Local speed of sound estimation in tissue using pulse-echo ultrasound: Model-based approach

Marko Jakovljevic,[1,a)] Scott Hsieh,[2] Rehman Ali,[1] Gustavo Chau Loo Kung,[3] Dongwoon Hyun,[1] and Jeremy J. Dahl[1]
[1]*Department of Radiology, Stanford School of Medicine, Stanford, California 94305, USA*
[2]*Department of Radiology, University of California Los Angeles, Los Angeles, California 90095, USA*
[3]*Department of Engineering, Pontificia Universidad Catolica del Peru, Lima, Peru*

(Received 7 February 2018; revised 1 June 2018; accepted 5 June 2018; published online 13 July 2018)

A model and method to accurately estimate the local speed of sound in tissue from pulse-echo ultrasound data is presented. The model relates the local speeds of sound along a wave propagation path to the average speed of sound over the path, and allows one to avoid bias in the sound-speed estimates that can result from overlying layers of subcutaneous fat and muscle tissue. Herein, the average speed of sound using the approach by Anderson and Trahey is measured, and then the authors solve the proposed model for the local sound-speed via gradient descent. The sound-speed estimator was tested in a series of simulation and *ex vivo* phantom experiments using two-layer media as a simple model of abdominal tissue. The bias of the local sound-speed estimates from the bottom layers is less than 6.2 m/s, while the bias of the matched Anderson's estimates is as high as 66 m/s. The local speed-of-sound estimates have higher standard deviation than the Anderson's estimates. When the mean local estimate is computed over a 5-by-5 mm region of interest, its standard deviation is reduced to less than 7 m/s. © *2018 Acoustical Society of America*.
https://doi.org/10.1121/1.5043402

[CCC] Pages: 254–266

I. INTRODUCTION

The speed of sound (SoS) in a medium influences several important aspects of ultrasound imaging including travel-time and strength of backscattered echoes. During (B-mode) image reconstruction, echoes that are recorded on the transducer elements are delayed and summed to focus the energy coming from a desired location. In order to sum the echoes in phase, focal delays (that are meant to account for echo travel times) are determined from the geometric distances between the focus and the transducer elements, assuming a constant speed of sound. In the presence of speed of sound inhomogeneities, different parts of the wavefront travel at different speeds and the focal delays based on a constant sound-speed result in an aberrated shape of the wavefront, causing incoherent summation of echoes across the aperture. The effects of phase aberration include a loss of image resolution and contrast and an increase in off-axis scattering and image clutter.[1–5] Knowledge of the local sound speeds along a wave propagation path would allow one to more accurately predict the arrival times of backscattered signals, and reconstruct a higher quality ultrasound image.

The speed of sound as a tissue property directly impacts scattering of waves (i.e., change in direction of wave propagation), including backscattering, which is recorded on receive elements. In particular, the speed of sound and the density of a medium determine its acoustic impedance, and the differences in acoustic impedance give rise to scattering. Echoes due to a single scattering event can be traced back to the location of target (i.e., tissue structure), while the waves that undergo many scattering events take a longer time to travel to transducer surface and are mistakenly assigned to larger depths. Reverberation of waves (i.e., multipathing) can lead to acoustic noise called clutter in the final image, which is another significant source of image degradation.[6–8]

The speed of sound in tissue has also been used as a biomarker to help differentiate between healthy and diseased tissues. For example, studies in excised human and animal livers have shown a markedly lower sound speed in livers that suffer from non-alcoholic fatty liver disease (NAFLD) compared to healthy livers.[9–16] In particular, Sehgal *et al.*[9] showed an average SoS of 1528 m/s in fatty liver, compared to 1567 m/s in normal liver. Furthermore, Tervola *et al.*[12] and O'Brien *et al.*[13] showed that the speed of sound decreased by an average of 2.3 to 2.5 m/s per % fat concentration in the liver. In a similar manner, Bamber *et al.*[17,18] showed in excised human livers that, on average, the speed of sound in hepatocellular carcinoma was 1.5% lower than the sound-speed in normal liver tissue. SoS measurements that are computed from the same pulse-echo data used to reconstruct a B-mode image could thus help improve early diagnosis and staging of the liver disease.

Current sound speed estimation techniques can be grouped into two categories: average sound-speed estimators, which estimate the average speed of sound between the face of the transducer and the focal depth, and local sound-speed estimators, which estimate the speed of sound in a localized target region. The average sound-speed estimators include techniques based on the apparent shift from two angles,[19,20] beam tracking,[21] transaxial compression,[22] phase

[a)]Electronic mail: marko.jakovljevic@stanford.edu

 0001-4966/2018/144(1)/254/13/$30.00

variance,[23] and echo arrival times at the receive aperture.[24–26] These methods have low accuracy in the presence of inhomogeneities, which makes them unsuitable for *in vivo* measurement through layers of subcutaneous fat and connective tissue. For example, the average SoS estimator by Anderson and Trahey[25] yields highly accurate estimates in homogeneous media (bias less than 0.2% and standard deviation less than 0.52%); however, in a two-layer phantom composed of water and agar-graphite, measured biases in the bottom layer exceeded 30 m/s while the standard deviation was approximately 10 m/s.

Sound speed estimators that can produce localized estimates include both through-transmission and pulse-echo based approaches. Ultrasound computed tomography (UCT) uses through-transmission to yield highly accurate and localized (i.e., with high spatial resolution) SoS estimates of complex media, but it requires the target tissue to be inside a ring of transducer elements, which narrows its applications.[27–31] Local SoS estimators based on pulse-echo acquisition geometry include the crossed-beam method,[32] the modified beam tracking method,[33] registered virtual detectors,[34] and ultrasound computed tomography in echo mode (CUTE).[35] All of these methods can be implemented on conventional ultrasound arrays, except for the modified beam-tracking method, which requires two transducers. However, while these local (pulse-echo-based) SoS estimators show reduced bias in the presence of inhomogeneities compared to the gross sound-speed estimators, their error is still too large to detect SoS changes in the range of 5–10 m/s, which limits their diagnostic use. For example, in Ref. 34, the method of registered virtual detectors is shown to have the bias as large as 16 m/s for a signal-to-noise ratio of 24 dB.

Here, we propose a model and method to estimate sound speed in a localized region of tissue with high accuracy that can be implemented on a pulse-echo ultrasound system. Our model relates the average SoS between the transducer and focus to the local SoS values along the wave propagation path. In this implementation of our local SoS estimator, we first compute the average SoS from arrival-time profiles using the method by Anderson and Trahey, and then we solve the proposed model for the local sound-speed using the method of gradient descent. We validate the proposed SoS model using the ideal and noisy arrival-times that are synthesized directly from known average SoS maps. We also demonstrate the local SoS estimator using synthetic aperture ultrasound signals from fullwave simulations and phantom experiments in homogeneous and two-layer media. In all cases, the standard deviation and bias of the local SoS estimates are measured and compared to the estimate statistics obtained for the Anderson method.

II. THEORY

A. Local speed of sound model

In ultrasound imaging, it is typically assumed that the speed of sound is constant, which results in a direct (i.e., interchangeable) relationship between travel time t and distance d of the received echoes ($d = ct$). This means that uniform sampling of the ultrasound signal in time implies uniform sampling of the signal in space. To derive our model for the local SoS, we start with the assumption that the signal traces recorded on individual elements are uniformly sampled in time, while the distances traveled between the samples might vary due to differences in the local speed of sound. This assumption is consistent with how most ultrasound scanners collect the data. The total distance traveled by the wave between the focus and transducer can then be expressed as

$$d_{\text{total}} = \sum_{i=1}^{N} d_i, \tag{1}$$

where d_i is the length of tissue segment i traveled during one sampling period T_s. Given that $d_i = c_i T_s$, where c_i is the local SoS, it follows that

$$c_{\text{avg}}\, N\, T_s = \sum_{i=1}^{N} c_i T_s, \tag{2}$$

$$c_{\text{avg}} = \frac{1}{N}\sum_{i=1}^{N} c_i. \tag{3}$$

Thus, the average speed of sound equals the arithmetic mean of the local sound-speed values sampled along a wave propagation path. The starting assumption for the proposed model in Eq. (1) is also depicted in Fig. 1.

The relationship in Eq. (3) can be generalized as a linear system of equations

$$\underline{c}_{\text{avg}} = A\underline{c}_{\text{local}} + \underline{\varepsilon}_{\text{meas}}, \tag{4}$$

where each row of model matrix A encodes a single measurement of average SoS and $\underline{\varepsilon}_{\text{meas}}$ is a vector of measurement errors. For average SoS measurements collected along the same direction of wave propagation (Fig. 1), A becomes a lower triangular matrix and Eq. (4) can be solved for $\underline{c}_{\text{local}}$ via direct inversion. In other scenarios, such as when multiple directions of wave propagation are included in the model, A does not have a clear structure to exploit, so numerical solvers may be necessary to solve for $\underline{c}_{\text{local}}$.

