrast ratio or CNR that is higher than its true value. Alternatively, other beamformers may fail to remove this sidelobe clutter, resulting in cysts of varying intensities appearing identical (e.g. an anechoic cyst measured as -40 dB), as demonstrated by Dei et al. with CRDR (Dei et al., 2017).

We provide several instances where these problems can occur clinically. First, renal cysts are quite common and are classified as either simple (benign, homogeneous, fluid-filled with a well-defined outline) or complex (indicative of disease, filled with non-homogeneous fluid and an irregular cyst wall with potential calcifications) (Eknoyan, 2009). However, complex cysts can be difficult to classify (Warren and McFarlane,

2005; Siegel et al., 1997), and many beamformers will display a -40 dB complex cyst identically to a anechoic simple cyst, either due to removing the weak signal from the complex cyst or failing to remove the sidelobe clutter from the simple cyst. A second example is the visibility of blood flow. Because the blood signal is typically much weaker than the surrounding tissue signal, this can lead to the blood signal being completely masked by sidelobe clutter, or worse the blood signal being completely removed along with the clutter.

Since the dark region artifact can appear in situations where nearby targets have severe differences in acoustic strength, it is an important consideration for high dynamic range applications. In previous papers, we introduced the aperture domain model image reconstruction (ADMIRE) method (Byram et al., 2015a; Dei and Byram, 2017; Byram and Jakovljevic, 2014) that uses a physics-based model to reduce strong sidelobe clutter and reverberation, which increased the dynamic range (Dei et al., 2017). However, it is susceptible to the dark region artifact (Dei and Byram, 2018), where the strong sidelobe clutter signal is preferentially fit by the model, resulting in the complete loss of the weaker underlying signal. In this paper we introduce iterative ADMIRE (iADMIRE) as a solution to that problem. In iADMIRE, sources of clutter are iteratively solved and subtracted from the input signal so that weaker signals can progressively be fit. This results in a more accurate estimate of the region of interest. We will show that this iterative approach can mitigate both sidelobe and dark region artifacts, improving CRDR in some cases.

3.2 Beamforming Algorithms

We include a brief explanation of ADMIRE in order to provide context for iterative ADMIRE. Additionally, several other adaptive beamformers are described for comparison, since many of these beamformers are susceptible to sidelobe artifacts, dark region artifacts, or both. All beamformers are implemented in MATLAB (The MathWorks, Natick, MA, USA).

3.2.1 Aperture Domain Model Image Reconstruction (ADMIRE)

Aperture domain model image reconstruction (ADMIRE) is a method for removing reverberation and off-axis clutter, as well as suppressing wavefront aberration. We discuss the method in detail in Chapter 2.

3.2.2 Iterative ADMIRE (iADMIRE)

Iterative ADMIRE (iADMIRE) is a modification to the ADMIRE algorithm that seeks to accommodate environments with high dynamic ranges by mitigating the shortcomings of the elastic-net (Schlunk et al., 2018). To provide some intuition for these problems, we include in Fig. 3.1, an example of how L1, L2, and elastic-net regularizations affect the estimate of a given coefficient. With an L1 estimate, coefficients that are sufficiently small will be ignored and set to zero, but larger coefficients will be fit relatively accurately,

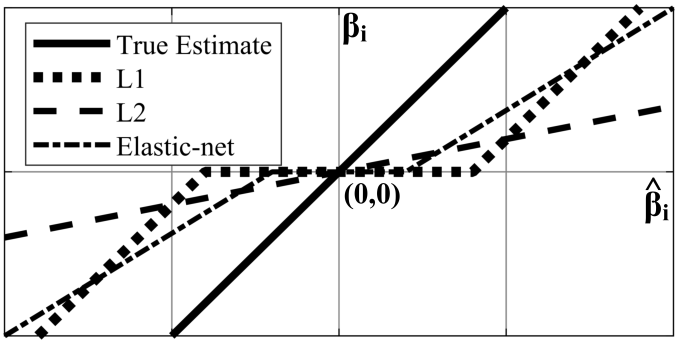

Figure 3.1: Qualitative example of the L1, L2, and elastic-net estimates for some specific model coefficient, β_i. In the L1 case, coefficients that are sufficiently small are zeroed in favor of larger ones. No coefficients are zeroed in L2, but instead there is significant coefficient shrinkage (loss of amplitude) for stronger components. ADMIRE uses the elastic-net, which balances the two extremes but includes features of both. Figure ©2021 IEEE.

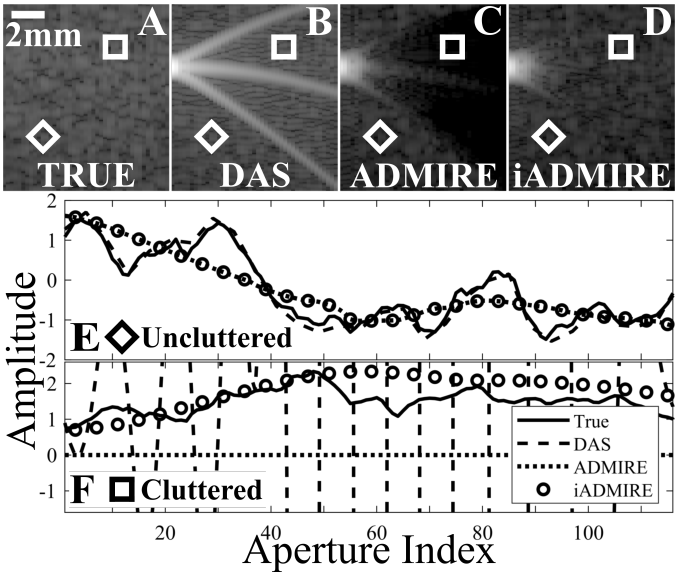

Figure 3.2: Example of the effect of iADMIRE on the dark region artifact. (A) DAS showing the uncluttered, True speckle background. (B) DAS of the same background with a simulated 100 dB bright scatterer introducing strong sidelobe clutter. (C) ADMIRE and (D) iADMIRE processing of the scatterer. ADMIRE demonstrates sidelobe reduction with a visible dark region artifact, and iADMIRE shows the same reduction without the artifact. (E) shows the aperture domain signals for the uncluttered region denoted by the diamond, and (F) shows the signals for the cluttered region denoted by the square. When no clutter is present, ADMIRE and iADMIRE match the True and DAS signals since there is no clutter to be removed (note that the slight difference in signal is due to the sparse fitting of coefficients and has marginal impact on the final image). With sidelobe clutter present, ADMIRE removes all signal in the region, while iADMIRE is able to restore the underlying signal. Figure ©2021 IEEE.

leading to coefficient selection. With L2, no coefficients are lost, but the cost is that larger coefficients undergo severe shrinkage and are underestimated. By incorporating characteristics of both, the elastic-net seeks to minimize the negatives of each method, but crucially does not eliminate them. Since ADMIRE uses the

Algorithm 1: Iterative clutter removal in ADMIRE

Given model predictors $X = [X_{\text{ROI}}\ X_{\text{clutter}}]$, aperture domain signal y_1, parameters α and λ, and $\delta > 0$

for $i = 1$ **do**

Solve (2.4) for model coefficients $\hat{\beta}_i$, given y_i, X

Compute clutter-only signal

$$y_{i,\text{clutter}} = X_{\text{clutter}}\hat{\beta}_{i,\text{clutter}} \tag{3.1}$$

Compute new aperture signal

$$y_{i+1} = y_i - y_{i,\text{clutter}} \tag{3.2}$$

Stop when $||y_{i+1} - y_i||_2^2 < \delta$

end for

Calculate $y_{\text{decluttered}} = X_{\text{ROI}}\hat{\beta}_{i,\text{ROI}}$

elastic-net and therefore has L1 features, we see that weaker signal components may be zeroed out in favor of stronger sources. This can result in a dark region artifact since only the strong clutter coefficients are fit, resulting in nothing from the region of interest being included. Fig. 3.2 shows an example of a speckle background (A) that is corrupted by a 100 dB scatterer (B) and highlights an uncluttered region (diamond) and a cluttered region (square). For the uncluttered region, the aperture domain signals for all four match closely since there is no clutter interference (note that the small differences between the DAS and ADMIRE signals is due to the elastic-net estimates of ADMIRE favoring the dominant signal components, and has little impact on the actual image). In contrast, the DAS signal in the cluttered region is completely different due to the sidelobe clutter, and ADMIRE predominantly fits that clutter resulting in a loss of the underlying signal and the creation of the dark region artifact. iADMIRE, however, first fits and removes that clutter signal, allowing it to recover the True underlying signal.

Algorithm 1 shows the process by which iADMIRE iteratively solves for the clutter sources in the signal and removes them, theoretically obtaining a more accurate estimate of the ROI signal. Rather than doing a single solve for the model coefficients and reconstructing only the ROI signal, iADMIRE computes the clutter-only signal y_{clutter} using (3.1) and the coefficients β_{clutter} corresponding to the clutter predictors in the model and subtracts that from the original signal with (3.2). By iteratively applying the elastic-net and removing the reconstructed clutter signal from the signal, strong clutter sources are continually removed until some threshold is reached. This new, less cluttered signal is then decomposed using the elastic-net one last time to produce the decluttered signal using (2.3).

We set $\delta = \max(\text{abs}(y))/1000$ so that convergence will depend on the input aperture domain signal. Fig. 3.3 shows a simple example of convergence for a dark artifact region. At zero iterations (normal ADMIRE), the dark region artifact is clearly visible, but after a single iteration the background appears to be mostly

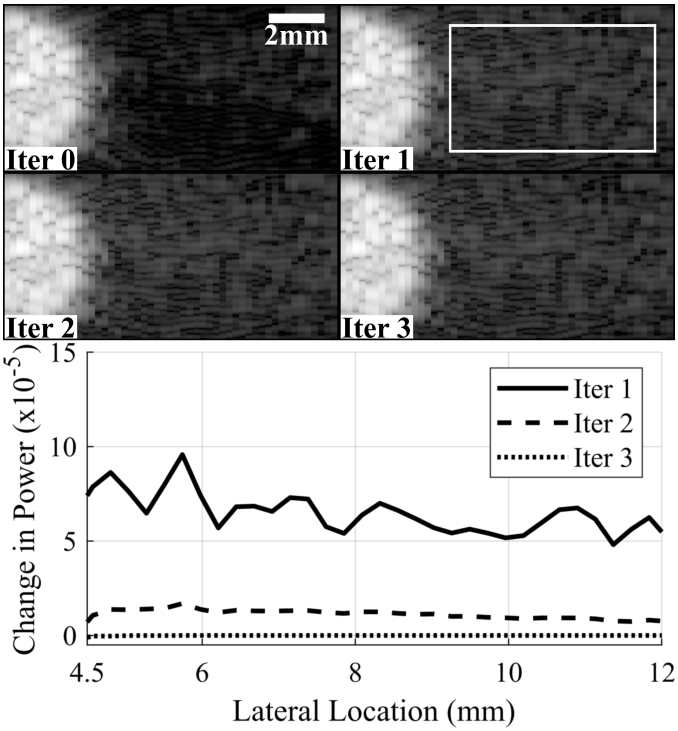

Figure 3.3: Example of iADMIRE after 0, 1, 2, and 3 iterations for a 60 dB cyst. The change in power between iterations of the indicated region is included. At 0 iterations (normal ADMIRE) the dark region artifact is clearly visible, but is almost entirely mitigated after the first iteration. In this example, iADMIRE has mostly converged by the third iteration. Figure ©2021 IEEE.

restored. Based on the change in power between iterations, the algorithm has functionally converged by the third iteration. For simple clutter scenarios where there is one main source of interference, it is likely that only one or a few iterations is required to remove that source. As a result, for this study we enforce a hard cutoff of three iterations to prevent wasted computation time. For ease of comparison, the other parameters that overlap with the standard ADMIRE algorithm are the same as those for ADMIRE (HDF) in the previous section. The balance between L1 and L2 fitting could be further tuned to adjust convergence and clutter removal, however, we've found that these values provide a fairly good balance to prevent overfitting (and failing to remove clutter) or underfitting (requiring more iterations).

Computation time for iADMIRE varies wildly depending on the number of iterations required. Since the elastic-net decomposition is computationally expensive, a window that requires many iterations (and therefore many repeated fits), requires nearly that much additional processing time. As a result, for any given signal being fit, it can take between one and three times as long to process as standard ADMIRE.

3.2.3 Delay-and-Sum (DAS)

The conventional standard beamformer, as defined in Section 1.4.2. In this study, we use rectangular weighting, unless otherwise specified.

3.2.4 Filtered Delay-Multiply-and-Sum (F-DMAS)

Filtered delay-multiply-and-sum (F-DMAS) is a simple adaptive beamformer proposed to improve contrast and resolution (Matrone et al., 2015). It works by combinatorially coupling and multiplying the delayed channel data before summing across the channels. The complete definition can be found in Section 1.4.3.

3.2.5 Generalized Coherence Factor (GCF)

Generalized coherence factor (GCF) is a weighting of the DAS image by the ratio of the energy in some low-frequency region to the total energy (Li and Li, 2003). The complete definition can be found in Section 1.4.4. We chose the low-frequency cutoff parameter $M_0 = 5$ as done in (Rindal et al., 2017) for this work.

3.2.6 Minimum Variance

Minimum variance (MV) beamforming is also an adaptively weighted DAS image, where the optimized apodization weights (Holfort et al., 2009), (Synnevåg et al., 2007) are calculated to minimize off-axis interference, improving lateral resolution. The complete definition can be found in Section 1.4.5. To get an invertible matrix, we used subarray lengths of L, equal to 50% of the total aperture length, to estimate the covariance matrix and diagonal loading defined as $\varepsilon = \Delta \cdot \mathrm{tr}(\hat{R})$, where $\Delta = 1/(10L)$, as recommended (Synnevåg et al., 2007) for this work.

3.2.7 Short-Lag Spatial Coherence (SLSC)

Short-lag spatial coherence (SLSC) is different from many other beamformers in that it only uses the spatial coherence of received echoes to form images. The coherence images are usually displayed on a linear scale. Introduced by Lediju et al. (Lediju et al., 2011), the foundation of SLSC is the van Citter-Zernike (VCZ) theorem, which demonstrates a prediction for the spatial coherence, or covariance, for the backscattered echoes (Mallart and Fink, 1991). The complete definition can be found in Section 1.4.6. For this study, we chose to use $L = 20$ lags (17% of our total number of channels), with a correlation kernel equivalent to 1 wavelength.

3.2.8 Gray Level Transformation (GLT)

For comparison purposes, we include the gray level transform (GLT) that is described by Rindal et al. as an example of a method that "fakes" improvements to CNR, but sacrifices the dynamic range (Rindal et al.,

2019). The GLT is a sigmoid function defined as

$$\hat{S}_{\mathrm{GLT}}(B) = \frac{1}{1+e^{-\alpha(B-\beta)}} \tag{3.3}$$

$$S_{\mathrm{GLT}}(B) = \frac{\hat{S}_{\mathrm{GLT}}(B) - \max(\hat{S}_{\mathrm{GLT}}(B))}{\varepsilon}, \tag{3.4}$$

where $B = 20\log_{10}(|S_{\mathrm{DAS}}|)$, $\alpha = 0.12$, $\beta = -40$, and $\varepsilon = 0.008$.

3.3 Methods

3.3.1 Contrast Target Phantom

Field II (Jensen, 1996; Jensen and Svendsen, 1992) was used to simulate 5 mm diameter cysts of known contrasts ranging from -50 dB to 70 dB relative to the background, plus an anechoic case. Each level of contrast was simulated with 6 independent realizations of speckle. Simulation parameters used in Field II are the same as those detailed in Table 2.2.

3.3.2 Simulated Reverberation Clutter

For each of the contrast phantoms, additional realizations were created that included simulated reverberation clutter. This reverberation clutter was simulated using the technique detailed by Byram and Shu (Byram and Shu, 2016b), (Byram and Shu, 2016a). The clutter was added directly to the channel data such that it satisfied the signal-to-clutter ratio (SCR) calculated by

$$\mathrm{SCR} = 10\log_{10}\left(\frac{P_{\mathrm{SOI}}}{\alpha^2 P_{\mathrm{clutter}}}\right), \tag{3.5}$$

where P_{SOI} is the power of the channel data of the original phantom, P_{clutter} is the power of the channel data of the reverberation clutter, and α is the scalar for the reverberation clutter in order to achieve the desired SCR. A region of background was used to calculate the power for each phantom so that the amount of clutter added was consistent across the different magnitudes of cysts. Equation (3.5) can then be rewritten to solve for α as

$$\alpha = \sqrt{\frac{P_{\mathrm{SOI}}}{P_{\mathrm{clutter}} 10^{\mathrm{SCR}/10}}}. \tag{3.6}$$

The scalar α is then calculated and the clutter channel data is scaled and combined with the phantom channel data. In this paper, the reverberation clutter is scaled against the background speckle to achieve realizations with 20, 10, 0, -10, and -20 dB SCR compared to the channel data. We believe these values cover a realistic range of *in vivo* possibilities, though we expect most clinical cases probably fall between 20 dB (easy-to-

image patients) and 0 dB (difficult-to-image patients). However, there is a significant lack of research into the quantification of reverberation, making it difficult to know the level of reverberation clutter for any given scenario. Some groups have investigated reverberation clutter, but mainly do so in the context of harmonic imaging (Pinton et al., 2011; Fatemi et al., 2019). Studies involving the addition of simulated reverberation clutter fall within our simulated range (Bell et al., 2015; Fatemi et al., 2019), and one *in vivo* study found that bladder images had bladder wall to clutter ratios between 30 and 0 dB for all sources of clutter including reverberation (Lediju et al., 2008), further supporting our choice.

3.3.3 Bright Scatterer Phantom

In order to measure the dark region artifact, we used Field II to simulate bright point targets at the focal depth with contrasts ranging from 40 dB to 120 dB relative to the background speckle. As before, 6 speckle realizations were created for each amplitude. The same parameters found in Table 2.2 were used for simulation.

3.3.4 *In vivo* Carotid Artery Data

We additionally captured *in vivo* carotid artery data from a healthy individual using a Verasonics Vantage Ultrasound System (Verasonics, Inc., Kirkland, WA) with a L12-5 linear transducer. A center frequency of 7.813 MHz was used to acquire 8 equally spaced angled plane waves spanning -9° to 9° at a PRF of 3000 Hz. Coherent compounding was used to generate a synthetically focused channel data set (Montaldo et al., 2009). The resulting frames were processed using either ADMIRE or iADMIRE and filtered using a 6th order Chebyshev filter with a cutoff frequency of 300 Hz to create a power Doppler image.

3.3.5 Contrast Ratio, CNR, and Resolution

The contrast ratio and contrast-to-noise ratio (CNR) were measured for each phantom in order to characterize the measured contrast and detectability of the simulated cyst. Contrast ratio and CNR were defined as

$$\text{contrast ratio} = 20\log_{10}\left(\frac{\mu_{\text{ROI}}}{\mu_{\text{background}}}\right) \tag{3.7}$$

$$\text{CNR} = \frac{|\mu_{\text{ROI}} - \mu_{\text{background}}|}{\sqrt{\sigma^2_{\text{ROI}} + \sigma^2_{\text{background}}}}, \tag{3.8}$$

where μ is the mean value and σ is the standard deviation of the enveloped, but not log compressed, data. Fig. 3.4 shows the chosen ROI and background regions for the (A) contrast and (B) bright scatterer phantoms. Additionally, the point spread function (PSF) was found for each bright scatterer and for each method. The mean point target width was measured across all scatterers (except 40 dB due to the noise floor) for each method at -6 dB (full width at half maximum), -10 dB, and -20 dB.

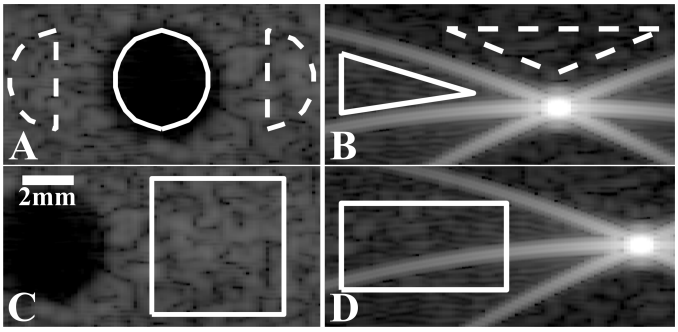

Figure 3.4: DAS images for hypoechoic contrast target phantoms (left) and bright scatterer phantoms (right) simulated in Field II. (A) and (B) indicate the target (solid) and background (dotted) regions used to calculate contrast ratio, CNR, and gCNR measurements. (C) and (D) show the regions used for computing the background speckle correlation of each beamformer. Figure ©2021 IEEE.

3.3.6 Contrast Ratio Dynamic Range (CRDR)

In an effort to better characterize true contrast in beamformers, Dei et al. proposed the contrast ratio dynamic range (CRDR) method (Dei et al., 2017). By applying a given beamformer to simulated cysts at varying intensities, in this case ranging from -50 to 70 dB plus an anechoic cyst, the measured contrast can be compared to the true contrast over a wide range. For a given realization, the measured contrast is then linearly interpolated to get a well sampled curve. The CRDR measurement can then be estimated from that measured versus true contrast curve by determining the range on the curve for which the measured contrast does not deviate from the true contrast (two-tailed t-test, $\alpha = 0.05$).

3.3.7 Generalized Contrast-to-Noise Ratio (gCNR)

An alternative to CNR for measuring detection probability was proposed by Rodriguez-Molares et al. called the generalized contrast-to-noise ratio (gCNR) (Rodriguez-Molares et al., 2018), (Rodriguez-Molares et al., 2020). It works by measuring the overlap of the probability density function between the ROI and the background, meaning it is unaffected by stretches or compressions in the dynamic range. The gCNR is then defined by

$$\text{gCNR} = 1 - \text{OVL}, \tag{3.9}$$

where OVL is the measured overlap between the two probability density functions. Therefore, gCNR varies between 0 and 1, where gCNR=1 indicates perfect discrimination. The gCNR is computed for the contrast target phantoms using the same ROI and background regions as for contrast ratio (Fig. 3.4). We used 100 bins as suggested in the original work.

3.3.8 Speckle Correlation

The 2-D correlation coefficient r was calculated by comparing each simulated phantom to its corresponding speckle realization in DAS (with no cyst or bright scatterer present) to determine the degree of speckle corruption from the sidelobe and dark region artifacts. The region used for the measurement is shown in Fig. 3.4. The correlation was calculated as

$$r = \frac{\sum_x \sum_z (S(x,z) - \mu_S)(D(x,z) - \mu_D)}{\sqrt{(\sum_x \sum_z (S(x,z) - \mu_S)^2)(\sum_x \sum_z (D(x,z) - \mu_D)^2)}}, \tag{3.10}$$

where S is the enveloped, uncompressed region of the data for the beamformer of interest, and D is short-hand for the reference DAS enveloped, uncompressed region of data. Values of r closer to 1 indicate more accurate speckle in the target region.

3.4 Results

3.4.1 Contrast Target Phantoms

The dynamic range for all beamformers in the presence of no added reverberation clutter and the computed CRDR for a range of signal-to-clutter ratios (SCR) is shown in Fig. 3.5. Since CRDR aims to quantify how accurately a beamformer represents contrast, a higher CRDR corresponds to a longer range over which a beamformer reports the true contrast. SLSC and GLT notably deviate from the true contrast line since they are designed to enhance CNR at the cost of CRDR, which they do as shown in Fig. 3.6B. Table 3.1 shows the measured CRDR for all beamformers at each level of added reverberation clutter from 20 dB (very low clutter) to -20 dB (very high clutter). We also include versions of DAS, ADMIRE (HDF), and iADMIRE summed using Hamming apodization since DAS especially benefits in this scenario. Though this does significantly improve the CRDR for DAS, it still does not match the performance of iADMIRE, apodized or unapodized. Overall, for signal-to-clutter ratios between 20 dB and 0 dB, iADMIRE continues to have the highest CRDR, but at more intense levels of reverberation all beamformers suffer poor CRDR.

One realization of a simulated 60 dB contrast target phantom without added reverberation clutter is shown in Fig. 3.6A. The images are displayed on a compressed log scale of 120 dB in order to highlight the impact of the strong sidelobes. We can qualitatively see the sidelobe artifacts in DAS, DMAS, GCF, and MV, which has a noticeable affect on the dynamic range as shown. We can also see the affect that SLSC and GLT have on the background speckle near the cyst, which provides a great boost to CNR, as reported in Fig. 3.6B. SLSC vastly outperforms the other methods due to this suppression of speckle in the background, with GLT performing second best for similar reasons. For all magnitudes of cysts, SLSC and GLT have better than or

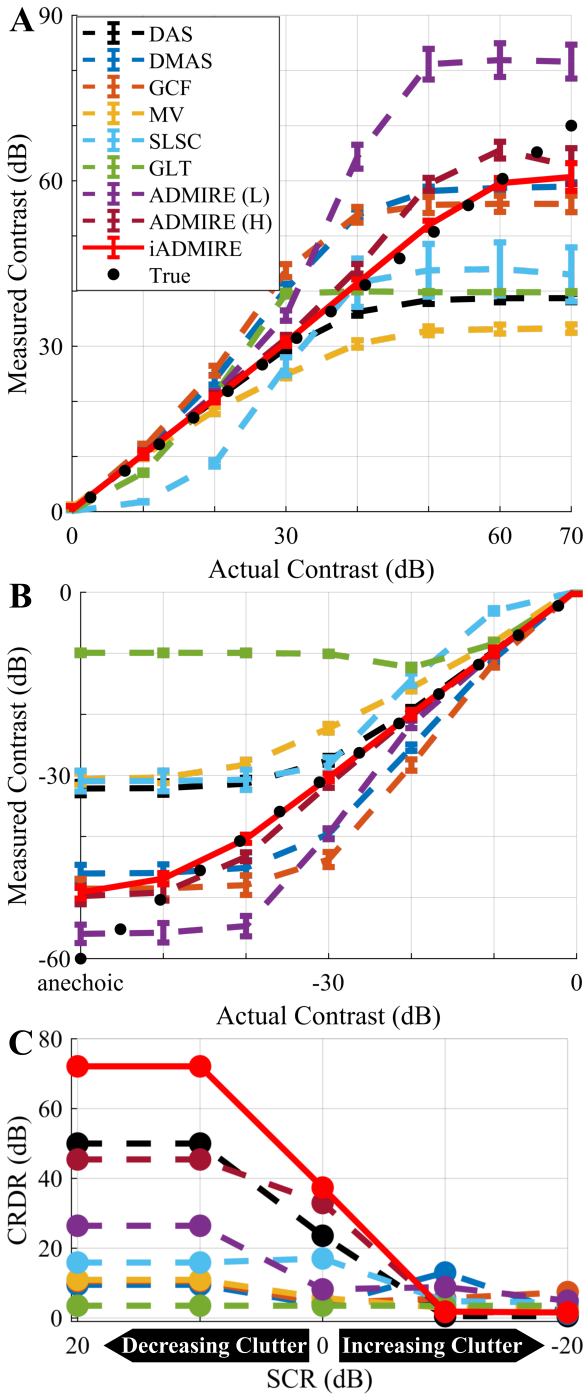

Figure 3.5: Dynamic range displayed as measured contrast versus actual contrast for (A) hyperechoic and (B) hypoechoic cysts. The dotted black line indicates the true contrast, and adherence to that line indicates more accurate reporting of measured contrast. (C) The computed contrast ratio dynamic range (CRDR) for a range of signal-to-clutter ratios (SCR). Included error bars indicate the standard deviation. Figure ©2021 IEEE.

equal CNR to the other beamformers. In comparison, for gCNR shown in Fig. 3.6C, all methods are fairly similar at all levels of contrast.

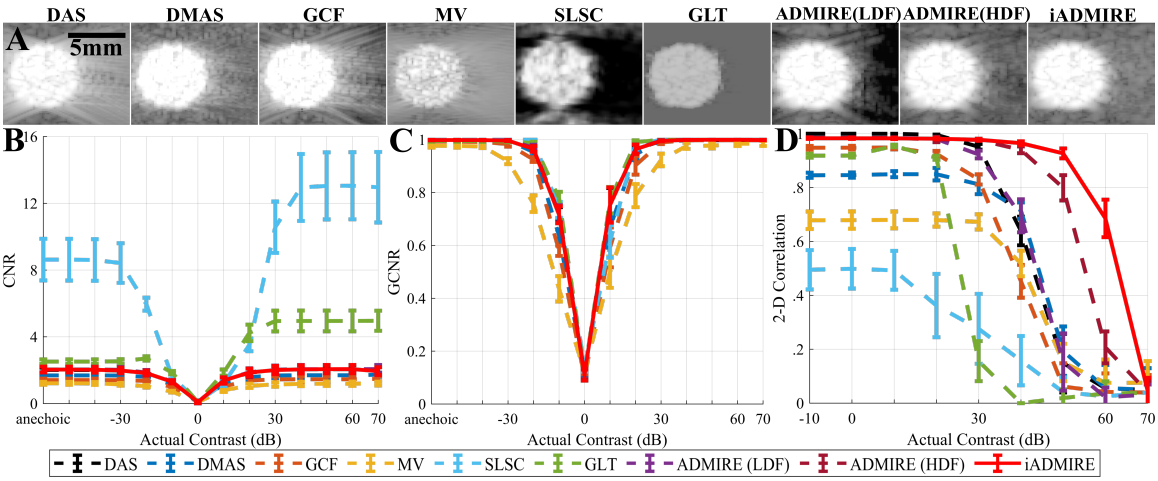

Figure 3.6: (A) Sample realization of each beamformer for a 60 dB hyperechoic cyst. All beamformers are displayed on a 120 dB log scale in order to highlight the strong sidelobes (where present), except SLSC, which is displayed linearly, as is the convention. (B) CNR, (C) gCNR, and (D) 2-D speckle correlation calculated for the full range of cysts without added reverberation clutter. Included error bars indicate the standard deviation. Images are normalized so that background speckle is at the same level across methods. Figure ©2021 IEEE.

Table 3.1: Contrast ratio dynamic range (CRDR) for different levels of added reverberation clutter. Table ©2021 IEEE.

	Contrast ratio dynamic range (dB)				
Signal-to-clutter ratio (dB)	20	10	0	-10	-20
DAS	50.0	50.0	23.6	0.5	0.5
DAS - Hamming	60.8	60.8	23.3	1.7	1.6
DMAS	9.5	9.5	3.6	**13.0**	1.2
GCF	10.5	10.5	4.3	5.6	**7.4**
MV	10.9	10.9	5.7	2.9	2.9
SLSC	15.9	15.9	17.0	4.8	4.4
GLT	3.5	3.5	3.5	3.5	3.5
ADMIRE (LDF)	26.4	26.4	8.2	8.8	4.9
ADMIRE (HDF)	45.4	45.4	32.9	1.7	1.6
ADMIRE (HDF) - Hamming	46.8	46.8	30.7	2.9	2.6
iADMIRE	72.1	72.1	37.3	1.8	1.6
iADMIRE - Hamming	**75.6**	**75.6**	**41.4**	2.6	2.6

Fig. 3.6D shows the 2-D correlation coefficient of the defined speckle region for all beamformers at all contrast levels compared to the corresponding speckle realization in DAS at 0 dB (no cyst), which should be the true speckle. Again, due to the suppression of speckle, SLSC and GLT have significantly reduced speckle accuracy, and all methods begin to lose accuracy at contrast levels greater than 30 dB due to the effect of the sidelobe artifact. In the low reverberation clutter cases, ADMIRE (HDF) has speckle accuracy above 0.90 for cysts up to 50 dB, and iADMIRE performs similar up to 60 dB. However, for increased levels of reverberation

all the methods are reduced to nearly 0 correlation as the reverberation clutter completely masks the original speckle.

3.4.2 Bright Scatterer Phantoms

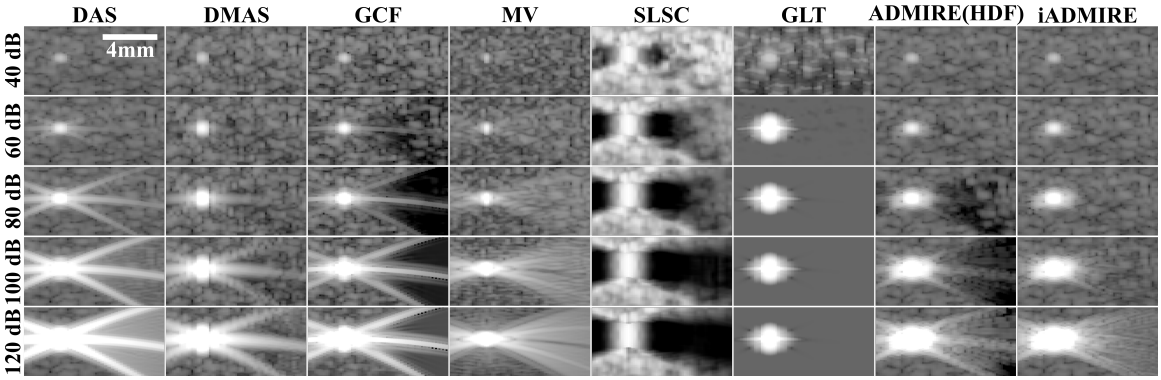

Figure 3.7: Sample realization of bright scatterer phantoms at 40, 60, 80, 100, and 120 dB. All beamformers are displayed on a 100 dB log scale to highlight sidelobes and background speckle, except SLSC, which is displayed linearly, as is the convention. Images are normalized so that background speckle is at the same level across methods. Figure ©2021 IEEE.

