Program Digest, 2019 IEEE International Ultrasonics Symposium (IUS)
Glasgow, Scotland, October 6-9, 2019

Delay Compression: Reducing Delay Calculation Requirements for 3D Plane-Wave Ultrasound

Brendan L. West
University of Michigan
Ann Arbor, MI, USA
westbl@umich.edu

Jian Zhou
Arizona State University
Tempe, AZ, USA
jzhou50@asu.edu

Chaitali Chakrabarti
Arizona State University
Tempe, AZ, USA
chaitali@asu.edu

Thomas F. Wenisch
University of Michigan
Ann Arbor, MI, USA
twenisch@umich.edu

***Abstract*—In 3D plane-wave ultrasound, computational requirements are directly proportional to the number of focal points in a volume. For a receive aperture of size of M_xM_y transducers, a beamforming aperture size of N_xN_y, and a depth of M_z focal points, $M_xM_yM_zN_xN_y$ round-trip delays must be computed. To reduce this requirement, we decompose the planar transmit distance into two parts: (1) from the plane-wave's origin to the first point in each focal line, and (2) from the plane when it is touching the first point in each focal line to each subsequent focal point along that line. The latter distance, as well as the reflection distances, are symmetric across beamforming apertures, and thus their computation can be shared. This decomposition results in up to a M_xM_y reduction in the number of unique delays while retaining full image quality. Using our technique, precomputing delays and storing them in look-up tables (LUTs) is now possible for 3D plane-wave ultrasound for the first time, opening new doors for computational architectures in this field. Our method works with 2D, 3D, and 3D-separable variants of plane-wave ultrasound.**

***Index Terms*—ultrasound, beamforming, accelerator, plane-wave**

I. Introduction

3D ultrasound is an imaging modality that has seen increased use in medical applications due to its low risk potential compared to X-ray and MRI. Ultrasound relies on the principle of sonic transmission and reflection, acting much like sonar and radar. For 3D variants, a 2D array of transducers transmits sound waves into the target volume by exciting the transducers with an electronic pulse. As the sound waves pass through the volume, partial reflections occur when the waves encounter tissue interfaces. These reflections are then sampled by the 2D transducer array; sampling is often done at 40MHz. The raw reflection signals collected by each transducer are used to reconstruct a volumetric image through a process called beamforming, wherein the signals are filtered to coherently sum reflections that originate from the numerous focal points within the image volume.

Beamforming is the most computationally expensive aspect of ultrasound imaging. Identifying the samples within each receive signal that correspond to each focal point is computationally expensive, as it requires numerous trigonometry calculations for each focal point. Moreover, there is substantial data sharing and reuse, as each sample may contribute to numerous voxels in the final image. Hence, although calculating each voxel is in principle embarrassingly parallel, performing these computations efficiently while exploiting data sharing is difficult.

Time-delay beamforming, one of the primary beamforming methods, calculates a time delay for the propagation path from each transmit element to each focal point to each receive element. For a target volume of size $M_xM_yM_z$ focal points and a receive array of N_xN_y transducers, the number of required delays, and the associated computational complexity, is $M_xM_yM_zN_xN_y$. After sampling the reflected signals, the samples are commonly interpolated to improve accuracy of the delay calculations without the power overhead of faster ADCs [1]. After interpolation, samples corresponding to each voxel are selected using the time delays. The selected values are then apodized to reduce side lobes; these apodized samples are finally summed to form the voxel.

To reduce the computational complexity of 3D ultrasound systems, researchers have proposed separable approaches to ultrasound beamforming [12]–[14]. Separable algorithms approximate traditional beamforming by splitting it into two stages: in the first stage beamforming is performed along the X-axis, while in the second stage it is performed along the Y-axis. The two-stage process allows for a significant reduction in computational complexity, reducing the 2D grid of contributing transducers for each voxel (multiplicative complexity) to two 1D arrays of contributing transducers (additive complexity), while losing little in terms of final image clarity.