It is interesting to note that if the ultrasound signal is sampled uniformly in space instead of in time, the expression $t_{\text{total}} = \sum_{i=1}^{N} t_i$ is used in place of Eq. (1). The alternative model for the average speed of sound is then derived as Eq. (5) instead of Eq. (3):

$$\frac{1}{c_{\text{avg}}} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{c_i}. \tag{5}$$

If slowness of a medium is defined as $\sigma = 1/c$, the slowness analog to Eq. (4) becomes

$$\underline{\sigma}_{\text{avg}} = A\underline{\sigma}_{\text{local}} + \underline{\varepsilon}_{\text{meas}}. \tag{6}$$

The alternative model in Eq. (6) however, is not typically representative of data acquired from pulse-echo ultrasound scanners, which rely on a CPU (central processing unit) clock to ensure uniform sampling in time, rather than space.[48]

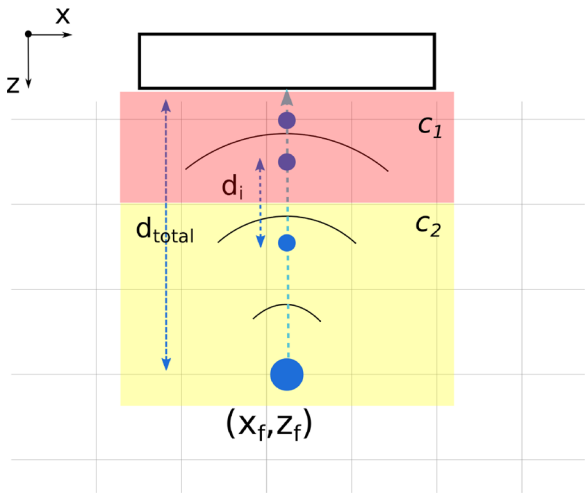

FIG. 1. (Color online) Schematics of the wave propagation from focus to the transducer surface under the assumption of uniform sampling in time. Due to differences in the local speed of sound (in this example $c_1 > c_2$), individual distances d_i traveled between the time samples are different.

III. METHODS

A. Average speed of sound measurements

In the following, the average SoS measurements have been obtained via the method by Anderson and Trahey,[25] although the proposed sound-speed model could use measurements by other average SoS estimators as well. With the Anderson method, the average SoS from the transmit focus to the transducer surface is computed using the corresponding times of arrival at the transducer elements.

Briefly, the arrival times at the individual elements are inferred from time delays between the neighboring element signals. Specifically, signals from a speckle target retain high degree of similarity among the neighboring elements and can be expressed as

$$s_{x_1}(t) = s_{x_2}(t + \delta_{x_1,x_2}) + \varepsilon_{x_1,x_2}, \tag{7}$$

where s_{x_1} and s_{x_2} are signal traces recorded at locations x_1 and x_2 in the aperture (typically not more than five elements apart), δ_{x_1,x_2} is the associated time delay, and ε is additive noise. To measure the delays, one-dimensional (1-D) normalized cross-correlation is applied to the corresponding signal traces around time t. An axial kernel of 481 samples and a search region of 541 samples were used to compute cross-correlations between the neighboring-element signals (which were up-sampled to 300 MHz). The arrival-time profile of the signals across the entire array is then formed from these time-delays; to this end, we have utilized the multi-lag least-squares method in a similar manner as outlined by Liu and Waag[36] and Dahl *et al.*[37] More detailed explanations of cross-correlation and the multi-lag least squares approach are provided in Appendix A.

For a medium with the uniform speed of sound c, the arrival times t can be also expressed in a closed-form, based on the geometric distance between the focus and a transducer element

$$t(x_e) = \frac{\sqrt{(x_e - x_f)^2 + {y_f}^2 + {z_f}^2}}{c}, \tag{8}$$

where x_f, y_f, and z_f are coordinates of the focal point, x_e is the x-coordinate of the transducer element and y_e and z_e are set to zero (i.e., the transducer runs parallel to the x axis and is centered at the origin). The expression in Eq. (8) can be rewritten as the equation for parabola

$$t^2(x_e) = p_1 x_e^2 + p_2 x_e + p_3, \tag{9}$$

with the coefficients

$$p_1 = \frac{1}{c^2}, \quad p_2 = -\frac{2x_f}{c^2}, \quad p_3 = \frac{x_f^2 + y_f^2 + z_f^2}{c^2}.$$

Therefore, the average speed of sound is estimated by fitting a parabola to the square of the arrival-time profile measured in Eq. (A4), followed by taking the square root of the inverse of the coefficient p_1.

As a variation of the technique by Anderson and Trahey, we use a multistatic synthetic transmit aperture sequence to collect the data,[38] and then utilize the principle of acoustic reciprocity to estimate the arrival-time profiles across the beamformed signals corresponding to different transmit elements as explained in Ref. 39. The resulting estimates are the same as those from the receive-element signals described previously. The use of beamformed data enables the estimation of arrival times (and average speed of sound) on a clinical scanner without significant modifications to the existing scanner architecture. In addition, the arrival-time profiles across the transmit elements are used to improve transmit focusing for subsequent iterations of the method without necessitating additional transmit events. This modified arrival-time estimation approach is described in greater detail in Ref. 40.

In the presented experiments, arrival-time profiles and the resulting average sound speed estimates are computed at the locations that are spaced half-a-wavelength apart in both axial and lateral dimensions over a 130×64 (axial x lateral) grid of points. The average SoS values are then used to solve for the local sound speeds in Eq. (4). The workflow of the whole estimation process is diagrammed in Fig. 2.

B. Solving for local sound speeds

Equation (4) was solved using gradient descent in order to demonstrate the use of a numerical solver for this type of problem. While analytic techniques, such as the pseudoinverse, can be used to solve for the local speed of sound in Eq. (4), numerical solvers can be advantageous when modeled propagation paths are more complex, causing the model matrix A to become too large to be inverted directly. Because we intend to extend our work to these more complex setups in the future, we decided to use an iterative solver.

The field-of-view (FOV) was first decoupled into distinct columns, with each column corresponding to a vertical ultrasound propagation path. Because only vertical propagation

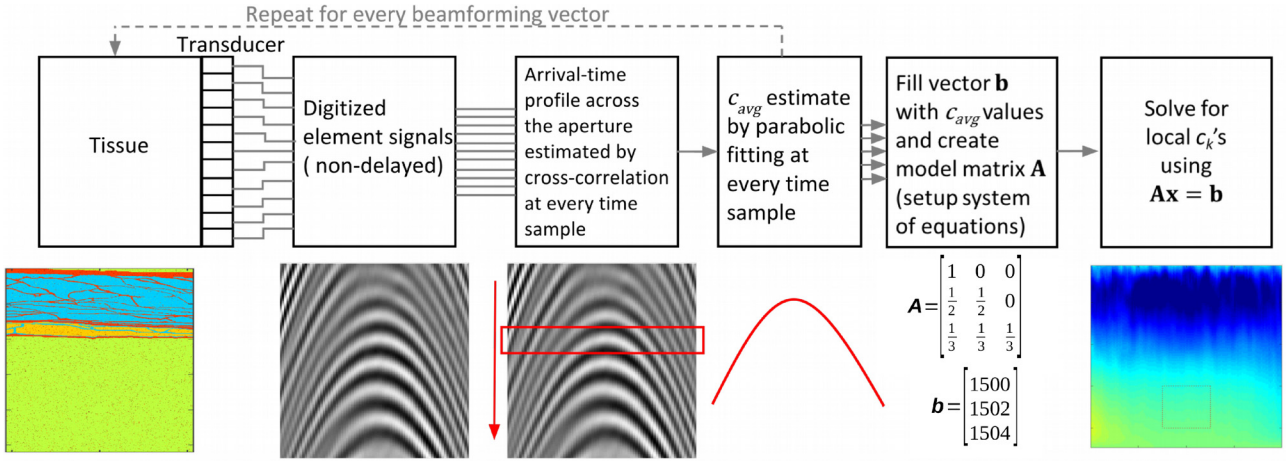

FIG. 2. (Color online) Processing flow of the local sound speed estimator. The raw element signals are captured and the average speed of sound is estimated at every point by parabolic fitting of the square of the arrival time profile. The average speed of sound values are then used to compute the local sound speed estimates from the model in Eq. (4). Note that the images are for illustration purposes and do not represent actual data.

paths were considered, there was no need to solve for the speed of sound across the entire FOV jointly. The system corresponding to each column was solved independently, via gradient descent. The regularization was a quadratic penalty minimizing the differences (in sound-speed) of nearby pixels in the least squares sense. This type of regularization was intended to suppress any high-frequency noise in the Anderson–Trahey estimates that could be caused by signal variations due to speckle. The regularization weights were derived from the output of an 11-tap low-pass filter based on Hamming-window design with a normalized cutoff frequency of 0.2. Step size was set to 0.001% of the initial guess for the sound-speed value (1540 m/s for all simulations and experiments) and the solver was run for 40 000 iterations until approximate convergence was achieved. Equations outlining the gradient descent algorithm used herein are presented in Appendix B.

C. Model validation

The models in Eqs. (9) and (4) were validated against ideal "synthetic" data obtained from known speed-of-sound maps. Two-layer SoS maps were used as a simple model of abdominal tissue; each map had a 15 mm-thick top layer with the SoS of 1480 m/s, and the bottom layer with SoS of 1520, 1540, and 1570 m/s, respectively.

The exact arrival times were computed for each point in the map (to the transducer surface) by solving the following Eikonal equation[41]:

$$\|\nabla T(x,y,z)\| = \frac{1}{c(x,y,z)}. \tag{10}$$

In Eq. (10), $c(x, y, z)$ is the local sound-speed and $T(x, y, z)$ is the shortest travel time to point (x, y, z) in space under a high-frequency approximation. The equation was solved for $T(x, y, z)$ numerically via the fast marching method.[42]

The exact arrival times were used to compute the average SoS via Anderson's method for each point in the map. The local SoS values were then inferred from the average SoS according to Eq. (4). The equation was solved for $\underline{c}_{\text{local}}$ numerically using the method of gradient descent as described in Sec. III B, except that no regularization was applied [$\lambda_2 = 0$ in Eq. (B5)] because the arrival-time measurements were exact. To assess inaccuracies in models (9) and (4) the mean and standard deviation were computed for both average and local sound-speed estimates.