A realization of the bright scatterer phantoms is shown in Fig. 3.7. Since ADMIRE (HDF) outperforms ADMIRE (LDF) in terms of mitigating the dark region artifact, we have only included the better variation. Sidelobe artifacts are present in DAS, DMAS, GCF, and MV, and begin to show in ADMIRE at higher bright scatterer contrasts. The dark region artifact is strongly present in the GCF, SLSC, and ADMIRE images, and somewhat present in the DMAS images. GLT does not obviously exhibit either artifact, but that is primarily due to the background speckle being completely suppressed by the operation. iADMIRE shows reduction in the sidelobe artifacts, but also the removal of the dark region artifact and restoration of the background speckle in cases up to 100 dB.

Fig. 3.8A shows the contrast in the area of the dark region artifact compared to the uncorrupted background, and Fig. 3.8B shows the calculated speckle correlation for the area impacted by the two artifacts. As expected from the included realization, DMAS demonstrates improved contrast in the dark artifact region due to somewhat suppressed sidelobes and minimal dark region artifact, though iADMIRE has better overall artifact mitigation. For the 120 dB case, all beamformers demonstrated severe artifacts either due to sidelobes or the dark region artifact. Several methods have speckle accuracy greater than 0.75 for cases at or less than 80 dB, but only iADMIRE has accuracy above 0.90 in all cases up to 100 dB. Finally, Fig. 3.8C shows the point spread function (PSF) for a realization of the 60, 80, and 100 dB bright scatterers, and Table 3.2 shows the corresponding mean point target width. For all levels, ADMIRE and iADMIRE have comparable or better resolution compared to DAS, but MV consistently has the best resolution among the beamformers.

Table 3.2: Mean Point Target Width at Varying Levels. Table ©2021 IEEE.

	Point Target Width (mm)		
Intensity (dB)	-6 (FWHM)	-10	-20
DAS	0.49±.00	0.68±.00	0.94±.00
DMAS	0.45±.00	0.59±.00	0.77±.00
GCF	0.35±.00	0.49±.00	0.77±.00
MV	**0.23±.00**	**0.30±.00**	**0.49±.00**
SLSC	2.04±.00	2.48±.00	3.14±.01
GLT	0.82±.01	0.99±.01	1.52±.01
ADMIRE (HDF)	0.48±.01	0.68±.00	0.89±.00
iADMIRE	0.50±.01	0.68±.01	0.89±.01

3.4.3 *In vivo* Carotid Artery

The top row of Fig. 3.9 shows the B-mode of the carotid artery beamformed using DAS, ADMIRE (HDF), and iADMIRE, as well as a difference image to show the changes between DAS and iADMIRE. The middle rows show the dashed white line region enlarged and beamformed with rectangular and Hamming apodization to highlight the high clutter in DAS with rectangular apodization. ADMIRE (HDF) removes this clutter entirely, presenting a dark region artifact in the same location, and iADMIRE lessens the clutter and preserves some of the underlying speckle. The difference images show iADMIRE with rectangular apodization compared to DAS with rectangular and Hamming apodization. Note that the middle rows are displayed on a smaller dynamic range to make the changes between speckle and clutter more visible. Finally, a power Doppler image is shown for each method.

3.5 Discussion

3.5.1 Contrast Target Phantoms

The contrast ratio dynamic range (CRDR) demonstrates the issue of saturation due to the interference of the sidelobe artifact, where cysts at different magnitudes appear identical on a more traditional compressed log scale. In the results, the dynamic range curves show that all beamformers eventually saturate for cysts above some level of contrast, both hyperechoic and hypoechoic, due to sidelobe clutter. In low reverberation clutter cases, iADMIRE is able to more accurately report the true contrast in the range of -40 to 60 dB cysts compared to other beamformers, improving upon ADMIRE and DAS. ADMIRE (HDF) and iADMIRE perform well primarily due to their ability to mitigate the sidelobe artifact without introducing a dark region artifact. However, at higher levels of sidelobe clutter, ADMIRE (HDF) begins to struggle to preserve the true background speckle resulting in a dark region artifact, while iADMIRE is able to differentiate between the clutter and the background speckle more effectively.

We characterized the effect of the sidelobe artifact on the background speckle by computing the 2-D correlation coefficient for all methods at all levels of contrast against an untouched speckle background (no cyst or additional clutter). For hypoechoic cysts, the background region is unaffected since there is no strong acoustic target. However, for the strong acoustic cysts at 30 dB and greater, many of the beamformers begin to suffer reduced speckle accuracy. Only ADMIRE (HDF) and iADMIRE manage to maintain above 0.90 speckle accuracy at 50 dB, and only iADMIRE preserves speckle accurately at 60 dB. As with CRDR, speckle correlation drops significantly as the level of added reverberation clutter rises above the background speckle.

From the CNR results, SLSC is the clear choice if cyst detectability is the primary focus. Since SLSC takes advantage of the lack of coherence in the sidelobe clutter, it is able to dramatically reduce the amplitude of the background in those areas, increasing the contrast and decreasing the variance. GLT functions similar, though to a lesser extreme. All the other beamformers have similar performance, with DAS and the ADMIRE variants performing nearly identically. gCNR, on the other hand, only compares the probability distribution of the magnitudes of the cyst versus the background, meaning that given a fixed speckle variance, there is some level of contrast above which all cysts will have a gCNR of 1, regardless of beamformer. Because of this, gCNR does not value the difference in contrast or the speckle variance beyond a certain level, and therefore does not fully differentiate between beamformers in this study. The gCNR results argue that the improved CNR that SLSC and GLT are getting from suppressing the background is not functionally improving detectability more than the other beamformers, which is fair since the cysts are extremely easy to differentiate from the background, regardless of which method is used.

The added reverberation clutter results show that iADMIRE is better able to correctly mitigate sidelobe clutter for reverberation cases up to 0 dB SCR, but once the reverberation clutter is sufficiently strong it begins to fail along with the other beamformers. At -10 and -20 dB SCR, all of the beamformers failed to represent true cyst contrast due to the complete corruption of the background speckle. This indicates that ADMIRE is more robust to high levels of sidelobe clutter than reverberation clutter due to the nature of the two sources. Specifically, off-axis signals have intrinsically lower correlation with the ROI signals compared to reverberation signals, which makes it more difficult to fully separate high amplitude reverberant sources (Byram et al., 2015c). This issue could potentially be avoided by using tissue harmonic imaging, which has been demonstrated many times to reduce reverberation (Choudhry et al., 2000; Pinton et al., 2011; Fatemi et al., 2019) and has previously been successfully implemented with ADMIRE (Byram et al., 2015a).

When ADMIRE is tuned towards an L1 fit (encouraging sparsity of coefficients), the high magnitude signals (reverberation) are preferentially fit, and the weaker signals are lost. However, with a more L2 fit (encouraging more nonzero coefficients), the model again prefers the high frequency predictors to achieve a lower error fit to the input signal, resulting in the underlying signal being misclassified as clutter. For

the severe reverberation clutter cases, iADMIRE requires many iterations to progressively remove the high magnitude components first without overfitting the region of interest, but this causes compounding of small model fitting errors due to the many iterations.

3.5.2 Bright Scatterer Phantoms

Qualitatively, the bright scatterer phantom results show that ADMIRE and iADMIRE almost entirely mitigate the sidelobe artifact that is present in the DAS, DMAS, GCF, and MV images. Furthermore, iADMIRE mitigates the dark region artifact that begins to appear in ADMIRE (HDF) at 80 and 100 dB, restoring the lost background speckle. Not only is the magnitude of the speckle restored, but the 2-D correlation coefficient shows that it is restoring the speckle accurately. In the 100 dB bright scatterer case, iADMIRE reports a correlation coefficient of 0.93, with MV and DMAS a distant runner-up with coefficients of 0.42 and 0.40, respectively. Though none of the beamformers were able to perform well in the 120 dB bright scatterer case, this shows that iADMIRE is able to mitigate both sidelobe and dark region artifacts in cases up to 100 dB.

3.5.3 In vivo Carotid Artery

The data obtained from the carotid artery shows an example of the dark region artifact *in vivo*. ADMIRE (HDF), in the presence of the sidelobe clutter from the tissue wall, completely removes the blood speckle along with the clutter, resulting in a complete loss of the blood signal in that region. In comparison, iADMIRE is able to restore that lost blood signal and largely preserve the underlying speckle, which can be seen by the improved speckle appearance in the highlighted cluttered region compared to DAS with rectangular apodization. Hamming apodization clearly helps to remove the clutter signal from the vessel, but seems to remove more signal compared to iADMIRE with rectangular apodization. Given the nature of the *in vivo* situation, it is difficult to know whether the Hamming apodization is removing too much signal or not.

3.6 Conclusion

We have presented iADMIRE as a modification to the original ADMIRE algorithm that is able to extend its dynamic range as well as mitigate the dark region artifact that results from adaptive beamformers. Using CRDR and gCNR, we confirmed the performance of iADMIRE compared to other common adaptive beamformers, while also demonstrating the issues with regards to sidelobe artifact saturation and true contrast.

In the no added reverberation clutter cases, iADMIRE had a CRDR of 72.1 dB, improving upon the next best methods DAS and ADMIRE (HDF) at 50.0 and 45.4 dB, respectively. The added reverberation clutter cases proved more difficult, where iADMIRE loses some of its relative improvements as the level of reverberation increases. This indicates that iADMIRE has difficulty fully differentiating these highly reverberant

sources from the ROI signals. As Byram et al. showed previously, the physical model of ADMIRE loses accuracy in the extreme near field (Byram et al., 2015a) and additionally sources at different depths from the ROI tend to have higher correlation with the ROI compared to same-depth, off-axis sources (Byram et al., 2015c). This means that while iADMIRE is robust to off-axis clutter, differentiating between near-field and ROI signals is more of a challenge in these highly reverberant cases.

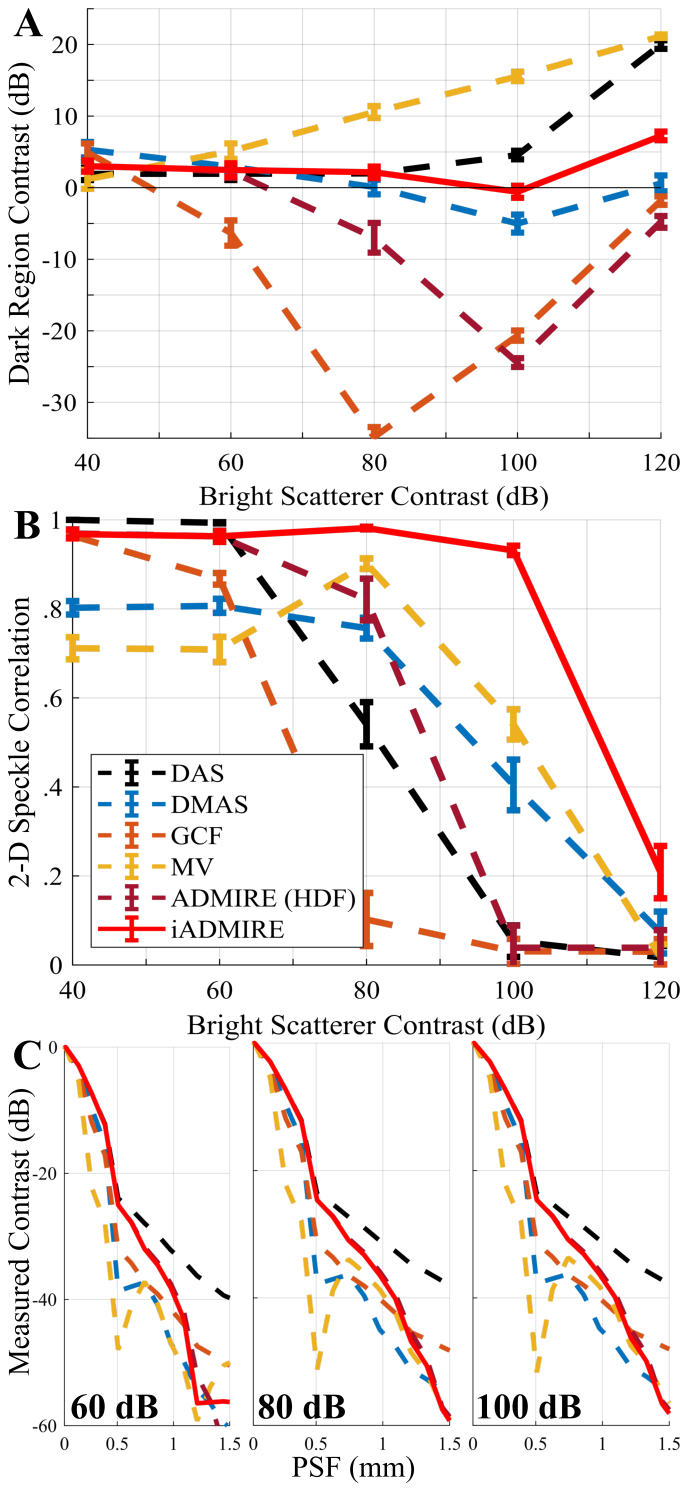

Figure 3.8: (A) Contrast of the dark region artifact versus the strength of the simulated bright scatterer. Negative contrast indicates the presence of a dark region artifact, while positive contrast indicates strong sidelobe clutter. Values close to 0 dB indicate low artifact interference. (B) 2-D speckle correlation in the area of the dark region artifact as compared to clean background speckle. Values close to 1 indicate better speckle accuracy. (C) Point spread functions (PSF) for the 60, 80, and 100 dB bright scatterers, displayed from the center of the target. Figure ©2021 IEEE.

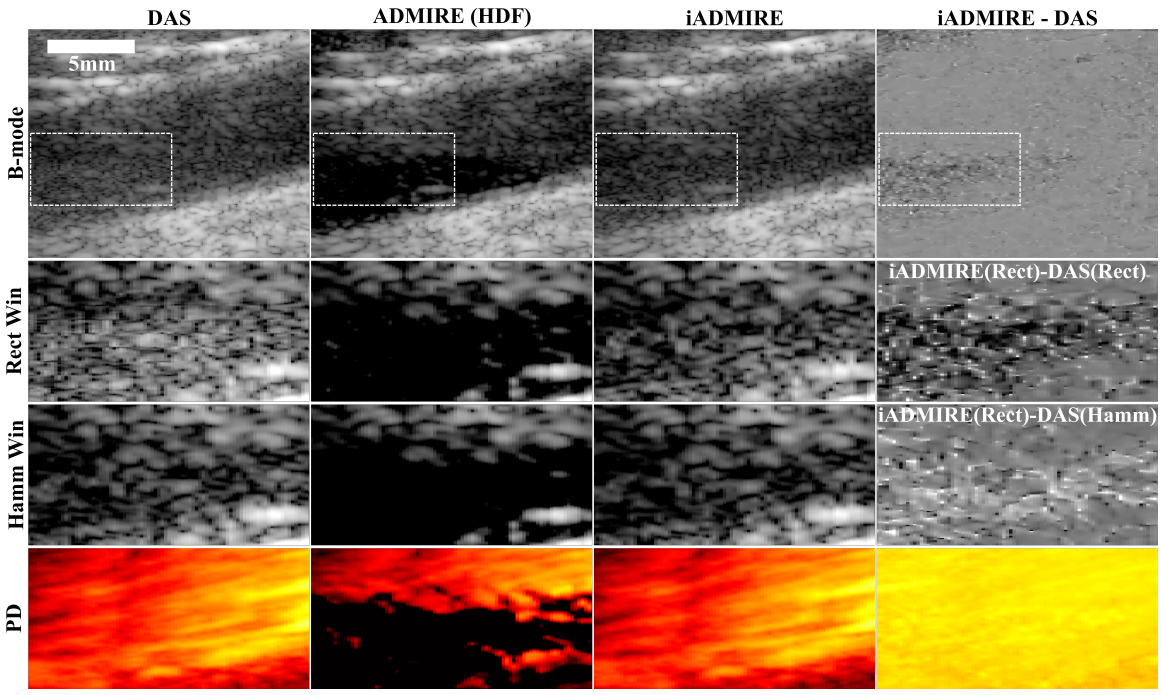

Figure 3.9: DAS, ADMIRE (HDF), iADMIRE, and the difference between iADMIRE and DAS images from an *in vivo* carotid artery from a healthy individual. The top row shows the standard B-mode images on a 70 dB log scale, and the middle and bottom rows show the dashed white line section from the B-mode enlarged. The middle rows show the same B-mode on a narrower 30 dB log scale using rectangular (Rect) and Hamming (Hamm) apodization, respectively, while the bottom row shows the power Doppler (PD) images on a 25 dB log scale. The lower dynamic range helps to highlight the improved speckle texture in the iADMIRE B-mode image. Figure ©2021 IEEE.

CHAPTER 4

Addressing Sizing Accuracy: Combining ADMIRE and MV

This chapter is adapted from work originally published in (Schlunk and Byram, 2022b), ©2022 IEEE. Reprinted, with permission, from Schlunk, S., & Byram, B. (2022). Combining ADMIRE and MV to Improve Image Quality. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 69(9), 2651–2662. https://doi.org/10.1109/TUFFC.2022.3194548.

4.1 Introduction

Many beamformers are designed with specific quality metrics in mind, such as improving contrast ratio or signal-to-noise ratio. Aperture domain model image reconstruction (ADMIRE) (Byram et al., 2015a; Dei and Byram, 2017; Byram and Jakovljevic, 2014) falls into this camp, with its use primarily intended for decluttering (reducing reverberation clutter, off-axis interference, and other sources of image degradation), which results in generally improved contrast ratio and signal-to-noise ratio. Other beamformers may seek to improve resolution, a well-known example of which is minimum variance (MV). MV is often an attractive beamformer, in part due to its rigorous mathematical basis, dating all the way back to Capon's original implementation in 1969 for use with seismic data (Capon, 1969). Though the applications of the original version were somewhat limited given the assumptions that are made, it has been adapted for use with ultrasound by many different groups (Sasso and Cohen-Bacrie, 2005; Synnevåg et al., 2007; Holfort et al., 2009; Synnevåg et al., 2009; Asl and Mahloojifar, 2010; Mehdizadeh et al., 2012; Deylami and Asl, 2018; Asl and Deylami, 2018; Diamantis et al., 2019; Wang et al., 2021; Salari and Asl, 2021), each looking to address various shortcomings of the original method. However, it is well known that MV still struggles in environments where signal-to-noise ratio (SNR) is low, and this is especially true in the case of reverberation clutter, as we will demonstrate in the next section.

We have previously shown in simple plane wave simulations with minimal reverberation that ADMIRE can be used prior to applying MV (Dei et al., 2016). ADMIRE preserves the dimensionality of the channel data, which means it can function as a pre-processing step, compared to most other beamformers that are strictly for post-processing. Since ADMIRE can improve SNR and remove troublesome sources of interference, this opens up the possibility that by processing with both ADMIRE and MV we can achieve a combined result that has both improved contrast ratio and lateral resolution. In this paper, we expand on our previous work (Dei et al., 2016; Schlunk and Byram, 2021b) by examining the assumptions of MV to shed some light on the issues particularly in high reverberation clutter environments, and implementing both a narrowband

and a broadband version of MV. We also include an expanded set of simulations with different sources of noise and more *in vivo* cases. We demonstrate that ADMIRE+MV can often produce better contrast ratio and better lateral resolution than ADMIRE or MV alone.

4.2 Background

4.2.1 Applying MV to Ultrasound Imaging

Minimum variance (MV) is mathematically designed to distinguish between a finite number of sources located in the far field (fewer than the number of array elements). For a chosen source, when the signal-to-noise ratio is high, the beamformer places nulls at the locations of other interfering plane waves (referred to as off-axis), minimizing their contribution and suppressing the interference (Capon, 1969; Van Trees, 2002). Though Capon's MV was formulated with passive sensing of seismic data in mind, conceptually the process is similar in ultrasound with proper beamforming. Echoes in ultrasound are caused by relatively near field scatterers compared to seismic signals, and so they appear as curved wavefronts at the transducer. But, by applying correct receive delays to the echoes near the transmit focus, the received echoes from these scatterers will take on the appearance of plane waves. Then the echoes from these scatterers can be differentiated based on the direction of propagation of the plane waves, as is done in the original work. In the ultrasound context, these off-axis plane waves are caused by scatterers at a similar depth to the focus, but not directly at the focus.

MV has proven itself a useful method for improving lateral resolution, though adaptation to ultrasound applications has required addressing some of the limitations and assumptions of the original method. Critics point out that Capon's MV is designed for narrowband applications in the far field, with low correlation between on and off-axis signals (Viola et al., 2008). Additionally, the covariance matrix, of which the inversion is required for the calculation, is often insufficiently robust when applied to ultrasound. That being said, these issues have been addressed and rectified by multiple groups.

Sasso and Cohen-Bacrie looked to solve the correlation problem by introducing a spatial smoothing preprocessing step to induce decorrelation, which works by analyzing subarrays and estimating a covariance matrix for each and averaging the set (Sasso and Cohen-Bacrie, 2005). The narrowband assumption does not *per se* invalidate the application (and many versions do not make considerations for it), but Holfort et al. specifically introduced a broadband implementation which breaks down the ultrasound signal in the frequency domain into a set of narrowband windows, over which the assumption is valid (Holfort et al., 2009). The far field assumption can be addressed with proper delaying of the output data, which flattens the incoming echoes. And finally, the lack of robustness of the covariance matrix has been addressed by many authors, and it usually consists of using spatial averaging and diagonal loading to guarantee an invertible matrix. Diagonal

loading has been demonstrated by many groups (Greening and Perkins, 2002; Cox, 2002; Stoica et al., 2003), while spatial averaging was initially used as mentioned for creating decorrelation. Eventually, Synnevåg et al. presented a robust version for both diagonal loading and spatial averaging, and discussed optimal parameters for each (Synnevåg et al., 2007).

4.2.2 Unresolved Problem – Reverberation Clutter

It is known that minimum variance performs worse as noise content increases, but we suggest here that reverberation clutter is particularly problematic. From the original Capon paper (Capon, 1969), it is assumed that incoming signals are unity amplitude monochromatic plane waves. When properly delayed, on and off-axis signals take the form of plane waves, which mostly satisfies this assumption, but reverberant signals (and other complications like phase aberration) do not. Reverberant signals originate from shallower depths than the target and are effectively time-delayed to the target time index by means of multi-path scattering. This means that the delays applied to reverberant signals will be mismatched, and those signals will not be flattened into plane waves, but rather, will remain spherical waves. In the case of phase aberration, this problem may be further exaggerated, as there will be no ideal plane waves and all signals will have spherical components. This is problematic, as we will demonstrate.

In the derivation in Capon's paper (Capon, 1969), he arrives at an important inequality,

$$b^2 >> \frac{R}{K}\frac{|B(\Delta \boldsymbol{k})|^2}{1-|B(\Delta \boldsymbol{k})|^2}. \tag{4.1}$$

b is related to the scattering amplitude of the chosen plane wave, R is the ratio of the incoherent noise power to the total power for K sensors, $|B(\Delta \boldsymbol{k})|^2$ is the beamforming array response pattern, and $\Delta \boldsymbol{k}$ is the difference between the vector wavenumber of the chosen and off-axis plane waves. Capon states that the inequality is satisfied when either R/K or $|B(\Delta \boldsymbol{k})|^2$ is small, effectively meaning either when the incoherent noise power is relatively low compared to the number of sensors, or when the difference of the vector wavenumbers $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ is sufficiently large. In the case of two plane waves that are sufficiently separated (and thus have sufficiently different vector wavenumbers), and given a sufficiently low noise level, this method will be able to resolve the two plane waves. Van Trees makes a similar assertion that the signal-to-noise ratio must be reasonably high and that the at most $K-2$ interfering plane waves must fall outside of the main lobe for proper nulls to be formed for each of the interfering signals (Van Trees, 2002).

However, spherical reverberant signals complicate things on multiple levels. It should be noted that in the Cartesian plane, a given plane wave will have a fixed wavenumber, the direction of which shows the direction of the wave. However, the Weyl expansion tells us that a spherical wave can be written as a linear

combination of infinite plane waves (Weyl, 1919; Chew, 1995; Kinayman and Aksun, 2005), specifically

$$\frac{e^{-jk_0 r}}{r} = \frac{1}{j2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-j(k_x x + k_y y)}\frac{e^{-jk_z|z|}}{k_z}dk_x dk_y. \tag{4.2}$$

A spherical wave with k_0 is then composed of a continuous set of plane waves with vector wavenumbers $\mathbf{k} = [k_x, k_y, k_z]$ that satisfy $k_0 = \sqrt{k_x^2 + k_y^2 + k_z^2}$. It is then apparent that there will be overlap between the fixed $\boldsymbol{k}_1$ of our target plane wave and the continuous set of possible values of $\boldsymbol{k}_2$ for the spherical reverberant wave. Therefore, some parts of the reverberant wave will be too similar to $\boldsymbol{k}_1$, and thus be unresolveable. Additionally, since the goal is to place nulls in the beampattern in the directions of interference (i.e. the plane waves corresponding to sufficiently different vector wavenumbers), this is more difficult in a case where there is a continuous set of plane waves. Thus the beamformer will struggle with rejecting the continuous set compared to well-defined discrete plane waves. We simulated an example of a point target and a reverberant target in Field II (Jensen, 1996; Jensen and Svendsen, 1992), which is included in Fig. 4.1 to show a visual example of how an on-target plane wave and reverberant spherical wave have vector wavenumbers with inherent overlap. Additionally, we see that MV is unable to properly reduce or eliminate the reverberant signal, resulting in somewhat worse resolution of the primary target, though still better than DAS alone. Fig. 4.1G shows an example of the MV beampattern when an off-axis target is present and the resulting null in the beampattern that is formed to suppress that off-axis interference. In comparison, the case where the off-axis target is replaced by a reverberant signal in the same location shows how the beampattern tries to generally suppress the reverberant region, but as from the b-mode images we see is less successful at suppressing that interference.

A final consideration is the ratio of the incoherent noise power, R/K. It is stated that as long as the noise contribution is relatively low, the two waves could still be distinguishable. Diffuse reverberation clutter takes on a speckle-like appearance in the final image and is relatively incoherent in the aperture domain (Dahl and Sheth, 2014; Pinton et al., 2011). Then, in the case of very strong diffuse reverberation clutter, it may have the effect of being strong enough to interfere with the MV beamformer. Supporting this theory, Austeng et al. showed that in the case of very strong phase aberration, MV has comparable performance to DAS with Hamming apodization, and slightly worse than DAS with rectangular apodization (Austeng et al., 2008). Here, we adopt the understanding that diffuse reverberation clutter originates from the complex tissue in the body wall (Dahl and Sheth, 2014; Robinson et al., 1994; Dahl et al., 2017).

Overall, reverberant signals and other non-plane wave signals pose an additional challenge for MV due to the inability to fully reject the overlap between the continuous vector wavenumbers present in spherical signals and the discrete vector wavenumber of an on-axis target. In these cases, or in general in the presence

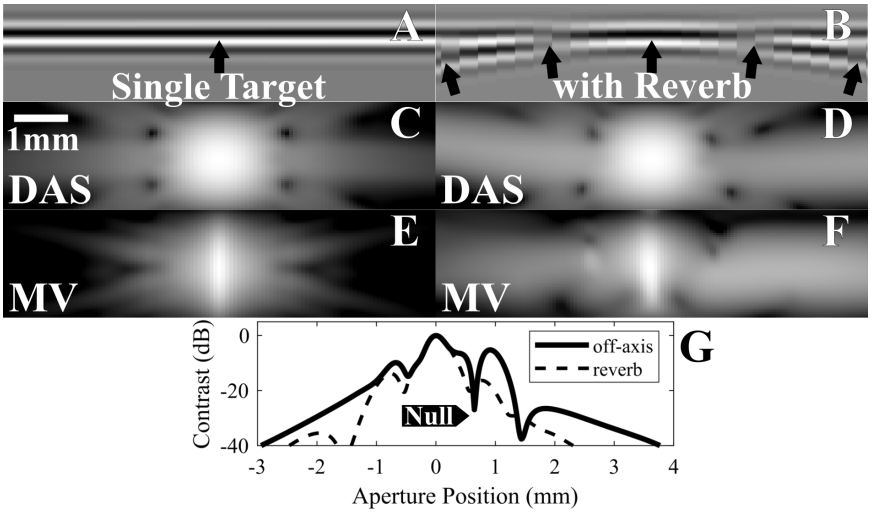

Figure 4.1: Field II simulations of an individual point target (A,C,E) without interference and (B,D,F) with reverberant signal interference. The reverberant signal is simulated at a much shallower depth than the point target and then time-delayed to appear at the same time index. (A,B) channel data at the depth of the point target. The arrows indicate the directions of the vector wavenumbers of the echoes, showing a subset of the possible wavenumbers in the reverberant spherical wave case. (C,D) DAS and (E,F) MV images of the point target. B-mode images are displayed on a 60 dB scale. (G) a high-resolution beamplot from MV of a separate simulation where the reverberant signal is replaced with an off-axis target, highlighting the null that is created in the off-axis case and the lack thereof in the reverberation case. Figure ©2022 IEEE.

of strong noise content, MV will struggle or fail to narrow the main lobe, reducing performance.

4.2.3 Proposed Solution - Pre-Processing with ADMIRE

The primary goal of this work is to improve the performance of MV in these high reverberation clutter or complex *in vivo* environments. To that end, we propose that a pre-processing step before MV to remove these sources of interference will do exactly that. Aperture domain model image reconstruction (ADMIRE) is a method for removing reverberation and off-axis clutter, as well as suppressing wavefront aberration (Byram et al., 2015a; Dei and Byram, 2017; Byram and Jakovljevic, 2014). Crucially for this work, ADMIRE is a somewhat unique beamformer in that when it processes channel data it does not alter the dimensionality of the input data. This means that we can subsequently process the decluttered data with other beamforming methods, in this case MV. We hypothesize that ADMIRE will be able to suppress these difficult sources of interference, removing the spherical components that MV struggles with and generally reducing the number of interfering waves and improving effective signal-to-noise ratio, making it possible for MV to perform better than without pre-processing. This gives us the possibility to create better images than with either method alone.

4.3 Beamforming Algorithms

We implemented all beamformers in MATLAB (The MathWorks, Natick, MA, USA).

4.3.1 Delay-and-Sum (DAS)

The conventional standard beamformer, as defined in Section 1.4.2. In this work, we will use both normal DAS with fixed rectangular apodization to show the raw data (DAS), and also DAS with Hamming apodization as a commonly used weighting scheme (DAS-Hamm).

4.3.2 Minimum Variance (MV)

We include the full definition here again for reference, since this chapter heavily discusses the minimum variance (MV) beamformer. MV can be thought of as an adaptively weighted and delayed DAS image where the weighting factor is optimized to improve lateral resolution by reducing off-axis clutter (Holfort et al., 2009; Synnevåg et al., 2007). Bold characters here and for the rest of the text indicate vectors. The optimized weights are defined as

$$\mathbf{w} = \frac{R^{-1}\mathbf{e}}{\mathbf{e}^H R^{-1}\mathbf{e}}, \tag{4.3}$$

where $\mathbf{e}$ is the steering vector, H is the conjugate transpose, and R is the covariance matrix defined as

$$R(x,z) = \mathrm{E}[\mathbf{s}(x,z)\mathbf{s}(x,z)^H], \tag{4.4}$$

where $\mathrm{E}[\cdot]$ denotes the expectation and $\mathbf{s}(x,z)$ is the delayed aperture signal at lateral index x and depth z. To ensure R is invertible we used subarray averaging and diagonal loading methods (Synnevåg et al., 2007). The recommendations provided are to use subarray lengths of $L = 0.5M$ and diagonal loading defined as $\varepsilon = \Delta \cdot \mathrm{tr}(\hat{R})$, where $\Delta = 1/(10L)$ (Synnevåg et al., 2007). The estimated MV signal is then defined as

$$\hat{\mathrm{S}}_{\mathrm{MV}}(x,z) = \frac{1}{M-L+1}\sum_{l=0}^{M-L} \mathbf{w}(x,z)^H \bar{\mathbf{s}}_l(x,z), \tag{4.5}$$

where $\bar{\mathbf{s}}(x,z)$ is the delayed channel data for a given subarray. We will call this version of minimum variance the narrowband version (NB), abbreviated as MVNB, but we additionally implemented the broadband version (BB) devised by Holfort et al. (Holfort et al., 2009), which we will denote as MVBB. For both implementations, the optimal parameters for subarray averaging and diagonal loading were estimated based on both simulated and *in vivo* data. While diagonal loading does have a small impact on image metrics, it is dwarfed by the impact of subarray averaging, so for consistency we use $\Delta = 1/(10L)$ unless stated otherwise. We include in the results multiple subarray averaging cases to demonstrate the effect it has on resolution and

sizing.

4.3.3 Aperture Domain Model Image Reconstruction (ADMIRE)

Aperture domain model image reconstruction (ADMIRE) is a method for removing reverberation and off-axis clutter, as well as suppressing wavefront aberration. We discuss the method in detail in Chapter 2.

4.3.4 Post ADMIRE Processing

Since ADMIRE returns decluttered channel data, we can, in theory, apply any beamforming method to data that has already been processed with ADMIRE. Most commonly, we simply sum the decluttered channel data like in DAS beamforming, but rather than write ADMIRE+DAS we simply refer to this as ADMIRE. For this work, we consider the additional possibility where we apply MV instead of DAS, producing two cases: ADMIRE+MVNB for narrowband MV and ADMIRE+MVBB for broadband MV, abbreviated as AD+MVNB and AD+MVBB, respectively.

