Speed-of-Sound Imaging Based on Reflector Delineation

Journal Article

Author(s): Sanabria, Sergio J.; Rominger, Marga B.; Goksel, Orcun

Publication date: 2019-07

Permanent link: https://doi.org/10.3929/ethz-b-000310433

Rights / license: In Copyright - Non-Commercial Use Permitted

Originally published in: IEEE Transactions on Biomedical Engineering 66(7), https://doi.org/10.1109/tbme.2018.2881302
Abstract—Objective: Speed-of-sound (SoS) has large potential for tissue and pathology differentiation. We aim to develop a novel Ultrasound Computed Tomography (USCT) technique that can reconstruct local SoS in tissue on conventional ultrasound machines with hand-held linear arrays. 

Methods: A passive reflector is placed opposite the tissue sample as an echogenic reference to measure the time-of-flight (ToF) of ultrasound wavefronts. A Dynamic Programming algorithm provides a robust ToF measurements based on global optimization of all transmit-receive echo data. An Anisotropically-Weighted Total Variation (AWTV) algorithm allows sharp delineation of focal lesions based on limited-angle USCT data.

Results: Inclusions, which are not visible in conventional ultrasound, could be delineated in SoS images. AWTV allows to reconstruct focal lesions with a contrast-ratio of 93.7% of their nominal value, compared to that of 31.5% with conventional least-squares based algebraic tomographic reconstruction. In full-wave simulations of realistic heterogeneous breast models, a high CR of 84.3% is observed, with the reconstruction filtering out background heterogeneity. In experiments, our proposed method quantifies SoS in a homogeneous background with an accuracy of 0.93 m/s, allowing to differentiate several tissue types.

Conclusion: We validate our method using numerical simulations with ray-tracing and full-wave models, and phantom and ex-vivo data. Preliminary in-vivo results show the potential of this new technique to detect and differentiate malignant and benign lesions in the breast.

Significance: Breast cancer is the most common cancer in women. Ultrasound B-mode only provides qualitative information about breast lesions, whereas USCT can provide quantitative tissue imaging biomarkers, such as SoS. The proposed method can potentially be implemented as a complementary modality to ultrasound for tissue and disease differentiation.

Index Terms—Ultrasound, sound-speed, limited-angle computed tomography, inverse problem, image reconstruction, total variation

I. INTRODUCTION

Breast cancer is the most common cancer in women. Worldwide 1.7 million new cases are diagnosed every year, leading to over half a million deaths [1]. The earlier the breast cancer is detected, the higher the chance for survival is and the less invasive the treatment can be. The conventional imaging option for early detection is X-ray mammography, subsequent to which suspicious lesions are biopsied as the histopathology being the gold-standard for diagnosis. Due to its ionizing radiation, X-ray mammography is prescribed for biannual screening only of women at high risk, over 50 years old, and >20% cancers occur in younger women [2]. Mammography can detect only 80% of cancerous tumors (and even less at 60% in dense breasts). For every cancer death avoided, 100 false alarms are raised in healthy women, and 5 unnecessary breast surgeries are performed [3]. Mammography requires significant breast compression, which translates into patient discomfort. Other techniques, such as magnetic-resonance (MR) imaging and positron emission tomography (PET), may provide a better sensitivity, however they are not widely accessible, also suffer from false positives and involve other complications, such as injection of contrast agent and long measurement sequences [4], [5].

Ultrasound (US) is a radiation-free, non-invasive, real-time and cost-effective imaging modality. Conventional B-mode ultrasound detects invasive cancers, which are not visible in X-ray mammography [6]. However, US shows low specificity for differentiating malignant and benign lesions, and it is not a quantitative modality, hence the performance depending on operator experience [7]. US is currently used to perform diagnostic imaging of suspicious lumps based on their geometric features. However, because of the elevated number of false positives US is not used as a screening modality [8]. Elastography methods have been developed, which aim at quantitative measurement of elastic properties by tracking displacements in US images as a response to tissue mechanical excitation [9]–[12]. Particularly, shear-wave elastography (SWE) provides spatially-resolved images of shear modulus based on shear waves generated deep inside the tissue remotely using acoustic radiation force from a focused ultrasound burst. Generated waves travel at low speeds (1-30 m/s), which can be tracked and measured from displacements observed in tissue. US imaging, which utilizes longitudinal waves traveling at speed-of-sound (SoS) of ~1540 m/s, thus can observe shear-wave propagation in several consecutive frames [13]. However, the specificity of elastography for differential diagnosis is limited [14], [15].

Ultrasound Computed Tomography (USCT) measures the acoustic properties of longitudinal waves themselves as quantitative imaging biomarkers for tissue malignancy – for instance, the speed-of-sound (SoS) in tissue. These systems have shown high potential for both breast tumor screening and differential diagnosis. Similarly to X-ray CT, USCT typically requires a transmission imaging setup to reference the received wave.
with respect to the source, together with lesion insonification from all tomographic projection directions, e.g. with circular or rotating transducer setups, for a well-determined tomographic reconstruction. This leads to a complicated setup and acquisition workflow, hindering the clinical translation of USCT [16–19].

There have been several attempts for tomographic US reconstruction based on conventional hand-held US, which is widely available in clinical practice. Typical US hardware consists of a hand-held array probe, which operates in pulse-echo mode, both transmitting US pulses into tissue as well as receiving the reflected echo signals. These echo signal traces are then beamformed to generate an image that is interpreted by the sonographer for navigation and diagnosis [20]. A USCT image promises great potential for providing additional and complementary diagnostic information. In one approach, a known-geometry surface could be used to reflect US as a reference for USCT measurements. Some prior works proposed to use the X-ray mammography plate as a reference [21]–[25] or multiple (opposing) linear transducers in transmission mode [25], [26] for USCT. Similarly to Digital Breast Tomosynthesis [27], the reconstruction is a limited-angle computed tomography problem and inherently ill-posed, such that conventional linear reconstruction algorithms based on least-squares error minimization, e.g. [21], [26], often fail to reconstruct lesions. Methods to alleviate this have been based on prior information or iterative thresholding [22]–[25]. However, these are only applicable for quantifying well-defined inclusions with known geometry, but not for imaging an unknown tissue domain. The use of a large size reflector such as the mammography plate limit flexibility in access to certain breast regions, for instance the tissue (hence lesions) close to the chest wall. Furthermore, with a larger reflector, wave path uncertainty (e.g., potential out-of-plane reflections) increase, complicating reflector delineation. Reconstructing SoS from local phase aberrations was proposed from tracked apparent tissue displacements from multiple observation directions [28], [29], the reconstruction performance of which however would depend on the existence of fully-developed speckle, tracking algorithm, and speckle correlation from different directions.

In this work, we introduce a novel hand-held setup and image reconstruction method, which overcomes the above limitations and allows for accurate lesion delineation and quantitative USCT reconstruction. Main physical components of our proposed setup in Fig. 1 include:

1) a conventional US system with a hand-held US probe,
2) a long, narrow reflector located on the opposite side of the breast as a timing reference,
3) an automatic reflector delineation algorithm,
4) a SoS reconstruction algorithm based on a novel anisotropically-weighted total-variation (AWTV) regularization approach to explicitly incorporate the available ultrasonic travel path information into the limited-angle reconstruction problem.

Preliminary results of this work were presented in [30]. Herein we (i) introduce extensions and improvements to the earlier reconstruction method, (ii) analyze imaging resolution and accuracy with a comprehensive in-silico noise study as well as full-wave ultrasound simulations, (iii) demonstrate results on a broad range of phantoms and ex-vivo tissue samples, and (iv) present preliminary in-vivo results from healthy and cancerous breast tissue.