To investigate error propagation in Eq. (4), the average SoS values were also generated directly from true sound speeds. The zero-mean, white, Gaussian noise with standard deviation of 5 m/s was added to mimic measurement variations due to thermal noise and speckle. The local SoS was then estimated from the synthetic average-SoS data using the method of gradient descent. Different levels of L_2-norm regularization were used in the numeric solver to reduce the noise in the final estimates. Specifically, the regularization weight λ_2 in Eq. (B5) was set to 0.0001% and 0.001% of the initial guess for c_{local} of 1540 m/s.

D. Fullwave simulations

We simulated signals received on the individual elements of a linear 1-D array from scanning homogeneous and two-layer phantoms. The signals were obtained using a fullwave simulation tool by Pinton *et al.*,[43] which models 2-D acoustic wave propagation including the effects of nonlinearity, attenuation, and multiple scattering (reverberation). In particular, a propagated medium is assigned a speed of sound, attenuation, density, and non-linearity map, which allows us to define a ground truth for speed-of-sound estimation from complex imaging targets.

The homogeneous phantoms were created with 1480, 1540, and 1600 m/s speed of sound. The two-layer phantoms each consisted of a top layer, which had 1480 m/s speed of sound, and bottom layers with speed of sound of 1520, 1540, and 1570 m/s, respectively. In order to simulate speckle, small local variations in speed of sound (average variation of 5% from the surroundings) were introduced throughout the modeled medium. The resulting point scatterers had a 40 um diameter and were randomly distributed with average

TABLE I. Simulated transducer properties.

Array width	19.2 mm
Number of elements	128
Element pitch	0.15 mm
Center frequency	5 MHz
Bandwidth at −6 dB	50%

density of twelve scatterers per resolution cell. The rest of the parameters were uniform and the same between all phantoms: the attenuation was 0.5 dB/MHz/cm, the density was 1000 kg/m^3, and the non-linearity (B/A) was 9, throughout the medium.

Having specified the acoustic properties of phantoms, a transmit event was simulated for each of 128 array elements to create synthetic transmit aperture data sets. For each transmit, the simulation code was run to numerically solve the full-wave equation (via finite-difference time-domain, i.e., FDTD method) giving the pressure field across the grid at all times. Initial conditions for solving the equation were set by prescribing the transmit waveform at each transmit element location. The transmit waveform was a two-cycle, Gaussian-weighted sinusoid at 5 MHz center frequency. Receive element signals were obtained by sampling the pressure field at the face of the transducer and convolving it with the axial transducer impulse response, which was set to yield fractional bandwidth of 0.5 at −6 dB level. The total aperture width was 1.92 cm. Specifications of the modeled array are summarized in Table I.

The receive-element signals from each transmit event were used to beamform 128 receive lines with $\lambda/2$ spacing. For each receive line, the principle of acoustic reciprocity was applied to compute the arrival times across transmit elements, which were then used to iteratively improve synthetic transmit focusing for subsequent arrival-time estimates, as explained in Sec. III A. The arrival-time profiles computed in this way after five iterations were used to estimate the average and local sound speeds following the procedure outlined in Fig. 2 and Sec. III B.

E. Phantom acquisitions

The experiments were also performed on a speckle generating phantom (model 534, ATS, Norfolk, VA) using a L12–3v linear array attached to a Verasonics Vantage 256 scanner. Multistatic synthetic aperture data were acquired with the transducer placed directly on the phantom and also through a 4 mm-thick layer of porcine tissue and through a 10 mm layer of bovine tissue. The data were acquired at a transmit center-frequency of 7 MHz. The speed of sound in the speckle phantom was calibrated to 1460 m/s by imaging a point target and maximizing its resolution and brightness. The degassed tissue samples contained skin, fat, muscle, and connective tissue and were used to mimic subcutaneous layers that induces speed of sound heterogeneity in the region of interest. The arrival time profiles were estimated across transmit elements (via the principle of acoustic reciprocity) similar as in Sec. III D. The average and local sound speeds were then estimated following the procedure outlined in Fig. 2 and Sec. III B.

IV. RESULTS

A. Model validation

An example of one-way arrival times computed via the Eikonal equation in Eq. (10) is shown in Fig. 3. In Fig. 3(a), the arrival time contours are overlayed on a two-layer speed of sound map to illustrate wavefronts originating from a focal point in the middle of FOV, at 30 mm depth. The speed of sound of the top layer is 1480 m/s and that of the bottom layer is 1520 m/s. The arrival times used to compute the average speed of sound via Anderson's method correspond to the cross-section of contours at the transducer surface. In Fig. 3(b), the arrival times at the transducer surface are plotted together with the square root of the corresponding parabolic fit in Eq. (9). For the two-layer medium used in this example, the model in Eq. (9) does not deviate significantly from true arrival-time profiles.

Average and local sound-speed estimates computed from the exact arrival times are displayed in Fig. 4 for the 1480/1520 m/s SoS map. The average (i.e., Anderson) SoS estimates gradually increase beyond 15 mm depth. The local SoS estimates show a clear separation between the two layers and are uniform within each layer. In this case, the arrival times are measured exactly and the models in Eqs. (9) and (4) yield local sound-speed estimates that are close to true SoS values. The statistics of the SoS estimates computed using the exact arrival times are presented in Table II

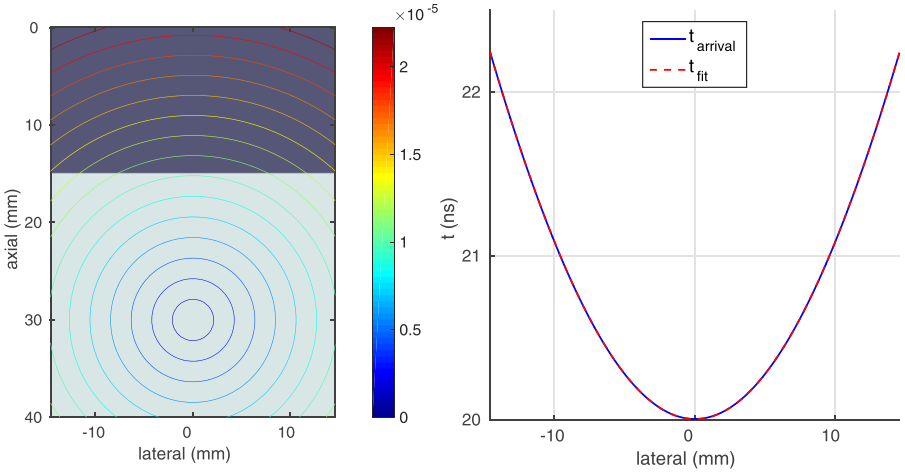

FIG. 3. (Color online) Example of one-way arrival times computed for a two-layer speed of sound map using the Eikonal equation in Eq. (10). The sound speed of the top and bottom layers is 1480 m/s and 1520 m/s, respectively. (a) Arrival-time contours for waves emanating from a focal point located in the center of FOV, at 30 mm depth. The color bar displays the color map used for contours and has units of seconds. (b) Arrival times corresponding to the cross-section of contours at the transducer surface. A second-order polynomial fit used to compute the average speed of sound [Eq. (9)] is an accurate model for the arrival-time profiles in this example.

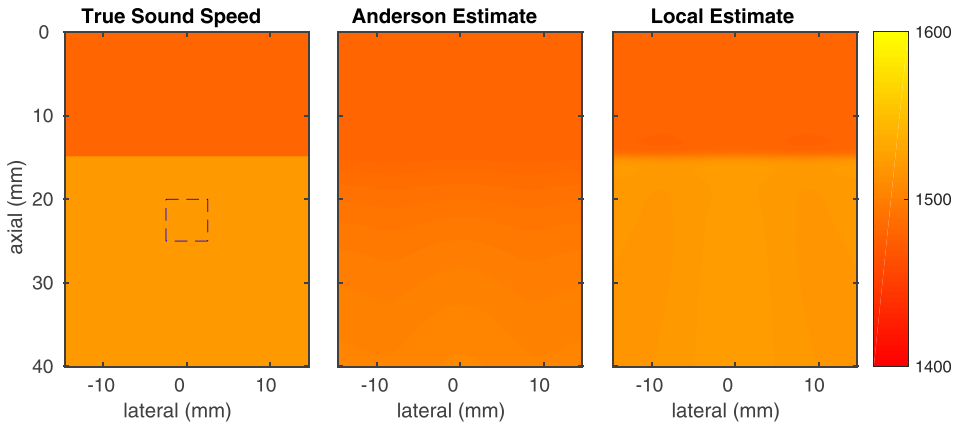

FIG. 4. (Color online) Example two-layer SoS maps used to validate the models in Eqs. (9) and (4). (Left) Ground-truth SoS map used to compute the exact one-way arrival times via the Eikonal equation. (Middle) The corresponding average SoS estimates computed from the arrival times via Anderson's method. (Right) The corresponding local SoS estimates. When the arrival times are known, the local sound-speed estimates are close to ground truth. A colorbar is used to show the dynamic range for SoS maps and has units of m/s. The region used to compute the mean and standard deviation of Anderson and local SoS estimates is denoted with a blue square. The estimate statistics for all SoS maps are summarized in Table II.

for all two-layer SoS maps. The region-of-interest (ROI) used to compute the estimate statistics is shown in the first map in Fig. 4 using a blue square. The bias of the Anderson estimates is not less than 26 m/s, and it increases as the difference in sound-speed between the two layers becomes larger. The local SoS estimates have a low bias (less than 5 m/s) and low standard deviation (0.3 m/s) that are consistent between the maps.