As we did with ADMIRE and MV individually, we tested multiple parameter choices for the combined methods AD+MVNB and AD+MVBB. We found that the default ADMIRE parameters continued to be optimal, and that changes to diagonal loading in MV were generally imperceptible, except *in vivo* where we were able to see measureable improvement to AD+MVNB by reducing the diagonal loading to $\Delta = 1/(100L)$. Similar to when applying MV alone, changing the length of the subarrays had a significant impact on the resulting image and metrics, so we included these results and some discussion in the following sections.

4.4 Methods

4.4.1 Simulated Speckle Texture for Measuring SNR

We mentioned that ADMIRE can improve signal-to-noise ratio (SNR), which would be beneficial to MV. To demonstrate this, we simulated a simple speckle texture in Field II (Jensen, 1996; Jensen and Svendsen, 1992) using the simulation parameters in Table 2.2. We then created 30 independent realizations of Johnson-Nyquist (thermal) noise by simulating normally distributed independent random noise using the randn function in MATLAB, to effectively create 30 "frames" of independent noise. We combined the thermal noise and speckle channel data such that it satisfied a specified SNR as defined by

$$\text{SNR} = 10\log_{10}\left(\frac{P_{\text{SOI}}}{\alpha^2 P_{\text{thermal}}}\right), \tag{4.6}$$

where P_{SOI} is the power of the channel data of the speckle phantom, P_{thermal} is the power of the channel data of the thermal noise, and α is the scalar for the thermal noise in order to achieve the desired SNR. For this simulation, we targeted an SNR of 0dB.

These simulations were then processed with ADMIRE and the SNR was calculated before and after ADMIRE. We calculated SNR using two methods. The first way, we used the equation

$$\text{SNR}_{power} = 10\log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right), \tag{4.7}$$

where P_{signal} is estimated from the average across all 30 frames, and P_{noise} was then estimated from the difference between each frame and the estimated signal. Since the signal is static across all frames, the signal power is calculated from the average of the estimated signal, and the noise power is calculated from the variance of all of the frames. The second method was to use the correlation between frames to estimate SNR, as done by Friemel et al. (Friemel et al., 1998). By calculating the correlation coefficient ρ between each successive frame, the SNR can also be estimated by

$$\text{SNR}_{frame} = 10\log_{10}\left(\frac{\rho}{1-\rho}\right). \tag{4.8}$$

4.4.2 Simulated Single Target Phantoms

We simulated a single point target using Field II to demonstrate how MV behaves when a primary target is receiving interference from a reverberant target. The simulation parameters are the same as in Table 2.2. The primary target was simulated at a depth of 3 cm, while the reverberation target was simulated at a depth of 0.5 cm and then time shifted to the same time index as our primary target, similar to how reverberation clutter has been simulated in previous work (Byram and Shu, 2016b,a). We could then observe how the point spread function (PSF) differs for MVNB and MVBB versus AD+MVNB and AD+MVBB without and with that reverberation clutter. We in part used the width of the PSF as an indicator for determining the optimal parameters for both MV methods and post-ADMIRE methods.

4.4.3 Simulated Cyst Phantoms

To represent the scenario of high noise power reverberation clutter, we used Field II to simulate 5 mm diameter anechoic cysts using the same parameters as in Table 2.2, creating six independent speckle realizations. We then simulated diffuse reverberation clutter using the method described by Byram and Shu (Byram and Shu, 2016b,a). We added the reverberation clutter such that it satisfied a 0dB signal-to-clutter ratio (SCR) compared to the cyst channel data. Though there has been relatively little research into classifying what low or high reverberation clutter levels are in terms of SCR, we have previously determined that 0dB SCR is a plausible representation of clutter corresponding to a difficult-to-image patient (Schlunk et al., 2021). This is backed up by other studies such as one where bladder wall (signal) to clutter ratios calculated from image data

were between 30 and 0 dB for all sources of clutter including reverberation (Lediju et al., 2008), suggesting our target for channel data is reasonable for *in vivo* scenarios.

To complement the reverberation clutter simulations, we also simulated Johnson-Nyquist (thermal) noise, which was added to the cyst channel data at the same 0dB ratios. We differentiate these thermal noise cases using the abbreviation signal-to-noise ratio (SNR).

4.4.4 *In vivo* Kidney Stone Data

We additionally captured *in vivo* kidney data from patients suffering from kidney stone disease using a Verasonics Vantage Ultrasound System (Verasonics, Inc., Kirkland, WA) with a C5-2 curvilinear transducer. A plane wave synthetic aperture acquisition was employed (Montaldo et al., 2009), where plane waves were transmitted at multiple angles to achieve focusing at all depths. A center frequency of 4.1667 MHz was used to acquire 64 angles spanning 37°. The data was acquired under protocols approved by the Vanderbilt University IRB (IRB# 170001).

4.4.5 Image Quality Metrics

For the simulated cyst phantoms we computed the contrast ratio and generalized contrast-to-noise ratio (gCNR) (Rodriguez-Molares et al., 2018, 2020), and as a resolution metric we included radial cyst edge width (Bottenus et al., 2020b). Contrast ratio was defined as

$$\text{contrast ratio} = -20\log_{10}\left(\frac{\mu_{\text{ROI}}}{\mu_{\text{background}}}\right) \tag{4.9}$$

where μ is the mean value calculated from the enveloped data, before log compression. gCNR is a generalized detectability metric which measures the overlap of the probability density function between the ROI and the background, making it more robust against stretches or compressions in dynamic range. We used the conventional 100 bins for the histogram estimation for this method. For these metrics, the ROI was defined as the interior of the cyst, and the background was an equivalently sized radial region surrounding the cyst.

In addition to calculating the width of the PSF in our point target simulation, we estimated radial cyst edge width similar to Bottenus et al. (Bottenus et al., 2020b) as a more practical measure of lateral resolution by radially averaging values of equal distance from the center of the cyst within $\pm 10^\circ$ of the lateral axis. This produces a function of intensity versus radius, which is normalized by subtracting by μ_{ROI} and scaling by $\mu_{\text{background}}$. Then the width is estimated as the rise distance on this curve. This is done for both the left and right sides of the anechoic cysts. An example showing two such radial intensity curves is shown in Fig. 4.2. For this work, we considered the rise distance from .25 to .75, representing a measurement similar to that of

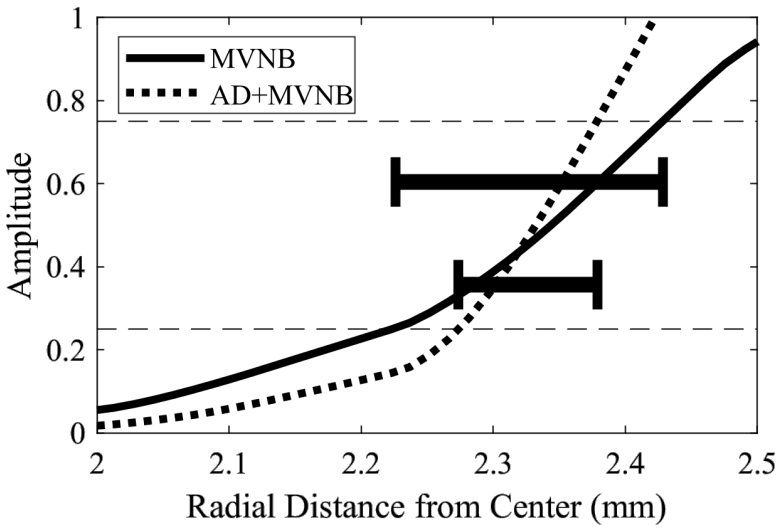

Figure 4.2: Example of MVNB and AD+MVNB radial intensity curves from which the cyst edge width is estimated as the distance it takes to rise from 0.25 to 0.75 (indicated with the dashed lines). Figure ©2022 IEEE.

the FWHM, as has been done with similar methods in previous work (Samei et al., 1998).

For the *in vivo* kidney stone cases, the stone region of interest was manually selected on a 30dB dynamic range b-mode image with a contour map as reference. A radial background region was then automatically calculated as 1.5 times the size of the selected stone, centered on the stone. This allowed us to calculate the contrast ratio and gCNR of the stone for all cases, and provided a lateral measure of the stone size. Since kidney stones behave as coherent targets (Schlunk et al., 2017; Hsi et al., 2018; Tierney et al., 2019), we consider smaller measures of stone size to be an improvement. We also estimate the SNR in these cases by implementing the robust version of the lag-one coherence method devised by Vienneau et al. (Vienneau et al., 2022) based on original work by Hyun et al. (Hyun et al., 2017) and Long et al. (Long et al., 2018). Similar to the simulations, this was estimated before and after processing with ADMIRE. In these *in vivo* cases, since we lack the multiple frames required to differentiate thermal noise from other clutter types, the SNR value estimated represents the combination of all types of noise: thermal, reverberation, or others.

To complete our parameter testing for MV and ADMIRE, we tested our suite of parameter choices on both the simulated cysts and some of the *in vivo* cases to determine if the optimal parameters change between simulations and the *in vivo* data. Examples of interesting parameters are included in the results section to demonstrate how these choices can impact our image quality metrics.

Table 4.1: Bright Target Simulation Metrics. Table ©2022 IEEE.

PSF Width at Various Amplitudes (mm)				
	No Reverb		With Reverb	
Amplitude	-6 dB	-30 dB	-6 dB	-30 dB
DAS	0.493	0.939	0.503	6.284
MVNB L=0.5M	0.080	0.631	0.174	4.155
MVNB L=0.75M	**0.059**	**0.411**	0.097	0.561
MVBB L=0.5M	0.110	0.863	0.216	5.097
MVBB L=0.75M	0.116	0.939	0.142	0.768
ADMIRE	0.463	0.939	0.471	1.187
AD+MVNB L=0.5M	0.153	0.868	0.084	0.671
AD+MVNB L=0.75M	0.119	0.752	**0.058**	**0.465**
AD+MVBB L=0.5M	0.160	0.895	0.284	1.123
AD+MVBB L=0.75M	0.204	0.939	0.129	0.658

4.5 Results

4.5.1 Simulated Speckle SNR Results

From the 30 realizations of independent thermal noise, both SNR_{power} and SNR_{frame} were calculated before and after ADMIRE. SNR_{power} increased from 0.984dB to 22.8dB after ADMIRE. Average SNR_{frame} increased from 0.715 ± 0.008dB to 22.653 ± 0.078dB after ADMIRE. Both methods show that ADMIRE produces a significant increase in SNR.

4.5.2 Point Target Simulation Results

The PSF of the point target simulation results is shown in Fig. 4.3 to demonstrate the impact of the subarray averaging parameter on MV. To summarize the differences caused by subarray averaging we focused on two cases, $L = 0.5M$ and $L = 0.75M$, for both MVNB and MVBB with and without ADMIRE. The PSF for each case is shown in Fig. 4.3B,D, while the widths of the PSF for each case at -6 and -30 dB are included in Table 4.1. From these results alone, MVNB benefits significantly from increasing the subarray size regardless of reverberation level. In comparison, MVBB prefers $L = 0.5M$ for the no reverberation case and $L = 0.75M$ for the with reverberation case. From the PSF plot in the figure, MVBB behaves rather erratically when the subarray size gets large, whereas MVNB has a more uniform improvement with increased subarray size. As a result, we generally include both parameter choices for MVNB while focusing on $L = 0.5M$ for MVBB in the simulated results.

Fig. 4.3 also shows that ADMIRE behaves almost identically to DAS, only slightly improving the sidelobes off-axis in the reverberation case. In comparison, MVNB at $L = 0.75M$ has by far the best performance when there is no reverberation clutter (or any significant noise source) present. Even at $L = 0.5M$ it still

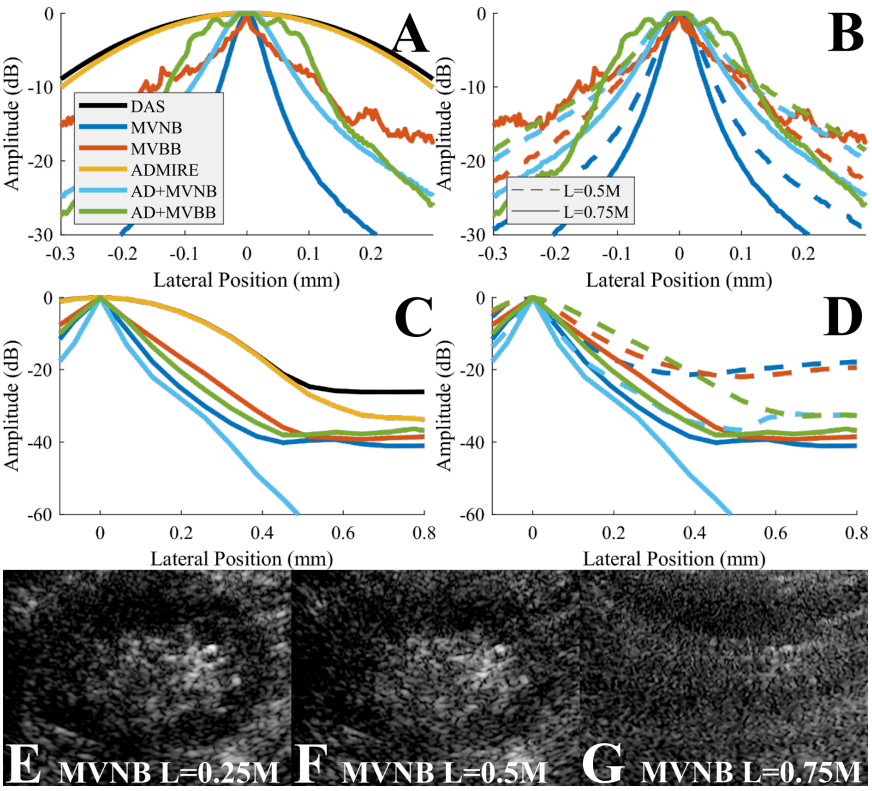

Figure 4.3: Parameter testing for the subarray length (L) versus the full aperture length (M) for both narrowband (MVNB) and broadband (MVBB) implementations. Point spread functions for (A,B) a bright target with no reverberation interference and (C,D) a bright target with an interfering reverberant signal. (E,F,G) show an *in vivo* example of MVNB with varying subarray lengths, demonstrating how increasing the subarray length can lead to general image quality degradation. Figure ©2022 IEEE.

outperforms all other beamformers, including the combined method AD+MVNB, though this still performs better than just ADMIRE by itself. However, in the presence of a strong reverberation signal, MVNB suffers substantially with the full-width at half maximum increasing by 64%. In this case, pre-processing with ADMIRE grants a significant increase in performance, with AD+MVNB performing in the reverberation case equivalently to how MVNB alone performed in the noise-free case. Strangely, AD+MVBB generally performed poorly, worse than running just MVBB alone, regardless of choice of parameters. However, regardless of narrowband or broadband, the combined method uniformly performed better than just ADMIRE alone, and by extension also DAS.

4.5.3 Anechoic Cyst Simulation Results

The MVNB cases for the simulated cysts in Fig. 4.4 show that increasing the subarray size can result in the degradation of the speckle in the background. As expected from this, Table 4.2 shows that cyst edge width and contrast ratio are improved with higher subarray sizes, though gCNR is lost as a result. This

means that the ideal subarray size in these cases may depend on which metrics are more important for a given application. Since both subarray size choices produce viable images, we included both in the tables and figures for MVNB.

The example cysts in Fig. 4.4 and the cyst edge width measurements in Table 4.2 show a general agreement with the bright target simulations. These simulations are split into a noise-free case, a reverberation clutter case (0dB SCR), and a thermal noise case (0dB SNR). In these cases, we measured cyst edge width as a complement to PSF width in the bright target simulations, and additionally showed more traditional image quality metrics: contrast ratio and gCNR. Unlike in the simple bright targets, in these cases AD+MVNB ($L = 0.75M$) always outperforms ADMIRE and MVNB alone in terms of both cyst edge width and contrast ratio, but loses some gCNR compared to ADMIRE to compensate. Even AD+MVNB ($L = 0.5M$) performs similar to ADMIRE in terms of contrast ratio and gCNR, and still demonstrates an improved cyst edge width compared to both methods. Interestingly, in the high reverberation clutter case (0dB SCR) AD+MVNB ($L = 0.5M$) has a better cyst edge width compared to AD+MVNB ($L = 0.75M$), suggesting that the loss of gCNR in that case did impact the cyst edge appearance. This is also reflected in the much lower variance in the reverberation case for AD+MVNB ($L = 0.5M$) compared to all other methods, showing consistency similar to the noise-free and thermal noise cases. A final observation is that AD+MVBB never demonstrates any significant improvement compared to just ADMIRE in any of the three cases. Overall, AD+MVNB generally outperforms all of the other methods, with the benefits being most readily apparent in the reverberation clutter and thermal noise cases.

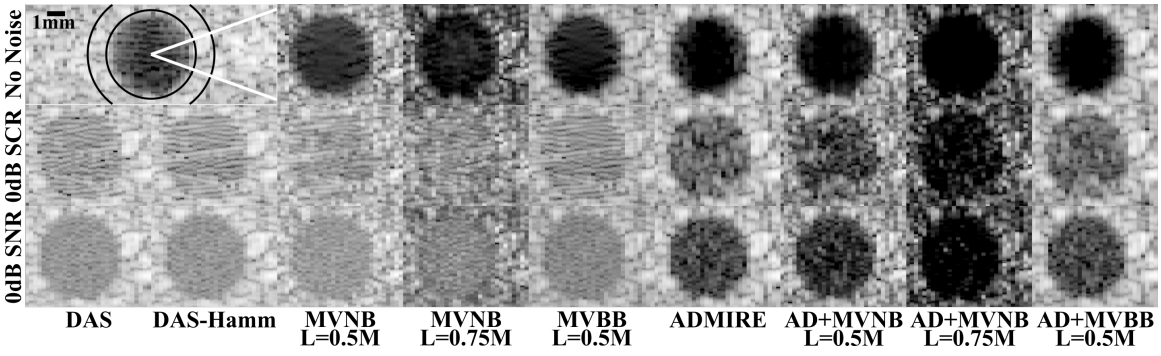

Figure 4.4: Sample cases of an anechoic cyst with no added noise, added reverberation clutter (0dB SCR), and added thermal noise (0dB SNR) displayed on a 50 dB dynamic range. The black solid circles indicate the true region of the cyst as well as a background speckle region surrounding it for use with our imaging metrics. The white solid lines show the radial region for which the cyst boundary width was estimated. Figure ©2022 IEEE.

Table 4.2: Anechoic Cyst Simulation Metrics. Table ©2022 IEEE.

	Cyst Edge Width (mm)		
Added Interference	None	0dB SCR	0dB SNR
DAS	0.189±0.025	0.216±0.154	0.171±0.015
DAS-Hamm	0.218±0.028	0.347±0.274	0.202±0.060
MVNB L=0.5M	0.305±0.230	0.641±0.623	0.303±0.291
MVNB L=0.75M	**0.101±0.037**	0.292±0.362	0.155±0.134
MVBB L=0.5M	0.204±0.029	0.437±0.395	0.259±0.282
ADMIRE	0.162±0.022	0.218±0.179	0.138±0.022
AD+MVNB L=0.5M	0.132±0.028	**0.140±0.054**	0.121±0.033
AD+MVNB L=0.75M	**0.098±0.026**	0.183±0.174	**0.105±0.021**
AD+MVBB L=0.5M	0.178±0.017	0.184±0.130	0.154±0.016
	Contrast Ratio (dB)		
Added Interference	None	0dB SCR	0dB SNR
DAS	-31.9±1.1	-12.1±0.7	-14.3±0.7
DAS-Hamm	-38.8±0.8	-11.4±1.1	-14.0±0.7
MVNB L=0.5M	-30.3±1.0	-7.97±1.1	-10.2±0.8
MVNB L=0.75M	-24.4±1.7	-4.83±1.4	-5.58±0.8
MVBB L=0.5M	-36.6±0.7	-10.0±0.9	-12.0±0.7
ADMIRE	-43.1±1.3	-21.5±1.5	-29.9±1.1
AD+MVNB L=0.5M	-42.1±1.4	-21.0±1.7	-30.4±1.4
AD+MVNB L=0.75M	**-46.1±1.5**	**-24.7±2.6**	**-33.5±2.4**
AD+MVBB L=0.5M	**-46.7±1.3**	-20.5±1.7	-29.5±1.2
	generalized Contrast-to-Noise Ratio (gCNR)		
Added Interference	None	0dB SCR	0dB SNR
DAS	0.992±0.003	0.774±0.032	0.932±0.015
DAS-Hamm	**0.999±0.002**	0.745±0.061	0.923±0.016
MVNB L=0.5M	0.969±0.014	0.489±0.055	0.725±0.026
MVNB L=0.75M	0.819±0.028	0.232±0.045	0.390±0.021
MVBB L=0.5M	0.994±0.002	0.649±0.055	0.854±0.020
ADMIRE	0.992±0.005	**0.879±0.030**	**0.982±0.011**
AD+MVNB L=0.5M	0.988±0.005	0.750±0.060	0.930±0.022
AD+MVNB L=0.75M	0.792±0.073	0.676±0.047	0.783±0.095
AD+MVBB L=0.5M	**0.997±0.003**	**0.863±0.046**	**0.983±0.009**

4.5.4 in vivo Kidney Stone Results

Fig. 4.3E,F,G shows different choices of subarray sizes for MVNB for *in vivo* Case 1. Compared to the simulated data, this case shows very obvious image degradation as the subarray size increases, to the point where the image is no longer recognizable as a kidney at $L = 0.75M$. Since image quality is critical to delineating the kidney stone, $L = 0.25M$ was better for all cases (necessary to identify the kidney stones correctly). The exception to this was AD+MVNB for which increasing subarray averaging to $L = 0.5M$ without losing image quality was possible.

Case 1 with the optimal MV parameters is shown in Fig. 4.5. Visually, an improvement in general im-

age contrast with ADMIRE-based methods is seen compared to non-ADMIRE-based methods. The general image clarity and boundaries of the kidney also appear the most defined in the AD+MVNB image, and the highlighted kidney stone (marked by the red arrow) appears sharper compared to ADMIRE alone (or any other method). Fig. 4.6 shows Case 1 as well as several others zoomed in specifically on the kidney stone of interest, with the stone region drawn in red. The measured stone size is indicated on the images by the blue line, and also included along with contrast ratio and gCNR in Table 4.3. In these *in vivo* cases, both AD+MVNB and AD+MVBB generally perform well, shrinking the apparent size of the stone and boosting contrast ratio and gCNR compared to ADMIRE, MV, or DAS, suggesting better prospects for AD+MVBB compared to the simulations. The only case where the combined methods do not improve upon ADMIRE or MV is Case 5, which is included in the figure. From the b-mode images, it would seem that the surrounding tissue complicates the case, such that even with improved resolution the distinction between stone and background tissue is ambiguous, resulting in no improvement to sizing. The contrast ratio and gCNR are still both improved. The estimated SNR for all noise sources based on the lag-one coherence improved on average in these cases from $-0.49 \pm 5.99dB$ before ADMIRE to $20.75 \pm 1.32dB$ after processing with ADMIRE.

Both the narrowband and broadband implementations of MV perform well *in vivo*, with AD+MVNB averaging slightly better statistics and appearing somewhat better in the b-mode images, corresponding to generally better gCNR. However, both methods do succeed at what they were designed to do. The two combined methods each measure the smallest stone size in three of the six cases, with AD+MVNB being slightly more consistent at improving contrast ratio and gCNR compared to ADMIRE or the base MV methods.

4.6 Discussion

AD+MVNB demonstrated an improvement over ADMIRE and MVNB in every case, simulated and *in vivo*, with the exception of the noise-free single bright target simulation. AD+MVBB was unimpressive in simulations, performing similar to ADMIRE, but showed similar improvements *in vivo* compared to the narrowband version. ADMIRE combined with MV overall showed improvements to resolution and general imaging quality metrics, especially in the cases with high reverberation clutter or thermal noise. Our primary hypothesis in this work was that ADMIRE could remove sources of reverberation and off-axis clutter that would otherwise degrade the performance of MV. Our simulations clearly demonstrate that high noise reduces the ability of MV to properly separate signals and improve resolution, though this has been well-known since the original implementations of MV. However, we did show that both versions of MV struggle significantly more when the interference is reverberation clutter compared to thermal noise, even when both are presented at similar signal-to-noise levels. While we can increase the subarray size to compensate for the increased noise presence, this comes at a severe cost of degraded image quality and does not fully mitigate the added noise. We

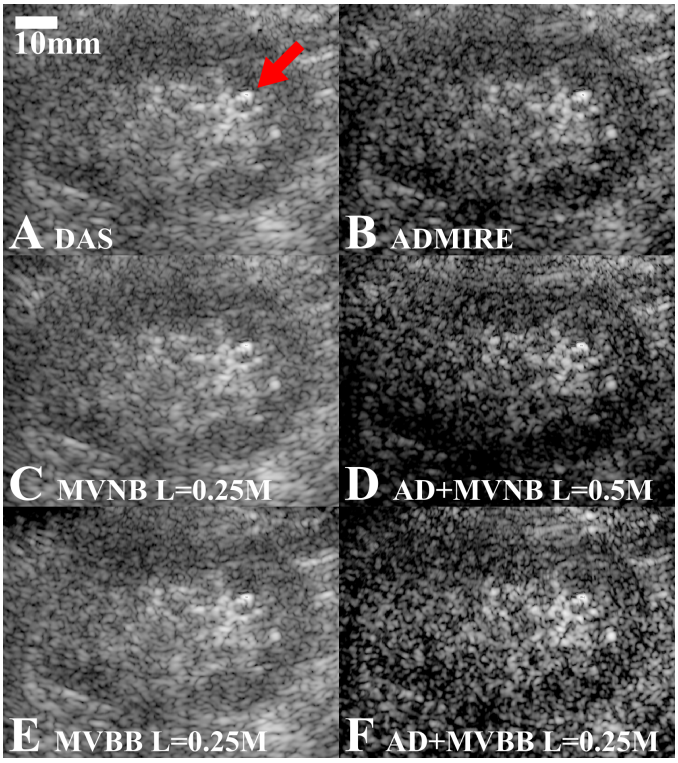

Figure 4.5: Comparisons of the b-mode images for Case 1 on a 60dB dynamic range for the minimum variance methods both individually (C,E) and after pre-processing with ADMIRE (D,F). The kidney stone is indicated by the red arrow in the DAS image. Figure ©2022 IEEE.

have shown that by pre-processing with ADMIRE, we can improve the resolution and contrast ratio compared to MV alone, and outperform ADMIRE in terms of resolution, and often in terms of contrast ratio as well.

The narrowband implementation of MV showed the most universal promise when complemented with ADMIRE, demonstrating noticeable and consistent improvements in the simulations and *in vivo*. AD+MVBB suffered somewhat in simulations, performing generally worse compared to ADMIRE, though still an improvement compared to MVBB alone. This may be due in part to some interactions with the way data is simulated and the similarity of processing that occurs for both ADMIRE and MVBB, since both take advantage of the frequency spectrum to process data in a narrowband fashion in similar ways. Whatever the exact cause, this phenomenon does not carry over to the *in vivo* results, where AD+MVBB generally shows improvements compared to both ADMIRE and MVBB. However, MVNB does slightly outperform MVBB, and considering the increased computation time the broadband variant requires, it is easy to recommend MVNB as the more effective complement to ADMIRE.

ADMIRE as a choice for pre-processing performed as expected. From the cyst simulations for the specific interference sources, ADMIRE was visually able to suppress both types and improve image quality metrics compared to DAS. We additionally confirmed in simulated speckle phantoms that ADMIRE could produce

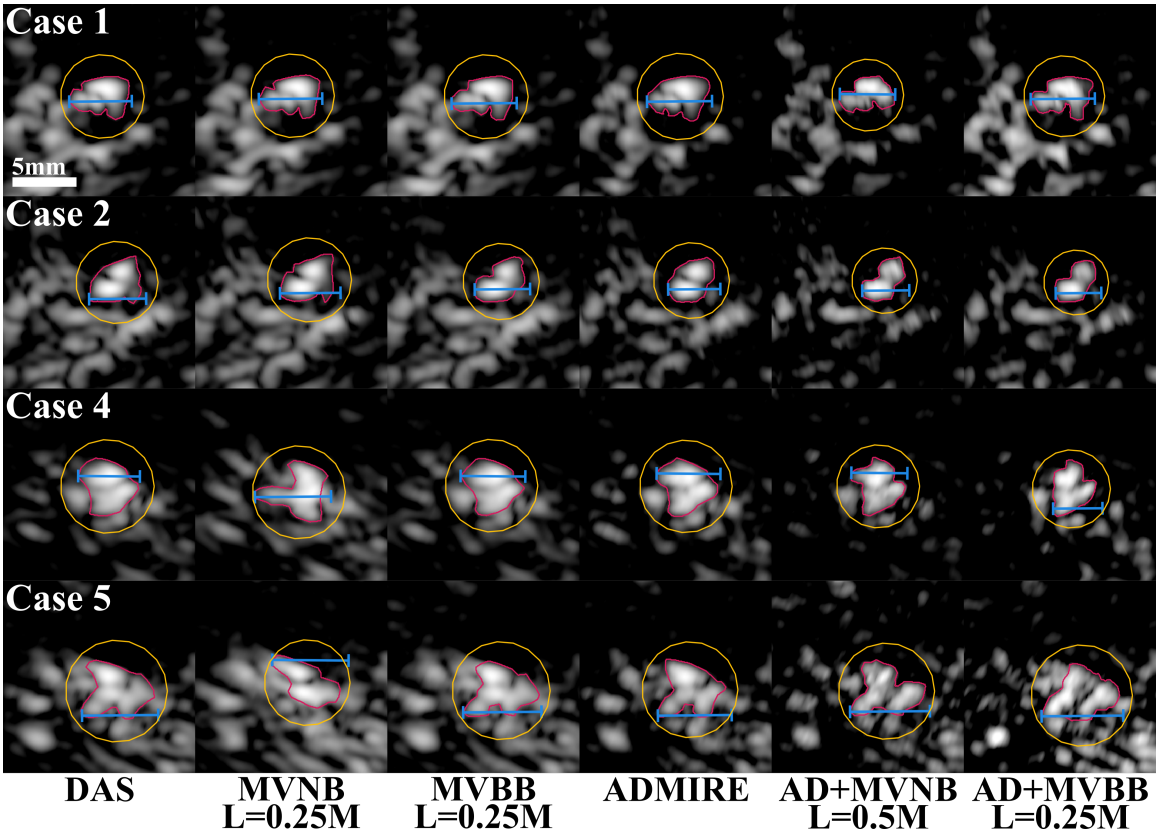

Figure 4.6: B-mode images on a 30dB dynamic range of a selection of the *in vivo* kidney stones. The individual stones are manually highlighted in red with the help of a contour map, and the background used for image quality metrics is shown in yellow based on the stone region selected. The blue line shows the lateral length of the stone region. Figure ©2022 IEEE.

significant improvements to SNR in the presence of thermal noise, and we similarly estimated SNR, but for all noise types, in the *in vivo* cases to demonstrate a similar level of improvement. Since MV performance is heavily dependent on the level of SNR of the image, this supports the idea that ADMIRE is a powerful option to improve MV in these low SNR cases. While we cannot specifically differentiate between reverberation clutter and other noise sources *in vivo* due to the lack of the additional frame data required, the simulations make it clear that reverberation clutter has a much stronger impact on MV performance compared to thermal noise.

Perhaps the main concern with implementing MV with ADMIRE, or MV in general, is the choice of parameters. We demonstrated that the subarray size by itself can have a significant impact on resolution and general image quality, depending on what one is willing to sacrifice to boost the other. In simulations and simple imaging scenarios, it may be easy to sacrifice visibility to improve resolution, but clearly in more complex *in vivo* cases this may not be true. Even as groups are investigating ways to automatically optimize these parameters, it is difficult to say how easy it will be to optimize these more complex cases. A consolation,

Table 4.3: Image Metrics for *in vivo* Kidney Cases. Table ©2022 IEEE.

	Measured Stone Size (mm)					
Case	1	2	3	4	5	6
DAS	0.489	0.445	0.814	0.482	0.595	0.535
DAS-Hamm	0.526	0.457	0.812	0.486	0.623	0.550
MVNB L=0.25M	0.495	0.473	0.813	0.593	0.595	0.550
MVBB L=0.25M	0.508	0.432	0.791	0.505	0.611	0.555
ADMIRE	0.508	0.407	0.750	0.498	**0.576**	0.495
AD+MVNB L=0.5M	**0.432**	0.367	0.734	0.437	0.623	**0.437**
AD+MVBB L=0.25M	0.501	**0.356**	**0.695**	**0.385**	0.638	0.471
	Contrast Ratio (dB)					
Case	1	2	3	4	5	6
DAS	19.0	14.6	9.68	13.9	11.5	11.4
DAS-Hamm	18.5	16.4	10.0	13.9	12.0	11.5
MVNB L=0.25M	18.3	14.8	9.74	13.6	10.9	11.2
MVBB L=0.25M	18.6	15.8	10.5	13.9	10.9	11.3
ADMIRE	21.6	19.3	8.14	14.9	12.7	14.0
AD+MVNB L=0.5M	**23.6**	**23.3**	9.71	**16.4**	**13.7**	16.6
AD+MVBB L=0.25M	21.8	21.2	**11.6**	15.5	10.6	**17.2**
	generalized Contrast-to-Noise Ratio (gCNR)					
Case	1	2	3	4	5	6
DAS	0.781	0.649	0.582	0.683	0.632	0.609
DAS-Hamm	0.781	0.726	0.609	0.721	0.661	0.600
MVNB L=0.25M	0.800	0.633	0.603	0.680	0.563	0.598
MVBB L=0.25M	0.793	0.716	**0.618**	0.676	0.603	0.593
ADMIRE	0.862	0.748	0.566	0.670	0.642	0.700
AD+MVNB L=0.5M	**0.895**	**0.893**	0.579	**0.749**	**0.707**	**0.798**
AD+MVBB L=0.25M	0.855	0.878	0.615	0.705	0.471	0.785

perhaps, is that regardless of what parameters we used in this work, AD+MVNB generated improvements compared to ADMIRE and MV, except in those *in vivo* cases where high subarray sizes degraded image quality so significantly that the image was no longer recognizable.