Some applications, such as 3D shear-wave elastography and 3D vector flow imaging, require image volume rates exceeding 1,000 volumes per second for tracking high-frequency motion. Plane-wave ultrasound has been shown to be a promising technique to achieve these high volume rates, but can lack resolution and result in low SNR due to the lack of transmit focusing. Prior works [8], [14] propose a plane-wave image compounding scheme, which increases resolution and SNR by performing multiple firings at different angles and summing the resulting images [8] and volumes [14]. However, the constraints of the algorithm in [14] required their proposed hardware design to store all of the data needed to construct a volume on-chip, iterating over the data multiple times to build the volume and thereby greatly decreasing the attainable volume rate.

In this paper, we build upon the prior work [12], [14], proposing a novel refinement of the 2D plane-wave imaging algorithm that enables drastic reduction in the number of unique round-trip computations per volume. We then extend this reduction to 3D for both non-separable and separable variants, demonstrating that this approach is not limited to the 2D case. Our time-delay decomposition approach dramatically decreases hardware complexity, increases performance, and lowers power requirements. In short, our key contributions are:

- A novel delay decomposition for plane-wave ultrasound, resulting in up to a 1024× reduction in the number of delays computed using our 3D system parameters.
- A two-part delay decomposition for separable 3D plane-wave ultrasound, resulting in an asymptotic reduction in the number of delays computed for the first and second stages—up to 1008× and 1006× reduction in the first and second stages when using our system parameters, respectively.

This work was supported in part by NSF CCF-1406739 and the Applications Driving Architectures (ADA) Research Center, a JUMP Center cosponsored by SRC and DARPA.

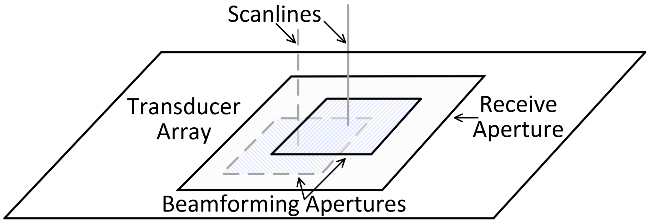


Fig. 1: 3D plane-wave transmit and receive setup; overlapping beamforming apertures "step" across the receive aperture, each creating a single scanline.

II. Plane-Wave Imaging Background

Plane-wave imaging is an ultrasound method that utilizes a flat-plane transmission scheme. A large 2D array of transducers receive and sample the reflected signals, with each transducer outputting a *channel* of samples. The receive array is divided into smaller, overlapping 2D sub-arrays called beamforming apertures, as seen in Figure 1. The channels within each beamforming aperture contribute to voxels along a single *scanline*, or column of voxels in the final volume. In this paper, we assume a receive array of M_xM_y transducers, beamforming apertures of N_xN_y transducers, and a depth of M_z, resulting in a final volume size of $M_xM_yM_z$ voxels.

Managing delay constants poses a significant challenge in many 3D ultrasound algorithms, and plane-wave is no different. For separable plane-wave imaging with angled compounding [14], which can produce high quality images with plane-wave transmits, each firing angle requires a combined $M_xM_zN_xN_y + M_xM_yM_zN_y$ delay constants for the two beamforming stages. In Equations (1), (2), (m_x, m_y, m_z) are the coordinates of each focal point, (n_x, n_y) are the coordinates of each receive transducer within the beamforming aperture, and (x_0, y_0, z_0) is an arbitrary point on the plane at its origin; in this formulation, there is little symmetry between beamforming apertures due to the varying firing angle of the plane.

$$d_{tx} = (m_x - x_0)\sin\alpha\cos\beta + (m_y - y_0)\sin\alpha\sin\beta + (m_z - z_0)\cos\alpha \tag{1}$$

$$d_{rx} = \sqrt{(m_x - n_x)^2 + (m_y - n_y)^2 + m_z^2} \tag{2}$$

The high number of required delay constants is one of the principle difficulties of 3D ultrasound in general, as it prohibits large beamforming apertures, receive volumes, and depth combinations. Prior works [2]–[5], [9], [10] have used specialized hardware to approximate delays or share small components of the computation between focal points to reduce the hardware requirements for beamforming, but all of these methods fall orders of magnitude short of the target 1,000+ volumes per second for plane-wave applications. Similarly, precomputing the delays and placing them in on-chip storage—or streaming them from off-chip storage—is also prohibitive due to sheer quantity required (millions of delays per volume for large receive apertures); none of the prior approaches are sufficient for high-volume-rate beamforming.