II. REFLECTOR-BASED SOS RECONSTRUCTION

Our reconstruction setup is illustrated in Fig. 2(a). A conventional B-mode ultrasound probe consists of an array of piezoelectric elements, each of which can individually transmit and receive ultrasound pulses. Typical probe geometries for two-dimensional US imaging are linear arrays (with elements arranged in a straight line) and convex arrays (with elements arranged along a curvilinear surface) [20]. For the sake of simplicity and without loss of generality, we focus on a linear array probe geometry. We employ full-matrix (also called multi-static or synthetic aperture) data acquisition. This operates by firing an ultrasound pulse using a single transducer element, followed by the simultaneous recording of echos by all available elements/channels; which is then repeated for a separate firing of each and every element sequentially. Hence, for an $N$-element transducer, a total of $P = N^2$ transmit-receive pairs are recorded [31]. Note that, assuming the linear regime in ultrasound interactions, signal-to-noise ratio of such multi-static recordings can be improved in practice by firing more elements at a time and linearly combining results later on. For instance, we apply Wals-Hadamard coding in our experimental setting [32].
In order to utilize a single transducer for the tomographic data (instead of two separate ones each for transmit and receive, which is commonly not readily accessible on standard US hardware), we introduce a reflector (reference surface) on the opposite side of the tissue to be imaged. Accordingly, a pulse transmitted by the transmitter Tx is coherently reflected by the reflector and propagated back to each receiver Rx. By accurately identifying the reflector echo in the received temporal US signal, the time-of-flight (ToF) for each Tx-Rx path is estimated and collocated in a vector \( \tau \in \mathbb{R}^P \). Each element of \( \tau \) is the integral SoS along a given wave propagation path, and can be split into a geometric component \( \tau_g \), given by a reference ToF between Tx and Rx for a nominal SoS \( \tau \), and a smaller perturbation \( \Delta \tau \), the latter of which contains the information about SoS heterogeneities. With SoS variations in soft tissues being small (<20%), we assume that a major portion of the wave follows a straight path (with minimal refractions), and accordingly approximate the wavefront propagation paths \( p \) as straight lines (rays), following Fermats shortest path principle. Without loss of generality, we focus on a flat reflecting surface, for which closed expressions from \( \tau_g \) can be calculated from trigonometric relations:

\[
\tau = \tau_g + \Delta \tau
\]

\[
\tau_g = \tau_g(Tx, Rx) = (\tau)^{-1} \left[ 4d^2 + (\xi_{Rx} - \xi_{Tx})^2 + 4\sin^2 \theta(\xi_{Tx}\xi_{Rx} - d^2) - 2\sin(2\theta)d(\xi_{Rx} + \xi_{Tx}) \right]^{0.5} \tag{1}
\]

where \( d \) is the reflector-probe distance, \( \theta \) the in-plane reflector inclination, and \( \xi_{Rx} \) and \( \xi_{Tx} \) are the lateral position of the Tx and Rx elements, respectively, as in Fig.2(a). Assuming that the delay perturbation vector \( \Delta \tau \) is zero-mean and negligible with respect to \( \tau_g \), we estimate the reflector location \((d, \theta)\) and average background SoS \( \tau \) by approximating \( \tau \approx \tau_g \) and fitting Eq.(1) to the measured time matrix \( \tau \). For numerical phantoms with various inclusion geometries in Fig.3(a), the simulated ToF data \( \Delta \tau \) can be seen in Fig.3(b) as examples. Having identified \((d, \theta)\), the only unknown in Eq.(1) is then the local speed-of-sound left to be identified by image reconstruction.

For a numerical solution, the field-of-view between the transducer and the reflector is discretized into a number \( C \) of rectilinear cells, for which the SoS deviations at each cell, expressed in slowness units \( \sigma [s/m] \), are related to the integral ToF deviations \( \Delta \tau \) as follows:

\[
\Delta \tau = \tau - \tau_g = \tau - LI_{C \times 1} \tau = L \sigma
\]

where \( L \) is a geometric matrix containing the differential lengths \( l_{p,c} [m] \) of path \( p \) within cell \( c \), which is defined by the probe and reflector geometries as in Eq.(1). For \( P \geq C \), Eq.(1) is an over-determined linear system. This can be satisfied using square cells with the size of transducer element pitch (approximately the imaging wavelength \( \lambda \)) when the reflector depth \( d \) is less than or equal to the probe width \( H \); i.e. when the aspect ratio \( d/H \leq 1 \), and larger cells for \( d/H > 1 \). Eq.(2) is a computed tomography (CT) reconstruction problem, similarly to X-ray CT [33], [34]. A popular algorithm is Filtered Back-Projection (FBP), which would “smear” ToF onto cells \( c \) along the corresponding propagation path. An alternative is Algebraic Reconstruction (ART), in which the linear system Eq.(2) is solved as an inverse problem. For a direct solution by factorization, the matrix inversion can be stabilized by preserving only the largest singular values of \( L \). In our tomographic reconstruction problem, ToF information is only available nearly axial (vertical) paths, up to a maximum available tomographic SoS projection angle \( \pm \phi_{max} \). For instance, for an aspect ratio \( d/H = 1 \), the angle \( \phi_{max} = 27^\circ \) at certain image locations and even further lower near image edges. This is then an ill-posed limited-angle tomographic reconstruction problem, similar to cone-beam CT reconstruction [27], [35]. Therefore, successful imaging using typical CT reconstruction algorithms would not be possible due to artifacts and low resolution, especially in the axial imaging direction lacking transversal projections.

Typical artifacts are exemplified for FBP in Fig. 3(c).

A. Anisotropically-weighted total variation (AWTV)

To overcome limited-angle artifacts, we introduce the knowledge from the imaging setup as regularization on the SoS reconstruction. Eq.(2) is then reformulated as an optimization problem:

\[
\sigma = \arg \min_{\sigma} \left\{ \| \Delta \tau - L \sigma \|^n_n + \lambda \| D \sigma \|^n_n \right\} \tag{3}
\]

where \( D \) is a geometric matrix containing the differential directions \( d_{ij} [m] \) of path \( d \) within cell \( i,j \), which is defined by the probe and reflector geometries as in Eq.(1). Having identified \((d, \theta)\), the only unknown in Eq.(1) is then the local speed-of-sound left to be identified by image reconstruction.

For a numerical solution, the field-of-view between the transducer and the reflector is discretized into a number \( C \) of rectilinear cells, for which the SoS deviations at each cell, expressed in slowness units \( \sigma [s/m] \), are related to the integral ToF deviations \( \Delta \tau \) as follows:

\[
\Delta \tau = \tau - \tau_g = \tau - LI_{C \times 1} \tau = L \sigma
\]

where \( L \) is a geometric matrix containing the differential lengths \( l_{p,c} [m] \) of path \( p \) within cell \( c \), which is defined by the probe and reflector geometries as in Eq.(1). For \( P \geq C \), Eq.(1) is an over-determined linear system. This can be satisfied using square cells with the size of transducer element pitch (approximately the imaging wavelength \( \lambda \)) when the reflector depth \( d \) is less than or equal to the probe width \( H \); i.e. when the aspect ratio \( d/H \leq 1 \), and larger cells for \( d/H > 1 \). Eq.(2) is a computed tomography (CT) reconstruction problem, similarly to X-ray CT [33], [34]. A popular algorithm is Filtered Back-Projection (FBP), which would “smear” ToF onto cells \( c \) along the corresponding propagation path. An alternative is Algebraic Reconstruction (ART), in which the linear system Eq.(2) is solved as an inverse problem. For a direct solution by factorization, the matrix inversion can be stabilized by preserving only the largest singular values of \( L \). In our tomographic reconstruction problem, ToF information is only available nearly axial (vertical) paths, up to a maximum available tomographic SoS projection angle \( \pm \phi_{max} \). For instance, for an aspect ratio \( d/H = 1 \), the angle \( \phi_{max} = 27^\circ \) at certain image locations and even further lower near image edges. This is then an ill-posed limited-angle tomographic reconstruction problem, similar to cone-beam CT reconstruction [27], [35]. Therefore, successful imaging using typical CT reconstruction algorithms would not be possible due to artifacts and low resolution, especially in the axial imaging direction lacking transversal projections.