4.7 Conclusion

We have shown in this work that because ADMIRE and MV seek to improve images in different ways, this allows us to combine these methods to create images of superior quality than either method alone. In particular, we had hypothesized that MV would benefit from a denoising pre-processing step based on the fundamentals of the MV method, and we hoped that ADMIRE would benefit from improved lateral resolution. We implemented both narrowband and broadband implementations of MV and found that AD+MVNB produced *in vivo* images that were better than images from either method alone. AD+MVNB was able to improve resolution compared to just MVNB while simultaneously improving the contrast ratio that ADMIRE provides.

So, not only did AD+MVNB attain the best image quality of both methods, it in fact managed to improve upon the best aspects of each as well. In future work we will continue to investigate both how we can better tune ADMIRE and MV to further promote improved image quality, as well as investigate other methods that could be incorporated into the processing pipeline.

CHAPTER 5

Addressing Sizing Accuracy: Extended ADMIRE Models

This chapter is adapted from work originally published in (Schlunk et al., 2022), ©2022 Ultrasonics. Reprinted, with permission, from Schlunk, S., Hsi, R., & Byram, B. (n.d.). Enhancing sizing accuracy in ultrasound images with an alternative ADMIRE model and dynamic range considerations. Ultrasonics, *In Review*.

5.1 Introduction

Sizing in ultrasound imaging can be a challenging task due to the width of the main lobe. Hypoechoic structures such as cysts and blood vessels may appear smaller and hyperechoic structures like calcifications may appear larger. A classic clinical task where this is relevant is in the diagnosis of kidney stone disease. Diagnosis of kidney stone disease is a two-part challenge for ultrasound, requiring both the correct detection and subsequently the correct sizing of a stone. Though ultrasound struggles with detection, coherence-based methods such as short-lag spatial coherence (SLSC) and mid-lag spatial coherence (MLSC) have shown promise (Schlunk et al., 2017; Hsi et al., 2018; Tierney et al., 2019; Jones et al., 2019), and color Doppler is frequently used to enhance stone "twinkling" (Shabana et al., 2009; Abdel-Gawad et al., 2016).

In comparison, sizing has demonstrated to be the more challenging task for ultrasound beamformers, especially since the size of the stone dictates whether or not surgical intervention is necessary. It is generally accepted that stones smaller than 5mm are able to pass naturally, while larger stones require intervention (Coe, 1977; Ueno et al., 1977). Ultrasound regularly overestimates the size of kidney stones by as much as 2-3 mm (Ulusan et al., 2007; Unal et al., 2003; Ray et al., 2010; Ganesan et al., 2017; Fowler et al., 2002; Dunmire et al., 2015; Sternberg et al., 2016), which is likely to lead to an unnecessary recommendation for intervention. It has been suggested that measuring the acoustic shadow of the kidney stone can improve accuracy (Dunmire et al., 2015, 2016; Schlunk et al., 2017), though the appearance of the shadow can depend on surrounding tissue and the transducer parameters (King et al., 1985; Moore and Scoutt, 2012), and is difficult to detect for small stones (May et al., 2016; Dunmire et al., 2016). Other semi-automatic measurement techniques have been proposed and tested *in vitro* (Dunmire et al., 2015; Hsi et al., 2018), but the ability to translate these methods to *in vivo* cases may be difficult due to the significantly increased complexity of surrounding tissue. As a result, the gold standard for sizing falls to computed tomography (CT) (Smith and Varanelli, 2000; Smith et al., 1996).

This work expands on our early efforts with expanded ADMIRE models in simulations (Schlunk and Byram, 2020) and will focus on the sizing task related to characterization of kidney stone disease, and specif-

ically we will have two goals to measure our success. First, the correct classification of small (less than 5mm) and large (greater than or equal to 5mm) kidney stones, which reflects a typical clinical scenario for choosing continued monitoring versus intervention. The second task is a more general sizing accuracy task, and our goal is to minimize the error between ultrasound measured sizes and the CT sizes, ideally attaining a margin of error for each stone that is less than or equal to $\pm 10\%$ of the CT reported size. In this work, we compare classic methods such as minimum variance (MV) and generalized coherence factor (GCF) to a modified version of aperture domain model image reconstruction (ADMIRE) that is designed to improve sizing accuracy. We additionally show that this new method outperforms ADMIRE combined with MV, which we have previously demonstrated enhances lateral resolution compared to either MV or ADMIRE alone (Schlunk and Byram, 2021b, 2022b). We include simulations to get a baseline understanding of our methods, but the primary focus is on the application of our methods to *in vivo* kidney stones.

We additionally consider how sizing accuracy might be related to the dynamic range of the image. Rather than strictly measure stones at traditional dynamic ranges, we show how the dynamic range of the image impacts the size of the target as measured by the beamformers presented.

5.2 Beamforming Algorithms

We implemented all beamformers in MATLAB (The MathWorks, Natick, MA, USA).

5.2.1 Delay-and-Sum (DAS)

The conventional standard beamformer, as defined in Section 1.4.2. Since our focus for this work is on improving sizing, particularly in the case of kidney stones using ultrasound, we opt to use rectangular apodization.

5.2.2 Minimum Variance (MV)

Minimum variance (MV) beamforming is also an adaptively weighted DAS image, where the optimized apodization weights (Holfort et al., 2009), (Synnevåg et al., 2007) are calculated to minimize off-axis interference, improving lateral resolution. The complete definition can be found in Section 1.4.5.

Since the goal is this work is improving spatial resolution to improve sizing accuracy, it should be noted that adjusting the subarray lengths and amount of diagonal loading changes the resolution of the beamformer. In particular, we generally want smaller amounts of diagonal loading, while increasing the size of the subarrays will result in improvements to resolution (Synnevåg et al., 2007). However, like others, we found in our previous work involving minimum variance (Schlunk and Byram, 2022b) that increasing the subarray size too much results in significant image quality degradation, which can render an image unrecognizable. As a

result, we will use moderate lengths of $L = 0.5M$ and $\Delta = 1/(10L)$ for simulations, while *in vivo* we drop the subarray sizes to $L = 0.25M$. The exception to this is when we apply minimum variance after pre-processing with ADMIRE or AD-Ex (described in Section 5.2.6), where we found we can safely increase the subarray sizes back to $L = 0.5M$ without loss of image quality (Schlunk and Byram, 2022b).

5.2.3 Generalized Coherence Factor (GCF)

Generalized coherence factor (GCF) is a weighting of the DAS image by the ratio of the energy in some low-frequency region to the total energy (Li and Li, 2003). The complete definition can be found in Section 1.4.4. We chose the low-frequency cutoff parameter $M_0 = 5$ as done in (Rindal et al., 2017) for this work.

5.2.4 Aperture Domain Model Image Reconstruction (ADMIRE)

Aperture domain model image reconstruction (ADMIRE) works by removing reverberation clutter and off-axis interference and suppressing wavefront aberration. We discuss the method in detail in Chapter 2.

5.2.5 ADMIRE-Extended (AD-Ex)

AD-Ex is an alternative model setup for ADMIRE originally implemented with iterative ADMIRE (Schlunk et al., 2018; Schlunk and Byram, 2020). Due to the interaction of the elastic-net regularization and our model, we occasionally run into the problem of certain signals not being fit well. Specifically, since the elastic-net balances between L1 and L2 regularization, we have some of the benefits and detriments of both. L1 pushes towards sparse, selective solutions, while L2 pushes to include all predictors when there is correlation among predictors. Predictors in close proximity can be highly correlated, and thus the use of L2 in the regularization results in many of these predictors being selected together, which can be an issue if they are located near the border of the ROI. The highly-sampled ROI in the model is attractive to the elastic-net for fitting, which can result in non-ROI signals sometimes being dragged into the ROI, and therefore not being correctly removed. This is especially problematic for signals originating from near, but not in, the ROI.

The model solution we chose is $\mathrm{X} = [\mathrm{X}_{\text{clutter}}\ \mathrm{X}_{\text{ROI}}\ \mathrm{X}_{\text{extended}}]$, shown in Fig. 5.1B. We designed a small extended clutter region surrounding the ROI that has a sampling rate that falls in between that of the ROI and the rest of the clutter region. The specific sampling and size of the extended region is included in Table 5.1. We consider both $\mathrm{X}_{clutter}$ and $\mathrm{X}_{extended}$ as clutter regions to be removed. This results in clutter signals that were being fit in the ROI being fit into this intermediate clutter region instead, resulting in the correct removal of these signal components without significantly increased computation time or image quality loss.

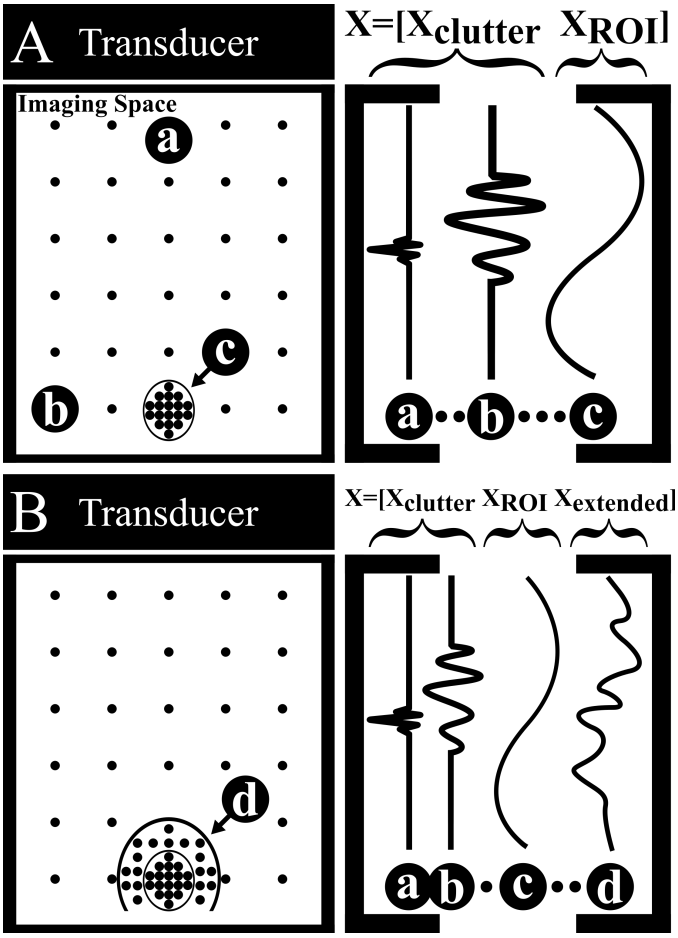

Figure 5.1: (A) Composition of an ADMIRE model X, composed of a set of estimated signals from locations considered to be clutter (sparsely sampled, e.g a and b), and a set from locations considered to be region of interest (highly sampled, e.g. c). By solving the matrix equation, we can estimate which physical sources in the modeled imaging space make up the received echo. Signals a and b represent reverberation clutter signals and off-axis signals, respectively, while c is an example of a signal from the ROI. (B) A modified model called here the AD-Ex model. It includes an extension, an intermediate clutter region $X_{extended}$ composed of signals represented by d. The proposed intermediate clutter region is sampled at half that of the ROI, but still more than the remaining clutter. This helps to ensure proper classification and removal of signals originating near, but not in, the ROI. Figure ©2022 Ultrasonics.

5.2.6 Post-ADMIRE Processing

We have shown in previous work that MV can benefit from pre-processing with ADMIRE for improving lateral resolution (Dei et al., 2016; Schlunk and Byram, 2021b, 2022b) due to the denoising that ADMIRE provides, so we include here MV applied after pre-processing with ADMIRE (AD+MV) to compare against AD-Ex. We also include both MV and GCF applied after pre-processing with AD-Ex (AD-EX+MV and AD-EX+GCF) to see how they compare against AD-Ex alone for completeness.

Table 5.1: ADMIRE and AD-Ex Model Space Parameters. Table ©2022 Ultrasonics.

Parameter	Value
α	0.9
λ	$(0.00189)y_{\text{RMS}}$
X_{clutter} Sampling	$3.577\text{res}_{lateral}, 7.154\text{res}_{axial}$
X_{clutter} Size	Full Imaging Space
X_{ROI} Sampling	$0.179\text{res}_{lateral}, 0.715\text{res}_{axial}$
X_{ROI} Ellipse Radii	$(0.5\text{res}_{lateral}, \text{res}_{axial})$
X_{extended} Sampling	$0.358\text{res}_{lateral}, 0.715\text{res}_{axial}$
X_{extended} Ellipse Radii	$(\text{res}_{lateral}, 2\text{res}_{axial})$

5.3 Methods

5.3.1 Simulated Phantoms

We used Field II (Jensen, 1996; Jensen and Svendsen, 1992) to simulate n=6 5mm hyperechoic cysts at amplitudes of 40, 50, 60, and 70 dB as rudimentary equivalents to kidney stones at our targeted size of interest. We additionally simulated single point targets with 60dB signal-to-noise ratio (SNR) of added Johnson-Nyquist noise to perform resolution comparisons between methods. The simulation parameters for both simulations are the same as detailed in Table 2.2.

We estimated the lateral amplitude profiles of the simulated cysts by radially averaging the amplitude at equal distances from the center of the cyst within $\pm 10^\circ$ of the lateral axis, as is shown in Fig. 5.2. For each method and cyst we calculated the lateral cyst size from the profiles at the 30dB and 6dB thresholds. The 30dB threshold represents how a sonographer would normally size from a b-mode image, since 30dB or lower thresholds all measure equivalently, as can be seen from the amplitude plot in the figure. The 6dB threshold is an alternative based on how the full width at half maximum (FWHM) is measured for resolution metrics. For the simulated point targets, the lateral point-spread functions were evaluated for each method.

5.3.2 *In vivo* Kidney Stone Data

We captured *in vivo* kidney data from patients with one or more kidney stones using a Verasonics Vantage Ultrasound System (Verasonics, Inc., Kirkland, WA) with a C5-2 curvilinear transducer and a curvilinear plane wave synthetic focusing acquisition (Montaldo et al., 2009), where synthetic transmit focuses are placed throughout the image to achieve focusing at all depths. A center frequency of 4.1667 MHz was used to acquire 64 angles uniformly spaced spanning 37°. For every patient and stone, a previous CT scan was available from which the stone size was measured, and the longest axis was determined. This allowed the physician to acquire the ultrasound data such that the longest axis was in the lateral dimension. We acquired n=8 kidney stone images across 6 different patient volunteers. For this work, we used the CT measurements as the gold

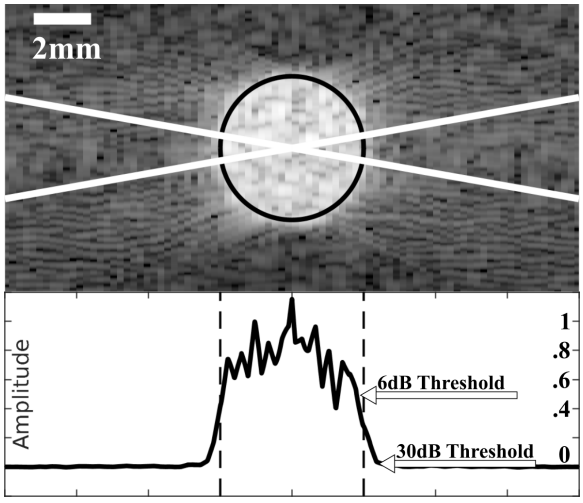

Figure 5.2: (Top) DAS of an example 40 dB simulated 5mm cyst, with the solid black line indicating the true cyst boundary. The white lines show the radial region where the cyst boundary was estimated. (Bottom) plot of radial average amplitude, with the dashed lines indicate the true edges of the cyst. We indicate two thresholds, 6dB and 30dB, to show how they would underestimate and overestimate cyst size, respectively. Figure ©2022 Ultrasonics.

standard, and we round stone sizes to the nearest millimeter strictly for classification (e.g. a 4.7mm stone is considered a large stone for diagnostic purposes, but a 4.4mm stone is not) mirroring the clinical procedure at our institution, so our exact threshold target is 4.5mm. In total, CT stone sizes ranged from 2.3 to 12.2 mm, comprising of two small stones (less than 4.5mm) and six large stones (greater than or equal to 4.5mm). The data was acquired under protocols approved by the Vanderbilt University IRB (IRB# 170001).

5.3.3 Kidney Stone Image Metrics

First, for every stone and method, we needed to segment the stone image for sizing and image quality purposes. We first had a physician manually size the DAS images of each stone, which was used alongside a contour map of each image as a reference to manually draw a stone ROI on the b-mode images displayed on a 30dB dynamic range. Since this work focuses solely on the sizing task of diagnosis, all stone locations were already known when measuring the images, but the specific CT measured sizes were not. From Fig. 5.2, sizing at 30dB or a more traditional 50dB should be comparable, so we chose to use 30dB for sizing for improved stone visibility. From this ROI, the stone size was automatically estimated as the longest distance within $\pm 10^{\circ}$ of the lateral axis, since it is difficult to ensure the longest stone dimension is perfectly lined up with the lateral axis. The background region is also automatically selected as a circular region centered on the middle of the stone ROI, with an area of 2.5 times the area of the stone ROI (to compensate for the non-standard shapes of the kidney stones). This process is challenging due to the extremely wide variety of shapes kidney stones can take, as shown in Fig. 5.3, which can make it difficult to differentiate a weirdly-

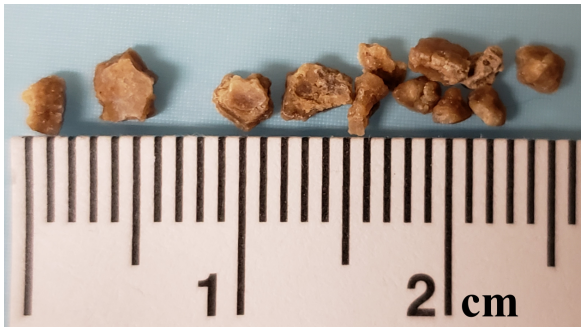

Figure 5.3: Examples of small *ex vivo* kidney stones, to demonstrate the wide variety of shapes they can take. Figure ©2022 Ultrasonics.

shaped large kidney stone versus a smaller stone with adjacent bright tissue. As a result, we try to balance the physician measurements with the contour maps to be as objective as possible with the ROI selections. All of the stones and the ROI selections can be found in the supplementary materials for transparency. An example of this process for DAS is shown in the top row of Fig. 5.5.

We finally used these manually drawn ROI's at the 30dB threshold to continuously estimate the stone ROI at thresholds down to 6dB, allowing us to plot the measured stone size versus threshold at which it was measured. This allowed us to see if there was a trend for any given method that indicates some threshold where the method is most accurate.

To evaluate the stone size measurements, all methods were compared to the gold standard CT measurements that were acquired before the ultrasound acquisitions. We calculated the percent stone error for each method compared to CT, and we additionally evaluated methods based on whether stones were correctly classified as small (< 4.5mm) or large (≥ 4.5mm). As mentioned in the previous section, we choose 4.5mm as our cutoff rather than 5mm since we are rounding to the nearest millimeter for classification.

For completeness, with the stone ROI selected and a background region chosen, we computed the contrast ratio and generalized contrast-to-noise ratio (gCNR) (Rodriguez-Molares et al., 2018, 2020) of each stone. Contrast ratio was defined as

$$\text{contrast ratio} = -20\log_{10}\left(\frac{\mu_{\text{ROI}}}{\mu_{\text{background}}}\right) \tag{5.1}$$

where μ is the mean value calculated from the enveloped data, before log compression. gCNR is a generalized detectability metric which measures the overlap of the probability density function between the ROI and the background, making it more robust against stretches or compressions in dynamic range. For this work, we used the standard 100 histogram bins for estimating the overlap.

Table 5.2: Measured 5mm Cyst Size at Various Amplitudes. Table ©2022 Ultrasonics.

Measured Size at 30 dB Threshold (mm)				
Cyst	40 dB	50 dB	60 dB	70 dB
DAS	6.23±0.22	6.29±0.22	6.29±0.23	6.43±0.25
MV	5.86±0.08	5.90±0.06	5.92±0.07	5.92±0.08
GCF	5.83±0.07	5.83±0.07	5.83±0.07	5.89±0.14
ADMIRE	5.87±0.06	5.90±0.06	5.90±0.08	5.98±0.17
AD+MV	5.74±0.08	5.78±0.06	5.79±0.06	5.83±0.09
AD-Ex	5.74±0.09	5.76±0.07	5.77±0.06	5.81±0.10
AD-Ex+MV	5.56±0.11	5.58±0.10	5.59±0.10	5.61±0.12
AD-Ex+GCF	5.69±0.12	5.71±0.11	5.71±0.11	5.75±0.13
Measured Size at 6 dB Threshold (mm)				
Cyst	40 dB	50 dB	60 dB	70 dB
DAS	4.77±0.34	4.78±0.33	4.78±0.33	4.72±0.38
MV	3.95±0.43	3.99±0.50	3.99±0.50	3.99±0.50
GCF	4.27±0.67	4.27±0.67	4.27±0.66	4.23±0.70
ADMIRE	4.77±0.35	4.78±0.34	4.78±0.34	4.72±0.39
AD+MV	4.17±0.64	4.18±0.64	4.18±0.64	4.17±0.65
AD-Ex	4.69±0.58	4.69±0.58	4.69±0.58	4.64±0.60
AD-Ex+MV	4.15±0.64	4.25±0.68	4.25±0.68	4.22±0.71
AD-Ex+GCF	4.53±0.64	4.53±0.64	4.53±0.64	4.34±0.73

5.4 Results

The average estimated widths of the 5mm simulated hyperechoic cysts are included in Table 5.2. The 30dB threshold measurements show that all methods tend to overestimate the size of the cyst by 0.5-1.5 mm, or by more than 10%. In comparison, at 6dB all methods underestimated the cyst size, implying that if we measured at some intermediate threshold we would size cysts accurately. However, each method is off by varying amounts, which means that the accurate threshold would be different for each method. For example, DAS and ADMIRE are fairly accurate at 6dB, but MV underestimates significantly. Additionally, at 30dB we see that methods tend to overestimate by greater amounts the brighter the cyst is, whereas at 6dB the measurements are fairly consistent across different intensities.

The point spread functions for the simulated single point targets are included in Fig. 5.4. The MV, AD+MV, and AD-Ex+MV methods show noticeable improvements at the apex of the PSF compared to other methods. AD-Ex shows improvement by lowering the amplitude of the side lobes, which is also reflected in the AD-Ex+MV and AD-Ex+GCF methods. AD-Ex+MV shows additional main lobe improvement compared to MV, AD-Ex, or AD+MV at all amplitudes.

Fig. 5.5 shows an example of the *in vivo* sizing process for DAS (top row) and the resulting region selections for the other methods (bottom row) for kidney stone Case 1.

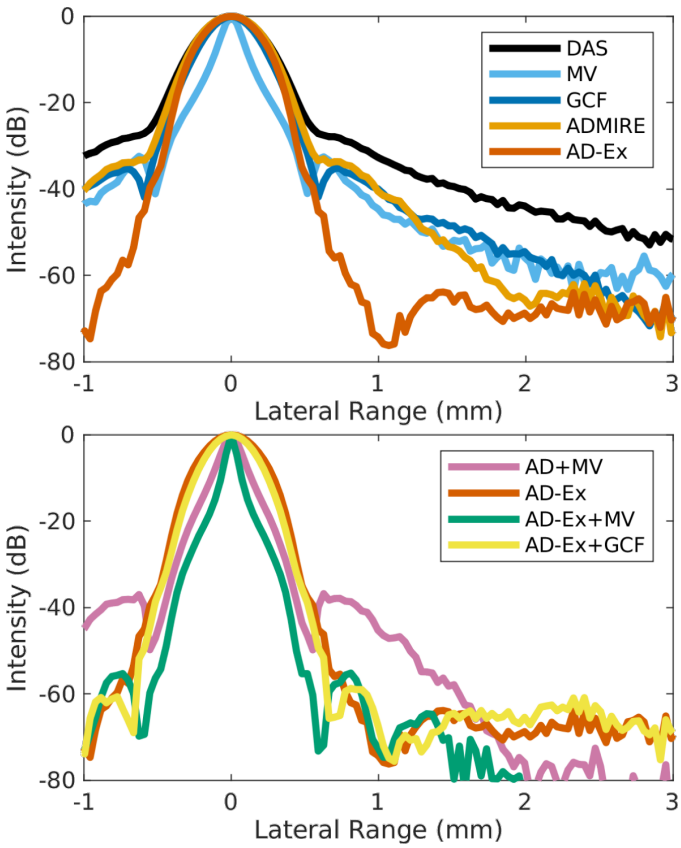

Figure 5.4: Point spread functions of simulations of a single point target with 60 dB SNR of added thermal noise. DAS, GCF, and ADMIRE all have similar profiles, while MV has a narrower profile in the range of -20 to 0 dB. AD-Ex has a similar profile to ADMIRE, but narrows at levels below -40 dB. AD+MV and AD-Ex+MV show a narrower profile compared to either method alone at all levels, while AD-Ex+GCF is nearly identical to AD-Ex. Figure ©2022 Ultrasonics.

In Fig. 5.6, the numbered triangles indicate the stone case and the threshold at which the stone is measured accurately, i.e. where the solid measurement line meets the dashed CT size line. DAS appears to be accurate mostly around -10dB and higher, while GCF and AD-Ex are much more varied. These plots generally agree with our cyst simulations, showing more overestimation towards -30dB, and less or even underestimation towards -6dB. It also supports our decision to make measurements at -30dB, since the curves flatten out considerably in that range for all methods and stones, meaning the sizing would be fairly similar to a more normal b-mode image at 50dB, though with improved visibility of the stones. However, the large variability of the optimal thresholds, especially between methods, makes it difficult to determine if there exists an optimal threshold for stone sizing. As a result, choosing the 30dB threshold for the rest of the measurements was the only objective way we could compare each method.

Table 5.3 compiles all of the sizing measurements for each stone case based on the stone ROI drawn on the 30dB dynamic range b-mode image for each method, and Fig. 5.5 shows an example for Case 1. The data

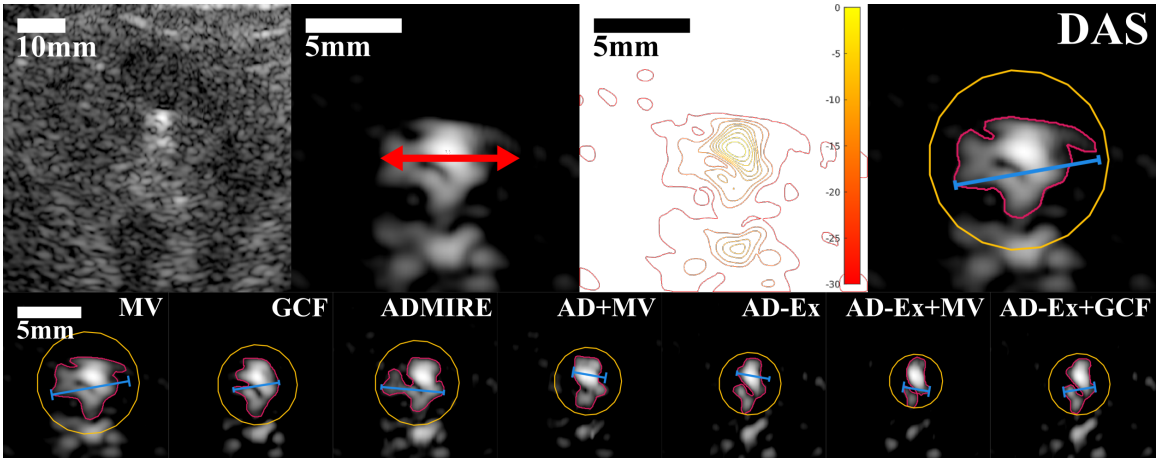

Figure 5.5: Example of the sizing process for *in vivo* kidney stone Case 1 (CT size of 2.3 mm). The top row shows the process for DAS, starting with the full kidney image on a 50 dB scale. The zoomed images are shown at 30 dB and were used for sizing. The red arrows show the physician-marked stone size, which was used in combination with the contour map to manually circle the stone region for measurements, marked as the red outline. The ultrasound images were acquired such that the stone's longest dimension is on the lateral axis, so we calculated the stone size as the longest distance within $\pm 10^{\circ}$ of that axis, marked as the blue line. The yellow line shows the region used for the background, which is automatically selected as a circular region double the area of the stone region. The bottom row shows the end results for the same case for MV, GCF, ADMIRE, AD+MV, AD-Ex, AD-Ex+MV, and AD-Ex+GCF. Figure ©2022 Ultrasonics.

Table 5.3: Size Measurements for *in vivo* Kidney Stones (Dashed line separates non-surgical vs. surgical cases). Table ©2022 Ultrasonics.

Measured Size (mm)								
Case	1	2	3	4	5	6	7	8
CT	2.3	3.4	4.7	5.4	5.9	6.3	7.1	12.2
DAS	7.5	5.4	10.2	6.2	12.0	8.5	14.6	14.1
MV	6.8	5.6	10.0	6.4	12.1	8.9	14.5	14.0
GCF	4.2	4.5	9.0	5.7	9.4	6.8	12.2	**12.6**
ADMIRE	5.7	5.5	9.0	6.8	8.5	7.4	14.1	**12.6**
AD+MV	3.4	4.5	7.5	7.1	7.8	5.3	10.0	12.8
AD-Ex	3.0	3.3	7.8	**5.5**	8.9	**6.3**	11.1	11.7
AD-Ex+MV	**2.4**	**3.4**	**6.7**	6.6	**5.6**	6.1	**9.7**	**11.8**
AD-Ex+GCF	2.8	3.3	7.6	7.9	8.9	6.1	11.1	**11.8**
Percent Error Compared to CT (%)								
Case	1	2	3	4	5	6	7	8
DAS	226.1	58.8	117.0	14.8	103.4	34.9	105.6	15.6
MV	195.7	64.7	112.8	18.5	105.1	41.3	104.2	14.8
GCF	82.6	32.4	91.5	5.6	59.3	7.9	71.8	**3.3**
ADMIRE	147.8	61.8	91.5	25.9	44.1	17.5	98.6	**3.3**
AD+MV	47.8	32.4	59.6	31.5	32.2	-15.9	40.8	4.9
AD-Ex	30.4	-2.9	66.0	**1.9**	50.8	**0.0**	56.3	-4.1
AD-Ex+MV	**4.3**	**0.0**	**42.6**	22.2	**-5.1**	-3.2	**36.6**	**-3.3**
AD-Ex+GCF	21.7	-2.9	61.7	46.3	50.8	-3.2	56.3	**-3.3**

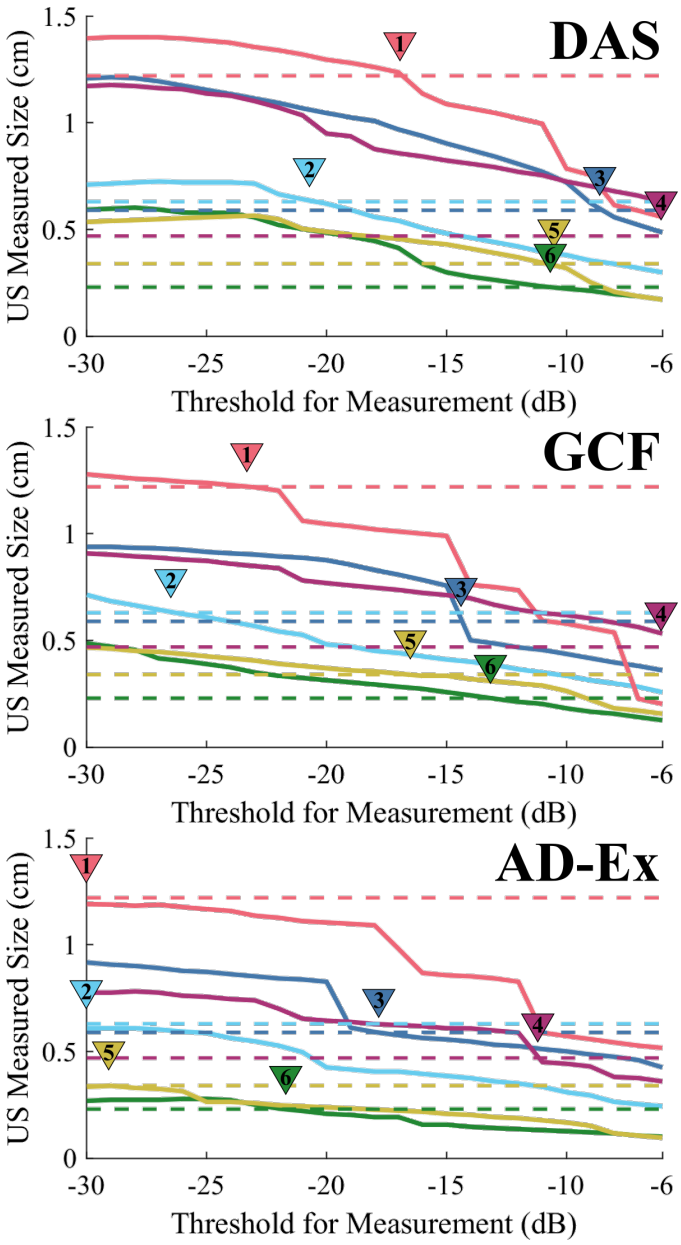

Figure 5.6: Examples of measured stone size for DAS, GCF, and AD-Ex for each stone at thresholds from -30 to -6 dB. The solid lines indicate the measured size at each threshold, and the dashed lines indicate the CT measured stone size. The colored arrows indicate the location where the solid and dashed lines intersect, i.e. the threshold where the US method measurement is equal to CT stone size. Figure ©2022 Ultrasonics.

shows that the stones were very rarely underestimated, and most instances of underestimation came from the larger stones. For the smaller stones under our diagnostic threshold of 4.5mm, AD-Ex correctly assigns both of them as being smaller than 4.5mm, though only one of them is measured with an error of less than $\pm 10\%$. AD-Ex+MV improves upon this by sizing both with an error under our threshold, though AD-Ex+GCF shows less of an improvement. GCF and AD+MV both correctly classify Case 1 as a small stone, though for Case 2

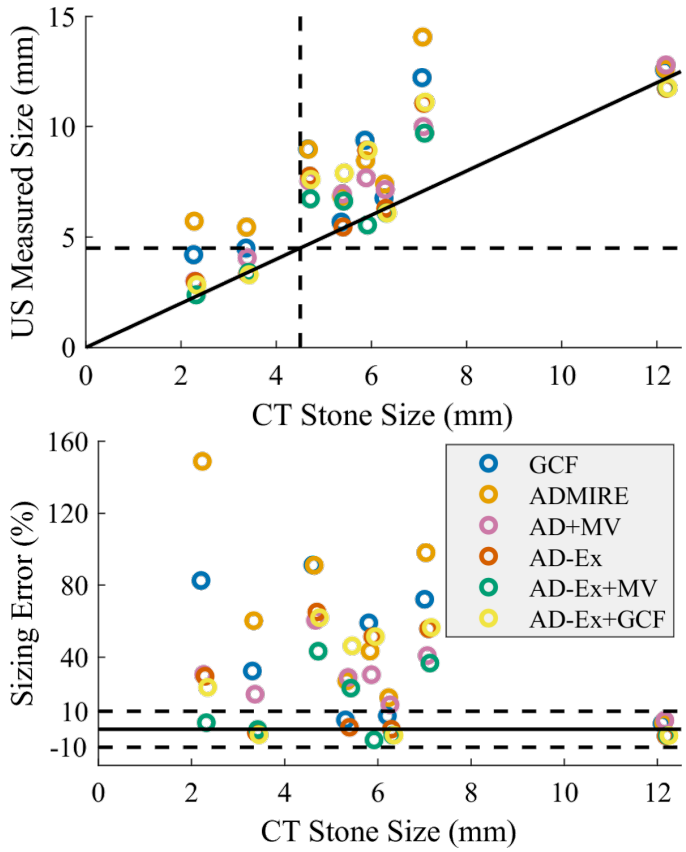

Figure 5.7: (Top) US measured stone size versus CT stone size for all cases for classifying small ($<$ 4.5mm) or large ($\geq$ 4.5mm) stones. Note that the methods are slightly staggered for visibility. The dashed lines are positioned at stone sizes of 4.5 mm, our targeted cutoff for diagnostic intervention. (Bottom) percent sizing error versus CT stone size for all cases. The dashed lines are positioned at $\pm$10%. Figure ©2022 Ultrasonics.

both measure exactly at our threshold for large stones. No other method produced a size of less than 4.5mm for either of these small cases.