III. Delay Compression

To combat the computational requirements of high-volume-rate beamforming, we propose *delay compression*. Delay compression recognizes additional symmetry between the round-trip times for all focal points, thereby allowing us to drastically reduce the total number of computations required for each volume. At a high level, delay compression works by breaking the transmission component of the round-trip distance calculation into two pieces, which we call d_{tx1} and d_{tx2}. d_{tx1} is a set of unique distances from the plane-wave's origin to the first focal point along each scanline, where $m_z = 0$, hereafter referred to as the *base* of each scanline, while d_{tx2} is a set of distances calculated from when the plane is touching the base of a single scanline to each focal point along that scanline. One d_{tx1} distance corresponds to each scanline, and the set of d_{tx2} distances is shared between all scanlines—the d_{tx1} distance acts as an offset for the d_{tx2} distances within each beamforming aperture. To thoroughly explain delay compression, we will work our way through its function in plane-wave's 2D, 3D, and finally 3D Separable [12], [14] variants.

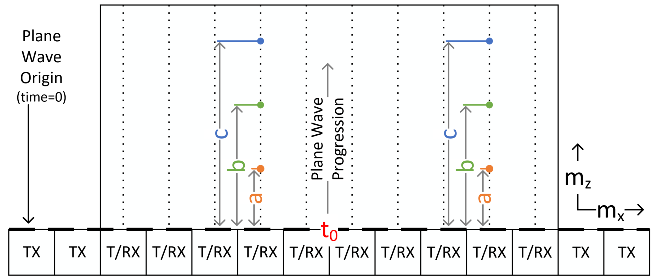


Fig. 2: X-Z slice of the image space showing a 2D component of delay compression. Unique transmit distances a, b, and c represent the shared d_{tx} distances from the plane to focal points m_z along each scanline.

A. Delay Compression in 2D

In its most simplistic form, plane-wave ultrasound fires a planar pulse directly into the target medium, with an elevational angle of 0°, $M_xM_zN_x$ round-trip delays, and a final image size of M_xM_z. In this case the delay decomposition is straight-forward, as all of the transmission and reflection distances are equal—the distance to the base of each scanline, d_{tx1}, is 0 for all scanlines. Equations (3), (4) demonstrate that the transmission distance does not vary across scanlines, so round-trip distance calculations can be shared—reducing the number of unique delays to M_zN_x.

$$d_{tx} = m_z - z_0 \tag{3}$$

$$d_{rx} = \sqrt{(m_x - n_x)^2 + m_z^2} \tag{4}$$

Figure 2 illustrates the symmetry. In the figure, the transmission distance is identical for each focal point that lies at the same depth m_z. With this in mind, distances a, b, and c can be computed only once for all scanlines, resulting in a single shared set of distance values. The same is true for the reflection distances: since each beamforming aperture is centered on the scanline it produces, the distances from each focal point on the scanline to each transducer within the beamforming aperture are identical across apertures.

However, whereas a planar pulse with an elevational angle of 0° leads to straight-forward symmetry for delay reduction, there are cases, such as improving image quality through coherent compounding, which require transmits at varying elevational angles—resulting in $M_xM_zN_x$ delay calculations. Equations (5), (6) show how the transmission and reflection distances are calculated for planar transmits with non-0° elevation angle α.

$$d_{tx} = (m_x - x_0)\sin\alpha + (m_z - z_0)\cos\alpha \tag{5}$$

$$d_{rx} = \sqrt{(m_x - n_x)^2 + m_z^2} \tag{6}$$

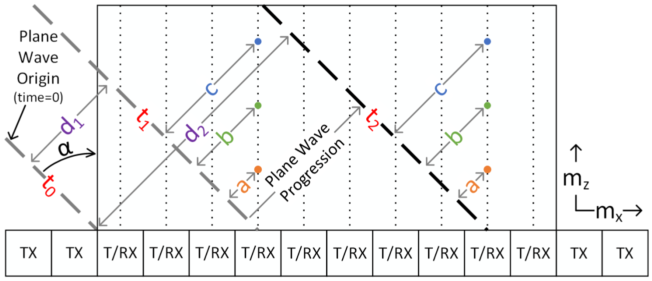
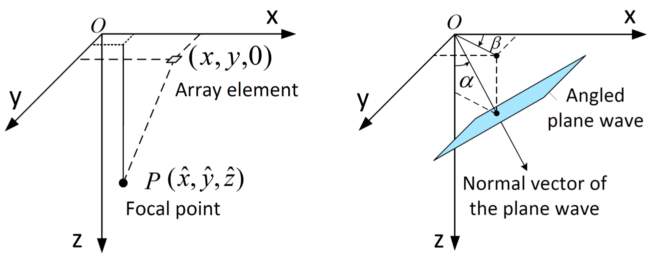