The plots in Fig. 5 show variations in the local SoS estimates caused by noise in the associated average SoS measurements. The average sound speeds are synthesized from true sound speed values using the forward model in Eq. (4). To facilitate interpretation of results, the SoS estimates are shown only along a single line. When there is no noise in the synthetic data, the average and local SoS estimates behave in a similar way as in Fig. 4; the average estimates increase gradually past the interface, and the local estimates are close to the true SoS values. When noise is added to the synthetic average SoS data, the local SoS estimates show increased error compared to the average SoS noise. In particular, when added noise has a standard deviation of 5 m/s, the standard deviation of the local SoS (computed between 20 and 40 mm depth) is 9.7 m/s. Increasing the weight of the l_2-norm regularization term used in the solver [λ_2 in Eq. (B5)] by a factor of 10 helps reduce the standard deviation in the local SoS estimates to 2.7 m/s, but at the expense of resolution.

B. Fullwave simulations

Examples of SoS maps estimated from the simulated ultrasound signals are shown in Figs. 6 and 7 for the homogeneous and two-layer media, respectively. For each medium, the SoS estimates are computed starting at 10 mm depth in order to omit noisy data close to the transducer surface. We report the SoS estimates averaged over a 5-by-5 mm ROI that is centered at 22.5 mm depth, and is denoted with a blue square in the first map of each figure. To obtain estimate statistics, we compute the mean and standard deviation of the ROI-averaged SoS estimates across six different scatterer distributions for each type of medium. These values are presented in Tables III and IV for the homogeneous and two-layer media, respectively. The standard deviation for an individual SoS estimate is also computed within each ROI, and is reported as "ROI SD" in both Tables.

In Fig. 6, the average (i.e., Anderson) and local SoS estimates are displayed for a simulated homogeneous medium with true sound speed of 1540 m/s. The local estimates show larger variation than the matching Anderson estimates, but on average, both are close to the true sound-speed value. These trends are confirmed by the estimate statistics in Table III. The local and Anderson estimates have comparable bias, which is less than 7 m/s for all maps. The standard deviations of the local estimates are smaller than 7 m/s, but they are larger than the standard deviations of the corresponding Anderson estimates (less than 2 m/s). The per-point (within ROI) standard deviation is also larger for the local estimates than for the Anderson estimates approximately by a factor of 2, which is similar to the trend observed in synthetic data in Fig. 5.

Examples of Anderson and local SoS estimates from simulated two-layer media are shown in Fig. 7. The corresponding ground-truth SoS maps are also shown in the left-most column, with the top layer of 1480 m/s and the bottom layers of 1520 and 1570 m/s for the first and second row of images, respectively. In both examples, the Anderson and local estimates obtained from the top layer appear close to the true sound-speed value (1480 m/s). Furthemore, in both examples, the Anderson estimates gradually increase beyond 15 mm depth resulting in a large negative bias in the bottom layers. The local SoS maps show a clear separation between the layers and the local estimates from the bottom layers are close to their respective true SoS values. The bottom layers

TABLE II. Bias and standard deviation of SoS estimates computed from the exact arrival times in two-layer media.

Sound speed in bottom layer	1520 m/s		1540 m/s		1570 m/s	
	Local	Anderson	Local	Anderson	Local	Anderson
Mean	1523.0	1493.8	1543.4	1500.3	1574.3	1509.9
Bias	3.0	−26.2	3.4	−39.7	4.3	−60.1
ROI SD	0.3	1.9	0.3	2.7	0.3	4.0

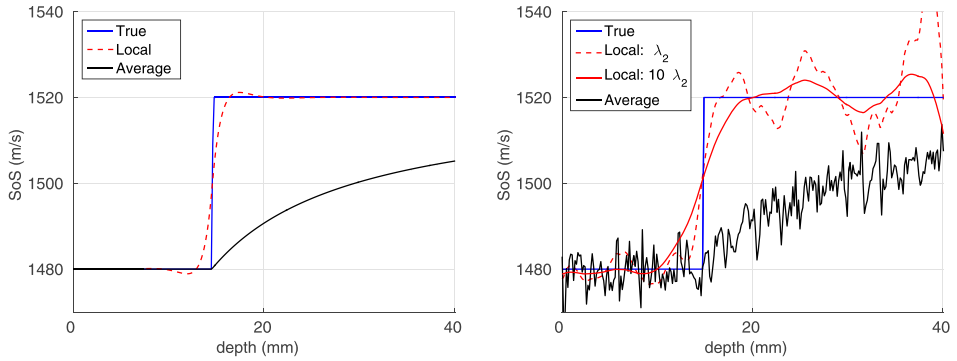

FIG. 5. (Color online) The impact of average sound speed measurement error on local sound speed estimates. When there is no noise in the synthetic data (i.e., standard deviation is $\sigma = 0$), the estimated local sound speed closely matches the true values. When Gaussian noise is added to the data ($\sigma = 5$ m/s), the gradient descent solver increases the noise in the estimates, especially at large depths. Increasing the weight of the regularization term in the solver [λ_2 in Eq. (B5)] by a factor of 10 suppresses the noise in the local SoS estimates, but blurs the boundary between the two layers. The standard deviation of the local SoS estimates at the bottom layer (between 20 and 40 mm) is computed to be $\sigma_{\lambda_2} = 9.7$ m/s and $\sigma_{10\lambda_2} = 2.7$ m/s for the two regularization weights, respectively.

in the local SoS maps appear noisier than the corresponding Anderson estimates.

The statistics of SoS estimates in Table IV support the observations in Fig. 7. The bias of Anderson estimates in the bottom layer is much larger than the bias of matching local estimates and increases with the difference in sound speed between the two layers. The smallest bias magnitude of the Anderson estimates is 33.2 m/s (computed for the 1480/1520 m/s map), while the largest bias of the local estimates is 4.7 m/s (computed for the same map). Furthermore, the standard deviations of the local estimates are larger than the standard deviations of the corresponding Anderson estimates, but none exceed 7 m/s. The mean SoS estimates computed from the simulated ultrasound signals (Table IV) are similar to the mean SoS estimates obtained from the synthetic arrival-time profiles (Table II).

C. *Ex vivo* experiments

The SoS estimates from the speckle phantom obtained with and without tissue layers in the wave propagation path

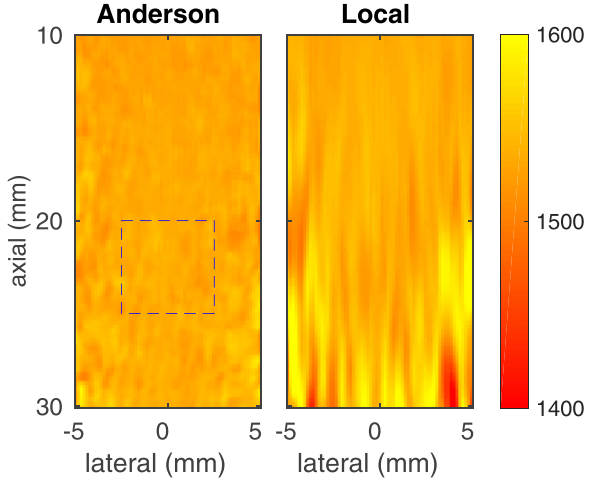

FIG. 6. (Color online) Sound speed estimates from a simulated homogeneous medium. (Left) The average SoS estimates computed via Anderson's method. (Right) The proposed local SoS estimator. The true sound speed in the medium is 1540 m/s. Color bar units are m/s. The local SoS estimates are computed starting at 10 mm depth in order to omit noisy data close to the transducer surface. The 5-by-5 mm region of interest used to compute the estimate statistics (Table III) is denoted with a blue square.

are shown in Fig. 8, and the corresponding estimate statistics is presented in Table V. In Fig. 8, the sound speed estimates are overlayed on top of the corresponding B-mode images. From left to right, sound speeds are estimated with the transducer placed directly on the phantom, with a 4 mm layer of porcine tissue, and with a 10 mm layer of bovine tissue placed between the transducer and the phantom. In the absence of tissue, both local and Anderson SoS estimates are close to the calibrated sound speed in the phantom (1460 m/s). When the ultrasound signals are acquired through the 4 mm tissue-layer, the average SoS estimates at 10 mm depth are above 1500 m/s and they gradually decrease with depth, while the matching local SoS estimates remain close to 1460 m/s. These trends are confirmed by the measured mean estimates in Table V. For the acquisition through the 10 mm-tissue layer, the Anderson measurements from 10 to 20 mm are noisy, and the corresponding B-mode image shows reverberation from the tissue-phantom interface at approximately 20 mm depth. In order to minimize the impact of reverberation on the local SoS estimates, the noisy Anderson measurements are omitted from the gradient descent solver and the local sound-speed is computed starting at 21 mm depth. The standard deviation of the local SoS estimates for this acquisition is 27.8 m/s.