Considering the six "large" cases, Cases 3-8, all methods correctly classify them as larger than 4.5mm, and GCF, AD-Ex, and AD-Ex+MV each size three of the cases within $\pm$10% of the CT measurement. Additionally, AD-Ex+GCF is within our margin of error for two cases. However, with the exception of Case 8, the other methods had generally poor accuracy compared to CT. Across all cases, AD-Ex+MV had the lowest average error at 11.8%, followed by AD-Ex at 24.8%, AD-Ex+GCF at 28.4%, and AD+MV at 28.8%. All other methods had average errors of over 40%.

The contrast ratio and gCNR calculated from the stone ROI for each case and method are included in Table 5.4, giving us a quantification of the visibility of these stones, with higher values implying easier visibility. Generally, GCF or AD-Ex+GCF had the greatest contrast ratio and gCNR of all the methods, and both perform consistently better than DAS and MV. AD-Ex+MV and AD-Ex+GCF had similar contrast

Table 5.4: Contrast Ratio and gCNR for *in vivo* Kidney Stones. Table ©2022 Ultrasonics.

Stone Contrast (dB)								
Case	1	2	3	4	5	6	7	8
DAS	15.51	11.35	14.92	7.52	13.03	14.83	11.66	12.54
MV	14.84	10.50	14.82	5.77	12.85	14.05	11.54	12.40
GCF	24.04	**23.29**	**23.33**	**13.92**	19.67	20.07	21.54	**19.13**
ADMIRE	19.71	14.23	18.32	5.85	14.99	18.44	15.20	15.16
AD+MV	23.61	17.37	19.54	6.10	14.01	17.12	13.62	14.83
AD-Ex	25.34	20.72	15.62	11.61	20.43	19.85	22.42	18.04
AD-Ex+MV	**27.29**	22.50	18.23	11.85	20.33	20.61	21.84	17.25
AD-Ex+GCF	26.34	22.21	15.71	11.15	**20.84**	**20.84**	**23.90**	18.41
Stone gCNR								
Case	1	2	3	4	5	6	7	8
DAS	0.78	0.72	0.77	0.59	0.65	0.73	0.54	0.70
MV	0.78	0.66	0.79	0.48	0.64	0.70	0.54	0.73
GCF	0.91	**0.91**	0.85	**0.63**	0.80	0.85	0.85	**0.81**
ADMIRE	0.86	0.76	0.84	0.44	0.75	0.82	0.72	0.76
AD+MV	0.90	0.74	**0.86**	0.38	0.19	0.81	0.04	0.64
AD-Ex	**0.94**	0.88	0.69	0.57	0.84	0.84	0.82	0.78
AD-Ex+MV	**0.94**	0.81	0.75	0.57	**0.90**	0.84	0.83	0.70
AD-Ex+GCF	**0.94**	0.88	0.69	0.54	0.84	**0.87**	**0.87**	0.77

ratio and gCNR compared to AD-Ex, with both on average having slightly improved contrast ratio. As is consistent with our previous work, AD-Ex+MV performed better than MV alone always boosting contrast ratio and often boosting gCNR, however, AD-Ex+GCF was not a improvement compared to GCF usually on par or worse in terms of contrast ratio and gCNR.

Considering both sizing accuracy and contrast ratio and gCNR, AD-Ex-based methods, AD+MV, and GCF reported better accuracy and visibility compared to the other methods, with AD-Ex+MV coming ahead of the others due to better sizing accuracy. Since MV primarily narrows the main lobe between -20 and 0 dB, it did not provide a noticeable improvement compared to DAS when measuring on a 30dB dynamic range, and when measuring at -6dB as we did for the simulated cysts, DAS was already underestimating true cyst size which makes MV underestimate by an even greater amount. Since clinical measurements would be made more in the range of 30dB or higher, it is reasonable to expect that MV would not demonstrate an improvement for sizing *in vivo*. However, when combined with ADMIRE or AD-Ex, the simulations showed that the main lobe is further narrowed over a much larger range, which does result in improvements *in vivo*.

5.5 Discussion

When considering the resolution of a beamformer, the FWHM is a common performance metric, but for clinical sizing applications where the dynamic range is much larger, it may not be as useful to gauge performance.

For example, ADMIRE's normal ROI is not designed to enhance the FWHM compared to DAS, but by reducing clutter and improving contrast ratio it generally performs better at sizing compared to even MV. To further improve ADMIRE's accuracy, we knew that addressing the off-axis interference caused by mis-fitting nearby signals was important. An obvious consideration was to only use L1 for regularization since L2 is a driving factor in this fitting problem, but this has significant drawbacks of degrading image quality. We also considered other solutions, such as increasing the sampling of the clutter region or adjusting the size of the ROI to compensate, but this often resulted in dramatically increased computation time or increased variance in the image or other unexpected or undesired image quality changes. AD-Ex demonstrated the best balance between computation time and image quality while still accomplishing our goal of improving resolution in the range of clinical measurements.

The point target simulations suggest that, in ideal conditions, MV improves the FWHM while AD-Ex improves main lobe performance at the base. It demonstrates the advantages of combining AD-Ex and MV, since they improve different aspects of the main lobe, and the combined method AD-Ex+MV carries over and improves both of these aspects. However, these simulations do not fully represent *in vivo* performance. For example, MV appears better than GCF in these simulations, while *in vivo* GCF demonstrates a significant improvement compared to MV. ADMIRE, and AD-Ex especially, generally reduce clutter to enhance the target while MV sharpens the target, which can result in different stone appearances based on content of the original image. This overall suggests that improving the clarity of the coherent target, such as by boosting coherence with GCF or decluttering with ADMIRE, is critical to improving sizing accuracy, rather than strictly improving resolution.

The cyst simulations show how ultrasound is overestimating hyperechoic targets. In particular, overestimation occurs when we measure the targets at traditional b-mode dynamic ranges (30-50 dB or more). Since methods like MV narrow the main lobe mostly in the range of -20 to 0 dB, this results in marginal improvements measured at 30dB. In comparison, measuring on an image with a dynamic range of 6dB (i.e. at the FWHM) results in underestimation, and a much larger difference for MV compared to DAS. This implies that there may be a theoretical dynamic range where a method is more likely to measure a target correctly, though the *in vivo* examples indicate that this threshold would vary between methods, and potentially depend on the amplitude of the target versus the background. In the *in vivo* cases, DAS overestimates consistently at -30dB, and mostly seems to be accurate between -6 and -12 dB. However, GCF and AD-Ex are substantially more variable for the same cases. Eventually this could potentially be a tool to improve ultrasound sizing accuracy, though it also follows that this threshold would have to be tuned for each method of interest.

For the sizing tasks *in vivo*, we continue to observe as in previous work that the ultrasound methods tend to overestimate size compared to CT. The general overestimation of stones resulted in all methods

correctly classifying large stones, but most also incorrectly classified the small stones as large, which in a clinical setting would likely result in a follow-up CT to confirm diagnosis. While pre-processing with AD-Ex resulted in correctly classifying every stone regardless of post-processing, there are still improvements to be made to decrease the percent error. It is, however, still a substantial improvement over traditional DAS, and demonstrated the most consistent improvement among the other methods. AD-Ex, AD-Ex+MV, AD-Ex+GCF, AD+MV, and GCF in most cases demonstrated improvements in sizing compared to DAS, which may partially be attributable to the increases in measured contrast ratio and gCNR. Though we do not focus on detection here as has been done elsewhere, these metrics subjectively related to the ease of drawing the ROI for those methods, especially compared to DAS and MV, where there were often challenges at differentiating surrounding tissue versus the stone, even at a dynamic range of 30dB.

Finally, we have demonstrated that AD-Ex is able to match and surpass the performance of AD+MV, which we have shown previously as a way to enhance lateral resolution beyond what ADMIRE or MV can achieve alone. This is especially valuable, since AD-Ex is still simply a pre-processing step, allowing us to further post-process the AD-Ex data with other beamformers. By achieving similar results with AD-Ex compared to AD+MV, we not only save the computation time that is normally required for MV, but we can realize further improvements depending on our choice of post-processing, as we have shown with AD-EX+MV, which outperformed all other methods.

5.6 Conclusion

AD-Ex shows promise as a pre-processing step for improving sizing accuracy in ultrasound. Alone, AD-Ex classified all stones correctly with an average error of 24.8%. However, the strength of AD-Ex compared to other methods is that we can perform additional post-processing, and in this work we have shown that post-processing with MV results in lowering the average error to 11.8%. Overall, methods that had AD-Ex pre-processing performed the best at our stone size classification task, correctly diagnosing all eight cases as small or large for intervention purposes, with AD+MV and GCF coming in a close second at seven cases, though with poorer average errors of 28.8% and 44.3%, respectively. While the mechanics for sizing are similar for hypoechoic structures, future work will involve more explicit investigation of these targets.

Finally, while our data does suggest that there may be a theoretical dynamic range that would produce more accurate ultrasound sizing measurements, it will require a larger data set. If in the future such a data set was available, this would potentially be a great target for a deep learning approach as well, where a machine might be able to tease apart some interactions or parameters that a practitioner can not, for example taking into account factors such as amount and type of surrounding tissue, transducer parameters, or even beamformer specific factors like general image quality.

CHAPTER 6

Considerations for a more robust implementation of gCNR

This chapter is adapted from work originally published in (Schlunk and Byram, 2022c), ©2022 IEEE. Reprinted, with permission, from Schlunk, S., & Byram, B. (2022). Robust generalized contrast-to-noise ratio. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, *In Review*.

As the primary goal of this dissertation was to create beamformers that accurately represented the imaging target, the robustness of image quality metrics was frequently considered. One such metric that has gained popularity in recent years is generalized contrast-to-noise ratio (gCNR), a lesion detectability metric designed to be a more robust version of contrast-to-noise ratio (CNR), both of which we included in the previous chapter. Though an improvement over traditional CNR, we found in our work that gCNR could be manipulated in the right circumstances. We lay out these discoveries and some considerations to improve its implementation in this chapter.

6.1 Introduction

Beamforming methods in ultrasound imaging are most frequently judged either qualitatively by experts or quantitatively with image quality metrics. Historically, image quality metrics have been proposed as quantitative measurements that are predictive of clinical performance. Patterson and Foster proposed the contrast-to-speckle ratio (CSR) as a metric that correlated well with subjective assessments of image quality (Patterson and Foster, 1983), which today we know as contrast-to-noise ratio (CNR). Smith et al. proposed a similar metric, CNR_{ψ}, which can be related to the optimal signal-to-noise ratio (SNR) by incorporating the size of the lesion and an ultrasound system's spatial resolution (Smith et al., 1983). By using a contrast/detail phantom and choosing a threshold value for detection of each cyst, this SNR calculation was able to predict cyst detection. This suggests that CNR, which can be related to SNR, is a valuable image quality metric for beamformer comparison as an increase in CNR predicts an increase in lesion detectability.

Over time, the use of metrics such as contrast ratio and CNR have become the de facto evaluation standard, as it is faster and more convenient than having a sonographer or radiologist evaluate dozens of images whenever a beamformer is modified. As a result, adaptive beamformers are often designed with these metrics in mind, and as these beamformers have become more common, so have the concerns over whether these beamformers are producing "real" improvements. To answer these questions, several new methods and metrics have been proposed. Contrast ratio dynamic range (CRDR) suggests that measuring relative contrast accurately is important, rather than arbitrarily increasing contrast ratio everywhere (Dei et al., 2017;

Schlunk et al., 2018, 2021). Histogram matching calculates the transformation between two methods using histograms, which allows a user to better discern true differences between methods (Bottenus et al., 2020a). Finally, and the topic of this work, generalized contrast-to-noise ratio (gCNR) is a lesion detectability metric designed to be robust against dynamic range alterations (Rodriguez-Molares et al., 2018, 2020).

However, as has specifically been demonstrated by the original gCNR works (Rodriguez-Molares et al., 2018, 2020), it is relatively easy to make a beamformer that artificially inflates CNR. Additionally, calculating CNR on log-compressed versus linear data results in different values, making it less ideal as a universal metric of lesion detectability. As a result, the value of a lesion detectability metric like gCNR, that is transform-resistant, is clear.

The gCNR metric determines the probability of detection by measuring the overlap between estimated probability density functions of a target and reference region of an image. In theory, the use of these probability functions makes gCNR resistant to dynamic range alterations. The standard implementation is to compare the histograms of the target and reference regions using 100 bins, and measure the overlap by the counts of the bins. The relative ease of this implementation of gCNR no doubt adds to the popularity of the metric, though Hyun et al. does suggest that gCNR when estimated this way is reliant on the choice of the number of bins of the histogram (Hyun et al., 2022). In this work, we seek to further expand on this line of thought by evaluating circumstances where gCNR may not work as intended, both with beamformers that can "trick" the gCNR metric as well as suboptimal implementations that can manipulate the metric. Further, we will consider and compare different implementations and suggest several that are more robust against these manipulations.

6.2 Methods

MATLAB (The MathWorks, Natick, MA, USA) was used for implementations of all simulation methods, beamforming algorithms, and image quality metrics.

We used Field II (Jensen and Svendsen, 1992; Jensen, 1996) to simulate lesions of varying amplitudes including anechoic, -20, -10, and 0 dB, with n=6 realizations for each. Lesions had a diameter of 5mm, and were acquired with a 117-element, 0.257mm pitch linear array transmitting at 3 MHz with a bandwidth of 60%, focused at the depth of the lesion at 30mm.

6.2.1 Beamforming Algorithms

As this work is focused on the analysis of gCNR, the discussion for the following beamformers is reserved for the original works, and we include here only specific implementation details.

6.2.1.1 Delay-and-Sum (DAS)

We implemented DAS with rectangular apodization as described in Section 1.4.

6.2.1.2 Minimum Variance (MV)

We implemented MV (Holfort et al., 2009; Synnevåg et al., 2007) as described in Section 1.4. Subarray lengths of $L = 0.5M$ and diagonal loading of $\varepsilon = \Delta \cdot \mathrm{tr}(\hat{R})$, where $\Delta = 1/(10L)$ were applied.

6.2.1.3 Generalized Coherence Factor (GCF)

We implemented GCF (Li and Li, 2003) with a cutoff frequency index of $M_0 = 5$, as described in Section 1.4.

6.2.1.4 Short-Lag Spatial Coherence (SLSC)

We implemented SLSC (Lediju et al., 2011) as in Section 1.4. The number of lags used was 20, corresponding to approximately 17% of the active aperture.

6.2.1.5 Aperture Domain Model Image Reconstruction (ADMIRE)

The high degree-of-freedom version of ADMIRE (Byram et al., 2015a; Dei and Byram, 2017; Byram and Jakovljevic, 2014) was implemented as in Chapter 2 and in our previous work (Schlunk et al., 2021).

6.2.1.6 Gray Level Transform (GLT)

The gray level transform (GLT) is a "fake" beamformer introduced by Rindal et al. (Rindal et al., 2019) as an example of a beamformer that merely transforms data with the specific goal of boosting CNR. The equation of the transform they introduced is

$$\hat{S}_{\mathrm{GLT}}(B) = \frac{1}{1 + e^{-\alpha(B-\beta)}} \tag{6.1}$$

$$S_{\mathrm{GLT}}(B) = \frac{\hat{S}_{\mathrm{GLT}}(B) - \max(\hat{S}_{\mathrm{GLT}}(B))}{\varepsilon}, \tag{6.2}$$

where $B = 20\log_{10}(|S_{\mathrm{DAS}}|)$, S_{DAS} is the DAS signal, $\alpha = 0.12$, $\beta = -40$, and $\varepsilon = 0.008$. We include this beamformer for completeness.

6.2.1.7 Envelope Power Transforms (EPT)

The envelope power transform (EPT) is similar to the GLT in spirit, but we have designed it specifically for beating the standard implementations of gCNR. We define it as a transform on the enveloped data, $|S_{\mathrm{DAS}}|$,

$$\mathrm{EPT_n}(|S_{\mathrm{DAS}}|) = |S_{\mathrm{DAS}}|^n, \tag{6.3}$$

where n is the desired power. For demonstration, we use $\text{EPT}_{.5}$, EPT_2, and EPT_4 in this work, which correspond to taking the square root, the second power, and the fourth power of the enveloped data. This result is log compressed as usual when displayed.

6.3 generalized Contrast-to-Noise Ratio (gCNR)

An in-depth explanation of generalized contrast-to-noise ratio (gCNR) can be found from the original authors (Rodriguez-Molares et al., 2018, 2020), but we include here the important aspects of the method to understand how manipulation can occur.

gCNR is presented as a more robust version of contrast-to-noise ratio (CNR), where the metric is calculated from the overlap of estimated probability density functions (PDFs) of the target area (e.g. a lesion) and some reference area (e.g. adjacent tissue). Fig. 6.1 shows an example of two such functions, and solving for the overlap only requires calculating the intersection between the two curves and integrating over the respective areas of each. The actual gCNR metric is then calculated as

$$\text{gCNR} = 1 - \text{OVL}, \tag{6.4}$$

where OVL is the area of overlap of the two curves. In practice, finding the actual PDF of either the lesion or background tissue can be difficult, so the original implementation estimates the PDFs using histograms with 100 bins, of which an example is shown in Fig. 6.1.

6.3.1 Potential Manipulations

Histograms have been around since the seventeenth century (Scott, 2010), and it is thought that a properly built histogram is an estimator of the underlying PDF (Scott, 1979). However, choosing the ideal binning of data for a histogram is an issue that has been debated for years, with Herbert Sturges writing about the optimal number of bins back in 1926 (Sturges, 1926). Using too few bins generally results in oversmoothing of the data, while too many will likewise lead to undersmoothing (Scott, 2010). Because of this, using a suboptimal number of bins can lead to misrepresentation of the data, obfuscating the PDF we are trying to estimate, and therefore interfering with our ability to accurately estimate the gCNR.

As a practical example of this, consider Fig. 6.2. Previously, Hyun et al. observed that gCNR is dependent on the number of bins used (Hyun et al., 2022), but here we show that it is related to the ratio of bins to number of pixels (or data points) used in the estimation, and varying either of those will result in the gCNR changing. By itself this does not necessarily seem like a probelm as long as there is consistency between the number of pixels and number of bins used when comparing beamformers; however, we show in Fig. 6.3 that varying

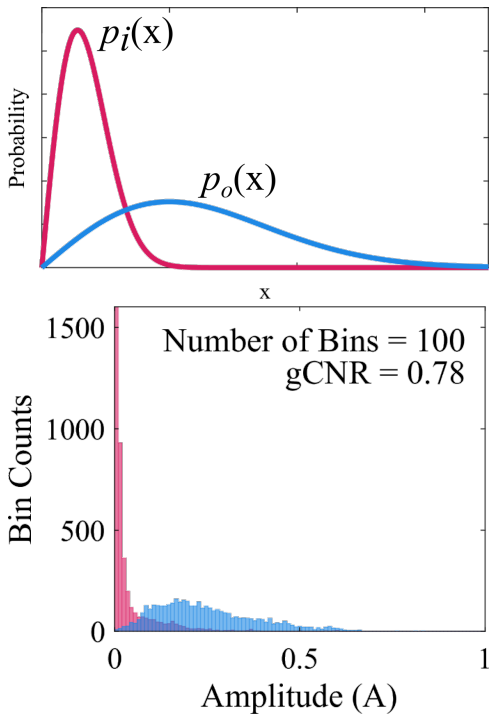

Figure 6.1: (Top) Example of two probability density functions, $p_i(\mathrm{x})$ and $p_o(\mathrm{x})$, representing the lesion and reference area, respectively. (Bottom) Histograms of an anechoic cyst simulation to mimic the functions in the above plot, implemented with the standard 100 bins. The gCNR estimate is then based on the overlap in each bin. Figure ©2022 IEEE.

the number of bins (or ratio of pixels/bins) used in the histograms can result in altering the rank-order of the evaluated beamformers. In this example, ADMIRE has the highest gCNR when estimating with 10 bins, while DAS is the best for 100 or 1000 bins. SLSC ends up being the best if an extreme number of bins (10000) is used.

A second manipulation can arise even if the number of pixels and bins are locked. Consider the EPT described before, where the enveloped data is taken to some chosen power. When the data is square rooted, squared, etc. the relative positioning of the data along the dynamic range can be manipulated. Fig. 6.4 shows DAS, $\mathrm{EPT}_{.5}$, and EPT_2 with the associated histograms for a -20dB lesion. Since all three methods are simply transforms of each other, the gCNR should be equal in all three cases, however, by manipulating the distribution of the data points along the data's range, the amount of overlap can change. This occurs due to the estimated PDFs being altered (even if the overlap should stay the same), and the binning scheme being unable to accurately represent the distributions. In the case shown here, as the skew of the data is intensified, the ability of the histogram to estimate and represent the PDF is diminished. The traditional calculation of CNR is included (Patterson and Foster, 1983) both on the linear and log-compressed data. Since the two transformations are based on powers, the powers cancel out such that the $\mathrm{CNR}_{\log}$ is the same across all three

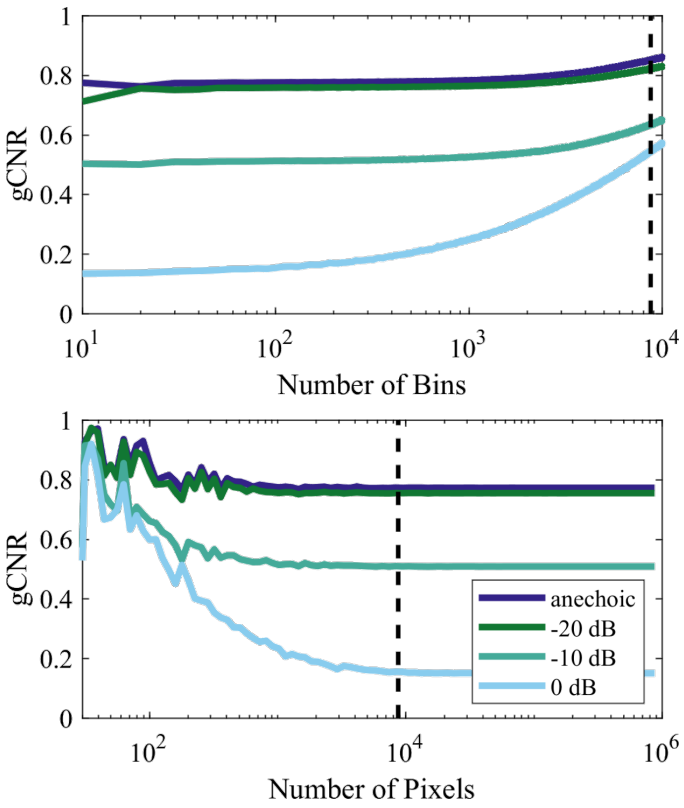

Figure 6.2: Plots of gCNR for set of simulated lesions (anechoic, -20, -10, 0dB) beamformed with DAS, where (Top) the number of bins used in the histograms was varied and (Bottom) the number of pixels was varied via interpolation. The dashed line in both plots represents the total number of pixels used between the target and reference regions. Note that if we continued to the extreme case of using 1 bin for all pixels, the gCNR would drop to 0 for all methods. Figure ©2022 IEEE.

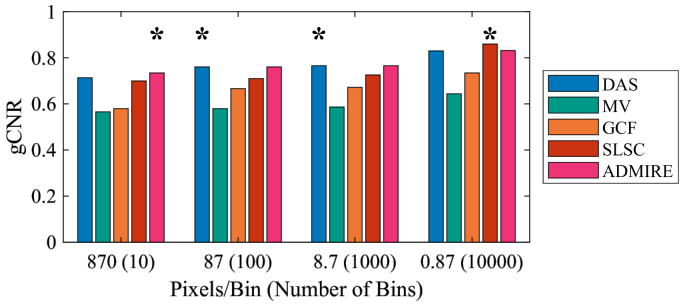

Figure 6.3: Average gCNR of -20 dB simulated lesions from the indicated beamforming methods. gCNR was calculated using different numbers of histogram bins. The asterisk for each grouping indicates the highest gCNR value. Of note, the ranking of the methods can change depending on the choice of pixels/bin. Figure ©2022 IEEE.

case, while gCNR and CNR increase and decrease together.

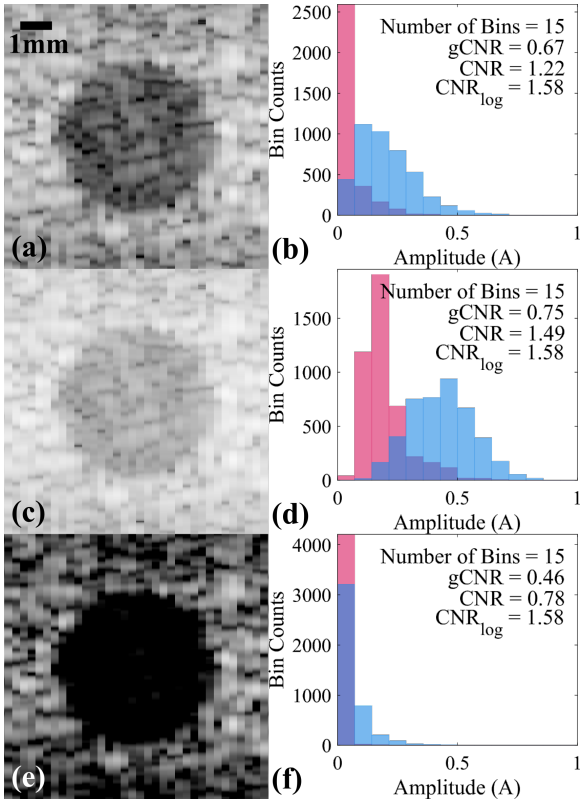

Figure 6.4: Simulated -20dB lesions beamformed with (a) DAS, (c) EPT$_{.5}$, and (e) EPT$_2$. Histograms used to calculate gCNR are included using Sturges' formula to determine the number of bins of the histogram. In these cases where the ratio of data points to bins is high, the overlap estimate can be wrong. Traditional CNR is included for each case, both calculated on linear and log-compressed data. Figure ©2022 IEEE.

6.3.2 Increasing robustness of the gCNR

To resolve these issues presented, careful consideration needs to be taken to make sure the histograms are accurately representing the data, ensuring a more robust estimation of the gCNR. We consider several possible options, detailed below.

6.3.2.1 Parameterized gCNR

Since the true goal of gCNR is to calculate the overlap between the PDFs of the lesion and reference material, a consideration can be made that estimating the distribution of the data by finding the actual PDF would be mathematically sound. Modeling different kinds of tissue in ultrasound for detection and identification purposes has been done (Mohana Shankar, 2000), and the specific model used was often related to the properties of the backscattered echo, divided between pre-Rayleigh, Rayleigh, and post-Rayleigh (Rician) (Wagner et al., 1983, 1987). Efforts to improve the modeling of ultrasound data led to the investigation of the K distribution (Shankar, 1995; Molthen et al., 1995) and later the homodyned K distribution (Dutt and Greenleaf, 1994). However, the K and homodyned K distributions are significantly more computationally complex

compared to previous distributions, making them relatively unappealing options. Instead, the Nakagami distribution (Nakagami, 1960) was suggested for use with ultrasound (Clifford et al., 1993; Gilman, 1997; Mohana Shankar, 2000) as a much simpler alternative, while still incorporating features of both Rayleigh and Rician models (Mohana Shankar, 2000).

For this work, we consider the Rayleigh, Rician, and Nakagami distributions as classic, and computationally efficient options for modeling ultrasound echoes. We additionally consider the gamma distribution, as an alternative choice that is related, but slightly different compared to Rayleigh and Nakagami distributions, and may be a better fit for certain beamformers. All four of these distributions are relatively straightforward and require only estimation of one or two parameters. Additionally, MATLAB has functions for fitting data using each of these choices, making implementation quick and efficient. Then, calculating the overlap can be done analytically or by using built in functions for the PDF and cumulative distribution function (CDF). For each distribution we merely need to calculate the intersection using the PDF and then integrate under the relative sections using the CDF.

6.3.2.2 Histogram-based gCNR

We can also consider methods to make histogram-based gCNR estimations more robust to data range manipulations. As previously stated, the choice of the number of bins has been considered for nearly a century, and as a result there are a multitude of methods available today. Though there are too many to all be considered in this work, we will include common choices for consideration.

Sturges' formula was perhaps the first method published, and suggests using $k = \lceil \log_2 n \rceil + 1$ (Sturges, 1926), for n data points. Since then, the Rice rule $k = \lceil 2\sqrt[3]{n} \rceil$ has been suggested as an alternative (Lane, 2007), and the square root rule $k = \lceil \sqrt{n} \rceil$ is commonly used in statistics software packages (Cameron, 2009). These methods, and the many others that have followed, generally are designed or are optimal for certain kinds of distributions of data, and are designed to be used with bins of a fixed width, which means we can of course potentially manipulate them all the same.

Instead, we consider variable bin widths. In general statistics, variable bin widths are often avoided due to the narrow bin widths that occur at high density locations of the data (Scott, 1992), and it is often specifically recommended not to use them without caution (Scott, 2010). However, for the purposes of estimating the overlap for gCNR, it can adapt well to unexpected distributions. If variable widths are applied, then data range manipulations simply cannot occur, as the widths of the bins will adjust as the density of the data is altered. For the implementation presented here, we calculate the number of bins desired as $k = \lceil 2n^{2/5} \rceil$ (Prins et al., 2003), and design bins widths such that each bin has an equal number of data points.

Another possible implementation we consider is to rank-order the data, and estimate the gCNR from

the list of rankings. This would entirely eliminate the possibility of manipulation via data range alteration, since the rank-order of the data would never change, and therefore the estimates of gCNR would be exactly identical. Performing statistical analysis on rank-order data is hardly a new concept, and can be used along side traditional metrics to gain more information about the behavior of the data (Conover, 2012). Additionally, rank-order data is used in many statistical tests such as those presented by Friedman (Friedman, 1937) and Wilcoxon (Wilcoxon, 1945, 1950). We implement rank-ordering by sorting the entire set of data (both lesion and reference) and separating the ranks into the original lesion and reference groups. For example, consider we have n total data points between the two regions. Then the list of rankings is $1,2,...,n$, and we separate the ranks based on which region the data point is from (e.g., $[x_1,x_2,...,x_{n/2}]$and$[x_{n/2+1},...,x_n]$). Then, we can simply estimate gCNR by using the usual histogram method on the two sets of ranks. For this, we use the square root rule to choose the number of bins, though the choice is more flexible since there will be a consistent number of data points per bin, regardless of choice.