Fig. 3: X-Z slice of the image space showing a 2D component of delay compression. Unique d_{tx1} distances are shown in d_1 and d_2, while a, b, and c represent the shared d_{tx2} distances from the plane to focal points m_z along each scanline.



(a) Coordinate system

(b) Firing plane angles

Fig. 4: Coordinate system and angle definition of 3D plane-wave system with coherent compounding.

For angled transmits, the transmit distance for each focal point at the same m_z depth depends on the scanline to which the focal point belongs. To recover the symmetry available in the 0° case, we use delay compression to decompose the transmission distance into two components, which we call d_{tx1} and d_{tx2}. d_{tx1} addresses the angle dependency by capturing the unique distance from the plane-wave's origin to the base of each scanline, from which the distances to each focal point are again equal across scanlines. By setting the value m_z in Equation (5) to 0, we find the distance from the plane-wave's origin to the first focal point in each scanline. We then set m_x equal to x_0 and z_0 equal to 0, essentially transforming the origin of the coordinate system to the base of the scanline. The 2D variant of the transmit distance decomposition can be seen in Equations (7), (8).

$$d_{tx1} = (m_x - x_0)\sin\alpha - z_0\cos\alpha \tag{7}$$

$$d_{tx2} = m_z\cos\alpha \tag{8}$$

Figure 3 again illustrates our delay decomposition, this time with an elevational angle of α. In this example, d_{tx1} distances d_1 and d_2 vary, but the d_{tx2} distances a, b, and c are constant across scanlines. By computing one d_{tx1} value for each scanline and a single set of d_{tx2} values, all of the unique transmission distances can be recomputed by simply adding the d_{tx1} values to the set of d_{tx2} values—reducing the number of unique delays to $M_x + M_zN_x$.

B. Delay Compression in 3D

3D plane-wave ultrasound requires drastically more round-trip calculations per volume than the 2D case. The coordinate system used for 3D plane-wave ultrasound includes both an elevational angle α and lateral angle β, as shown in Figure 4. This additional angle results in the number of delay calculations increasing to $M_xM_yM_zN_xN_y$.

TABLE I: Parameters of the target 3D plane-wave system.

Property	Value
Speed of sound (tissue), m/s	1540
Pitch, µm	192.5
Transmit aperture size, transducers	128 × 96
Receive aperture size, transducers	32 × 32
Beamforming aperture size, transducers	32 × 32
Number of scanlines per stage	32 × 32
Stage 1 scanline output length (M_z), points	2,089
Stage 2 scanline output length (M_z), points	1,679
Maximum imaging depth, cm	4
Center frequency, MHz	4
6 dB transducer bandwidth, MHz	2
ADC sampling rate, MHz	40

We implement 3D delay compression much like 2D, decomposing the distance calculation into a base component and a distance along the scanline, sharing the latter across all scanlines. First, we again set m_z to 0 in order to calculate the distance d_{tx1}. We then set m_x to x_0, m_y to y_0, and m_z to 0—translating the coordinate system to the base of the scanline and facilitating calculation of d_{tx2}. This breakdown can be seen in Equations (9), (10).

$$d_{tx1} = (m_x - x_0)\sin\alpha\cos\beta + (m_y - y_0)\sin\alpha\sin\beta - z_0\cos\alpha \tag{9}$$

$$d_{tx2} = m_z\cos\alpha \tag{10}$$

With this decomposition, the unique delays are reduced to $M_xM_y + M_zN_xN_y$—nearly a 1024x reduction under our system parameters.