The degradation of ultrasound signal from the speckle phantom due to tissue layers in Fig. 8 is measured in terms of"lag-one" correlation[44] (i.e., the normalized cross-correlation between the neighboring-element signals), the RMS error of the parabolic fit to the squared arrival-times, the RMS value of the aberrator, and signal-to-clutter ratio (SCR). These metrics are summarized in Table VI. Phase aberration profiles are measured so to exclude the gross sound-speed error; in other words, they account only for the second-order variations of the wavefronts. SCR is computed using Eq. (9) from,[45] with average lag-one correlation from the no-tissue acquisition used as true cross-correlation between the channel signals. All metrics are computed at 23 mm depth, except for the SCR, which is estimated at 5 mm below the phantom surface. For the 4 mm tissue acquisition, all metrics have similar values as for the no-tissue case, indicating low level of acoustic noise. For the acquisition through the 10 mm tissue-layer, low lag-one correlation

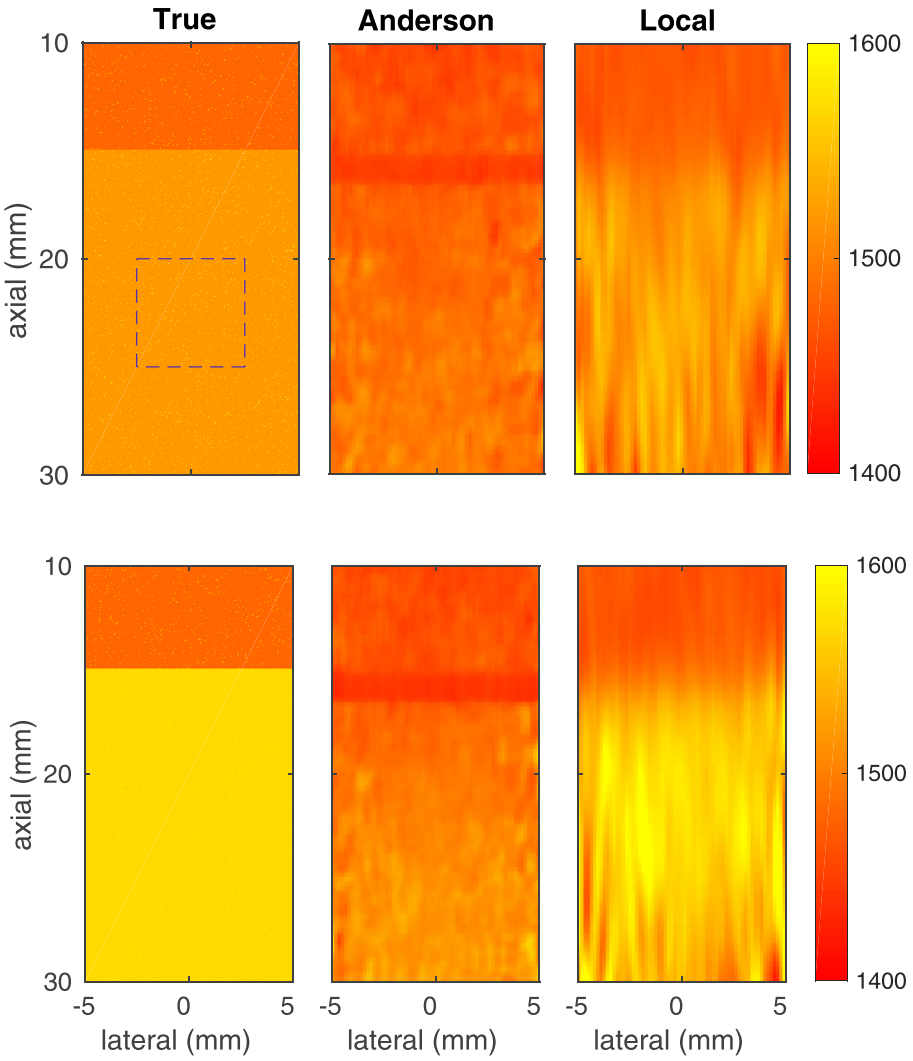

FIG. 7. (Color online) Example SoS maps of the simulated two-layer media. (Left column) Ground-truth SoS maps. (Middle column) The corresponding Anderson estimates. (Right column) The corresponding local sound-speed estimates. The true sound-speed in the bottom layer is 1520 m/s in the first example (top row) and 1570 m/s in the second example (bottom row). Color bar units are m/s. In both examples, the local SoS estimates from the bottom layer indicate lower bias than the corresponding Anderson estimates. The region of interest used to compute the estimate statistics is denoted with a blue square. The estimate statistics for all two-layer simulations is summarized in Table IV.

and low SCR indicate that ultrasound signal is corrupted by strong reverberation clutter, which is estimated to be two times higher than clutter levels from the 4 mm tissue-layer.

V. DISCUSSION

The proposed SoS estimator, as implemented in this paper is suitable for quantitative ultrasound applications in the liver, where a single-number measurement can assist with the diagnosis and staging of fatty liver disease. We envision that the method would be applied during a normal ultrasound scan, in a similar manner as a shear-wave velocity estimator typically used in the clinic. Relying on a B-mode image for guidance, the user would select a region of homogeneous tissue several centimeters underneath the liver capsule to avoid reverberation noise from the superficial layers that can impact the estimate. The method would then produce an estimate of the local speed of sound within the selected region of interest. We illustrate such approach with the sound-speed estimates from the speckle phantom in Fig. 8.

We validated the proposed SoS model and the associated local sound-speed estimator using the two-layer media as an approximation to abdominal tissue. Specifically, when average sound speeds are computed using the exact arrival times, the local sound speeds that are inferred from the model in Eq. (4) appear almost identical to true SoS maps (Fig. 4). This can be explained by the following reasons. The

TABLE III. Bias and standard deviation of estimates in simulated homogeneous media.

True sound speed	1540 m/s		1480 m/s		1600 m/s	
	Local	Anderson	Local	Anderson	Local	Anderson
Mean	1542.2	1535.9	1486.2	1475.2	1600.2	1598.4
Bias	2.2	−4.1	6.2	−4.8	0.2	−1.6
SD	4.0	0.8	6.5	1.4	4.7	1.5
ROI SD	11.1	5.0	12.8	6.3	17.6	8.7

TABLE IV. Bias and standard deviation of estimates in simulated two-layer media.

Sound speed in bottom layer	1520 m/s		1540 m/s		1570 m/s	
	Local	Anderson	Local	Anderson	Local	Anderson
mean	1524.7	1486.8	1538.7	1493.1	1571.5	1503.3
bias	4.7	−33.2	−1.3	−46.9	1.5	−66.7
SD	6.8	1.2	3.2	1.2	3.9	1.4
ROI SD	12.3	7.3	11.9	7.0	11.5	7.5

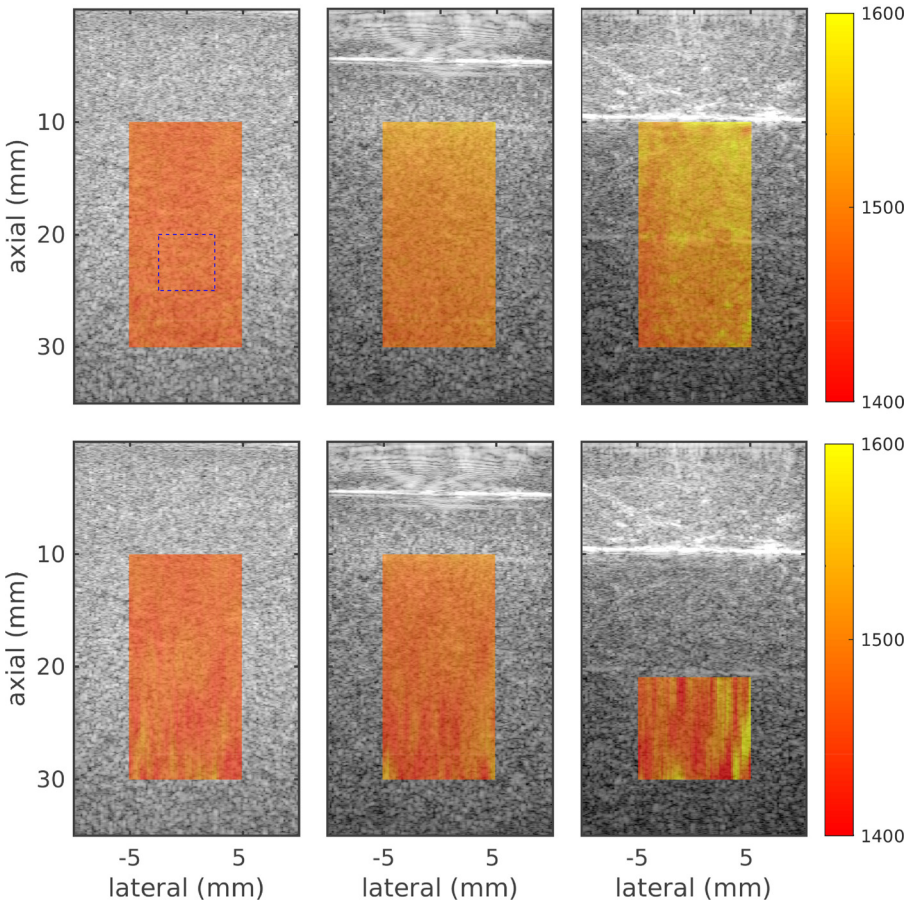

FIG. 8. (Color online) Sound speed estimates and the corresponding B-mode images of the speckle phantom, with and without tissue layers in the acoustic path. (Top row) The Anderson sound-speed estimates. (Bottom row) The proposed local SoS estimator. From left to right the data were acquired directly on the phantom, through a 4 mm-thick layer of porcine tissue, and through a 10-mm-thick layer of bovine tissue. The Anderson estimates obtained through the tissue layers are higher than the Anderson estimates for the no-tissue case, while the local SoS estimates are similar between the acquisitions. The 5-by-5 mm region of interest used to compute the estimate statistics (Table V) is denoted with a blue square. For the acquisition through the 10-mm-tissue layer, the Anderson estimates are noisy down to 21 mm depth, and the B-mode image shows reverberation from the tissue-phantom interface at approximately 20 mm depth. The local SoS estimates are therefore computed starting at 21 mm depth in order to omit the data corrupted by reverberation clutter. The ROI for this acquisition is adjusted to be centered at 25 mm depth. Color bars are shown for the speed-of-sound maps in units m/s. B-mode images are displayed using a dynamic range of 50 dB.