6.3.3 Empirical gCNR

However, we can avoid both a parametric and histogram-based approach via the use of empirical cumulative distribution functions (eCDFs). We include a section at the end of this chapter with a proof showing that gCNR can be estimated from the difference between the CDFs of the target and reference regions, without the use of any additional information. Specifically, and as visually explained in Fig. 6.5, gCNR can be estimated based on the amplitudes of the local maximums and minimums of the difference function, $H = \text{eCDF}_T - \text{eCDF}_R$. We prove below in Section 6.3.3.1, that with ideal functions,

$$\text{gCNR} = \sum \mathbb{H}_{MAX} - \sum \mathbb{H}_{MIN}, \tag{6.5}$$

where $\mathbb{H}_{MAX}$ and $\mathbb{H}_{MIN}$ are the set of local maximums and minimums, respectively.

Our implementation of empirical gCNR uses MATLAB functions for ease of use. First, the ecdf function (©1993-2021 The MathWorks, Inc.) calculates the empirical CDFs for both the target and reference regions. We then use some basic interpolation so that the x-axis data points are consistent across the two eCDFs. Then the H function is simply calculated from the difference of the two eCDFs, and the local maximums and minimums are calculated using the islocalmax and islocalmin functions (©2017-2021 The MathWorks, Inc.). This then gives us the necessary information to compute the estimated gCNR as in the above equation.

A note on implementation, the function for calculating the eCDF is based on the non-parametric Kaplan-Meier estimator (Kaplan and Meier, 1958), which is used to calculate the survival function from a set of data, which in turn is the complement of the CDF. Additionally, upper and lower confidence bounds can

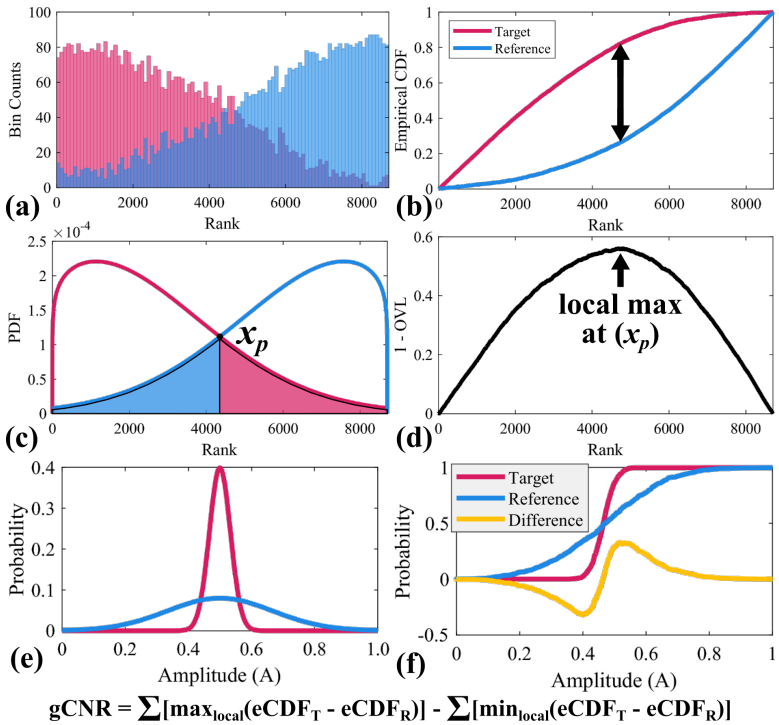

Figure 6.5: Visual example of computing the empirical gCNR from empirical cumulative distribution functions (eCDFs). (a) Histogram representation of the target and reference PDFs. (b) eCDF of the two regions. (c) Theoretical PDF of the two regions (not accurate, only for demonstration). (d) Difference function $\text{eCDF}_T - \text{eCDF}_R$. Local maximums and minimums occur at intersections of the PDFs, so the integration of the regions of the PDFs is related to the overlap, and therefore the gCNR as given in the equation. (e-f) Show a separate example of two normal (Gaussian) distributions PDFs and the related eCDFs. We can see in this example that there are multiple intersections, which results in both local maximums and minimums. Figure ©2022 IEEE.

be calculated using Greenwood's formula, which is all considered in the MATLAB implementation (default is 95% confidence, which we use here) based on prior work (Cox and Oakes, 1984; Lawless, 2002). With our data, the main challenge is that the noise in the data can create many spurious local maximums and minimums, which will slightly offset our gCNR estimate from the expected value. To minimize this error, we used a minimum prominence of 0.01 with the local peak functions, though a similar result could likely be achieved either with smoothing or curve-fitting. A potential advantage of using this method, is that the upper and lower confidence bounds give us a better ability to discern whether the estimated gCNR of two images or beamformers is sufficiently different to make a conclusion about ranking of the methods.

6.3.3.1 Proof of Empirical Generalized Contrast-to-Noise Ratio

Generalized contrast-to-noise ratio is defined as $1 - \text{OVL}$, where OVL is the overlap of two probability distribution functions (PDFs). We will demonstrate here that the use of the cumulative distribution function (CDF) allows us to accurately estimate the overlap between two unknown PDFs.

Assumptions

Let's consider idealized CDFs, where they are continuous and differentiable everywhere. By definition, a CDF F is related to the PDF f by $f(x) = dF(x)/dx$, and additionally F is non-decreasing. For simplicity, we consider CDFs that are scaled from 0 to 1 on both axes, such that $F(0) = 0$ and $F(1) = 1$.

Lemma 1. *The set of points $\mathbb{Y} = \{x_{p_1}, ..., x_{p_k}\}$ that correspond to the local maximums and minimums of the difference between two CDFs is the set of intersection points between the two corresponding PDFs, and there exists at least one such point for any two distinct CDFs.*

Proof. Consider two known CDFs, F and G defined as in our assumptions, with corresponding unknown PDFs, f and g. Let $\mathbb{X} = \{x_1, ..., x_n\}$ be the complete set of points where the CDFs are equal, i.e., for each $x_i \in \mathbb{X}$, $F(x_i) = G(x_i)$. By definition, this set contains $n \geq 2$ points, as $F(0) = G(0) = 0$ and $F(1) = G(1) = 1$ by our assumptions.

Consider any interval (x_i, x_{i+1}) for $x_i \in \mathbb{X}$, then since $\mathbb{X}$ is the complete set, F and G are never equal on this interval. Let's choose $F(x) > G(x)$ for all $x_i < x < x_{i+1}$. Then, for some small $\varepsilon > 0$,

$$\lim_{\varepsilon \to 0} F(x_i + \varepsilon) > \lim_{\varepsilon \to 0} G(x_i + \varepsilon) \quad \text{and} \quad \lim_{\varepsilon \to 0} F(x_{i+1} - \varepsilon) > \lim_{\varepsilon \to 0} G(x_{i+1} - \varepsilon). \tag{6.6}$$

Since $F(x_i) = G(x_i)$ and $F(x_{i+1}) = G(x_{i+1})$, then since the functions are continuous,

$$\frac{d}{dx}F|_{x=x_i} > \frac{d}{dx}G|_{x=x_i} \quad \text{and} \quad \frac{d}{dx}F|_{x=x_{i+1}} < \frac{d}{dx}G|_{x=x_{i+1}}. \tag{6.7}$$

As we have defined the CDFs to be continuous, this implies that on the interval of (x_i, x_{i+1}) there exists at least one point x_p such that

$$\frac{d}{dx}F|_{x=x_p} = \frac{d}{dx}G|_{x=x_p}. \tag{6.8}$$

Let's define $H(x) = F(x) - G(x)$, the difference function of the two CDFs. Then, $\frac{d}{dx}H$ is the slope of the difference function, and we define the set of points $\mathbb{Y} = \{x_{p_1}, ..., x_{p_k}\}$ as the set of local maximums and minimums of H on the interval (x_i, x_{i+1}), i.e. where the derivative of H changes sign. Then for each $x_p \in \mathbb{Y}$, $\frac{d}{dx}H|_{x=x_p} = 0$, and by extension

$$\frac{d}{dx}F|_{x=x_p} = \frac{d}{dx}G|_{x=x_p}. \tag{6.9}$$

From Equations 6.7 and 6.8, we know that $\mathbb{Y}$ contains at least one point where the slope of H changes sign, which means $k \geq 1$. We can then relate this understanding to the PDF as

$$f(x = x_p) = g(x = x_p), \tag{6.10}$$

since by definition the derivative of the CDF is the PDF. Then the set of points $\mathbb{Y} = \{x_{p_1}, ..., x_{p_k}\}$ is also the set of intersection points of f and g on the interval (x_i, x_{i+1}). □

Corollary 2. *If there are m such local maximums on an interval as defined in Lemma 1, then there are $m-1$ local minimums, which means there is an odd number k of total local maximums and minimums. In particular, for $j \in [1, ..., k]$ when j is odd then x_{p_j} is a local maximum, and when j is even then x_{p_j} is a local minimum.*

Definition 1. *The overlap (OVL) between two PDFs f and g is the integral of the minimum of the two functions. For the interval (x_i, x_{i+1}), we have $x_p \in \mathbb{X} = \{x_{p_1}, ..., x_{p_k}\}$, the set of intersections between f and g. Then OVL on this interval is defined as*

$$OVL = \int_{x_i}^{x_{p_1}} min(f,g) + \int_{x_{p_1}}^{x_{p_2}} min(f,g) + ... + \int_{x_{p_k}}^{x_{i+1}} min(f,g). \tag{6.11}$$

Lemma 3. *The overlap, OVL, between two unknown PDFs f and g can be calculated from the difference of the corresponding known CDFs F and G, $H = F - G$, as*

$$OVL = 1 + \sum \mathbb{H}_{MIN} - \sum \mathbb{H}_{MAX}, \tag{6.12}$$

where $\mathbb{H}_{MIN}$ and $\mathbb{H}_{MAX}$ are the set of local minimums and maximums of the H function, respectively.

Proof. Consider the same interval (x_i, x_{i+1}) from Lemma 1. We previously chose $F(x) > G(x)$ for all $x_i < x < x_{i+1}$, and found the set of points $\mathbb{Y} = \{x_{p_1}, ..., x_{p_k}\}$ as the intersections of f and g. Then from Equation 6.7 we know that $f(x) > g(x)$ for all $x_i < x < x_{p_1}$, meaning $\min(f,g) = g$ on this interval. Furthermore, since $\mathbb{Y}$ is the set of intersection points, we know that the minimum of f and g swaps at every intersection. Since there are an odd number k of intersections from Corollary 2, Equation 6.11 becomes

$$\text{OVL} = \int_{x_i}^{x_{p_1}} g + \int_{x_{p_1}}^{x_{p_2}} f + \int_{x_{p_2}}^{x_{p_3}} g + ... + \int_{x_{p_k}}^{x_{i+1}} f. \tag{6.13}$$

By definition of the CDF, we can expand this equation into

$$\text{OVL} = [G(x_{p_1}) - G(x_i)] + [F(x_{p_2}) - F(x_{p_1})] + [G(x_{p_3}) - G(x_{p_2})] + ... + [F(x_{i+1}) - F(x_{p_k})], \tag{6.14}$$

and then rearrange the equation as

$$\text{OVL} = [F(x_{i+1}) - G(x_i)] + [G(x_{p_1}) - F(x_{p_1})] + [F(x_{p_2}) - G(x_{p_2})] + ... + [G(x_{p_k}) - F(x_{p_k})], \tag{6.15}$$

which groups the differences by the locations of the local maximums and minimums from Lemma 1. Then for all $j \in [1,...,k]$, when j is odd we have terms in the form $G(x_{p_j}) - F(x_{p_j})$ and when j is even we have terms in the form $F(x_{p_j}) - G(x_{p_j})$. From Corollary 2, the odd values of j are maximums and the even values are minimums, which means if we rearrange each term to be in the form $F - G$, the local maximums are being subtracted while the minimums are added, giving us

$$\text{OVL} = [F(x_{i+1}) - G(x_i)] + [\sum_{x_i}^{x_{i+1}} \mathbb{H}_{MIN}] - [\sum_{x_i}^{x_{i+1}} \mathbb{H}_{MAX}]. \tag{6.16}$$

Now we can consider the entire range of the CDFs, [0,1]. As from Lemma 1, we divide this interval into $[x_1,x_2],[x_2,x_3],...,[x_{n-1},x_n]$, where $x_1 = 0$ and $x_n = 1$. It follows that the overlap of the entire range [0,1] is the overlap of each interval summed together. That is, $\text{OVL}([0,1]) = \text{OVL}([x_1,x_2]) + \text{OVL}([x_2,x_3]) + ... + \text{OVL}([x_{n-1},x_n])$. Then we can use Equation 6.16 to get

$$\begin{aligned}
\text{OVL}([0,1]) = {} & [F(x_2) - G(x_1)] + [\sum_{x_1}^{x_2} \mathbb{H}_{MIN}] - [\sum_{x_1}^{x_2} \mathbb{H}_{MAX}] \\
& + [F(x_3) - G(x_2)] + [\sum_{x_2}^{x_3} \mathbb{H}_{MIN}] - [\sum_{x_2}^{x_3} \mathbb{H}_{MAX}] \\
& + ... \\
& + [F(x_n) - G(x_{n-1})] + [\sum_{x_{n-1}}^{x_n} \mathbb{H}_{MIN}] - [\sum_{x_{n-1}}^{x_n} \mathbb{H}_{MAX}].
\end{aligned} \tag{6.17}$$

Note that the ordering of F and G here is does not matter, as $[x_1,...,x_n]$ were defined as the points where $F = G$ and the signs of the maximums and minimums will compensate, meaning that this equation further simplifies by canceling many of the intermediate terms

$$\begin{aligned}
\text{OVL}([0,1]) = {} & [-G(x_1)] + [\sum_{x_1}^{x_2} \mathbb{H}_{MIN}] - [\sum_{x_1}^{x_2} \mathbb{H}_{MAX}] \\
& + [\sum_{x_2}^{x_3} \mathbb{H}_{MIN}] - [\sum_{x_2}^{x_3} \mathbb{H}_{MAX}] \\
& + ... \\
& + [F(x_n)] + [\sum_{x_{n-1}}^{x_n} \mathbb{H}_{MIN}] - [\sum_{x_{n-1}}^{x_n} \mathbb{H}_{MAX}].
\end{aligned} \tag{6.18}$$

and, as defined, $G(x_1) = 0$ and $F(x_n) = 1$, which gives us

$$\text{OVL}([0,1]) = 1 + [\sum_{x=0}^{1} \mathbb{H}_{MIN}] - [\sum_{x=0}^{1} \mathbb{H}_{MAX}] \tag{6.19}$$

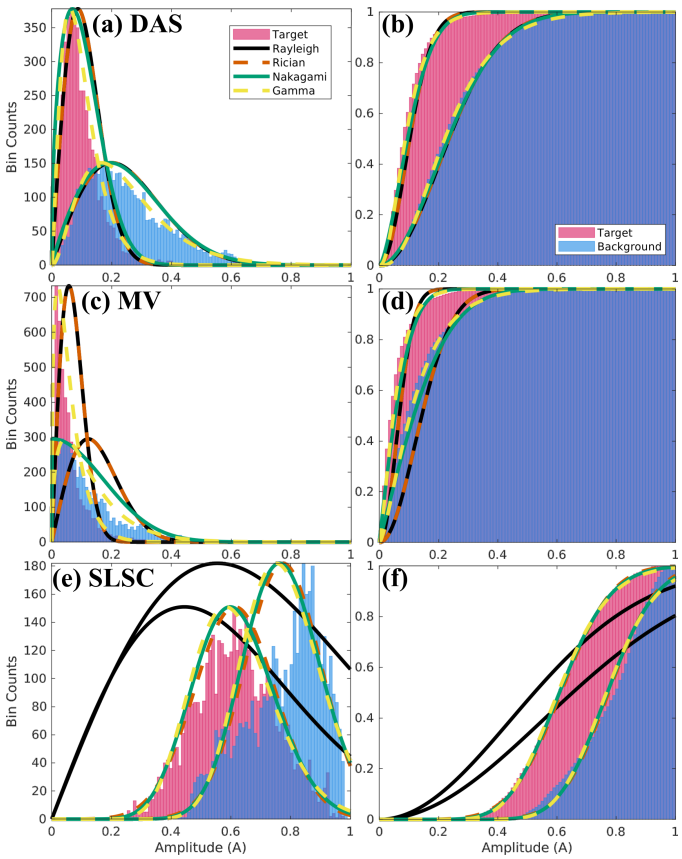

Figure 6.6: Simulated -10dB lesions beamformed for (a-b) DAS, (c-d) MV, and (e-f) SLSC. The target and background data is shown as both a probability distribution function (PDF) and cumulative distribution function (CDF), with the indicated distribution fits included. The Nakagami distribution for the target region of MV failed to fit, and therefore there is no curve shown. Figure ©2022 IEEE.

□

Theorem 4. *The gCNR metric can be estimated from the known CDFs F and G as*

$$gCNR = \sum \mathbb{H}_{MAX} - \sum \mathbb{H}_{MIN}, \tag{6.20}$$

where $H = F - G$ and $\mathbb{H}_{MAX}$ and $\mathbb{H}_{MIN}$ and the local maximums and minimums of the function, respectively.

Proof. Since gCNR $= 1 -$ OVL, it follows clearly from Lemma 3. □

6.4 Results

Some examples of modeling the lesion data for a -10dB lesion are included in Fig. 6.6. For the DAS data, all models seem relatively similar, matching the shape of the data in the histogram. However, for MV and SLSC some distributions are poor fits or outright fail to fit. The gamma distribution here is a slightly better fit for

Table 6.1: Average Estimated gCNR for all beamformers and implementations. The dashed line separates histogram-based and parametric-based implementations. Table ©2022 IEEE.

	Default (k=100)	Square Root (k=94)	Sturges' (k=15)	Rice Rule (k=42)	Variable (k=76)	Rank-order (k=94)	Empirical (eCDF)	Rayleigh	Rician	Nakagami	Gamma
	Anechoic Lesion										
DAS	0.78	0.78	0.75	0.78	0.78	0.78	0.78	0.72	0.73	0.67	0.76
GLT	0.77	0.77	0.77	0.77	0.77	0.77	0.77	**0.67**	0.79	0.77	0.80
EPT .5	0.78	0.78	0.78	0.78	0.78	0.78	0.78	**0.58**	0.78	0.76	0.80
EPT 2	0.77	0.77	**0.51**	0.71	0.78	0.78	0.78	0.80	0.80	**0.56**	**0.34**
EPT 4	**0.31**	**0.30**	**0.07**	**0.18**	0.78	0.78	0.78	**0.28**	**0.34**	**0.00**	**0.00**
	-10 dB Lesion										
DAS	0.51	0.51	0.49	0.51	0.51	0.51	0.51	0.51	0.51	0.49	0.49
GLT	0.52	0.52	0.52	0.52	0.52	0.52	0.52	**0.38**	0.51	**0.44**	**0.43**
EPT .5	0.51	0.51	0.51	0.51	0.51	0.51	0.51	**0.29**	0.50	0.49	0.48
EPT 2	0.51	0.51	0.48	0.48	0.51	0.51	0.51	**0.74**	**0.73**	**0.43**	0.49
EPT 4	**0.30**	**0.29**	**0.07**	**0.18**	0.51	0.51	0.51	**0.27**	**0.40**	**0.00**	**0.00**
	0 dB Lesion										
DAS	0.15	0.15	0.13	0.14	0.16	0.16	0.13	0.11	0.11	0.12	0.12
GLT	0.15	0.15	0.13	0.14	0.15	0.16	0.12	**0.06**	0.11	0.11	0.11
EPT .5	0.16	0.15	0.14	0.15	0.16	0.16	0.13	**0.06**	0.12	0.12	0.12
EPT 2	0.15	0.15	0.12	0.14	0.16	0.16	0.13	0.18	0.18	0.11	0.12
EPT 4	0.13	0.13	**0.07**	0.11	0.16	0.16	0.13	**0.30**	**0.30**	**0.06**	**0.07**

MV, while Nakagami introduces an infinite value that makes us unable to plot the PDF of the target lesion data. In comparison, the Rayleigh distribution does not fit SLSC at all, while the other three methods seem decent, but are clearly still suboptimal.

Fig. 6.7 shows a comparison of estimated gCNR for the normal histogram implementation versus variable bin widths and rank-ordering as the ratio of pixels to bins changes, as well as the average empirical gCNR estimates with the included upper and lower bounds. Fig. 6.8 shows visual examples of several of the interesting beamformer cases for an anechoic lesion, and includes the histograms for the default, variable bin widths, rank-order, and empirical gCNR methods we suggested. Table 6.1 includes the average estimated gCNR values for DAS and the transformation-based beamformers using all suggested implementations for the anechoic, -10, and 0 dB lesions. Histogram-based methods are presented to the left of the dashed line, and parametric methods are to the right, with empirical gCNR in the middle. In this table, bold face is used to indicate clearly deviant values, as all beamformers should produce the same gCNR.

6.5 Discussion

To start, we should make it clear that in the majority of reasonable cases with "normal" beamformers, the default implementation of gCNR using $k = 100$ bins and sufficiently large ROIs works perfectly well. Assuming there is enough data to populate the standard number of bins, Table 6.1 makes it clear that gCNR is not "failing" under normal circumstances. However, we have demonstrated that there exists a possibility for methods to, willfully or not, manipulate gCNR estimates by doing something as innocent as using a target

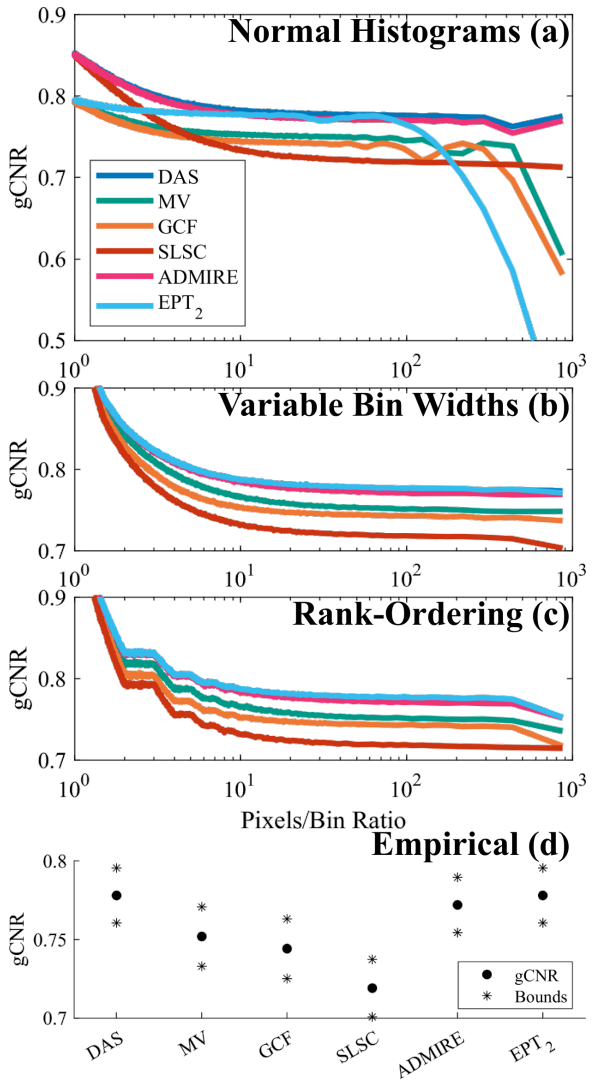

Figure 6.7: gCNR for anechoic lesions with various beamformers estimated with (a) traditional uniform bin widths, (b) variable bin widths, and (c) rank-order uniform bin widths. For these cases, the number of pixels used is constant and the number of bins is varied, and we report the ratio of pixels per bin. (d) Plot of the average gCNR and the upper and lower 95% confidence bounds using the empirical gCNR method. Figure ©2022 IEEE.

region of interest that is too small, or using some unreasonable number of bins. As a result, the discussion of how to increase the robustness of the implementation of gCNR is warranted.

There is a certain attractiveness to being able to model the ultrasound scatterers using parametric methods, and it is used in several applications to characterize images. However, the issue lies in determining what models are appropriate fits, and this issue is compounded by the common presence of adaptive beamformers. While DAS (raw b-mode data) has a well-understood distribution which makes fitting the data often reliable, the same is clearly not true for the other adaptive methods, where the results may vary wildly or fail to be fit

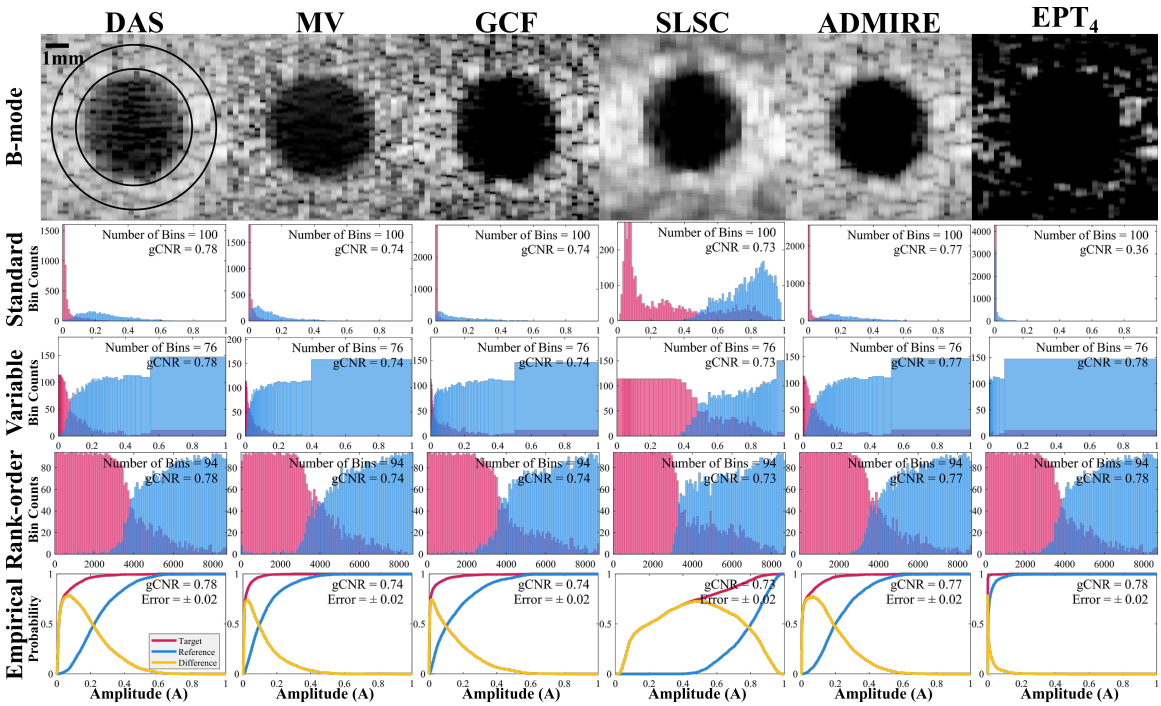

Figure 6.8: B-mode images for a selection of beamforming methods for an example simulated anechoic lesion. The circled regions on the DAS image indicate the lesion and background regions for gCNR calculation. Log-compressed methods are displayed with a 50 dB dynamic range. Histograms used for estimating gCNR using the default (k=100), variable bin widths (k=76), and rank-order (k=94) methods are included for each case. We additionally include the empirical gCNR plots of the eCDFs for these cases, with the error representing the average distance to the upper and lower bounds of the empirical gCNR estimate. Figure ©2022 IEEE.

at all in some cases.

For ease of use, a histogram-based implementation may still be best as it is straightforward and requires less consideration compared to parametric methods. In this case, it seems reasonable that a best-practices choice would be, at a minimum, to have some specific method for choosing the number of bins, and declaring that method and reporting the number of bins used, since this gives the reader some sense of the ratio of the number of data points versus the number of bins used in the analysis. This is important, since as we and others have shown, the number of bins, and by extension, the size of the data set, can alter not just the gCNR estimates for each method, but potentially the rank of the methods being compared.

Perhaps the most reliable histogram-based method is to use variable bin widths or rank-ordering, as it removes much of the uncertainty around choosing the number of bins. Fig. 6.8 shows examples for each of these methods, demonstrating how the variable widths change dramatically depending on the beamformer in question, while the rank-order histograms are identical for methods that are true transforms of each other. With histograms, these two implementations were the only two tested that were able to correctly match

the results for DAS, GLT, and the EPT methods to within an error of 0.01 in all circumstances. Both of these implementations would seem to make gCNR estimations via histograms resistant to dynamic range alterations. It also does not prevent or restrict the use of gCNR, or increase the complexity of the interpretation of the metric, which falls in line with the goals laid out by the authors in the original work (Rodriguez-Molares et al., 2020).

However, as we have shown, we can produce estimates of gCNR without the use of parametric methods or histograms using the eCDF, which we simply call empirical gCNR. In the same Fig. 6.8, the empirical probability plots show that not only is the estimate relatively intuitive to calculate, but the upper and lower bounds produce a result with a small overall error and it works consistently with all the beamformers to produce nearly identical results to the robust histogram-based implementations. In relatively simple cases, like the simulated lesions here, the gCNR estimate is trivial to calculate from these plots, though even in more complex cases the full formulation we presented in Equation 6.5 is sufficient. Additionally, we could even apply this empirical method on the rank-ordered lists we described to produce the rank-ordered version of this metric.

A final note, there is still a question of how different two measurements should be to consider the performance of the two methods to be different. With histograms, it is reasonable to expect that there is some inherent variability in the estimation (as with any estimator), such that, for example, a gCNR of 0.77 is not sufficiently different than 0.78 to be considered worse. Empirical gCNR does produce upper and lower bounds, which is potentially valuable exactly for this purpose, and in Fig. 6.7, we can assert that DAS and ADMIRE are of equivalent performance, while MV and GCF are similar, but lower compared to the other two. This may be a valuable area to consider for future work.

6.6 Conclusion

gCNR is a powerful tool, resistant to transformations and dynamic range alterations when properly implemented. While gCNR implementations using histograms with 100 bins work properly in many circumstances, we have shown that manipulation can occur by altering the ratio of the pixels per bin or with some extreme transformation-based beamformers. Though a parametric approach is attractive, the inability to model an arbitrary adaptive beamformer makes it all but impossible without significantly more complex models, making a histogram-based or non-parametric solution much more appealing. In testing, making the histograms more robust is straightforward, and at a minimum the ratio of pixels per bin and the method to determine the number of bins used should be reported. We can further improve the robustness of the histograms by either using variable bin widths (where the number of data points per bin is approximately equal), or by rank-ordering the data and estimating gCNR from the rank-ordered data. Alternatively, we can use empirical CDFs to produce

an estimate that relies on neither parametric methods or histograms, while still being robust in all the cases we considered here. All three of these methods are much more resistant to transformations, and empirical gCNR in particular requires no special parameter consideration, making it arguably the most consistent of the three methods.

CHAPTER 7

Considerations for the Application of Histogram Matching

This chapter is adapted from the supplemental material originally published in (Schlunk and Byram, 2022b), ©2022 IEEE. Reprinted, with permission, from Schlunk, S., & Byram, B. (2022). Combining ADMIRE and MV to Improve Image Quality. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 69(9), 2651–2662. https://doi.org/10.1109/TUFFC.2022.3194548. Additional content is adapted from an International Ultrasonics Symposium presentation (Schlunk and Byram, 2022a), ©2022 IEEE. Reprinted, with premission, from Schlunk, S., & Byram, B. (2022). Breaking and Fixing gCNR and Histogram Matching. 2022 IEEE International Ultrasonics Symposium (IUS), 1–4.

We demonstrate the ability of iterative ADMIRE to improve the contrast accuracy in images in Chapter 3, and the combination of ADMIRE and MV to improve resolution in Chapter 4. However, we have yet to use histogram matching to demonstrate that these methods are providing "true" improvements over DAS, despite the value of the method we laid out in Section 1.2.4. In this chapter, we will take a brief look at some of the results from the chapters mentioned, and demonstrate why we have been cautious to implement and interpret histogram matching thus far.

7.1 Introduction

Histogram matching is a method that is designed to provide a "level playing field" for imaging metrics by calculating the transform between a target image and a reference image (Bottenus et al., 2020a). This is intended to make qualitative comparisons more fair and to prevent transformation-dependant improvements to specific metrics, similar to how gCNR is designed to produce a CNR-like metric that is transform independent. For example, consider Fig. 7.1, where an *in vivo* kidney image is processed with some adaptive beamformer designed to boost contrast ratio. A mapping, h, is calculated between the two images and applied to DAS, showing that the improvements, if any, in the adaptive image are less obvious.

Histogram matching can be applied using a full matching method, where the cumulative distribution functions of the two images are compared, or with partial matching, where a scaling and shifting transformation is calculated based on the histograms of the distributions of the two images. The authors suggest that full matching using an ROI of homogenous speckle is the optimal choice when the speckle fully captures the dynamic range of the image, however, in high dynamic range cases where there are strong hypoechoic or hyperechoic targets present, partial matching is needed to reduce errors due to extrapolation.