Typical artifacts are exemplified for FBP in Fig. 3(c).

A. Anisotropically-weighted total variation (AWTV)

To overcome limited-angle artifacts, we introduce the knowledge from the imaging setup as regularization on the SoS reconstruction. Eq.(2) is then reformulated as an optimization problem:

\[
\sigma = \arg \min_{\sigma} \left\{ \| \Delta \tau - L \sigma \|^n_n + \lambda \| D \sigma \|^n_n \right\} \tag{3}
\]

Although one can choose the fitting and regularization norms independently, these terms often scale better using equal norms \( n \), which we also observed in our preliminary experiments and therefore considered herein. Having solved this optimization problem in Eq.3, the quantitative SoS values are then calculated from the slowness values \( \sigma \) and the average background SoS \( \tau \) as: SoS = \((\tau^{-1} + \sigma)^{-1} \). \| D \sigma \|^n_n regularizes the gradients in the reconstructed image, and \( \lambda \) is the regularization constant.

ART typically uses \( n = 2 \) (i.e. \( \ell_2 \)-norm, Tikhonov regularization), for which different methods exist to solve Eq. 3. This is robust to zero-mean noise, but since smooth gradients are favored, no resolution improvement is observed with respect to FBP. With \( n = 1 \) (\( \ell_1 \)-norm, total-variation, TV), sharp and smooth gradients are equally weighted and sparsity is enforced [36] as illustrated in Fig. 2(b).

In our reconstruction problem, \( \ell_1 \)-norm is observed to yield improved image quality compared to \( \ell_2 \)-norm [37], especially given our noisy simulations and real data. From the possible implementations of \( \| D \sigma \|^n_n \), we herein use axis-aligned finite-differencing as \( \| D \sigma \|^n_n = \sum_{i,j} | \sigma_{i+1,j} - \sigma_{i,j} | + | \sigma_{i,j+1} - \sigma_{i,j} | \). In our hand-held USCT problem, this successfully prevents streaking artifacts and helps better delineate inclusion geometries, cf. Fig. 3(d). However, the limited-angle acquisition still yields poor axial resolution.

We extend the smoothness regularization to explicitly incorporate the available angular information for each ray path illustrated in Fig.2(c). Since wave propagation paths are defined by the reflector setup, angular tomographic orientations available for each cell are known a priori. We therefore weight the gradients in different angular directions according to the available ray information in each of these directions. We refer to this as anisotropically-weighted total variation (AWTV). A
basic form (axis-aligned AWTV), introduces a constant $\kappa$ to balance horizontal and vertical gradients as:

$$\|D\sigma\|_{\text{AWTV}} = \sum_{i,j} |\sigma_{i+1,j} - \sigma_{i,j}| + (1 - \kappa)|\sigma_{i,j+1} - \sigma_{i,j}|$$

(4)

A single value $\kappa = 0.9$ for all reconstruction cells already results in improvement for a variety of inclusion geometries as seen in Fig. 3(e).

This is further extended using multiple gradient directions (multi-angle AWTV or MA-AWTV) as follows:

$$\|D\sigma\|_{\text{MA-AWTV}} = \sum_a \sum_{c,a} \kappa_{c,a} \|D_a\sigma\|$$

(5)

where $\|D_a\sigma\| = D\sigma \cdot e_a$ is the directional derivative along the unit vector $e_a$ with inclination angle $\alpha$. Given the maximum available angle $\phi_{\text{max}} = \arctan(0.5H/d)$ in our hand-held SoS imaging setup, the angular range $[-\phi_{\text{max}}, \phi_{\text{max}}]$ is divided linearly into a set of $N_\alpha$ gradient directions $\{\alpha\}$. The weights $\kappa_{c,a}$ at each cell $c$ are calculated with the following algorithm:

1) Initialize all weights $\{\kappa_{c,a}\} = \{0\}$

2) For each wave path $p$, for each cell $c$ it crosses

a) calculate the inclination of wave path $\phi_{c,p} = \pm \arctan(0.5(\xi_{Rx} - \xi_{Tx})/d)$,

b) increase $\kappa_{c,a}$ for the gradient direction $\{\alpha\}$ closest to $\phi$ by the geometric overlap $l_{p,c}$ of path $p$ within cell $c$

3) Define a single $\{\kappa_{c,a}\} = \frac{1}{c} \sum_{c} \{\kappa_{c,a}\}$ list for the full reconstruction domain as the mean over all cells and normalize such that $\sum_{a} \kappa_{c,a} = 1$.

For most inclusion geometries, MA-AWTV yields significantly improved results compared to AWTV; cf. Fig. 3(f). AWTV with $\kappa = 0.9$ is indeed a special case of MA-AWTV for $N_\alpha = 2$ and $H = d$.

### B. Incorporating prior information

If geometric information about anatomical structures is available a priori, then piecewise constant SoS values can be sought within some (not necessarily all) regions of the reconstructed image. This can substantially improve the robustness and accuracy of reconstructed quantitative values. For instance, given a B-mode US image of breast tissue, a target region of a suspected breast lesion may be delineated manually, and for this tissue region a homogeneous-equivalent SoS value can then be reconstructed as a malignancy marker (e.g. for BIRADS classification [8]). For this, in Eq.(3) all $R$ values $\{\sigma_1, \sigma_2, \ldots, \sigma_R\}$ within a target region-of-interest (ROI) are grouped as a single value $\sigma_{\text{ROI}}$. Accordingly, the corresponding columns of the imaging matrix $L$ is aggregated as $l_{p,1}\sigma_1 + l_{p,2}\sigma_2 + \ldots + l_{p,R}\sigma_{\text{ROI}} = (\sum_{c=1}^{R} l_{p,c})\sigma_{\text{ROI}} = l_{p,\text{ROI}}\sigma_{\text{ROI}}$. The columns of gradient matrix $D$ corresponding to $\sigma_{\text{ROI}}$ are also aggregated similarly, which leads the derivatives inside the grouped ROI to vanish and only regularization along its circumference to remain. Note that meanwhile the image regions outside such (one or multiple) ROI can still be reconstructed unconstrained, i.e. without imposing any piecewise-constancy assumption to allow to capture variation outside these ROIs.

### C. Robust reflector delineation with Dynamic Programming

The recorded full-matrix dataset $A(t_m, Rx, Tx)$, with $Rx = 1 \ldots N$ and $Tx = 1 \ldots N$, is a 3D matrix consisting of $N \times N \times N$...
radio-frequency (RF) lines in function of echo time \( t_m \), where \( m = 1 \ldots M \) are temporal discretization indices. In each RF line, a single ToF value corresponding to the echo from the reflector surface is sought \( \tau = \tau(Rx, Tx) \), which delineates the reflector surface across multiple (Tx,Rx) combinations.

Each RF line \( A(t_m) \) is a modulated waveform with an oscillatory pattern as seen in Fig. 4(a.1). Due to the heterogeneous SoS distribution, which leads to interference and weak scattering effects, the waveform shape may change significantly for different paths. In extreme cases, the reflected ultrasound signals may fall below the system noise level (fading) at certain paths and a ToF measurement then becomes not possible along all RF lines. Therefore, individual RF line analysis, e.g., by picking the peak signal amplitude, leads to ToF ambiguities as in Fig. 4(a.2). Indeed, in preliminary experiments even more sophisticated correlation-based or wavelet-based methods led to incorrect local maxima be selected among several (Tx-Rx) pairs. Envelope signal, such as waveform demodulation with Hilbert transform, was also not reliable alone and led to loss of temporal resolution, due to reverberations following the initial reflection, which distort the measurement of small perturbations \( \Delta \tau \). Therefore, in the literature ToF matrices are heavily post-processed for these outliers [23, 38, 39].