average SoS measurements herein are obtained along a single direction of propagation, so the model matrix A, which encodes averaging (i.e., integration) of the local SoS values is a lower triangular matrix. Such a system of equations has a unique solution. In the absence of noise, the local sound speeds can be then estimated exactly, via direct differentiation of the adjacent measurements, or numerically (e.g., via gradient descent) with a high level of precision. When echo arrival-times are computed from the speckle targets, the Anderson SoS measurements are noisy, but the estimated local sound speed still conveys the structure of insonified media without significant artifacts (Figs. 6 and 7). In particular, the local SoS maps of simulated two-layer media display a clear boundary between the layers (Fig. 7). Regularization in the gradient descent solver is used to restrict the noise level of the final local SoS estimates (Figs. 5).

The local sound-speed estimates exhibit low bias in the presence of inhomogeneities compared to the matching Anderson estimates. In the homogeneous media, the bias of local SoS estimates is not larger than 6.2 m/s and is comparable to the bias of Anderson estimates (Table III). In the two-layer media, the bias of the local SoS estimates in the bottom layers is lower by up to 60 m/s than the bias of matching Anderson estimates and is generally unaffected by the sound-speed of the top layer (Tables II and IV). This trend can be explained by Eq. (4), which models Anderson (i.e., average) sound-speed as an average of local sound-speeds along the propagation path. Therefore, the bias of Anderson estimates from the bottom layers is expected to increase with the difference in sound-speed between the two layers. The local sound-speed is computed by taking a difference of average sound-speed measurements, which is expected to annul their individual biases (caused by preceding layers in the propagation path). However, the bias of a local SoS

TABLE V. Bias and standard deviation of SoS estimates in the speckle phantom with and without tissue layers in the acoustic path.

	Homogeneous		4 mm		10 mm[a]	
Tissue layer	Local	Anderson	Local	Anderson	Local	Anderson
Mean	1459.9	1466.1	1465.8	1494.5	1463.6	1510.7
Bias	−0.1	6.1	5.8	34.5	3.6	50.7
ROI SD	13.6	5.0	11.1	6.5	27.8	9.7

[a]For this acquisition, the ultrasound signals from 0 to 21 mm depth are corrupted by clutter and are omitted from the local SoS computation.

TABLE VI. Degradation of ultrasound signal from the speckle phantom measured with and without tissue layers in the acoustic path. Metrics include lag-one normalized cross-correlation (ρ), RMS error of parabolic fit to squared arrival-times, RMS value of the aberrator, and signal-to-clutter ratio (SCR). SCR is estimated at 5 mm below the phantom surface, while other metrics are computed at 23 mm depth.

Tissue layer	Homogeneous	4 mm	10 mm
ρ mean	0.974	0.965	0.912
ρ SD	0.008	0.010	0.014
RMS error (s^2)	1.25×10^{-13}	1.57×10^{-13}	5.14×10^{-13}
RMS phase ab. (s)	4.0×10^{-9}	5.10×10^{-9}	17.0×10^{-9}
SCR (dB)	Inf	16.7	8.5

estimate is low and independent of preceding inhomogeneities only if the neighboring measurements of average SoS are accurate. If inhomogeneities along the propagation path cause strong reverberation and noise in the average SoS measurements, the resulting local SoS estimates may suffer from increased bias and standard deviation (Fig. 8).

The proposed local sound-speed estimator has a higher standard deviation compared to Anderson's average SoS estimator. Specifically, the standard deviation (SD) of the estimates (denoted as "ROI SD" in Tables III, IV, and V) increases between the average and local SoS estimates in our simulation and phantom experiments by roughly a factor of 2. As stated earlier, because the model matrix A in Eq. (4) is an integration operator, solving for the local SoS implies differentiation of average SoS measurements, which increases the noise in the estimates. If no regularization is employed and the average SoS measurements are independent, the standard deviation of the resulting local estimates can be predicted as $\sqrt{2}N_A S$, where N_A is the number of average SoS measurements sampled along a path and S is the standard deviation of an individual measurement; meaning, the proposed local estimator accumulates measurement error with depth. In practice, vertically adjacent Anderson measurements show modest correlation as their associated wave-propagation paths overlap, which causes a decrease in standard deviation of local estimates from the levels predicted above. In addition, correlations of vertically adjacent Anderson measurements cause streaks of error in the local SoS maps (see Figs. 6 and 7), where positive and negative errors show as vertical stripes due to the correlations of measurement noise.

To reduce the standard deviation of local estimates and enable their potential use in clinic as a quantitative metric, we reported the mean local estimate over a 5-by-5 mm ROI. This size of ROI was chosen empirically to achieve the necessary precision for fatty liver diagnosis and staging, while at the same time, allowing for an easy capture of homogeneous regions in the tissue (a typical ROI selected during the liver scan to measure shear wave velocity is 1–2 cm.). Indeed, the standard deviation of the local SoS averaged over the ROI (denoted as SD) is two to three times smaller than the standard deviation of individual local estimates (Tables III and IV). Under such conditions, the local estimator is sensitive to sound-speed changes of 5–10 m/s, making it a promising method to distinguish between different levels of fat concentration in the liver.[12,13] However, the extent of error-reduction achieved by ROI-averaging is specific to the acquisition setup employed and depends on factors such as the amount of beam overlap from adjacent average SoS measurements and regularization weight applied along the axial dimension. To use local sound speed for phase aberration correction, further refinements of the method are needed that would reduce the standard deviation of individual local SoS estimates and help preserve spatial resolution of SoS maps.

While the proposed local estimator produces accurate sound-speed estimates in the presence of mild noise, its performance can be compromised when the ultrasound signal is corrupted by strong reverberation clutter. In particular, when the ultrasound signal is acquired through a 4 mm tissue layer, the level of acoustic noise is low (Table VI) and the local SoS estimates from the speckle phantom show similar mean and standard deviation as when the data are collected directly on the phantom (Fig. 8 and Table V). However, in the presence of a 10 mm tissue layer, estimated signal-to-clutter ratio decreases by a factor of 2 (with respect to the levels for the 4 mm tissue-layer; Table VI) and the standard deviation of local SoS estimates is 15 m/s higher than in the no-tissue case (Table V). As the level of acoustic noise increases, the arrival-time estimates become less accurate, which in turn diminishes accuracy of Anderson and local SoS estimates. The large error in estimated arrival times for the 10 mm-tissue case is directly reflected in low lag-one correlation values and decreased quality of the parabolic fit to the squared arrival-time profiles (Table VI). To improve the method's performance in the presence of strong clutter, clutter reduction techniques that preserve the individual-element signals, such as ADMIRE[46] and spatial prediction filtering[47] may be used. In addition, a focused transmit may be employed instead of synthetic-transmit-aperture acquisition to improve the ratio of signal to electronic noise at larger depths, where data contains less reverberation clutter and can be used to compute average and local SoS.

The proposed sound-speed model and estimator face several limitations due to the physics of wave propagation in inhomogeneous media. First, the model assumes wave propagation along a straight line (see Fig. 1), not accounting for the lateral extent of the ultrasound beam, or for the effects of diffraction and refraction as the wave passes through media with different sound speeds. Second, the existing implementation of the local estimator uses the method by Anderson and Trahey to measure average SoS; this method assumes propagation through a homogeneous medium in order to model (the square of) arrival-times as a second-order polynomial [Eqs. (8) and (9)]. In two-layer media, these imperfections in the model(s) are not significant enough to compromise the accuracy of local sound-speed estimates. For example, the exact arrival times computed for the two-layer SoS maps (by solving the Eikonal equation) show little deviation from their parabolic fit that is used to find the Anderson estimates (see Fig. 3). As a result, the local sound-speeds look almost identical to the ground-truth (Fig. 4).

When the ultrasound data are acquired through more complex media, the effects of refraction, diffraction, and a loss of focus may result in average sound-speed measurements at locations that deviate from the straight path of interest (along which local sound-speeds are estimated). In addition, because the adjacent beams are likely to overlap, the corresponding average sound-speed measurements may be correlated across the FOV in both lateral and axial dimensions. Finally, geometries more complex than a two-layer medium may produce strong deviations of arrival-times from a parabola, further diminishing the accuracy of average SoS measurements. One way to mitigate these problems is by omitting the Anderson measurements from inhomogeneous layers and computing local sound speeds at larger depths, as was done for the phantom acquisition through the 10 mm-layer of bovine tissue (Fig. 8). Furthermore, the model in Eq. (4) can be modified to include multiple directions of wave propagation (including paths at non-zero angles) and to account for the correlation between the adjacent Anderson measurements. That way, local sound speeds can be estimated

jointly across the FOV using a larger number of average sound-speed measurements to further reduce noise in the estimates.