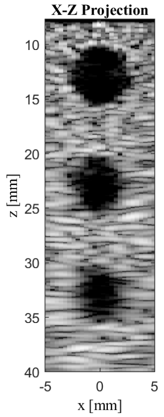
We have also implemented delay compression in separable 3D plane-wave imaging [12], [14], a technique to reduce calculations by splitting beamforming into two sequential steps—first performing beamforming along the X-axis, and then along the Y-axis. The naive implementation of separable 3D plane-wave imaging entails a computational complexity of $M_xM_zN_xN_y$ for the first stage and $M_xM_yM_zN_y$ for the second stage. Although this reduction is already significant, delay compression can further reduce the number of required calculations to only $M_xM_y + M_zN_x$ for the first stage and $M_xM_y + M_zN_y$ for the second stage.

C. Tetris *Hardware Accelerator*

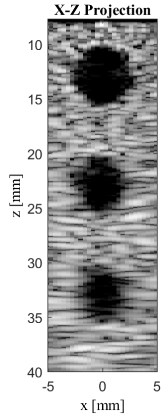
Delay compression opens new possibilities for hardware acceleration of beamforming. Previously, high-volume-rate beamforming was precluded by high bandwidth and processing requirements. In [11], we proposed Tetris, a hardware accelerator that implements delay compression to perform single-pass separable plane-wave volume creation at physics-limited rates—its volume acquisition rate is limited not by computational restrictions, but rather the propagation speed of sound in human tissue. Tetris comprises 1024 12-bit pipelines, each processing a single scanline. Using delay compression and a compact pre-computed delay representation, Tetris can perform physics-limited beamforming in a system power budget of ~2.5W.

IV. Methodology & Results

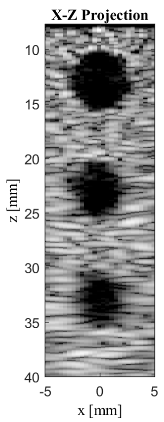
To validate that delay compression does not reduce image quality, we recreate the simulations described by Yang et al. [14] with our new delay compression method in Field II [6], [7]. The system parameters used for this simulation are given in Table I; the firing angles are $(\alpha, \beta) \in \{(0°, 0°), (3°, 0°), (6°, 0°), (3°, 90°), (6°, 90°), (3°, 180°), (6°, 180°), (3°, 270°), (6°, 270°)\}$.



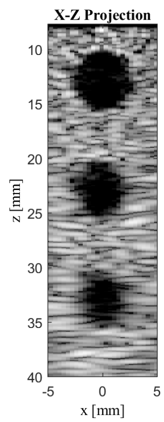
(a) Non-separable; 1,760,559,104 delay constants per angle



(b) Non-separable delay compression; 1,720,320 delays per angle



(c) Separable; 123,469,824 delay constants per angle



(d) Separable delay compression; 122,624 delay constants per angle

Fig. 5: 2D slices of simulated 3D cyst phantoms; non-separable and separable plane-wave beamforming; 9-angle coherent compounding.

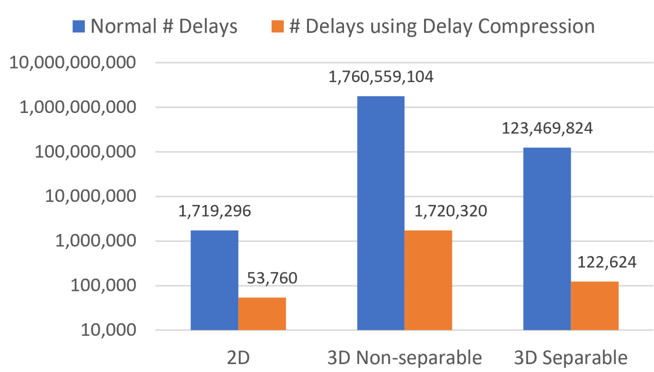


Fig. 6: Delay savings across algorithm variants.

A. Image Quality Analysis

We simulate three anechoic cysts at depths of 13mm, 23mm, and 33mm. We obtain identical images as the baseline [14] for non-separable beamforming. For separable beamforming, results differ slightly due to changes in edge-case rounding relative to the baseline [14], but image quality is unchanged. Figure 5 demonstrates image quality, while Figure 6 compares the number of unique delays required for each method. As seen in Figure 6, delay compression reduces calculation requirements by ~1024x for all methods. In terms of per-firing-angle storage or bandwidth savings (assuming 14-bit delay values), delay compression saves ~2.5MB for 2D, ~2.6GB for 3D Non-separable, and ~185MB for 3D Separable implementations under our system parameters. This drastic reduction enables on-chip storage of delay constants for 3D applications for the first time, allowing new architectures and approaches to be considered.