In contrast, we herein propose a global optimization approach, which simultaneously considers all Tx-Rx traces and minimizes an energy function to calculate the optimum delay matrix \( \tau \). This approach reduces timing ambiguities and also provides a continuous \( \tau \) surface in the transceiver-time space of \( A \). Dynamic Programming (DP) was proposed for the delineation of bones [40] and vessel walls [41] in ultrasound. Herein we introduce an algorithm using DP for detecting reflector time-delay surface across multiple RF lines.

a) Reflector in B-scan: In a first step, our algorithm tracks the reflector in a 2D image (B-scan), with the horizontal coordinate showing the echo time \( m \) and the vertical coordinate a list of successive RF lines \( l \) corresponding to adjacent Tx-Rx pairs with the same lateral Tx-Rx separation \( s=Rx-Tx \) (cf. Fig. 2). For instance, Fig. 4(a.1) shows the ray paths and a B-scan for \( s=0 \) (same element used for both Tx and Rx) and Fig. 4(b) shows the ray paths for larger \( s \) values.

For each B-scan \( A_{m,l} \), the algorithm then cumulatively builds a global cost matrix \( C_{m,l} \) along successive RF lines \( l \) for each possible timing candidate \( m \). A search window \( w = -W/2 \ldots W/2 \) of \( W+1 \) samples is successively defined for each candidate in each line \( l \) to its adjacent line. Also, a memory matrix \( E_{m,l} \) records discrete timing decisions for each \( l \) and \( m \). The optimum reflector delineation minimizes the cumulative cost, and the ToF profile \( \tau(l) = t_M(l) \) can be drawn by traversing \( E_{m,l} \) backwards as follows:

\[
C_{m,l} = \min_w \{ C_{m+w,l-1} - f_1(A_{m,l}, A_{m+w,l-1}) \} - f_0(A_{m,l})
\]

\[
E_{m,l} = \begin{cases} 
\arg\min_m C_{m,l} & l = L, \\
E_M(l+1),l+1 & l = \{L - 1, \ldots, 1\}.
\end{cases}
\]

The cost is evaluated with the echo likelihood \( f_0 \), where \( f_1 \) is the smoothness function of the current \( A_{m,l} \) and adjacent \( A_{m+w,l-1} \) RF locations of reflector delineation. This general framework can be used to introduce regularization into the energy function, for instance, in terms of ToF continuity between adjacent Tx-Rx index pairs, and/or constraints with respect to allowed reflector positions and orientations. For instance, in the current implementation, \( f_0 \) is formulated as a weighted sum of non-linear terms such that

\[
f_o \propto A_{m,l}, f_{rel}(A_{m,l}), f_{osc}(A_{m,l}), \quad (7)
\]

where \( f_{rel} \) and \( f_{osc} \) are binary step functions that are respectively activated if \( A \) shows a relative maxima at \( m \) or if an oscillatory pattern can be identified around \( m \). Undesired phase jumps in the delineation are avoided by using

\[
f_1 \propto |A_{m,l} - A_{m+w,l}|, f_{pj}(A_{m,l}, A_{m+w,l}), w, \quad (8)
\]

where \( f_{pj} \) is a binary step function that is activated if a phase jump greater than \( \tau \) occurs between \( A_{m,l} \) and \( A_{m+w,l} \).

b) Extension to all Tx-Rx pairs: Eq.(6) can be extended to simultaneously evaluate the 3D dataset \( A(t_m,Tx,Rx) \) by reorganizing the RF lines in a graph of relations between array elements. Adjacency relations (Tx,Rx), (Tx±s,Rx±s) between adjacent RF lines with same separation \( s \) have been described above for B-scan DP. In Fig. 4(b.2), individual evaluation of B-scans \( A_{m,l} \) for different \( s \) is shown. As \( s \) increases, the number of available RF lines in the B-scan \( L = N - s \) decreases. Considering their physical equivalence, reciprocal paths (Tx,Rx) and (Rx,Tx) can be considered in a single step by joining their terms in the energy function. Furthermore, regularization across reflector delineations at different contiguous \( s \) can also be introduced.

Fig. 4(b.3) illustrates a graph strategy to build \( C_{m,l,s} \) in one step for our 3D dataset, by additionally exploiting contiguity relations given by pairs (Tx,Rx), (Tx,Rx±1), (Tx±1,Rx). Cost matrix \( C_{m,l,s} \) can then be actualized with averages of the current \( s \) and contiguous \( s+1 \) B-scans. A graph strategy is defined, in which the averaged graph edges are guaranteed to cross the same number of nodes. Edge values in red in Fig. 4(b.3) indicate the number of crossed nodes. Several starting points (edge value = 4 in Fig. 4b.3) are then available for calculating \( M(L,s) \) in Eq. 9. At the branching points (e.g. the pair [3,3] in Fig. 4b.3) where one of the two preceding edges (edge value = 3) needs to be selected in order to accordingly read the memory matrix \( E_{M(l+1),l+1} \), the branch with the smaller cost function is chosen. This 3D algorithm can be written as follows:

\[
C_{m,l,s} = \arg\min_{w} \{ C_{m+w,l-1,s} + C_{m+w,l-1,s} - f_1(A) \} - f_0(A)
\]

\[
E_{m,l,s} = \arg\min_{w} \{ C_{m+w,l-1,s} + C_{m+w,l-1,s} - f_1(A) \}
\]

\[
M(l,s) = \left\{ \begin{array}{ll}
\arg\min_{m} C_{m,l,s} & l = L, \\
E_{M(l+1),l+1} & C_{m,l+1,s} \geq C_{m,l,s+1}, \\
E_{M(l+1),l+1,s+1} & C_{m,l+1,s} < C_{m,l,s+1}.
\end{array} \right. \quad (9)
\]

As a post-processing step, time readings below a given SNR (typically <10 dB) or not fitting to an oscillation relative maxima are filtered out and not included in the reconstruction. Discrete DP time values are then interpolated at the oscillation relative maxima to refine ToF estimations.
D. Implementation

All algorithms were implemented in Matlab®. For \( n = 1 \), Eq.(3) is a convex Second Order Cone Programming (SOCP) problem, which was solved with the CVX package using the solver Mosek v8.0—a high-performance optimizer for large-scale SOCP problems using an interior-point method [37], [42]. For \( N=128 \) number of transceivers and a reconstruction grid of \( 128 \times 128 \) pixels, the MA-AWTV reconstruction in CVX currently runs in \(<30\) s on an Intel® Core (TM) i7-4770k CPU@3.5 GHz with 16 GB RAM, while the 3D-dimensional reflector tracking runs in \(<1.5\) s. Typically the optimization takes 10-30 interior-point iterations to converge. We observed optimal reconstructions with the regularization constant \( \lambda \) scaling linearly with the cell grid resolution \( h \) and the image aspect ratio, while the algorithm was found to be robust against up to an order of magnitude in its variation. Too large \( \lambda \) leads to over-regularized images and axial resolution loss, while too low \( \lambda \) yields noisy background. This constant was empirically set to \( \lambda = 0.051 \) in most experiments to optimize Contrast Ratio fidelity for our discretization as \( \sqrt{H/d} = 1 \) and \( h \approx 300 \mu m \) the transducer pitch, where L1-norm, i.e. \( n = 1 \), was used in both fitting and regularization terms. This regularization weight was adjusted for different \( h, H \) or \( d \) values according to \( \lambda \propto h \sqrt{H/d} \).