VI. CONCLUSION

We have presented a model and method to estimate sound speed in a localized region of tissue with high accuracy that can be implemented on a clinical pulse-echo ultrasound scanner. The model relates the average SoS between the transducer and focus to the local SoS values along the wave propagation path. In this implementation of our local estimator, the average SoS was measured from arrival-time profiles using the method by Anderson and Trahey, and the proposed model was then solved for the local sound-speed via gradient descent. We tested our estimator using two-layer media as a simple representation of abdominal tissue. In a series of simulation and *ex vivo* phantom experiments, the estimated local SoS maps clearly conveyed the extent of two layers and the local estimates from the bottom layers had a significantly lower bias than their matched Anderson estimates. One of the potential applications of the proposed local estimator is in human livers, where the local SoS estimates could be used to help differentiate between fatty and healthy liver tissues. To use the local speed of sound for phase aberration correction, the standard deviation of our estimator has to be reduced further, and one way to achieve this would be to include multiple directions of wave propagation in the SoS model.

ACKNOWLEDGMENTS

This work is supported by the National Institute of Biomedical Imaging and Bioengineering through Grant No. R01-EB015506.

APPENDIX A

A. Multi-lag least-squares arrival-time estimator

To compute the arrival-time profiles across the aperture, we assume that the neighboring-element signals have a high degree of similarity. This assumption is expressed in Eq. (7), which is rewritten here for convenience,

$$s_{x_1}(t) = s_{x_2}(t + \delta_{x_1,x_2}) + \varepsilon.$$

Here, s_{x_1} and s_{x_2} are signal traces recorded on elements x_1 and x_2, δ_{x_1,x_2} is the associated time delay, and ε is additive noise. The delay δ corresponds to the peak of the normalized cross-correlation function between the two signals

$$\hat{\delta}_{x_1,x_2} = \arg\max_{\tau} r(\tau), \tag{A1}$$

where

$$r(\tau) = \frac{\sum_{n=n_1}^{n_2} s_{x_1}(nT)\, s_{x_2}(nT+\tau)}{\sqrt{\sum_{n=n_1}^{n_2} s_{x_1}^2(nT) \sum_{n=n_1}^{n_2} s_{x_2}^2(nT+\tau)}}.$$

In Eq. (A1), T is the sampling period of the element signals, n is a sample number with the difference $n_2 - n_1$ determining the size of axial kernel over which cross correlation is measured, and τ is the amount of kernel shift over the search region. In this implementation, the individual element signals are up-sampled to 300 MHz, the axial kernel is 481 samples long, and τ is allowed to vary from -9 to $+9$ ns (± 30 samples). To reduce peak-hopping, the signal within the search region is tapered using a decaying exponential with the time constant of $1/(3f_0)$.

Time delays estimated via cross-correlation can also be expressed in terms of the arrival-times at individual elements,

$$d_{x_1,x_2} = \hat{\delta}_{x_1,x_2} = t_{x_1} - t_{x_2} + \varepsilon_{x_1,x_2}. \tag{A2}$$

In Eq. (A2), d_{x_1,x_2} is the estimated delay, t_{x_1} and t_{x_2} are desired arrival times, and ε_{x_1,x_2} is the associated estimation error. The "hat" symbol is used to denote the estimated value. The Eq. (A2) can be written in the matrix form for all neighboring receive-element-pairs

$$\underbrace{\begin{bmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & -1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & -1 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & \cdots & 0 & 1 & -1 \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ \vdots \\ t_{N-1} \end{bmatrix}}_{\underline{t}} + \underbrace{\begin{bmatrix} \varepsilon_{1,1} \\ \varepsilon_{1,2} \\ \varepsilon_{1,3} \\ \vdots \\ \varepsilon_{N-1,L} \end{bmatrix}}_{\underline{\varepsilon}} = \underbrace{\begin{bmatrix} d_{1,1} \\ d_{1,2} \\ d_{1,3} \\ \vdots \\ d_{N-1,L} \end{bmatrix}}_{\underline{d}}, \tag{A3}$$

where each row of the model matrix M encodes a pair of receive elements, N is the number of elements in the array, vectors $\underline{d}$ and $\underline{\varepsilon}$ contain all measured delays and the associated errors, respectively, and the vector $\underline{t}$ contains the arrival times. The vector $\underline{t}$ is finally estimated as the least-squares solution

$$\hat{\underline{t}} = (M^T M)^{-1} M^T \underline{d}. \tag{A4}$$

In Eq. (A3), M was designed to encode pairs of the neighboring elements that were no more than five elements apart on the receive aperture (i.e., the maximum lag of five). Equations (A1) and (A4) were solved to compute the arrival-time profile at each location in a 130×64 (axial $\times$ lateral) grid, spaced by half-a-wavelength in both axial and lateral dimensions.

APPENDIX B

A. Gradient descent solver for the local sound speed estimates

To compute the local speed of sound, the FOV was first decoupled into distinct columns, with each column corresponding to a vertical ultrasound propagation path. Equation (4) was then solved for each column independently by minimizing the following cost function:

$$C\{\underline{c}_{\text{local}}\} = \|A\,\underline{c}_{\text{local}} - \underline{c}_{\text{avg}}\|^2 + \lambda \|\Gamma\,\underline{c}_{\text{local}}\|^2. \tag{B1}$$

In Eq. (B1), the first quadratic term is square of the L_2-norm of the forward propagation error and the second term is regularization, with Γ derived to minimize the differences of adjacent estimates (in the least-square sense) and λ is used to set relative weighting on the regularizer. Because average SoS measurements were sampled along the same direction of wave propagation, A was posed as a lower triangular matrix,

$$A = \begin{bmatrix} 1 & & & 0 \\ \frac{1}{2} & \frac{1}{2} & & \\ \vdots & \vdots & \ddots & \\ \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \end{bmatrix},$$

where N is the number average SoS measurements.

In this implementation of the local sound speed estimator, which considers only a wave propagating normal to the transducer surface, the model matrix A is simple and direct matrix inversion could be an effective way to compute the estimates. However, when multiple directions of wave propagation are included in the estimator, A can become large, in which case direct inversion would be time-consuming or even computationally unfeasible. In such cases, numerical solvers (like the method outlined below) become necessary, because they reduce the cost function iteratively and without the use of expensive operators.

The gradient descent solver involved the following key steps:

(1) setting the initial local SoS estimates to 1540 m/s

$$\underline{\hat{c}}_{\text{local}}(0) = 1540\,\text{m/s}; \tag{B2}$$

(2) evaluating the gradient of the forward-propagation-error term

$$G(n) = 2A^T(A\,\underline{\hat{c}}_{\text{local}}(n) - \underline{c}_{\text{avg}}), \tag{B3}$$

where n refers to iteration number;

(3) evaluating the regularization term

$$R(n) = \Gamma\,\underline{\hat{c}}_{\text{local}}(n) = LPF\{\underline{\hat{c}}_{\text{local}}(n)\} - \underline{\hat{c}}_{\text{local}}(n), \tag{B4}$$

where LPF denotes a low-pass filter used to convolve the local SoS estimates at each iteration;

(4) updating the local SoS estimates

$$\underline{\hat{c}}_{\text{local}}(n+1) = \underline{\hat{c}}_{\text{local}}(n) + \lambda_1 G(n) + \lambda_2 R(n); \tag{B5}$$

(5) repeating steps (2)–(4) until the convergence is achieved.

In Eq. (B4), an 11-tap low-pass filter based on Hamming-window design was used with a normalized cutoff frequency of 0.2. In Eq. (B5), λ_1 and λ_2 can be adjusted relative to each other to adjust the smoothness of the solution; both were set to 0.001% of the initial c_{local} value of 1540 m/s. The algorithm was run for 40 000 iterations. A procedure similar to the one outlined above can be applied to solve for the local slowness under the alternative model in Eq. (6).