V. Conclusion

In this work, we proposed *delay compression*, a novel delay decomposition for plane-wave ultrasound that drastically reduces the number of delay calculations required for the 2D, 3D, and 3D Separable variants of the plane-wave algorithm. Under our system parameters, we reduce the number of unique delay calculations for non-separable 3D by nearly 1024×, and the separable 3D variant's first and second stages by 1008× and 1004×, respectively.

References

[1] R. Cobbold. *Foundations of Biomedical Ultrasound*. Biomedical Engineering Series. Oxford University Press, 2007.

[2] P. A. Hager, A. Bartolini, and L. Benini. Ekho: A 30.3w, 10k-channel fully digital integrated 3-d beamformer for medical ultrasound imaging achieving 298m focal points per second. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 24(5):1936–1949, May 2016.

[3] P. A. Hager, P. Vogel, A. Bartolini, and L. Benini. Assessing the area/power/performance tradeoffs for an integrated fully-digital, large-scale 3d-ultrasound beamformer. In *2014 IEEE Biomedical Circuits and Systems Conference (BioCAS) Proceedings*, pages 228–231, Oct 2014.

[4] A. Ibrahim, P. Hager, A. Bartolini, F. Angiolini, M. Arditi, L. Benini, and G. D. Micheli. Tackling the bottleneck of delay tables in 3d ultrasound imaging. In *2015 Design, Automation Test in Europe Conference Exhibition (DATE)*, pages 1683–1688, March 2015.

[5] A. Ibrahim, P. A. Hager, A. Bartolini, F. Angiolini, M. Arditi, J. Thiran, L. Benini, and G. D. Micheli. Efficient sample delay calculation for 2-d and 3-d ultrasound imaging. *IEEE Transactions on Biomedical Circuits and Systems*, 11(4):815–831, Aug 2017.

[6] J. A. Jensen. FIELD: A program for simulating ultrasound systems. *10th Nordicbaltic Conference on Biomedical Imaging, Vol. 4, Supplement 1, Part 1:351–353*, pages 351–353, 1996.

[7] J. A. Jensen and N. B. Svendsen. Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, 39(2):262–267, Mar. 1992.

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

[9] R. Sampson, M. Yang, S. Wei, C. Chakrabarti, and T. F. Wenisch. Sonic millip3de: A massively parallel 3d-stacked accelerator for 3d ultrasound. In *2013 IEEE 19th International Symposium on High Performance Computer Architecture (HPCA)*, pages 318–329, Feb 2013.

[10] W. Simon, A. C. Yüzügüler, A. Ibrahim, F. Angiolini, M. Arditi, J.-P. Thiran, and G. D. Micheli. Single-fpga, scalable, low-power, and high-quality 3d ultrasound beamformer. In *2016 26th International Conference on Field Programmable Logic and Applications (FPL)*, pages 1–2, Aug 2016.

[11] B. L. West, J. Zhou, R. G. Dreslinski, J. B. Fowlkes, O. Kripfgans, C. Chakrabarti, and T. F. Wenisch. Tetris: A streaming accelerator for physics-limited 3d plane-wave ultrasound imaging. In *Proceedings of the 56th Annual Design Automation Conference 2019*, DAC '19, pages 189:1–189:6, New York, NY, USA, 2019. ACM.

[12] M. Yang, R. Sampson, S. Wei, T. F. Wenisch, and C. Chakrabarti. High frame rate 3-d ultrasound imaging using separable beamforming. *J. Signal Process. Syst.*, 78(1):73–84, Jan. 2015.

[13] M. Yang, R. Sampson, S. Wei, T. F. Wenisch, and C. Chakrabarti. Separable beamforming for 3-d medical ultrasound imaging. *IEEE Transactions on Signal Processing*, 63(2):279–290, Jan 2015.

[14] M. Yang, R. Sampson, S. Wei, T. F. Wenisch, B. Fowlkes, O. Kripfgans, and C. Chakrabarti. High volume rate, high resolution 3d plane wave imaging. In *2014 IEEE International Ultrasonics Symposium*, pages 1253–1256, Sept 2014.