III. NUMERICAL SIMULATIONS

A. Ray tracing simulations. Noise study

Ray-tracing simulations were run to compute the ToF matrix \( \tau \) for a US linear array with \( N = 128 \) elements and a pitch separation of 0.3 mm (width \( H = 38.1 \) mm). Reconstructions were then performed on an \( N \times N \) pixels square grid (i.e., \( H=d \)). Thirteen artificial inclusions of 1570 m/s were placed as in Fig. 3(a), on homogeneous background of 1554 m/s for an SoS contrast of 1%. Fig. 3(b) shows the corresponding delay perturbation vectors \( \Delta \tau \). With this, a 5 mm inclusion introduces 0.1 \( \mu \)s perturbation, which is realistic and satisfies the assumptions for applying Eq. 1 for estimate the reflector position and the background SoS, since it is two orders of magnitude below typical reflector delays \( \tau \) seen in Fig. 4(a).

Time-delay errors were synthesized and added to \( \Delta \tau \), following a Gaussian distribution with zero mean and a standard deviation expressed as percentage with \{0.8, 3.2, 6.5, 12.6\} % of the peak \( \Delta \tau \) for each example, hereafter referred to as noise levels.

Total Variation was observed to outperform conventional \( \ell_2 \) regularization (FBP), with streaking artifacts being eliminated, and closed inclusion geometries being delineated. AWTV provides a major improvement with respect to TV in axial resolution, achieving vertical inclusion separation (P3, P4). The improvement between MA-AWT and AWT is incremental and mainly affects oblique edges, for instance in P6 and P7, also improving axial resolution, i.e. the separability of vertical inclusions (P13). At all noise levels, rounder (P1, P2, P5-P9) or vertically elongated (P11 with an aspect ratio of 3:1) inclusions are well delineated, and the results have high lateral resolution, i.e. the inclusions can be separated horizontally (cf. P12, separated by array pitch). Not surprisingly considering the limited-angle problem, the most challenging reconstruction cases are vertically-separated inclusions (P13, separation of half-the-diameter) as well as laterally-spread inclusion geometries (P10 with aspect ratio of 1:3). It can be seen in Fig. 3b that \( \Delta \tau \) of P10 is easily confused with a circular inclusion (P2) and that P13 is very similar to a single vertical oriented inclusion (P11). Yet, below a noise level of 3.2% (Fig. 3g) all inclusions are still well-resolved. For noise >10% (Fig. 3h), smearing in P3, P4, P10 and P13 is observed, while the rest of inclusions are still delineated reasonably.

Next, we quantify potential time-delay errors with respect to the signal-to-noise ratio (SNR) of recorded RF lines. We first
collected experimental RF data in water, which has very low (negligible) noise (SNR = 54 dB). We then simulated different SNR levels by adding Gaussian noise on these RF lines, cf. Fig. 6a. Subsequently, our reflector delineation algorithm in Eq. (9) was applied to these noisy datasets and the noise statistics of $\Delta \tau$ were analyzed. The calculated timing-errors in $\Delta \tau$ follow a normal distribution and are, in logarithmic scale, linearly related to the SNR of the RF lines as seen in Fig. 6b. For moderate SNR (20 dB) in the RF lines, the timing-delay errors (3.5%) are not too large, thus ensuring a good reconstruction performance as seen in Fig. 3g. Although the water dataset is a relatively simple reflector delineation case, the advantages of the global optimization algorithm (3D vs. 2D and vs. individual RF line analysis) are apparent, especially for lower SNR, e.g. below 25 dB. The reconstruction of $\sigma$ is unaffected by the SNR of the RF lines (Fig. 6c), with an uncertainty of 0.1% for an SNR of 0 dB.

Quantitative reconstruction metrics are shown in Fig. 5 for the simulations in Fig. 3 [43]:

1) Background noise $\sigma_{bg}$ calculated as the SoS standard-deviation in the background region

2) Root-Mean-Square Error as

$$\text{RMSE} = \sqrt{\frac{1}{C} \sum_{c=1}^{C} (v^*_c - v_c)^2},$$  \hspace{1cm} (10)

where $v^*_c$ and $v_c$ are the estimated and ground-truth SoS values at image cell $c$

3) Contrast Ratio in [%] as

$$\text{CR} = \left| \frac{\mu_{inc}^* - \mu_{bg}^*}{\mu_{bg}^*} \right|$$  \hspace{1cm} (11)

where $\mu_{inc}$ and $\mu_{bg}$ are the mean values of SoS estimations in the inclusion and background

4) Dice coefficient of inclusion as

$$\frac{2|c_{inc}^* \cap c_{inc}|}{|c_{inc}^*| \cup |c_{inc}|}$$  \hspace{1cm} (12)

where $c_{inc}^*$ and $c_{inc}$ are the regions of the reconstructed and ground truth inclusions. To delineate the inclusion in the reconstructed images, we used a SoS cut-off value defined from the ground-truth as the mid-value between the preset background and inclusion values.

The RMSE represents a global quantitative assessment of image reconstruction quality, while $\sigma_{bg}$ gives the noise floor. CR has a quantitative diagnostic meaning, describing the SoS contrast fidelity with which inclusions can be reconstructed. The Dice coefficient assesses the geometric fidelity of inclusion delineation.

With respect to the ground-truth CR value of 1% used in our simulations, at a noise level of 0.8%, AWTV improves median CR from 0.31% (FBP) and 0.57% (TV) to 0.87% (AWTV) and 0.94% (MA-AWTV), which indicates the superior reconstruction efficacy of contrast of potential tumorous lesions with respect to tissue background. In contrast to previous hand-held SoS imaging works, which only aim at lesion detection [21], [26], [28], our method is suitable for both lesion detection and lesion quantification. The contrast is well-preserved for time-delay errors of 3.2% (CR = 0.82%). In agreement to previous works [22], [23], [25], if geometric information about an inclusion is available a priori, the contrast can be kept high (CR = 0.93%) even at higher (12.6%) noise levels.

**B. Full-wave ultrasound simulation**

In order to assess the influence of physical ultrasound wave propagation effects (diffraction, refraction, scattering) in reconstruction of heterogeneous SoS distributions, full-wave numerical simulations were performed. A finite-difference time-domain model (FDTD) developed in a separate work implements elastic wave propagation in terms of particle velocity $v_p$ and stress tensor $\sigma_{pq}$, which leads to a system of equations

$$\partial_t v_p = \rho^{-1} \partial_q \sigma_{pq}$$

and $\partial_t \sigma_{pq} = C_{pqr,s} \partial_s v_r$ in function of the density $\rho$, the elasticity tensor $C_{pqr,s}$, and the corresponding stiffness properties [44]. A pixel size $h_{\text{FDTD}} = 5 \mu m$ and a time step $h_t = 1.5 \text{ ns}$ allowed to neglect dispersion effects and provided accurate $\Delta \tau$ values down to 5 ns. The array design and 5 MHz excitation pulse simulate our experimental settings detailed later in Sec.IV. In order to simulate the
Fig. 6. Relation between time-delay errors and signal-to-noise ratio (SNR) of RF lines. An experimental dataset in water from Fig. 10 (SNR = 54 dB) was used for the analysis. Random noise was added to the RF lines to simulate different SNR levels. a) RF lines, reflector delay \( \tau \), and delay perturbation \( \Delta \tau \) vectors. b) Time-delay error \( \Delta \tau \) in function of SNR. The time-delay error units are equivalent to the benchmarks of Fig. 3. c) Error in estimation of background SoS (\( \tau \)) with Eq. (1).

synthetic aperture dataset, \( N=128 \) simulations were run, each for an individual Tx excitation. For each simulation, the \( N \) received echo traces were calculated by extracting the simulated pressure waves at the defined Rx positions. Tissue and inclusion were modeled as fluid media with SoS 1515 m/s and 1548.3 m/s simulating a tumor with 2.2% SoS contrast (similarly to the experimental examples in Sec. IV-B.C.). The reflector material (Poly(methylmethacrylate), \( v_{\text{REF}} = 2735 \) m/s) was chosen to avoid any coupling of inhomogeneous waves, which lead to phase distortion and therefore introduce errors in the reflector delineation. Given Snell’s law, inhomogeneous waves are coupled above the angle of total refraction, and the speed of sound in the reflector should be \( v_{\text{REF}} < \frac{\tau}{\sin(\phi_{\text{max}})} \).