[1]G. E. Trahey, P. D. Freiburger, L. F. Nock, and D. C. Sullivan, "*In Vivo* measurements of ultrasonic beam distortion in the breast," Ultrason. Imag. **13**(1), 71–90 (1991).
[2]G. E. Trahey and S. W. Smith, "Properties of acoustical speckle in the presence of phase aberration part I: First order statistics," Ultrason. Imag. **10**(1), 12–28 (1988).
[3]L. M. Hinkelman, D.-L. Liu, L. A. Metlay, and R. C. Waag, "Measurements of ultrasonic pulse arrival time and energy level variations produced by propagation through abdominal wall," J. Acoust. Soc. Am. **95**(1), 530–541 (1994).
[4]L. M. Hinkelman, D.-L. Liu, R. C. Waag, Q. Zhu, and B. D. Steinberg, "Measurement and correction of ultrasonic pulse distortion produced by the human breast," J. Acoust. Soc. Am. **97**(3), 1958–1969 (1995).
[5]M. O'Donnell and S. W. Flax, "Phase aberration measurements in medical ultrasound: Human studies," Ultrason. Imag. **10**(1), 1–11 (1988).
[6]B. Byram, K. Dei, and D. Dumont, "An improved acoustic clutter model and direct *In Vivo* assessment of off-axis and multipath clutter energy in the liver," in *Ultrasonics Symposium (IUS), 2014 IEEE International* (2014), pp. 531–534.
[7]M. Lediju, M. J. Pihl, J. J. Dahl, and G. E. Trahey, "Quantitative assessment of the magnitude, impact, and spatial extent of ultrasonic clutter," Ultrason. Imag. **30**(3), 151–168 (2008).
[8]G. F. Pinton, G. E. Trahey, and J. J. Dahl, "Sources of image degradation in fundamental and harmonic ultrasound imaging: A nonlinear, fullwave, simulation study," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **58**(6), 1272–1283 (2011).
[9]C. M. Sehgal, G. M. Brown, R. C. Bahn, and J. F. Greenleaf, "Measurement and use of acoustic nonlinearity and sound speed to estimate composition of excised livers," Ultrasound Med. Biol. **12**(11), 865–874 (1986).
[10]T. Lin, J. Ophir, and G. Potter, "Correlations of sound speed with tissue constituents in normal and diffuse liver disease," Ultrason. Imag. **9**(1), 29–40 (1987).
[11]C. F. Chen, D. E. Robinson, L. S. Wilson, K. A. Griffiths, A. Manoharan, and B. D. Doust, "Clinical sound speed measurement in liver and spleen *in vivo*," Ultrason. Imag. **9**(4), 221–235 (1987).
[12]K. M. U. Tervola, M. A. Gummer, J. W. Erdman, Jr., and W. D. O'Brien, Jr., "Ultrasonic attenuation and velocity properties in rat liver as a function of fat concentration: A study at 100 MHz using a scanning laser acoustic microscope," J. Acoust. Soc. Am. **77**(1), 307–313 (1985).
[13]W. D. O'Brien, Jr., J. W. Erdman, Jr., and T. B. Hebner, "Ultrasonic propagation properties (@100 MHz) in excessively fatty rat liver," J. Acoust. Soc. Am. **83**(3), 1159–1166 (1988).
[14]H. Hachiya, S. Ohtsuki, and M. Tanaka, "Relationship between speed of sound in and density of normal and diseased rat livers," Jpn. J. Appl. Phys. **33**(Part 1, No. 5B), 3130–3133 (1994).
[15]T. Matsuhashi, N. Yamada, H. Shinzawa, and T. Takahashi, "An evaluation of hepatic ultrasound speed in injury models in rats: Correlation with tissue constituents," J. Ultrasound Med. **15**(8), 563–570 (1996).
[16]G. Ghoshal, R. J. Lavarello, J. P. Kemmerer, R. J. Miller, and M. L. Oelze, "*Ex vivo* study of quantitative ultrasound parameters in fatty rabbit livers," Ultrasound Med. Biol. **38**(12), 2238–2248 (2012).
[17]J. Bamber and C. Hill, "Acoustic properties of normal and cancerous human liver—I. dependence on pathological condition," Ultrasound Med. Biol. **7**(2), 121–133 (1981).
[18]J. C. Bamber, C. R. Hill, and J. A. King, "Acoustic properties of normal and cancerous human liver–II: Dependence on tissue structure," Ultrasound Med. Biol. **7**(2), 135–144 (1981).
[19]D. E. Robinson, F. Chen, and L. S. Wilson, "Measurement of velocity of propagation from ultrasonic pulse-echo data," Ultrasound Med. Biol. **8**(4), 413–420 (1982).
[20]J. F. Krücker, J. B. Fowlkes, and P. L. Carson, "Sound speed estimation using automatic ultrasound image registration," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **51**(9), 1095–1106 (2004).
[21]J. Ophir, "Estimation of the speed of ultrasound propagation in biological tissues: A beam-tracking method," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **33**(4), 359–368 (1986).
[22]J. Ophir, T. Moriya, and Y. Yazdi, "A single transducer transaxial compression technique for the estimation of sound speed in biological tissues," Ultrason. Imag. **13**(3), 269–279 (1991).
[23]C. Yoon, Y. Lee, J. H. Chang, T.-k. Song, and Y. Yoo, "*In vitro* estimation of mean sound speed based on minimum average phase variance in medical ultrasound imaging," Ultrasonics **51**(7), 795–802 (2011).

[24]N. Hayashi, N. Tamaki, M. Senda, K. Yamamoto, Y. Yonekura, K. Torizuka, T. Ogawa, K. Katakura, C. Umemura, and M. Kodama, "A new method of measuring *in vivo* sound speed in the reflection mode," J. Clin. Ultrasound **16**(2), 87–93 (1988).
[25]M. E. Anderson and G. E. Trahey, "The direct estimation of sound speed using pulse-echo ultrasound," J. Acoust. Soc. Am. **104**(5), 3099–3106 (1998).
[26]F. R. Pereira, J. C. Machado, and W. C. A. Pereira, "Ultrasonic wave speed measurement using the time-delay profile of rf-backscattered signals: Simulation and experimental results," J. Acoust. Soc. Am. **111**(3), 1445–1453 (2002).
[27]C. Li, N. Duric, P. Littrup, and L. Huang, "*In vivo* breast sound-speed imaging with ultrasound tomography," Ultrasound Med. Biol. **35**(10), 1615–1628 (2009).
[28]A. Hormati, I. Jovanovic, O. Roy, and M. Vetterli, "Robust ultrasound travel-time tomography using the bent ray model," in *Proceedings of the SPIE Medical Imaging*, LCAV-CONF-2010-001 (SPIE-Int. Soc. Optical Engineering, Bellingham, WA, 2010).
[29]T. P. Matthews, K. Wang, C. Li, N. Duric, and M. A. Anastasio, "Regularized dual averaging image reconstruction for full-wave ultrasound computed tomography," IEEE Trans. Ultrason., Ferroelectr., Freq. Control **64**(5), 811–825 (2017).
[30]Z. Zhang, L. Huang, and Y. Lin, "Efficient implementation of ultrasound waveform tomography using source encoding," Proc. SPIE **8320**, 832003 (2012).
[31]G. Sandhu, C. Li, O. Roy, S. Schmidt, and N. Duric, "Frequency domain ultrasound waveform tomography: Breast imaging using a ring transducer," Phys. Med. Biol. **60**(14), 5381 (2015).
[32]M. Kondo, K. Takamizawa, M. Hirama, K. Okazaki, and K. Iinuma, "An evaluation of an *in vivo* local sound speed estimation technique by the crossed beam method," Ultrasound Med. Biol. **16**(1), 65–72 (1990).
[33]I. Céspedes, J. Ophir, and Y. Huang, "On the feasibility of pulse-echo speed of sound estimation in small regions: Simulation studies," Ultrasound Med. Biol. **18**(3), 283–291 (1992).
[34]B. C. Byram, G. E. Trahey, and J. A. Jensen, "A method for direct localized sound speed estimates using registered virtual detectors," Ultrason. Imag. **34**(3), 159–180 (2012).
[35]M. Jaeger, G. Held, S. Peeters, S. Preisser, M. Grünig, and M. Frenz, "Computed ultrasound tomography in echo mode for imaging speed of sound using pulse-echo sonography: Proof of principle," Ultrasound Med. Biol. **41**(1), 235–250 (2015).
[36]D.-L. Liu and R. C. Waag, "Time-shift compensation of ultrasonic pulse focus degradation using least-mean-square error estimates of arrival time," J. Acoust. Soc. Am. **95**(1), 542–555 (1994).
[37]J. J. Dahl, D. A. Guenther, and G. E. Trahey, "Adaptive imaging and spatial compounding in the presence of aberration," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **52**(7), 1131–1144 (2005).
[38]J. T. Ylitalo and E. Helmut, "Ultrasound synthetic aperture imaging: Monostatic approach," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **41**(3), 333–339 (1994).
[39]N. Bottenus and K. F. Üstüner, "Acoustic reciprocity of spatial coherence in ultrasound imaging," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **62**(5), 852–861 (2015).
[40]G. Chau, M. Jakovljevic, R. Lavarello, and J. Dahl, "Robust minimum variance beamformer using locally adaptive phase aberration correction," in *Ultrasonics Symposium (IUS), 2016 IEEE International* (2016), pp. 1–4.
[41]J. W. Goodman, *Introduction to Fourier Optics*, 2nd ed. (McGraw-Hill, New York, 1996).
[42]J. A. Sethian, "A fast marching level set method for monotonically advancing fronts," Proc. Natl. Acad. Sci. **93**(4), 1591–1595 (1996).
[43]G. F. Pinton, J. J. Dahl, S. Rosenzweig, and G. E. Trahey, "A heterogeneous nonlinear attenuating full-wave model of ultrasound," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **56**(3), 474–488 (2009).
[44]W. Long and G. E. Trahey, "Spatial coherence as a predictor of image quality," in *Ultrasonic Imaging and Tissue Characterization Symposium*, Arlington, VA (2017).
[45]D. Hyun, L. Abou-Elkacem, V. A. Perez, S. M. Chowdhury, J. K. Willmann, and J. J. Dahl, "Improved sensitivity in ultrasound molecular imaging with coherence-based beamforming," IEEE Trans. Med. Imag. **37**, 241–250 (2018).
[46]B. Byram, K. Dei, J. Tierney, and D. Dumont, "A model and regularization scheme for ultrasonic beamforming clutter reduction," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **62**(11), 1913–1927 (2015).
[47]J. Shin and L. Huang, "Spatial prediction filtering of acoustic clutter and random noise in medical ultrasound imaging," IEEE Trans. Med. Imag. **36**(2), 396–406 (2017).
[48]In fact, most hardware beamformers assume constant speed of sound when converting between pulse-echo travel time and distance. When speed of sound in a medium is not constant and the true travel distance between the samples varies, the beamformed signals might not correspond to the uniformly spaced axial positions that are displayed on the scanner screen.