To demonstrate the work from the previous chapters, we choose to match to iADMIRE or ADMIRE since

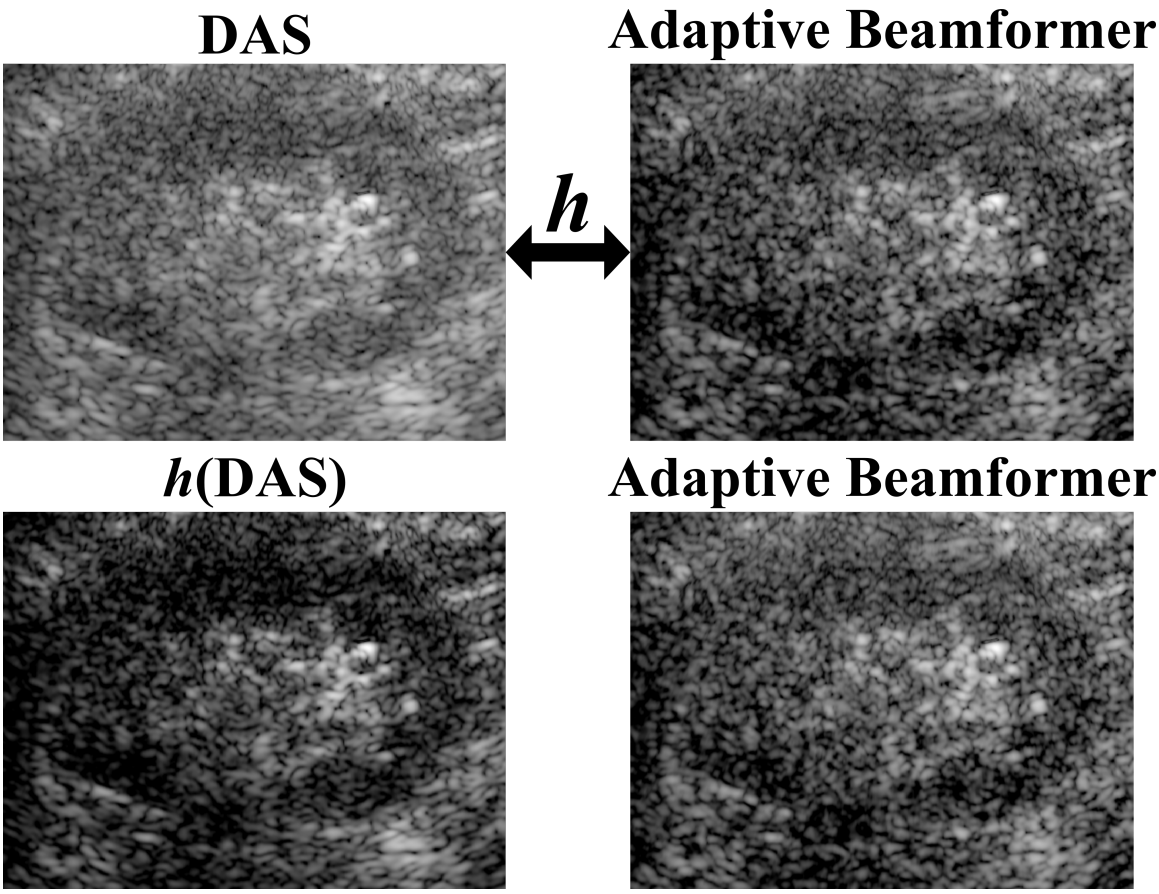

Figure 7.1: Example *in vivo* kidney image displayed with delay-and-sum (DAS) and an adaptive beamformer designed to improve contrast ratio. Histogram matching using a full matching method is used to calculate a mapping h between the two images. Applying the mapping to the DAS image reveals less obvious improvements in the adaptive image. Figure ©2022 IEEE.

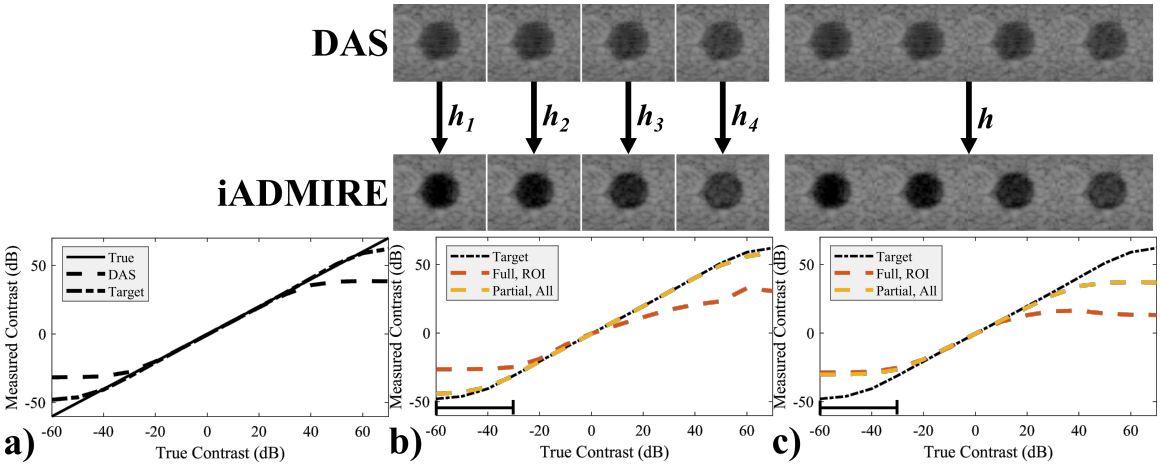

Figure 7.2: Comparison between DAS and iADMIRE (Target) using contrast ratio dynamic range (CRDR) to compare contrast accuracy of the two beamformers. The anechoic, -50, -40, and -30 dB cysts are shown for both beamformers. a) shows the original results from Chapter 3, showing iADMIRE can extend the range of accurate contrast past what DAS is capable of. b) shows the result when each lesion is individually matched to iADMIRE. c) shows the result when all lesions are simultaneously matched to iADMIRE. The results from b) show that when matching any individual cyst, DAS can be stretched to appear identical to iADMIRE, which incorrectly suggests that DAS is on par with iADMIRE's performance. However, when the full contrast curve is considered, here by matching all the cysts simultaneously in the same figure, it shows that DAS cannot differentiate some of the cases, since histogram matching would try to match the anechoic, -50, -40, and -30 dB iADMIRE cases all to the same value as in the DAS figures. Figure ©2022 IEEE.

these methods generally have strong contrast ratio and high dynamic range. Unfortunately, since our work is mostly centered on high dynamic range situations (both anechoic cysts and kidney stones generally fall into this category), we use partial matching. For the cyst simulations, the entire simulation image area is used for matching. For the kidney cases, the matching region was chosen as the entirety of the kidney.

7.2 Results

7.2.1 iADMIRE (Chapter 3)

Fig. 7.2 shows the original contrast ratio dynamic range (CRDR) results for DAS and iADMIRE from Chapter 3, as well as cases where DAS is matched to iADMIRE and the CRDR is calculated on the matched images. Matching was performed both with full matching and a homogeneous ROI, as well as with partial matching using the full image space for demonstration. When matching the individual images, as in Fig. 7.2b, DAS appears to attain similar performance to iADMIRE. However, these individual mappings are misleading, as when the full array of lesions are matched together, as in Fig. 7.2c, DAS gains no noticeable improvement, and iADMIRE has the clearly superior CRDR.

7.2.2 ADMIRE+MV (Chapter 4)

Tables 7.1 and 7.2 show the anechoic cyst image metrics and *in vivo* metrics, respectively, for both the original data and when histogram matching is used to match each of the methods to ADMIRE. Figures 7.3 and 7.4 show a representative example case for the anechoic cysts and *in vivo* cases, respectively.

In the simulations for cyst edge width, the reverberation clutter case of 0dB SCR is unchanged in terms of ranking of the methods before and after histogram matching, and the no interference case is mostly unchanged. However, the ranking of the methods in the presence of thermal noise (0dB SNR) is significantly different. The general observation that AD+MVNB is better than ADMIRE remains true, but MVNB (L=0.75M) is better relative to the other methods after matching. There are similar rearrangements in method ordering for contrast ratio, though gCNR is largely unchanged, which makes sense since it should be transform independent. *In vivo*, while contrast ratio and even gCNR see some significant changes, the kidney stone sizes measured are the same (with minor variances), which does not impact our perception of the performance of the beamformers.

Table 7.3 shows all combinations for histogram matching targets for average cyst edge width for the anechoic cyst simulations. We present this data to demonstrate a potential concern with histogram matching, since it is quickly apparent that the "best" method can change depending on the matching. For example, in the reverberation clutter case AD+MVNB (L=0.5M) is the best method in most of the cases, but for thermal noise AD+MVNB (L=0.75M) is the best without matching, and DAS and MVNB (L=0.75M) are the best depending on the matching used.

7.3 Discussion

Histogram matching has the potential to be a powerful tool for equalizing methods to ensure certain metrics are fair. However, the main issues at present are how to "fairly" match methods as well as how to interpret the results in the context of the original results. We presented two different situations where we would like to use histogram matching, but challenges in the implementation and interpretation of the results makes it difficult.

In the iADMIRE work from Chapter 3, in fairness histogram matching had not been published at the time of the work, but we can see that the initial instinct to apply histogram matching to each image and remake the CRDR metric would be misleading. DAS, in these simulated cysts, is incapable of distinguishing between the anechoic, -50, and -40 dB cysts. However, if we match each image separately, the mapping can easily map that same amplitude to whichever target contrast it needs to, implying that DAS somehow could mimic the performance of iADMIRE. Mapping all the cysts simultaneously shows that DAS indeed cannot differentiate these cysts, so the mapping simply has to choose some single amplitude target, resulting in the same CRDR curve as before matching. This demonstrates that iADMIRE is actually producing real

contrast improvements compared to DAS, but more importantly that applying histogram matching requires more thought and consideration than other metrics might.

In the work with ADMIRE+MV from Chapter 4 specifically, our primary metric was focused on resolution and size, which is often related to the point spread function (PSF) and is significantly impacted by the dynamic range of the image and the amplitude at which the measurement is made. Since histogram matching specifically manipulates an image's amplitude to "match" methods together, there may be unintended and undesired impacts on these metrics that rely on the PSF.

The issue of the matching target used may not matter when only a couple of methods are being compared, but in a work such as this one, choosing a single reference for all methods may be an issue. Table 7.3 shows that, at least for certain metrics, there may be significant variability in how the results can be perceived depending on the matching reference. This issue may be exaggerated somewhat in this work since resolution and sizing measurements can be influenced heavily by amplitude, but it is nonetheless the case that it makes it almost impossible to draw a conclusion about the data.

Interpreting the results is also potentially challenging, even assuming one is confident that the process up to this point is infallible. Consider the thermal noise case, where AD+MVNB (L=0.75M) produces a width of 0.105mm, compared to DAS which produces a width of 0.171mm. When both are matched to ADMIRE, they produce the same width of 0.130mm. This would presumably imply that the two methods are equivalent in this particular case, but the absolute measurement for AD+MVNB (L=0.75M) is better when no matching is applied. And indeed, when matched to AD+MVNB (L=0.75M), DAS does not perform equivalently.

This suggests to us that, at least in the case of resolution, histogram matching will mostly provide a sense of the ranking of methods, rather than give an absolute metric of performance. Though, this is more a fault of using the width of the PSF as a metric of resolution or size, which heavily depends on how the contrast may be scaled or shifted. Otherwise, there is clearly something more complex as to how to choose matching targets optimally, but that is simply out of the scope of this work for the time being.

7.4 Conclusion

In these two examples, contrast plays an important role in measuring the success of the results of our beamformers. Therefore, it is no surprise that histogram matching, which manipulates that contrast, can have unintended or unexpected side effects. Resolution in all forms is some measure of size at some amplitude, but since histogram matching targets the amplitude to calculate the transformation, this has a significant impact on perceived resolution. We think there is significant merit to the use of histogram matching, but we equally believe that there is much left to figure out with the method before it can be used generally. We hope that future works will further investigate the issues we have presented here, to make the method more robust.

Table 7.1: Anechoic Cyst Simulation Metrics - with Histogram Matching to ADMIRE. Table ©2022 IEEE.

	Cyst Edge Width (mm)					
	Original Data			Matched to ADMIRE Data		
Added Interference	None	0dB SCR	0dB SNR	None	0dB SCR	0dB SNR
DAS	0.189±0.025	0.216±0.154	0.171±0.015	0.148±0.023	0.181±0.150	0.130±0.027
DAS-Hamm	0.218±0.028	0.347±0.274	0.202±0.060	0.188±0.027	0.213±0.142	0.150±0.025
MVNB L=0.5M	0.305±0.230	0.641±0.623	0.303±0.291	0.392±0.406	0.459±0.402	0.235±0.268
MVNB L=0.75M	**0.101±0.037**	0.292±0.362	0.155±0.134	**0.087±0.033**	0.282±0.362	**0.116±0.049**
MVBB L=0.5M	0.204±0.029	0.437±0.395	0.259±0.282	0.176±0.028	0.308±0.306	0.145±0.049
ADMIRE	0.162±0.022	0.218±0.179	0.138±0.022	0.162±0.022	0.218±0.179	0.138±0.022
AD+MVNB L=0.5M	0.132±0.028	**0.140±0.054**	0.121±0.033	0.129±0.024	**0.146±0.050**	0.125±0.035
AD+MVNB L=0.75M	**0.098±0.026**	0.183±0.174	**0.105±0.021**	0.105±0.031	0.197±0.157	0.130±0.026
AD+MVBB L=0.5M	0.178±0.017	0.184±0.130	0.154±0.016	0.191±0.017	0.183±0.130	0.151±0.018
	Contrast Ratio (dB)					
	Original Data			Matched to ADMIRE		
Added Interference	None	0dB SCR	0dB SNR	None	0dB SCR	0dB SNR
DAS	-31.9±1.1	-12.1±0.7	-14.3±0.7	-43.7±1.3	-18.0±2.5	-26.1±2.8
DAS-Hamm	-38.8±0.8	-11.4±1.1	-14.0±0.7	**-46.3±1.4**	-17.4±3.0	-25.7±2.7
MVNB L=0.5M	-30.3±1.0	-7.97±1.1	-10.2±0.8	-39.3±1.5	-11.1±2.4	-19.3±2.4
MVNB L=0.75M	-24.4±1.7	-4.83±1.4	-5.58±0.8	-30.3±2.1	-5.67±1.7	-8.57±1.4
MVBB L=0.5M	-36.6±0.7	-10.0±0.9	-12.0±0.7	-44.7±1.3	-14.8±2.5	-23.2±2.5
ADMIRE	-43.1±1.3	-21.5±1.5	-29.9±1.1	-43.1±1.3	**-21.5±1.5**	**-29.9±1.1**
AD+MVNB L=0.5M	-42.1±1.4	-21.0±1.7	-30.4±1.4	-43.3±1.7	-18.8±2.2	-27.7±2.0
AD+MVNB L=0.75M	**-46.1±1.5**	**-24.7±2.6**	**-33.5±2.4**	-40.0±1.6	-17.9±3.2	-25.0±3.0
AD+MVBB L=0.5M	**-46.7±1.3**	-20.5±1.7	-29.5±1.2	-44.3±1.3	**-21.0±1.7**	**-29.8±1.3**
	generalized Contrast-to-Noise Ratio (gCNR)					
	Original Data			Matched to ADMIRE		
Added Interference	None	0dB SCR	0dB SNR	None	0dB SCR	0dB SNR
DAS	0.992±0.003	0.774±0.032	0.932±0.015	0.988±0.005	0.772±0.032	0.916±0.019
DAS-Hamm	**0.999±0.002**	0.745±0.061	0.923±0.016	**0.998±0.003**	0.742±0.059	0.905±0.025
MVNB L=0.5M	0.969±0.014	0.489±0.055	0.725±0.026	0.968±0.014	0.487±0.054	0.702±0.049
MVNB L=0.75M	0.819±0.028	0.232±0.045	0.390±0.021	0.622±0.077	0.218±0.047	0.367±0.057
MVBB L=0.5M	0.994±0.002	0.649±0.055	0.854±0.020	0.994±0.002	0.642±0.049	0.796±0.069
ADMIRE	0.992±0.005	**0.879±0.030**	**0.982±0.011**	0.992±0.005	**0.879±0.030**	**0.982±0.011**
AD+MVNB L=0.5M	0.988±0.005	0.750±0.060	0.930±0.022	0.989±0.004	0.755±0.060	0.929±0.023
AD+MVNB L=0.75M	0.792±0.073	0.676±0.047	0.783±0.095	0.919±0.028	0.689±0.029	0.854±0.056
AD+MVBB L=0.5M	**0.997±0.003**	**0.863±0.046**	**0.983±0.009**	**0.998±0.002**	**0.863±0.046**	**0.983±0.009**

As a result, we have included some histogram matching results for the reader to peruse, but we refrain from drawing any strong conclusions from the data.

Table 7.2: Image Metrics for *in vivo* Kidney Cases - with Histogram Matching to ADMIRE. Table ©2022 IEEE.

	Measured Stone Size (mm)											
	Original Data						Matched to ADMIRE Data					
Case	1	2	3	4	5	6	1	2	3	4	5	6
DAS	0.489	0.445	0.814	0.482	0.595	0.535	0.466	0.396	0.784	0.446	0.586	0.516
DAS-Hamm	0.526	0.457	0.812	0.486	0.623	0.550	0.495	0.425	0.800	0.476	0.623	0.536
MVNB L=0.25M	0.495	0.473	0.813	0.593	0.595	0.550	0.462	0.414	0.780	0.532	0.595	0.516
MVBB L=0.25M	0.508	0.432	0.791	0.505	0.611	0.555	0.492	0.403	0.767	0.470	0.608	0.541
ADMIRE	0.508	0.407	0.750	0.498	**0.576**	0.495	0.508	0.407	0.750	0.498	**0.576**	0.495
AD+MVNB L=0.5M	**0.432**	0.367	0.734	0.437	0.623	**0.437**	**0.436**	0.367	0.734	0.439	0.626	**0.437**
AD+MVBB L=0.25M	0.501	**0.356**	**0.695**	**0.385**	0.638	0.471	0.501	**0.356**	**0.695**	**0.385**	0.644	0.471
	Contrast Ratio (dB)											
	Original Data						Matched to ADMIRE Data					
Case	1	2	3	4	5	6	1	2	3	4	5	6
DAS	19.0	14.6	9.68	13.9	11.5	11.4	**26.6**	22.3	13.7	20.3	17.4	17.1
DAS-Hamm	18.5	16.4	10.0	13.9	12.0	11.5	25.8	23.4	13.7	**20.4**	**18.1**	17.5
MVNB L=0.25M	18.3	14.8	9.74	13.6	10.9	11.2	26.5	**25.1**	14.2	19.6	16.0	**17.9**
MVBB L=0.25M	18.6	15.8	10.5	13.9	10.9	11.3	23.7	21.2	**15.3**	19.3	14.7	15.4
ADMIRE	21.6	19.3	8.14	14.9	12.7	14.0	21.6	19.3	8.14	14.9	12.7	14.0
AD+MVNB L=0.5M	**23.6**	**23.3**	9.71	**16.4**	**13.7**	16.6	19.8	19.7	8.70	14.6	12.2	14.1
AD+MVBB L=0.25M	21.8	21.2	**11.6**	15.5	10.6	**17.2**	18.1	18.0	10.0	13.8	9.42	16.0
	generalized Contrast-to-Noise Ratio (gCNR)											
	Original Data						Matched to ADMIRE Data					
Case	1	2	3	4	5	6	1	2	3	4	5	6
DAS	0.781	0.649	0.582	0.683	0.632	0.609	0.942	0.874	0.635	**0.791**	0.732	0.753
DAS-Hamm	0.781	0.726	0.609	0.721	0.661	0.600	0.933	0.873	0.632	0.786	**0.742**	0.763
MVNB L=0.25M	0.800	0.633	0.603	0.680	0.563	0.598	**0.948**	**0.929**	0.650	0.764	0.647	0.776
MVBB L=0.25M	0.793	0.716	**0.618**	0.676	0.603	0.593	0.885	0.836	**0.673**	0.758	0.644	0.684
ADMIRE	0.862	0.748	0.566	0.670	0.642	0.700	0.862	0.748	0.566	0.670	0.642	0.700
AD+MVNB L=0.5M	**0.895**	**0.893**	0.579	**0.749**	**0.707**	**0.798**	0.884	0.913	0.584	0.745	0.701	**0.801**
AD+MVBB L=0.25M	0.855	0.878	0.615	0.705	0.471	0.785	0.843	0.866	0.612	0.699	0.466	0.781

Table 7.3: Anechoic Cyst Simulation Metrics for Alternative Histogram Matching Targets. Table ©2022 IEEE.

Average Cyst Edge Width (mm)

0dB SCR (Reverberation Clutter)										
Matching to...	Original	DAS	DAS-Hamm	MVNB L=0.5M	MVNB L=0.75M	MVBB L=0.5M	ADMIRE	AD+MVNB L=0.5M	AD+MVNB L=0.75M	AD+MVBB L=0.5M
DAS	0.216	0.216	0.217	0.202	0.188	0.213	0.181	**0.129**	0.121	0.184
DAS-Hamm	0.347	0.345	0.347	0.274	0.219	0.337	0.213	0.159	0.144	0.214
MVNB L=0.5M	0.641	0.660	0.663	0.641	0.467	0.657	0.459	0.448	0.435	0.462
MVNB L=0.75M	0.292	0.384	0.388	0.371	0.292	0.379	0.282	0.225	0.221	0.283
MVBB L=0.5M	0.437	0.437	0.440	0.352	0.318	0.437	0.308	0.298	0.289	0.310
ADMIRE	0.218	0.251	0.252	0.239	0.221	0.246	0.218	0.167	**0.107**	0.219
AD+MVNB L=0.5M	**0.140**	**0.187**	**0.189**	**0.175**	**0.157**	**0.183**	**0.146**	0.140	0.130	**0.149**
AD+MVNB L=0.75M	0.183	0.229	0.230	0.216	0.202	0.224	0.197	0.187	0.183	0.198
AD+MVBB L=0.5M	0.184	0.257	0.260	0.248	0.229	0.257	0.183	0.174	0.121	0.184
0dB SNR (Thermal Noise)										
Matching to...	Original	DAS	DAS-Hamm	MVNB L=0.5M	MVNB L=0.75M	MVBB L=0.5M	ADMIRE	AD+MVNB L=0.5M	AD+MVNB L=0.75M	AD+MVBB L=0.5M
DAS	0.171	**0.171**	**0.171**	**0.168**	**0.148**	**0.172**	0.130	0.125	0.125	0.133
DAS-Hamm	0.202	0.199	0.202	0.196	0.165	0.205	0.150	0.143	0.134	0.149
MVNB L=0.5M	0.303	0.307	0.395	0.303	0.264	0.401	0.235	0.227	0.206	0.234
MVNB L=0.75M	0.155	0.437	0.436	0.350	0.155	0.441	**0.116**	**0.111**	**0.101**	**0.116**
MVBB L=0.5M	0.259	0.177	0.254	0.176	0.157	0.259	0.145	0.141	0.134	0.145
ADMIRE	0.138	0.208	0.210	0.205	0.176	0.213	0.138	0.128	0.114	0.137
AD+MVNB L=0.5M	0.121	0.183	0.183	0.182	0.154	0.184	0.125	0.121	0.108	0.125
AD+MVNB L=0.75M	**0.105**	0.175	0.177	0.173	0.151	0.178	0.130	0.117	0.105	0.132
AD+MVBB L=0.5M	0.154	0.215	0.216	0.214	0.187	0.220	0.151	0.141	0.118	0.154

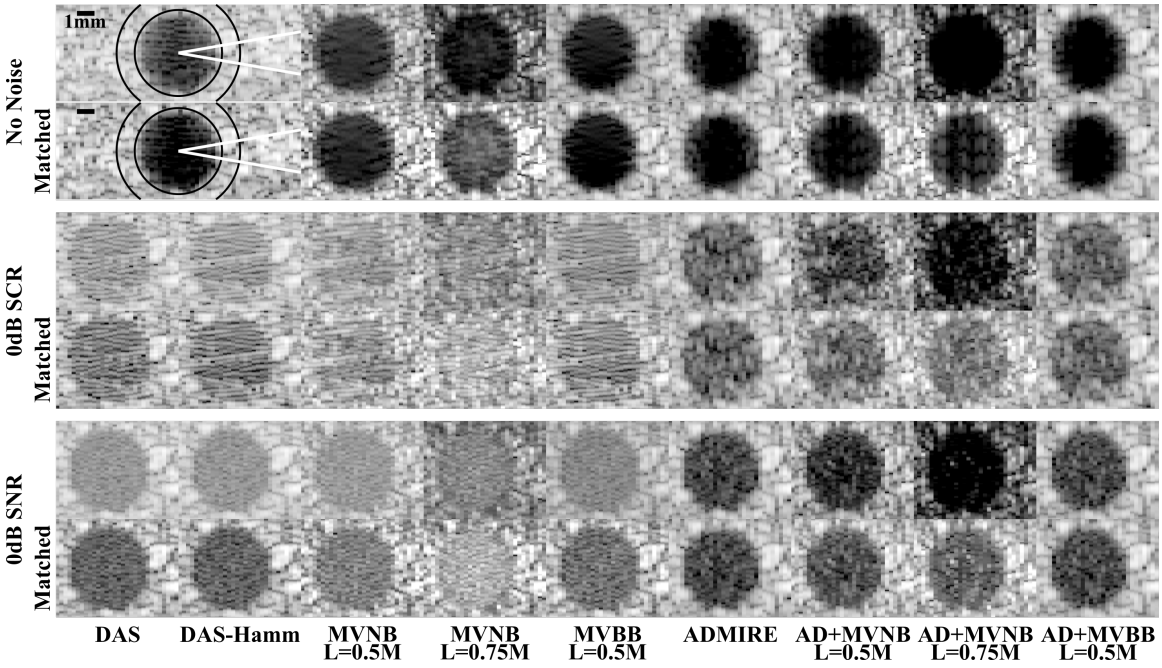

Figure 7.3: B-mode images of one of the anechoic cyst cases shown on a 50 dB dynamic range. Top row shows the No Noise case, middle row shows the added 0dB SCR reverberation clutter case, and bottom row shows the added 0dB SNR Johnson-Nyquist (thermal) noise case. Within each row (top) shows the original b-mode image and (bottom) shows the image matched to ADMIRE. Figure ©2022 IEEE.

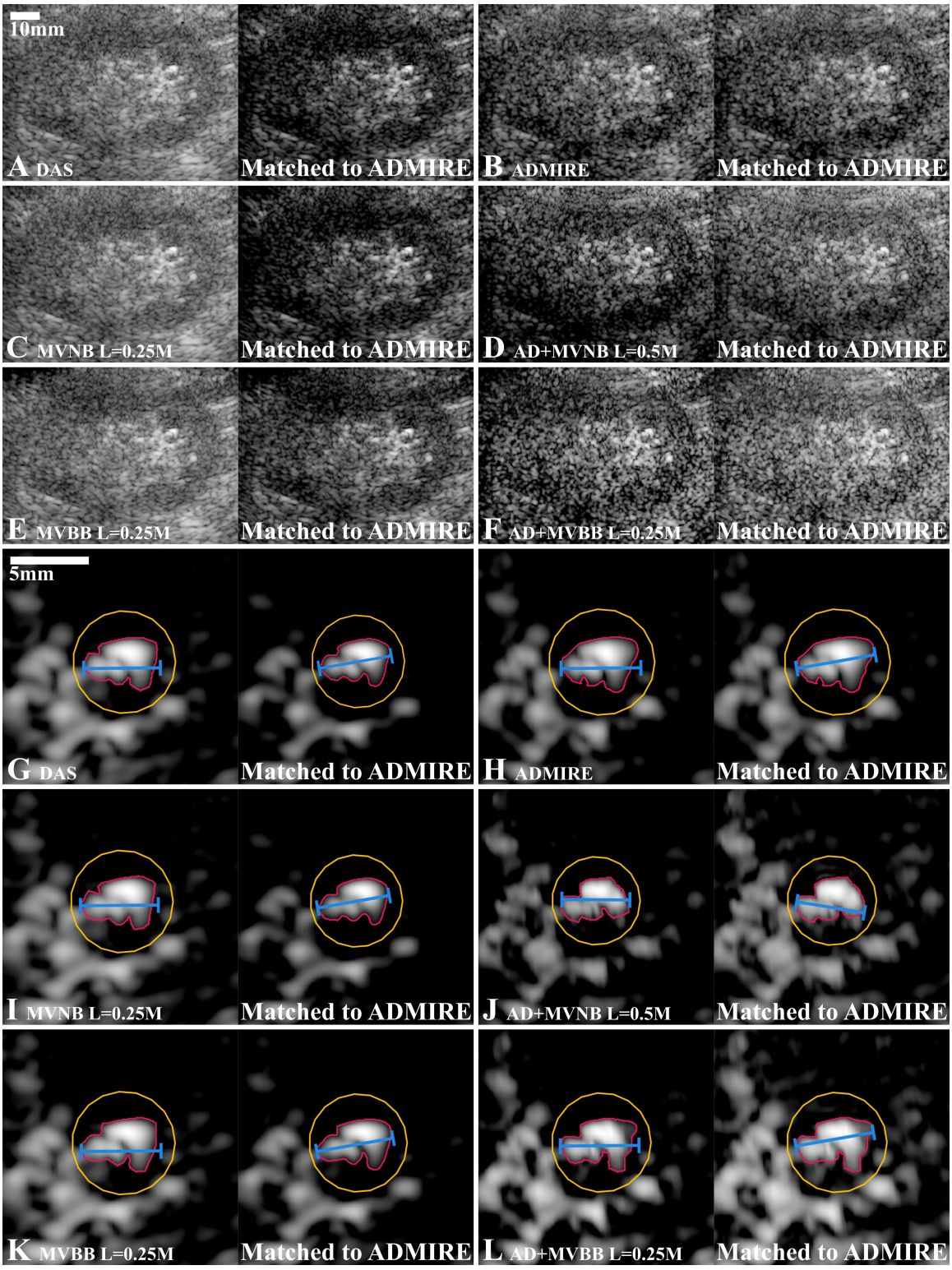

Figure 7.4: B-mode images of *in vivo* kidney Case 1. (A-F) show the full kidney on a 50 dB dynamic range, while (G-L) show the kidney stone with ROI drawn and length measured. Each image includes both the original (left) and matched to ADMIRE (right) versions. Figure ©2022 IEEE.

CHAPTER 8

Conclusions and Future Work

The clinical value of ultrasound imaging is primarily from its low cost and rapid, real-time acquisitions, allowing it to be used frequently for initial screening for many applications. However, for many of these applications the quality of the ultrasound is sufficiently poor that additional imaging with other modalities is required. For example, we previously mentioned that one study found that ultrasound alone for kidney stone disease diagnosis would result in an incorrect diagnosis in 22% of cases (Ganesan et al., 2017). Additionally, increasing obesity rates in the United States (Flegal et al., 1998, 2002, 2010) suggest that this issue is likely to continue to worsen, as these difficult-to-image patients have increased amounts of reverberation clutter, leading to further image quality degradation (Dahl et al., 2017). This highlights the value of ultrasound beamformers that can improve image quality and, more importantly, improve the clinical value of these images.

Ultrasound beamformers have historically set out to improve quantifiable image quality metrics, as these metrics are believed to correlate with subjective clinician assessments of image quality. Improving the contrast ratio and contrast-to-noise ratio to boost lesion detectability or tissue boundary definition, for example, is often the goal of a modern adaptive beamformer. However, depending on the application, there are potentially more valuable ways we can improve the quality of an image besides just blindly improving metrics. After all, boosting the contrast of a lesion or kidney stone may improve our ability to detect it, but it may not improve our ability to make an informed diagnosis.

In this work, we have laid out our efforts to improve ultrasound image accuracy, using the model-based method ADMIRE as a foundation. Image accuracy, especially contrast accuracy, can work counter to traditional beamformer metrics. In the case of difficult-to-classify renal cysts, it may be more valuable to represent the contrast of these cysts accurately to aid diagnosis, rather than with high contrast to aid detectability. In Chapter 3, iterative ADMIRE was presented as a method designed specifically to generate images with accurate contrast representation. It still provides the benefits of reducing off-axis and reverberation clutter of normal ADMIRE, while correcting errors related to improper fitting in the algorithm leading to improved image accuracy. This is especially valuable for hypoechoic targets, like lesions, cysts, and blood vessels, where the important information is at a lower intensity compared to surrounding tissue. Image degradation caused by off-axis clutter from the nearby tissue or reverberation clutter leads to obfuscation of the low-amplitude content, and attempts to declutter this noise can result in a loss of the true underlying signal. Iterative ADMIRE corrects this, allowing us to reconstruct the target data, such as in the blood vessel case we showed

in Fig. 3.9, where clutter masked the blood content in the DAS image, and ADMIRE improperly removed both the clutter and underlying blood content. This suggests iterative ADMIRE may be useful for these high dynamic range scenarios, where the target of interest may be masked by image quality degradation.