Fig. 7a1 shows simulation snapshots of the pressure fields. A Tx element acts like a point source, initiating a circular wave that interacts with the inclusion and is reflected backwards at the reflector. At the inclusion, most of the incident waveform transmits forward, but also additional waves are generated (white arrows) for waves reflected backwards and diffraction around the inclusion edges. These modes are particularly important for rays that are incident (quasi)parallel to the inclusion edges, for which, according to Snell’s law, strong refractions occur. These effects are also visible in simulated B-scan (Tx = Rx), where at the inclusion boundaries mode interference is observed (Fig. 7a2, white arrows), which explain the fading effects observed below the edges of the inclusion in (Fig. 4a).

Fig. 7b1 compares ray-based delay perturbation vectors \( \Delta \tau \) with those evaluated with the FDTD simulations (Fig. 7b2-3) using the method of Fig. 1. Due to the high transducer frequency, the ray-tracing algorithm provides a fairly good approximation of the physical delays, however, at the inclusion boundaries, the full-wave simulations show a more continuous \( \Delta \tau \) transition pattern. Since the inverse problem is solved with the ray-tracing approximation, some distortion is also observed in the SoS reconstruction images, with small lobular artifacts at the horizontal inclusion boundaries (white arrows). These effects become more significant for later waveform phases 1 and 2, for which later wave arrivals overlap result in interfering modes. It is thus preferable to use the initial waveform phases for the reflector delineation. It was observed that by augmenting regularization (4-fold \( \lambda \) with respect to Fig. 3) the artifacts were significantly reduced. Overall, the influence of full-wave effects in the reconstructed SoS values in the inclusion is small (<0.2% SoS error) and the ray tracing model is a computationally-efficient imaging model for this reconstruction problem.

Additional FDTD simulations were performed to model breast imaging of women with different breast densities. In particular, we investigate the impact of high background/inclusion SoS contrast and tissue heterogeneity in the SoS reconstructions. Fig. 8 shows a model of a purely fatty tissue background (1470 m/s [45]) with a very hard tumorous lesion (1585 m/s), leading to an SoS contrast of 7.8%. This emulates a carcinoma within a fatty breast of category ACR a, classified by the American College of Radiology. As visible in the B-scan, the reflection/diffraction effects at the inclusion are more significant here than in Fig. 7, and strong mode interference is observed at the inclusion boundaries, so that the phase transition between inclusion and background tissue in the time perturbation matrix becomes blurred (Fig. 8c.2). In the extreme case, the waveform distortion in the inclusion may introduce one-period ambiguity in the tracked oscillation in inclusion with respect to the background (Fig. 8d.1), when the delineation algorithm attempts to keep phase continuity between the interfering modes. Overall, these effects lead to a slight decrease in Contrast Ratio: With FDTD, 84.3% CR w.r.t. given value is obtained, compared to 88.5% CR with ray-traced forward model.

Fig. 9 shows the same inclusion embedded in a heterogeneously dense (ACR c) breast. For this purpose, the anatomically realistic SoS breast model of [45] was used. The model is based on segmentation of tissue layers (skin, fat, glandular tissue) in real magnetic resonance imaging (MRI) breast datasets, for each of which different SoS values (1650 m/s, 1470 m/s, 1515 m/s) were assigned. Based on this model, we simulated the density distribution between ultrasound and reflector at a retromamillary segment [46]. Since the FDTD simulations are two-dimensional, an invariant density distribution was assumed along the transducer elevation in order to highlight in-plane heterogeneity. Fig. 9a shows the resulting simulation setup. Comparison of simulated B-scans (Fig. 9d.1) with the previous cases show a higher delay heterogeneity outside the inclusion region. The received echoes show also longer trails (Fig. 9d.2) due to signal reverberation in the heterogeneous breast structure, yet the reflector wave trace is still well-
Fig. 7. Full-wave ultrasound simulation. a) FDTD simulation setup. a.1) Time snapshots $t_1, t_2, t_3$ of the pressure fields radiated from element $Tx_1$. a.2) B-scan (comparable to Fig. 4a.1), a.3) RF line b). b) Time perturbation vectors $\Delta \tau$. c) SoS imaging results. Lobular artifacts are observed (white arrows), corresponding to interferences between diffracted and refracted waves at horizontal inclusion boundaries.

Fig. 8. Full-wave ultrasound simulation of tumorous lesion in fatty breast. a) Original setup. b) SoS imaging results. c) Time perturbation vectors $\Delta \tau$ (black pixels correspond to excluded values as in Fig. 4a.4). For comparison, SoS reconstructions were calculated based both on ray tracing (b.1, c.1) and FDTD (b.2, c.2) forward models. d.1) shows a FDTD B-scan (comparable to Fig. 8d.1) and d.2) a RF line. The high inclusion contrast (SoS 7.8%) introduces one-period phase ambiguity at the background-inclusion transition, which does not significantly degrade the reconstruction.

defined due to the high acoustic impedance contrast between breast and reflector. Fat layers are flat structures, which lie typically orthogonal to the imaging direction, similar to the skin layers. Consequently, as in P10 in Fig. 3 these structures are vertically smeared and eventually filtered out from the limited-angle reconstruction (Fig 9b.2). The reconstructed SoS value in the background (1497 m/s) is quantitatively equivalent to the SoS average of the heterogeneous breast (1503 m/s), with $\sigma_{bg} = 22.9$ m/s for the heterogeneous setup, 5.7 m/s for reconstruction with ray tracing forward model, and 8.6 m/s for reconstruction with FDTD forward model. On the other hand, the tumorous inclusion is satisfactorily delineated, with some vertical resolution loss (Dice = 0.80) caused by the regularization of tissue heterogeneity, and CR = 96.3% of given value for ray tracing forward model and 84.3% for FDTD model. Both ray tracing (Fig 9b.1) and full-wave simulations (Fig 9b.2) provide qualitatively similar reconstruction results, showing that the ray tracing approximation satisfactorily models the ray-front propagation for SoS reconstruction even in heterogeneous tissue.

IV. EXPERIMENTS

A 128-element 5 MHz linear ultrasound array (L14/5.38), with a pitch of 0.3 mm between elements and an element elevation of 7 mm, was used to generate ultrasound waves through a variety of materials (water, gelatin phantoms, ex vivo animal tissue and in vivo breast tissue) and measure the ultrasound echoes reflected by a thin reflector stripe (length $\times$ width $\times$ depth $= 70$ mm $\times 7$ mm $\times 5$ mm). Both transducer and reflector were attached to a positioning frame to ensure co-planarity. The reflector design constrains received echoes to the desired imaging plane, and allows for flexible access to different breast locations. The flat reflector geometry is simple to manufacture and easy to track in US data (Eq 1).

Conventional B-mode ultrasound images and synthetic aperture data were acquired using a SonixTouch ultrasound machine (Ultrasonix Medical Corporation, Richmond, BC, Canada). A multi-channel digitizer (SonixDAQ, Ultrasonix) collects raw pre-beamformed data in parallel at a sampling frequency of 40 MHz and 12 bits per sample. The process was repeated for all Tx elements, so that a total of $N^2$ time traces were recorded in $\leq$0.1 s (about 100 MB/frame).

A. Water calibration

Given the temperature in pure water, SoS can be known with an accuracy down to 0.03 m s$^{-1}$ [47]. We used distilled water to calibrate our quantitative SoS imaging setup. For this purpose, the reflector was delineated at different $\tau$ values (1485 to 1517 m s$^{-1}$) for a fixed reflector position ($d =$...
Fig. 10. Water calibration. a) Setup. b) Reflector delay $\tau$ and fit residuals $\Delta \tau$. c.1) Calibration of measured SoS with respect to nominal values in function of water temperature. Both transducer and reflector remain fixed during the experiment and the water is homogeneously cooled down. Non-physical reflector-probe distance $d$ (c.2) and inclination $\theta$ variations c.3) estimated with Eq. 1 is used to quantify $d, \theta$ estimation uncertainty.

From an initial state at temperature 21 °C, water was heated in a cooker and two cooling processes 29 °C to 24 °C and 33 °C to 29 °C were monitored with a temperature probe (HI98505, Hanna Instruments Inc., Italy), for which $\tau$ was measured in 0.1 °C steps. Temperature variations within the measurement window were <0.1 °C. $d$ was estimated with an uncertainty (RMSE) of 5 μm and $\theta$ with an uncertainty of 0.003°. The time residuals $\Delta \tau$ show an RMSE of 4.1 ns, corresponding to a noise floor of 0.1 m s$^{-1}$. The time residuals consist of a systematic component, which seems to be related to the surface imperfections of array crystals and reflector, and a random component of 0.7 ns. As for $\tau$, the RMSE is 0.93 m s$^{-1}$ (0.06%) with respect to the theoretical values [47]. Uncertainties are also associated to the limited accuracy of the temperature sensor (±0.2 °C).

### B. Tissue-mimicking materials

First, a breast elastography phantom (Model 0459, CIRS Inc., Norfolk, VA, USA) was tested (Fig. 11a). The phantom is fabricated with a tissue-mimicking material (Zeradine\textsuperscript{TM}) and approximates a real breast geometry, incorporating both skin layers and glandular tissue, together with cystic (Water) and hard lesions (low B-mode contrast and embedded microcalcifications). SoS images are represented in color scale and in relative units to highlight inclusion contrast. The background SoS value (1471 m/s) is in good agreement with the manufacturer’s nominal value (1475 m/s). The hard inclusion showed a 48 m/s SoS increase (3.3% contrast with respect to the background SoS) compared to the 16 m/s SoS increase of the cyst (0.9% contrast), both separated from the background noise $\sigma_{bg} = 1.0$ m/s (0.07%). In comparison, the manufacturer nominal values are -10 m/s SoS increase for cystic masses and 40 m/s SoS increase for hard inclusions. The hard inclusion shows the same lobular artifacts as predicted in the full-wave simulation in Fig. 7, which are associated to refraction effects at lateral inclusion edges.

Tissue-mimicking phantoms were next made of gelatin (9 g/100 mL water), mixed with flour to simulate typical B-mode texture (speckle). SoS inclusions were defined by adding a higher amount of gelatin (13 g/100 mL water) in well-defined phantom regions. To make the inclusions invisible to B-mode, the same amount of flour was used in tissue background and hard inclusions, so that both show the same echogenicity. Fig. 11b shows an example. The SoS images here reveal two invisible 5 mm inclusions. The SoS contrast of the hard inclusions is small (13.7 m/s, 1%), however, the lesions are satisfactorily delineated, with $\sigma_{bg} = 0.6$ m/s. Fig. 11c shows larger diameter inclusions at different depths. In this case, similar gelatin concentration as in Fig. 11b was used, but the flour amount was increased in the inclusions to reveal their position in B-mode. The lower aspect ratio $w/d = 0.56$ in Fig. 11c compared to Fig. 11b and Fig. 11a ($w/d = 1$) reduces the available angular range for reconstruction, however, there is not a significant impact in the imaging results until $w/d < 0.5$. Both hard and soft inclusions are satisfactorily resolved, with a Dice of 0.71 computed based on delineation in B-mode. The SoS contrast of stiff inclusions is consistent with the one observed in Fig. 11b. In Fig. 11c, the SoS artifacts are observed close to the reflector surface, due to the presence of impurities (air bubbles) at the reflector-tissue interface.

### C. Ex-vivo animal tissues

A variety of ex-vivo animal tissues were tested to illustrate the general applicability of the SoS imaging method to reconstruct focal lesions (Fig. 12). SoS inclusions (contrast scale 3% in Fig. 12d) were simulated by cutting off small parts from each tissue type, ablating (submerged in 250 mL water for 6 min at 700 W in microwave), and reinserting in the corresponding tissue. The inclusions are non-echogenic and acoustically attenuating, introducing an acoustic shadow in the B-mode images below their actual positions (Fig. 12b). Fig. 12c shows the recorded RF lines. At acoustically shadowed reflector positions, the signal amplitude drops, decreasing the signal-to-noise ratio. Nevertheless, our reflector delineation is able to filter out Tx-Rx pairs where such fading occurs. In all cases, the inclusion signature is visible as time perturbation (Fig. 12d) and the SoS images (Fig. 12e) show contrast in a ±3% scale at the expected inclusion positions. The inclusion delineation is piecewise homogeneous for bovine kidney, porcine liver and porcine myocardium and shows more heterogeneous patterns for chicken breast and bovine muscle. Lobular artifacts are observed as in Fig. 11a for the inclusions with highest SoS contrast (liver, myocardium). The average background SoS values (muscle: 1567 m/s, kidney: 1539 m/s, liver: 1585 m/s, myocardium: 1552 m/s) show a good quantitative agreement with reported values in human tissue (respectively, 1561 m/s, 1554 m/s, 1586 m/s, 1561 m/s) [48]. $\sigma_{bg}$ ranges from 1.8 m/s (bovine kidney) to 3.3 m/s (chicken breast).
D. In-vivo breast imaging

In order to illustrate the clinical applicability of our method, we herein present the first in-vivo results in female breast for tumor diagnosis. In-vivo data was recorded for a 58 year old patient showing a cancerous tumor, which the biopsy confirmed as Invasive Lobular Carcinoma (ILC); a 79 year old patient with a benign solid lesion, biopsy-confirmed as a fibroadenoma (FA); and a 27 year old healthy volunteer with a benign cystic mass. The preliminary data was obtained from a single-institution study at the University Hospital of Zurich, with the approval from the institutional review board and the local ethics committee [46]. While the subject sat in tripod position, the sonographer placed the ultrasound probe on the region-of-interest, and adjusted the reflector to achieve acoustic contact. An experienced breast radiologist annotated all lesions in the B-mode images as a reference.

Similarly to the previous ex-vivo results with ablated inclusions, B-mode shows tissue distortion but low echogenicity contrast at the inclusion region, with acoustic shadow behind. However, the SoS image (Fig. 13b) shows strong contrast at the tumor region (1585 m/s) with respect to the background 1485 m/s value, corresponding to a 6.7% SoS increase. As in Fig. 11c1, SoS artifacts are observed close to the reflector surface, which are associated with imperfect coupling between reflector and breast skin. The inclusion also appears vertically elongated, likely due to the smoothing effect from the time uncertainties caused by the relatively more noisy RF data. The echo profiles (Fig. 13c) and time perturbation vector (Fig. 13d) appear more heterogeneous than those in the previous phantom and ex-vivo cases, which may deteriorate the reconstruction. The SNR is also lower than that of previous experiments, due to higher acoustic attenuation. Nevertheless, the reflector echo is still well-defined (Fig.13c:2), which allows a consistent tracking of arrival time. Despite these artifacts, the reconstructed SoS image highlights well the inclusion position with respect to the homogeneous background, demonstrating the potential of our technique. The fibroadenoma lesion (Fig.14) shows lower SoS = 1521 m/s than that of the ILC, with a contrast of 4.3% with respect to the 1457 m/s background. As for the cystic mass, while it shows a strong hypoechogenic contrast in the B-mode image (Fig.15a), it presents only 1.1% contrast with SoS = 1502 m/s with respect to its 1485 m/s background. The SoS images may thus allow for lesion identification and quantitative differentiation, with potential diagnostic value for breast imaging.