On the opposite end of the spectrum, hyperechoic targets that are much brighter than the surrounding tissue are unlikely to be lost or obscured since they are typically at a higher amplitude compared to the noise, but they face an alternative problem: sizing accuracy. In ultrasound, the size of bright targets is often overestimated, with an increased differential between the target and background increasing the overestimation. This is related to the issue of resolution, and more specifically the lateral point spread function of the beamformer. Though the point spread function will be consistent regardless of amplitude, the intersection of the function with the background will change as the amplitude changes. As a target increases in amplitude, the point spread function will be raised relative to the background, resulting in an apparently broader base, increasing the apparent size of the target. With this in mind, we proposed in Chapter 4 combining ADMIRE with MV, a method known for improving lateral resolution, and in Chapter 5 modifying the ADMIRE model to further improve the point spread function by narrowing the base. By narrowing the mainlobe and making it more consistent at different amplitudes, we both improved sizing estimates and made it less dependent on the amplitude of the imaging target. We demonstrated this with *in vivo* kidney stones, which we previously mentioned are often overestimated by as much as 2-3 mm (Ulusan et al., 2007; Unal et al., 2003; Ray et al., 2010; Ganesan et al., 2017; Fowler et al., 2002; Dunmire et al., 2015; Sternberg et al., 2016), which compared to the surgical intervention cutoff size of 5mm is quite significant. We demonstrated that ADMIRE-Ex combined with MV can greatly reduce sizing error, which would lead to more accurate diagnosis of these stones in the screening process, reducing reliance on more expensive and time-consuming modalities that place undue burden on patients, doctors, and hospital facilities.

Overall, we have shown in this dissertation that we can adapt ADMIRE in different ways to increase image accuracy, and therefore quality, in high dynamic range situations, regardless of whether the targets are at higher or lower amplitudes. Though the specific issues are different in these two different cases, it is promising that ADMIRE can be adapted to different medical applications, as well as for different desired outcomes.

However, there are some outstanding challenges that continue to make bringing the full power of ADMIRE into a clinical setting difficult. Namely, ADMIRE continues to be difficult to implement in real time. Though there is currently an implementation of ADMIRE that leverages the power of GPU parallel processing to achieve real time status (Khan et al., 2019, 2020, 2021), there is a trade-off that has to be made to make it possible: the size of the model has to be severely reduced in order to fit current hardware limitations. In this case, it is achieved by the use of dimensionality reducing techniques to shrink the size of the model to a much

smaller footprint (Dei et al., 2019). This has two side effects. First, the model, in which the signals relate to specific echoes based on the linear wave propagation physics, is now reduced, and the signals no longer have specific, physical meaning. Second, and more importantly to this work, the reduced models are incompatible with the techniques laid out in this dissertation. This is due to the model-fitting interactions being different when the reduced model is being used, as opposed to the full physical model used in this work. This is a combination of the model size being significantly different, but also the model space being less well-sampled. To be compatible with iterative ADMIRE and ADMIRE-Ex, both of which are quite computationally intensive, we would either need significantly more advanced hardware that can handle the full models, or some alternative means of reducing the models that does not impact the functionality.

Computationally, the most intensive part of ADMIRE is the model-fitting process, and it scales with the size of the model. Due to the repeated iterative fits that iterative ADMIRE requires for processing, it is unlikely we would be able to achieve true real time processing that would be clinically viable, since the repeated fits would likely cut the frame rate in half or worse. On the other hand, ADMIRE-Ex could potentially be adapted to current hardware, since it is functionally more of an adapted model than an alternative fitting process. By reducing the model size but keeping the meaning of the different model sections intact, it is possible that ADMIRE-Ex could be run in real time, which would be necessary for it to be used for screening kidney stones. This will require future efforts into adapting the model for GPU processing, as well as experimentation to determine what model reduction techniques will allow us to maintain performance.

Though untested with iterative ADMIRE and ADMIRE-Ex, preliminary work has been done on evaluating different ways to reduce the size of the model without using dimensionality reducing methods. In a recent proceedings, we examined ways we can reduce sampling of both clutter regions and the ROI, and evaluated image quality versus computation time (Schlunk and Byram, 2021a). The use of dimensionality reduction was demonstrated to be effective, but we also saw that in many cases we could dramatically reduce the sampling of the model without greatly impacting image quality. In one case, only the reduced-sampled version of the ROI was used, and it was still able to reduce thermal and reverberation clutter, while processing the images 75 times faster than the default model. This particular model was actually smaller than the models normally used in GPU parallel processing, suggesting that we may be able to use non-reduced models in real time if the right sacrifices to the model are made.

Another possibility to reach real time performance could be to restrict the ROI to be processed. Computing iterative ADMIRE or ADMIRE-Ex on an entire image is expensive, but if we restrict an ROI for processing, for example only the area directly around a kidney stone, we would free up a large amount of GPU computing power, potentially allowing us to use larger, better performing models than we otherwise could. This would allow a sonographer to harness the full power of the model, which would be sufficient for

small targets such as kidney stones. Future work could even consider looking into detection methods, and AI or algorithmic processing to automatically choose small regions to run the advanced model on, requiring less manual input from a clinician.

A second ongoing concern with ADMIRE is the optimal way to tune parameters. Early work has suggested there are generally robust choices for the elastic-net parameters (α and λ), though it is clear that these parameters do have impact on the character of the image. Additionally, as we have seen with ADMIRE-Ex, there are different ways we can formulate the model itself. Finally, as we discussed briefly in Chapter 4 with ADMIRE combined with MV, there may be optimal ways to tune ADMIRE for use with other beamformers that differs than when we use ADMIRE alone.

Early in the development process of iterative ADMIRE, we had considered methods for automatically optimizing the elastic-net parameters. In particular, we had found that using cross-validation techniques could be used, using a portion of the signal for training and the rest for testing the choices. This showed promise, and indicated that certain aspects of an image (e.g. tissue versus blood) had different optimal parameter choices, however cross-validation increased computation time far beyond any reasonable amount, as we often required a magnitude or more fits per signal to perform the cross-validation. As a result, we did not choose to continue pursuing this route, though there are certainly other options available. Future work should consider investigating if there are global image features (overall contrast, variance in the image, estimated levels of noise, etc...) that may predict more optimal choices, which would only slightly increase processing overhead. For example, lag-one coherence is a metric that gives a measure of thermal and acoustic noise in an image (Long et al., 2018), which may provide a fast way to estimate the initial quality of the image. There are additionally methods for measuring overall image signal-to-noise ratio (Friemel et al., 1998), and even more exact spatiotemporal coherence methods that can detect specific levels of certain kinds of noise, such as reverberation versus thermal noise (Vienneau et al., 2022), which would be relatively easy to implement on a clinical system.

Tuning the model itself is vastly more complex, though there are similar ways we might be able to streamline the computation. If we are able to measure the level of reverberation clutter in an image using the spatiotemporal coherence method mentioned before, we can adjust the sampling of the region of the model related to that clutter source (i.e. near field predictors). In cases with especially low reverberation, we might be able to entirely remove the near field component, significantly reducing the size of the model. Otherwise, it would be feasible that we could prepare model parameters in advance for different kinds of imaging applications. For example, if we need to size hyperechoic targets, we can choose to use the ADMIRE-Ex model, and otherwise we use the default model.

Finally, tuning for use with other beamformers such as MV is difficult. ADMIRE+MV will certainly work

with the default, normally robust options, but there are simply too many different applications and parameters to realistically optimize for each. Not to mention, as we looked at in detail in Chapter 4, there are potentially ways to tune the beamformer we use after ADMIRE as well. In that case, we found that there are situations where we can increase the subarray averaging parameter of MV when used with ADMIRE, which normally we can not. Needless to say, the amount of work that would be required to clinically evaluate and optimize all the different parameters in these situations would need to be done in a case-by-case basis, as we did with ADMIRE+MV applied specifically to kidney stone data acquired with plane wave synthetic aperture on a curvilinear transducer.

Real time implementations and parameter tuning are the primary challenges that obstruct our ability to implement these methods in a practical clinical setting. While technically these methods could be applied for offline processing that would still reduce the amount of additional imaging required, the modern value of ultrasound imaging is the fact it can be used point-of-care, with immediate results provided to the clinician. However, if the above challenges can be resolved, whether with the considerations suggested or some other methods, ADMIRE and ADMIRE-based methods will be powerful clinical tools. ADMIRE is already known to reduce or eliminate common image degradation sources, and as these sources become increasingly common in the clinical space, these advanced beamformers will become more valuable. Additionally, as this dissertation has demonstrated, ADMIRE is immensely flexible and can be adapted to highly specific applications. Iterative ADMIRE can be used to assist with imaging low amplitude targets, while ADMIRE-Ex can improve sizing accuracy of high amplitude targets.

In addition to those presented here, these techniques could be valuable for other imaging applications. Iterative ADMIRE could be valuable in blood flow applications, where the ability to denoise the image without introducing additional artifacts is important as the blood content is vital for measuring blood flow with Doppler imaging. Additionally, it may be useful in scenarios where bright artifacts cause significant clutter in an image, which can be common in transthoracic cardiac imaging, where the lungs and rib cage can each introduce such noise, or functional brain imaging, where the skull is a major source of image degradation. ADMIRE-Ex could be used for imaging and accurately visualizing stones in the gallbladder or bladder, but additionally may be useful for characterizing calcifications in other areas of the body, such as the cardiac valves in the heart. It would also complement the use of lithotripsy, using high-frequency focused ultrasound shock waves to break up calcifications, such as in the cases listed.

References

Abdel-Gawad, M., Kadasne, R. D., Elsobky, E., Ali-El-Dein, B., and Monga, M. (2016). A Prospective Comparative Study of Color Doppler Ultrasound with Twinkling and Noncontrast Computerized Tomography for the Evaluation of Acute Renal Colic. *Journal of Urology*, 196(3):757–762.

Anderson, M. E. and Trahey, G. E. (2006). A seminar on k-space applied to medical ultrasound.

Asl, B. M. and Deylami, A. M. (2018). A low complexity minimum variance beamformer for ultrasound imaging using dominant mode rejection. *Ultrasonics*, 85:49–60.

Asl, B. M. and Mahloojifar, A. (2009). Minimum variance beamforming combined with adaptive coherence weighting applied to medical ultrasound imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 56(9):1923–1931.

Asl, B. M. and Mahloojifar, A. (2010). Eigenspace-based minimum variance beamforming applied to medical ultrasound imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 57(11):2381–2390.

Austeng, A., Bjastad, T., Synnevaag, J. F., Masoy, S. E., Trop, H., and Holm, S. (2008). Sensitivity of Minimum variance beamforming to tissue aberrations. *Proceedings - IEEE Ultrasonics Symposium*, pages 1072–1075.

Bell, M. A. L., Dahl, J. J., and Trahey, G. E. (2015). Resolution and brightness characteristics of short-lag spatial coherence (SLSC) images. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 62(7):1265–1276.

Bottenus, N., Byram, B., and Hyun, D. (2020a). Histogram matching for visual ultrasound image comparison. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 3010(c):1–1.

Bottenus, N., Le Fevre, M., Cleve, J., Crowley, A. L., and Trahey, G. (2020b). Resolution and Speckle Reduction in Cardiac Imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 3010(c):1–1.

Byram, B., Dei, K., Tierney, J., and Dumont, D. (2015a). A Model and Regularization Scheme for Ultrasonic Beamforming Clutter Reduction. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 62(11):1913–1927.

Byram, B., Dei, K., Tierney, J., and Dumont, D. (2015b). A model and regularization scheme for ultrasonic beamforming clutter reduction. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 62(11):1913–1927.

Byram, B. and Jakovljevic, M. (2014). Ultrasonic Multipath and Beamforming Clutter Reduction: A Chirp Model Approach. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 61(3):428–440.

Byram, B. and Shu, J. (2016a). A pseudo non-linear method for fast simulations of ultrasonic reverberation. *Medical Imaging 2016: Ultrasonic Imaging and Tomography*, 9790(April 2016):1–7.

Byram, B. and Shu, J. (2016b). Pseudononlinear ultrasound simulation approach for reverberation clutter. *Journal of Medical Imaging*, 3(4).

Byram, B., Shu, J., and Dei, K. (2015c). Nonlinear beamforming of aperture domain signals. *IEEE International Ultrasonics Symposium (IUS)*, pages 1–6.

Cameron, A. C. (2009). EXCEL 2007: Histogram.

Capon, J. (1969). High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8):1408–1418.

Chew, W. C. (1995). *Waves and Fields in Inhomogenous Media*. IEEE Press Series on Electromagnetic Wave Theory. Wiley.

Choudhry, S., Gorman, B., Charboneau, J. W., Tradup, D. J., Beck, R. J., Kofler, J. M., and Groth, D. S. (2000). Comparison of Tissue Harmonic Imaging with Conventional US in Abdominal Disease 1. *Radio-Graphics*, 20(4):1127–1135.

Clifford, L., Fitzgerald, P., and James, D. (1993). Non-rayleigh first-order statistics of ultrasonic backscatter from normal myocardium. *Ultrasound in Medicine & Biology*, 19(6):487–495.

Cobbold, R. S. C. (2007). *Foundations of Biomedical Ultrasound*. Oxford University Press, New York City.

Coe, F. L. (1977). The Natural History of Calcium Urolithiasis. *JAMA: The Journal of the American Medical Association*, 238(14):1519.

Conover, W. J. (2012). The rank transformation-an easy and intuitive way to connect many nonparametric methods to their parametric counterparts for seamless teaching introductory statistics courses. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(5):432–438.

Cox, D. and Oakes, D. (1984). *Analysis of Survival Data*. Chapman and Hall/CRC.

Cox, H. (2002). Adaptive beamforming in non-stationary environments. In *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, 2002.*, volume 1, pages 431–438. IEEE.

Dahl, J. J., Hyun, D., Li, Y., Jakovljevic, M., Bell, M. A. L., Long, W. J., Bottenus, N., Kakkad, V., and Trahey, G. E. (2017). Coherence beamforming and its applications to the difficult-to-image patient. *IEEE International Ultrasonics Symposium (IUS)*.

Dahl, J. J. and Sheth, N. M. (2014). Reverberation Clutter from Subcutaneous Tissue Layers: Simulation and in Vivo Demonstrations. *Ultrasound in Medicine & Biology*, 40(4):714–726.

Dei, K. and Byram, B. (2017). The Impact of Model-Based Clutter Suppression on Cluttered, Aberrated Wavefronts. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 64(10):1450–1464.

Dei, K. and Byram, B. (2018). A Robust Method for Ultrasound Beamforming in the Presence of Off -Axis Clutter and Sound Speed Variation. *Ultrasonics*, 89:34–45.

Dei, K., Luchies, A., and Byram, B. (2017). Contrast ratio dynamic range: a new beamformer performance metric. *IEEE International Ultrasonics Symposium (IUS)*, pages 1–4.

Dei, K., Schlunk, S., and Byram, B. (2019). Computationally Efficient Implementation of Aperture Domain Model Image Reconstruction. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 66(10):1546–1559.

Dei, K., Tierney, J., and Byram, B. (2016). Plane wave image quality improvement using ADMIRE algorithm. *IEEE International Ultrasonics Symposium, IUS*, 2016-Novem(Mv):0–3.

Deylami, A. M. and Asl, B. M. (2018). Iterative Minimum Variance Beamformer with Low Complexity for Medical Ultrasound Imaging. *Ultrasound in Medicine and Biology*, 44(8):1882–1890.

Diamantis, K., Anderson, T., Butler, M. B., Villagomez-Hoyos, C. A., Jensen, J. A., and Sboros, V. (2019). Resolving ultrasound contrast microbubbles using minimum variance beamforming. *IEEE Transactions on Medical Imaging*, 38(1):194–204.

Dunmire, B., Harper, J. D., Cunitz, B. W., Lee, F. C., Hsi, R., Liu, Z., Bailey, M. R., and Sorensen, M. D. (2016). Use of the Acoustic Shadow Width to Determine Kidney Stone Size with Ultrasound. *Journal of Urology*, 195(1):171–177.

Dunmire, B., Lee, F. C., Hsi, R. S., Cunitz, B. W., Paun, M., Bailey, M. R., Sorensen, M. D., and Harper, J. D. (2015). Tools to Improve the Accuracy of Kidney Stone Sizing with Ultrasound. *Journal of Endourology*, 29(2):147–152.

Dutt, V. and Greenleaf, J. F. (1994). Ultrasound Echo Envelope Analysis Using a Homodyned K Distribution Signal Model. *Ultrasonic Imaging*, 16(4):265–287.

Eknoyan, G. (2009). A Clinical View of Simple and Complex Renal Cysts. *Journal of the American Society of Nephrology*, 20:1874–1876.

Fatemi, A., Berg, E. A. R., and Rodriguez-Molares, A. (2019). Studying the Origin of Reverberation Clutter in Echocardiography: In Vitro Experiments and In Vivo Demonstrations. *Ultrasound in Medicine and Biology*, 45(7):1799–1813.

Feynman, R. P. (2011). The Feynman Lectures on Physics, Volume 1, Chapter 47: Sound. In *The Feynman Lectures on Physics Vol. 1: New Millennium Edition*, chapter 47. Basic Books.

Flegal, K. M., Carrol, M. D., Kuczmarski, R. J., and Johnson, C. L. (1998). Overweight and obesity in the United States: Prevalence and trends, 1960–1994. *Int J Obes Relat Metab Disord*, 22(1):39–47.

Flegal, K. M., Carrol, M. D., Ogden, C. L., and Curtin, L. R. (2010). Prevalence and trends in obesity among US adults, 1999–2008. *JAMA*, 303(3):235–241.

Flegal, K. M., Carrol, M. D., Ogden, C. L., and Johnson, C. L. (2002). Prevalence and trends in obesity among U.S. adults, 1999–2000. *JAMA*, 288(14):1723–1727.

Fowler, K. A. B., Locken, J. A., Duchesne, J. H., and Williamson, M. R. (2002). US for Detecting Renal Calculi with Nonenhanced CT as a Reference Standard. *Radiology*, 222(1):109–113.

Friedman, M. (1937). The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *Journal of the American Statistical Association*, 32(200):675–701.

Friemel, B. H., Bohs, L. N., Nightingale, K. R., and Trahey, G. E. (1998). Speckle decorrelation due to two-dimensional flow gradients. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 45(2):317–327.

Ganesan, V., De, S., Greene, D., Torricelli, F. C. M., and Monga, M. (2017). Accuracy of ultrasonography for renal stone detection and size determination: is it good enough for management decisions? *BJU International*, 119(3):464–469.

Gilman, L. (1997). First-order statistics of pulsed-sinusoid backscatter from random media: basic elements of an exact treatment. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 44(4):798–804.

Greening, M. V. and Perkins, J. E. (2002). Adaptive beamforming for nonstationary arrays. *The Journal of the Acoustical Society of America*, 112(6):2872–2881.

Holfort, I. K., Gran, F., and Jensen, J. A. (2009). Broadband minimum variance beamforming for ultrasound imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 56(2):314–325.

Hsi, R. S., Schlunk, S. G., Tierney, J. E., Dei, K., Jones, R., George, M., Karve, P., Duddu, R., and Byram, B. C. (2018). Feasibility of non-linear beamforming ultrasound methods to characterize and size kidney stones. *PLoS ONE*, 13(8):1–14.

Hyun, D., Crowley, A. L. C., and Dahl, J. J. (2017). Efficient Strategies for Estimating the Spatial Coherence of Backscatter. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 64(3):500–513.

Hyun, D., Kim, G. B., Bottenus, N., and Dahl, J. J. (2022). Ultrasound Lesion Detectability as a Distance Between Probability Measures. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 69(2):732–743.

Jensen, J. A. (1991). A model for the propagation and scattering of ultrasound in tissue. *Journal of the Acoustical Society of America*, 89(1):182–190.

Jensen, J. A. (1996). Field: A Program for Simulating Ultrasound Systems. *Paper presented at the 10th Nordic-Baltic Conference on Biomedical Imaging Published in Medical & Biological Engineering and Computing*, 34:351–353.

Jensen, J. A. and Svendsen, N. B. (1992). Calculation of Pressure Fields from Arbitrarily Shaped, Apodized, and Excited Ultrasound Transducers. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 39:262–267.

Jones, R., Byram, B. C., Schlunk, S., Tierney, J. E., and Hsi, R. (2019). The impact of mid lag spatial coherence parameters on coherent target detection. In Ruiter, N. V. and Byram, B. C., editors, *Medical Imaging 2019: Ultrasonic Imaging and Tomography*, number March 2019, page 22. SPIE.

Kaplan, E. L. and Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, 53(282):457.

Khan, C., Dei, K., and Byram, B. (2019). A GPU-Based Implementation of ADMIRE. In *2019 IEEE International Ultrasonics Symposium (IUS)*, volume 2019-Octob, pages 1501–1504. IEEE.

Khan, C., Dei, K., Schlunk, S., Ozgun, K., and Byram, B. (2020). Real-Time, Simultaneous DAS, ADMIRE, and SLSC Imaging Using GPU-Based Processing. In *2020 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4. IEEE.

Khan, C., Dei, K., Schlunk, S., Ozgun, K., and Byram, B. (2021). A Real-Time, GPU-Based Implementation of Aperture Domain Model Image REconstruction. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 3010(c):1–1.

Kinayman, N. and Aksun, M. I. (2005). *Modern Microwave Circuits*. Artech House microwave library. Artech House.

King, W., Kimme-Smith, C., and Winter, J. (1985). Renal stone shadowing: an investigation of contributing factors. *Radiology*, 154(1):191–196.

Lane, D. M. (2007). Introduction to Statistics.

Lawless, J. F. (2002). *Statistical Models and Methods for Lifetime Data*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA.

Lediju, M. A., Pihl, M. J., Dahl, J. J., and Trahey, G. E. (2008). Quantitative assessment of the magnitude, impact and spatial extent of ultrasonic clutter. *Ultrasonic Imaging*, 30(3):151–168.

Lediju, M. A., Trahey, G. E., Byram, B. C., and Dahl, J. J. (2011). Short-lag spatial coherence of backscattered echoes: Imaging characteristics. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 58(7):1377–1388.

Li, P.-c. and Li, M.-L. (2003). Adaptive Imaging Using the Generalized Coherence Factor. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 50(2):128–141.

Long, W., Bottenus, N., and Trahey, G. E. (2018). Lag-One Coherence as a Metric for Ultrasonic Image Quality. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 65(10):1768–1780.

Mallart, R. and Fink, M. (1991). The van Cittert–Zernike theorem in pulse echo measurements. *The Journal of the Acoustical Society of America*, 90(5):2718–2727.

Matrone, G., Savoia, A. S., Caliano, G., and Magenes, G. (2015). The Delay Multiply and Sum Beamforming Algorithm in Ultrasound B-Mode Medical Imaging. *IEEE Transactions on Medical Imaging*, 34(4):940–949.

May, P. C., Haider, Y., Dunmire, B., Cunitz, B. W., Thiel, J., Liu, Z., Bruce, M., Bailey, M. R., Sorensen, M. D., and Harper, J. D. (2016). Stone-Mode Ultrasound for Determining Renal Stone Size. *Journal of Endourology*, 30(9):958–962.

Mehdizadeh, S., Austeng, A., Johansen, T. F., and Holm, S. (2012). Eigenspace based minimum variance beamforming applied to ultrasound imaging of acoustically hard tissues. *IEEE Transactions on Medical Imaging*, 31(10):1912–1921.

Mohana Shankar, P. (2000). A general statistical model for ultrasonic backscattering from tissues. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 47(3):727–736.

Molthen, R., Shankar, P., and Reid, J. (1995). Characterization of ultrasonic B-scans using non-rayleigh statistics. *Ultrasound in Medicine & Biology*, 21(2):161–170.

Montaldo, G., Tanter, M., Bercoff, J., Benech, N., and Fink, M. (2009). 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.

Moore, C. L. and Scoutt, L. (2012). Sonography first for acute flank pain? *Journal of Ultrasound in Medicine*, 31(11):1703–1711.

Nakagami, M. (1960). The m-Distribution—A General Formula of Intensity Distribution of Rapid Fading. In *Statistical Methods in Radio Wave Propagation*, pages 3–36. Elsevier.

Patterson, M. and Foster, F. (1983). The improvement and quantitative assessment of B-mode images produced by an annular array/cone hybrid. *Ultrasonic Imaging*, 5(3):195–213.

Pinton, G. F., Trahey, G. E., and Dahl, J. J. (2011). Sources of image degradation in fundamental and harmonic ultrasound imaging using nonlinear, full-wave simulations. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 58(4):754–765.

Prins, J., McCormack, D., Michelson, D., and Horrell, K. (2003). Chi-square goodness-of-fit test.

Ray, A. A., Ghiculete, D., Pace, K. T., and Honey, R. J. D. (2010). Limitations to Ultrasound in the Detection and Measurement of Urinary Tract Calculi. *Urology*, 76(2):295–300.

Rindal, O. M. H., Austeng, A., Fatemi, A., and Rodriguez-Molares, A. (2019). The effect of dynamic range alterations in the estimation of contrast. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 66(7).

Rindal, O. M. H., Austeng, A., Torp, H., Holm, S., and Rodriguez-Molares, A. (2016). The dynamic range of adaptive beamformers. *IEEE International Ultrasonics Symposium (IUS)*, pages 1–4.

Rindal, O. M. H., Rodriguez-Molares, A., and Austeng, A. (2017). The Dark Region Artifact in Adaptive Ultrasound Beamforming. *IEEE International Ultrasonics Symposium (IUS)*, pages 1–4.

Robinson, B. S., Shmulewitz, A., and Burke, T. M. (1994). Waveform aberrations in an animal model. *Proceedings of the IEEE Ultrasonics Symposium*, 3:1619–1624.

Rodriguez-Molares, A., Rindal, O. M. H., D'hooge, J., Masoy, S.-E., Austeng, A., Lediju Bell, M. A., and Torp, H. (2020). The Generalized Contrast-to-Noise Ratio: A Formal Definition for Lesion Detectability. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 67(4):745–759.

Rodriguez-Molares, A., Rindal, O. M. H., D'Hooge, J., Måsøy, S.-E., Austeng, A., and Torp, H. (2018). The Generalized Contrast-to-Noise Ratio. *IEEE International Ultrasonics Symposium (IUS)*, pages 1–4.

Salari, A. and Asl, B. M. (2021). User Parameter-Free Minimum Variance Beamformer in Medical Ultrasound Imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 68(7):2397–2406.

Samei, E., Flynn, M. J., and Reimann, D. A. (1998). A method for measuring the presampled MTF of digital radiographic systems using an edge test device. *Medical Physics*, 25(1):102–113.

Sasso, M. and Cohen-Bacrie, C. (2005). Medical ultrasound imaging using the fully adaptive beamformer. In *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, volume 2, pages 489–492. IEEE.

Schlunk, S. and Byram, B. (2020). Expanded beamforming models for high dynamic range scenarios. In *2020 IEEE International Ultrasonics Symposium (IUS)*, volume 2020-Septe, pages 1–4. IEEE.

Schlunk, S. and Byram, B. (2021a). On Model Space Sampling in ADMIRE for Image Quality and Computational Efficiency. In *2021 IEEE International Ultrasonics Symposium (IUS)*, number 3, pages 1–4. IEEE.

Schlunk, S. and Byram, B. (2022a). Breaking and Fixing gCNR and Histogram Matching. In *2022 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4.

Schlunk, S. and Byram, B. (2022b). Combining ADMIRE and MV to Improve Image Quality. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 69(9):2651–2662.

Schlunk, S. and Byram, B. (2022c). Considerations for a more robust implementation of gCNR. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, In Review.

Schlunk, S. and Byram, B. C. (2021b). Using ADMIRE to improve minimum variance performance in the presence of reverberation clutter. In Ruiter, N. V. and Byram, B. C., editors, *Medical Imaging 2021: Ultrasonic Imaging and Tomography*, number 1, page 1. SPIE.

Schlunk, S., Dei, K., and Byram, B. (2018). Iterative ADMIRE for High Dynamic Range B-Mode. *IEEE International Ultrasonics Symposium (IUS)*, 2018-Octob:1–4.

Schlunk, S., Dei, K., and Byram, B. (2021). Iterative Model-Based Beamforming for High Dynamic Range Applications. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 68(3):482–493.

Schlunk, S., Hsi, R., and Byram, B. (2022). Enhancing sizing accuracy in ultrasound images with an alternative ADMIRE model and dynamic range considerations. *Ultrasonics*, In Review:1–10.

Schlunk, S., Tierney, J., George, M., Karve, P., Duddu, R., Hsi, R. S., and Byram, B. (2017). Non-linear beamforming approaches for sizing and detecting large calcifications. In *2017 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4. IEEE.

Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika*, 66(3):605–610.

Scott, D. W. (1992). *Multivariate Density Estimation*. Wiley Series in Probability and Statistics. Wiley.

Scott, D. W. (2010). Histogram. *WIREs Computational Statistics*, 2(1):44–48.

Shabana, W., Bude, R. O., and Rubin, J. M. (2009). Comparison Between Color Doppler Twinkling Artifact and Acoustic Shadowing for Renal Calculus Detection: An In Vitro Study. *Ultrasound in Medicine & Biology*, 35(2):339–350.

Shankar, P. M. (1995). A model for ultrasonic scattering from tissues based on the K distribution. *Physics in Medicine and Biology*, 40(10):1633–1649.

Siegel, C. L., McFarland, E. G., Brink, J. A., Fisher, A. J., Humphrey, P., and Heiken, J. P. (1997). CT of Cystic Renal Masses. *American Journal of Roentgenology*, 169:813–818.

Smith, R. C. and Varanelli, M. (2000). Diagnosis and Management of Acute Ureterolithiasis. *American Journal of Roentgenology*, 175(1):3–6.

Smith, R. C., Verga, M., McCarthy, S., and Rosenfield, A. T. (1996). Diagnosis of acute flank pain: value of unenhanced helical CT. *American Journal of Roentgenology*, 166(1):97–101.

Smith, S. W., Wagner, R. F., Sandrik, J. M., and Lopez, H. (1983). Low Contrast Detectability and Contrast/Detail Analysis in Medical Ultrasound. *IEEE Transactions on Sonics and Ultrasonics*, 30(3):164–173.

Sternberg, K. M., Eisner, B., Larson, T., Hernandez, N., Han, J., and Pais, V. M. (2016). Ultrasonography Significantly Overestimates Stone Size When Compared to Low-dose, Noncontrast Computed Tomography. *Urology*, 95:67–71.

Stoica, P., Zhisong Wang, and Jian Li (2003). Robust Capon beamforming. *IEEE Signal Processing Letters*, 10(6):172–175.

Sturges, H. a. (1926). The Choice of a Class Interval. *Journal of the American Statistical Association*, 21(153):65–66.

Synnevåg, J.-F., Austeng, A., and Holm, S. (2007). Adaptive Beamforming Applied to Medical Ultrasound Imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 54(8):1606–1613.

Synnevåg, J. F., Austeng, A., and Holm, S. (2009). Benefits of minimum-variance beamforming in medical ultrasound imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 56(9):1868–1879.

Szabo, T. L. (2014). *Diagnostic Ultrasound Imaging: Inside Out*. Academic Press, second edition.

Tibshirani, R. J. and Taylor, J. (2012). Degrees of Freedom in Lasso Problems. *The Annals of Statistics*, 40(2):1198–1232.

Tierney, J. E., Schlunk, S. G., Jones, R., George, M., Karve, P., Duddu, R., Byram, B. C., and Hsi, R. S. (2019). In vitro feasibility of next generation non-linear beamforming ultrasound methods to characterize and size kidney stones. *Urolithiasis*, 47(2):181–188.

Ueno, A., Kawamura, T., Ogawa, A., and Takayasu, H. (1977). Relation of spontaneous passage of ureteral calculi to size. *Urology*, 10(6):544–546.

Ulusan, S., Koc, Z., and Tokmak, N. (2007). Accuracy of sonography for detecting renal stone: Comparison with CT. *Journal of Clinical Ultrasound*, 35(5):256–261.

Unal, D., Yeni, E., Karaoglanoglu, M., Verit, A., and Karatas, O. F. (2003). Can Conventional Examinations Contribute to the Diagnostic Power of Unenhanced Helical Computed Tomography in Urolithiasis? *Urologia Internationalis*, 70(1):31–35.

Van Trees, H. L. (2002). Optimum Waveform Estimation. In *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*, chapter 6, pages 428–709. John Wiley & Sons, Inc., New York, USA.

Vienneau, E. P., Ozgun, K. A., and Byram, B. C. (2022). Spatiotemporal Coherence to Quantify Sources of Image Degradation in Ultrasonic Imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 69(4):1337–1352.

Viola, F., Ellis, M. A., and Walker, W. F. (2008). Time-domain optimized near-field estimator for ultrasound imaging: Initial development and results. *IEEE Transactions on Medical Imaging*, 27(1):99–110.

Wagner, R., Smith, S., Sandrik, J., and Lopez, H. (1983). Statistics of Speckle in Ultrasound B-Scans. *IEEE Transactions on Sonics and Ultrasonics*, 30(3):156–163.

Wagner, R. F., Insana, M. F., and Brown, D. G. (1987). Statistical properties of radio-frequency and envelope-detected signals with applications to medical ultrasound. *Journal of the Optical Society of America A*, 4(5):910.

Wang, W., Yan, S., Mao, L., and Guo, X. (2021). Robust Minimum Variance Beamforming with Sidelobe-Level Control Using the Alternating Direction Method of Multipliers. *IEEE Transactions on Aerospace and Electronic Systems*, 57(5):3506–3519.

Warren, K. S. and McFarlane, J. (2005). The Bosniak classification of renal cystic masses. *BJU International*, 95(7):939–942.

Weyl, H. (1919). Ausbreitung elektromagnetischer Wellen über einem ebenen Leiter. *Annalen der Physik*, 365(21):481–500.

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1(6):80.

Wilcoxon, F. (1950). SOME RAPID APPROXIMATE STATISTICAL PROCEDURES. *Annals of the New York Academy of Sciences*, 52(6):808–814.

Yang, B. (2008). A Study of Inverse Short-Time Fourier Transform. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3541–3544.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320.