V. DISCUSSION

The reflector delineation approach is robust for both phantom and a wide range of animal ex-vivo data, which presently allows to implement the method fully automatically. For reflector delineation, earlier wave oscillations were seen less...
influenced by interference from refraction artifacts (Fig. 7) but they also do present lower SNR. The method has been implemented to two-dimensional ultrasound probes, which allows to minimize the reflector footprint and improve the lesion accessibility. However, it can be generalized to 3D probes [49]. Although observed refraction artifacts are minimal, further refining of the reconstruction may be achieved with iterative reconstruction methods [50]. The reflector delineation is able to filter out transmit-receive (Tx-Rx) pairs where signal fading occurs. These are then not incorporated into the reconstruction, since our reconstruction algorithm can operate without necessarily having all Tx-Rx time-of-flight readings.

Artifacts, even in principle considered a limitation, can in practice provide useful patterns for trained radiologists to recognize and differentiate inclusions from heterogeneous tissue backgrounds. For instance, the current World Federation for Ultrasound in Medicine and Biology (WFUMB) guidelines for breast elastography [51] suggest malignancy scores when breast lesions appear larger in elastography than B-mode images. Also, the “bull’s eye” artifact in elastograms, which is characterized by a bright central signal within a darker outer region and a bright region posterior to the lesion, was demonstrated to be highly specific for benign simple and complicated cysts [51].

Similarly, for reflector-based SoS imaging, we have observed characteristic lobular artifacts for hard inclusions in figures 11a,3, 12e, and 13b, which all consist of an artefactual low velocity region to the sides of a high SoS inclusion. These artifacts are connected to the increased refractivity at the inclusion horizontal edges, as shown in Fig. 7c.

It is known from available in-vivo breast USCT literature that average SoS values in breast approximately range from 1430 m s\(^{-1}\) (very fatty) to 1520 m s\(^{-1}\) (very dense), with an average value of 1450 m s\(^{-1}\) [52]. Cysts show typically similar SoS values to water, while benign tumors (fibroadenoma) show on average lower SoS contrast (1513 m s\(^{-1}\)) compared to malign tumors (carcinoma), which have SOS from 1520 m s\(^{-1}\) to 1620 m s\(^{-1}\) [18]. Another full-angle USCT system implementation [53] measured an average of 1545 m s\(^{-1}\) for healthy glandular tissue and 1595 m s\(^{-1}\) for carcinoma. Our preliminary observations are consistent with these values. The heterogeneous reflector profiles (Fig. 13c-d) are associated with density variations in the healthy breast tissue, which were reproduced in the full-wave ultrasound simulations (Fig. 9). In a clinical study using a similar device for breast density classification [46], it was observed that such variations are particularly significant for heterogeneously dense breasts. The AWTV algorithms tend to smear the background SoS variations, highlighting high-contrast SoS inclusions. Reversal variations are also observed due to reflections at the breast skin layer [46], which are here visible in figures 13c, 14c, 15c, and marked in Fig. 15c.2. A custom reflector plate was successfully designed such that a well-defined coherent echo can be separated from spurious tissue reflections. Full-wave numerical simulations show that SoS values in inclusions can be quantitatively reconstructed in heterogeneous breast models.

As shown in our ex-vivo experiments, ablated tissue has good SoS contrast, so SoS may be used for post-ablation quality control. Additionally, since tissue SoS is a function of temperature, SoS imaging may also be used for real-time monitoring of ablation as in [54], e.g. by tracking incremental
SoS changes from a reference initial SoS measurement. Similarly to shear-wave elastography, speed-of-sound imaging provides quantitative tissue biomarkers. However, each of these methods measure different mechanical properties. Shear wave speed (SWS) correlates with Shear-Wave Modulus $G$, while speed-of-sound SoS correlates with the Bulk modulus $K$. Previous literature suggest that SWS is strongly associated to palpable stiffness and mechanically loading state of the tissue, whereas SoS is more associated to density and tissue microstructure. For instance, [55] showed that SoS can outperform SWS for differentiation of common phantom and ex-vivo tissues. However, a large clinical study comparing SWS and SoS for breast lesion differentiation is still not available. SoS also discriminates fatty and fibrous tissues, and has been applied to breast density classification [46], [52], liver steatosis diagnosis [56] and assessment of fatty musculoskeletal disease [57]. Simple hand-held SoS methods as proposed in this work are hoped to contribute to an increase in the available clinical evidence on SoS.

The main purpose of this paper is to establish the new SoS image reconstruction methodology and evaluate accuracy and resolution in numerical simulations and well-controlled phantom/ex-vivo experiments. The in-vivo results show the potential applicability of the SoS method for breast diagnostics, which needs to be confirmed with further clinical studies. A recent clinical paper [46] measuring N = 106 women with the same hardware setup for breast density classification has shown that the setup is practically applicable for women of different breast densities and cup sizes. Measurements are fast (<2 min/breast) and well-tolerated. The SoS-US showed significantly better comfort scores than mammography, due to the localized and minimal breast compression of the SoS-US setup. Measurements were performed in both retromammary, inner, and outer breast segments. There might be some reflector access issues at the periphery of the breast. However, this is a common limitation of all transmission based approaches (including X-ray mammography and USCT), and, in comparison with the latter, our smaller reflector size could facilitate a flexible access to most breast positions. In clinical practice, reflector and ultrasound probe would be separated up to the maximum frame opening, a conventional hand-held ultrasound scan would be performed, and if suspicious lesions are found the reflector would be adjusted until there is acoustic coupling for SoS evaluation.

VI. CONCLUSIONS

A passive add-on to conventional US systems with array transducers has been shown to be sufficient to obtain quantitative SoS images of focal lesions. A passive reflector used as a timing reference enables integral time-of-flight measurements, from which a limited-angle CT problem is defined based on ray assumption of wavefront propagation. The reflector delineation is performed fully automatically with a global optimization method that takes the full ultrasound dataset into account and handles signal fading, which is often present in experimental data. We have introduced a smoothness regularization term in the inverse limited-angle CT problem, which weights gradients according to available measurement paths, and by using Total Variation to achieve edge-preserving regularization. This provides sharp delineation of focal lesions and allows for quantitative SoS reconstruction in both tissue background and inclusions. Ray tracing simulations show the generality of the method for a wide range of inclusion geometries. Additionally, full-wave simulations show that a ray tracing model can be sufficient for accurate SoS reconstruction, apart from small characteristic lobular artifacts at inclusion lateral edges for SoS contrast > 3%. The collimation improves for the first received wave oscillations. The method performs well for SNR >20 dB and allows to reconstruct non-echogenic SoS inclusions in breast-mimicking phantoms and ex-vivo animal tissues. Preliminary results in breast tissue shows contrast at a cancerous tumor with SoS values consistent with previous literature.

The proposed method is easy to implement and can contribute in bringing USCT advances to a wider clinical practice. In terms of multi-parametric diagnosis, it can potentially contribute to improve the differentiation of breast lesions in the diagnostic setting. Being computationally inexpensive, it could also potentially be used as a complementary ultrasound modality in hand-held breast examination or for surgical guidance.

ACKNOWLEDGMENT

This work was funded by the Swiss National Science Foundation and an ETH Zurich Pioneer Fellowship. We acknowledge Corin Otesteanu for phantom preparation, Simon Studerus for the GPU implementation of the FDTD simulations, and Dr. Richard Rau and Dr. Valery Vishnevsky for running additional simulations for the paper revisions. PD Dr. med. Konstantin Dedes and Dr. med. Denise Vorbürgen are acknowledged for the acquisition of the preliminary clinical datasets. Martin Vogt from the ETH workshop is also acknowledged for the useful discussions to optimize the reflector frame design.

REFERENCES
