Spatial Domain Reconstruction for Imaging Speed-of-Sound with Pulse-Echo Ultrasound: Simulation and In-Vivo Study

Sergio J. Sanabria*, Ece Ozkan*,
Marga Rominger, and Orcun Goksel

*Both first authors contributed equally.

S. J. Sanabria, E. Ozkan, and O. Goksel are with the Computed-assisted Applications in Medicine Group, ETH Zurich, Switerland.

M. Rominger is with the Department of Diagnostic and Interventional Radiology at the University Hospital Zurich, Switzerland.

E-mail: `ogoksel@vision.ee.ethz.ch`

Abstract

Despite many uses of ultrasound, some pathologies such as breast cancer still cannot reliably be diagnosed in either conventional B-mode ultrasound imaging nor with more recent ultrasound elastography methods. Speed-of-sound (SoS) is a quantitative imaging biomarker, which is sensitive to structural changes due to pathology, and hence could facilitate diagnosis. Full-angle Ultrasound Computed Tomography (USCT) was proposed to obtain spatially-resolved SoS images, however, its water-bath setup involves practical limitations. To increase clinical utility and for widespread use, recently, a limited-angle Fourier-domain SoS reconstruction was proposed, however, it suffers from significant image reconstruction artifacts. In this work, we present a SoS reconstruction strategy, where the forward problem is formulated using differential time-of-flight measurements based on apparent displacements along different ultrasound wave propagation paths, and the inverse problem is solved in *spatial-domain* using a proposed total-variation scheme with spatially-varying anisotropic weighting to compensate for geometric bias from limited angle imaging setup. This is shown to be robust to missing displacement data and easily allow for incorporating any prior geometric information. In numerical simulations, SoS values in inclusions are accurately reconstructed with 90% accuracy up to a noise level of 50%. With respect to Fourier-domain reconstruction, our proposed method improved contrast ratio from 0.37 to 0.67 for even high noise levels such as 50%. Numerical full-wave simulation and our preliminary in-vivo results illustrate the clinical applicability of our method in a breast cancer imaging setting. Our proposed method requires single-sided access to the tissue and can be implemented as an add-on to conventional ultrasound equipment, applicable to a range of transducers and applications.

1. Introduction

Ultrasound (US) imaging is a low-cost, portable, easily accessible, real-time, and non-invasive technology. Conventional B-mode ultrasound images show the reflectivity of

tissue interfaces and scattering within the body. Although for many tissues this may present contrast, yet it fails to differentiate others. In particular, for breast cancer imaging, B-mode shows higher sensitivity than the gold standard X-ray mammography to early-stage invasive cancers. However, it provides low specificity to differentiate benign and malignant tumors, and thus cannot prevent unnecessary biopsies. Moreover, B-mode image interpretation is still not a quantitative approach, with its outcome very much dependent on operator experience (Arribas et al. 2016).

Biomechanical properties of soft tissues can be altered by pathological processes. This may provide complementary information for clinical diagnosis to detect tissue abnormalities. Since pathological tissues may cause changes in the biomechanical properties of tissues, tissue elasticity has been used in medicine for centuries as a diagnosis marker. Manual palpation is still commonly used for screening breast and prostate tumors. However, this qualitative method is limited to superficial regions.

Quantification of tissue biomechanical properties using ultrasound has a high potential for diagnosis and staging of diseases. Elastography is the "computerized palpation", where elastic properties of the tissue, e.g. Young's or shear modulus or shear-wave speed (SWS), are estimated from the displacement field measurements in ultrasound images as a response to tissue excitations or perturbations (Ophir et al. 1991, Eskandari et al. 2008, Sarvazyan et al. 2011, Goksel et al. 2013). However, the specificity of elastography for differential diagnosis is limited, with measurements varying strongly with the tissue mechanical loading state (Cosgrove et al. 2013).

An independent biomechanical intrinsic material characteristic is bulk modulus, which relates to the propagation of longitudinal waves in the material, measurable by ultrasound as the speed-of-sound (SoS). Measuring SoS has been shown to outperform SWS for differentiation of common phantom and *ex-vivo* tissues (Glozman & Azhari 2010), and has a strong potential for benign and malignant tumor differentiation (Bamber & Hill 1981, Jeong et al. 2008, Li et al. 2009, Duric et al. 2010, Zografos et al. 2015). A large clinical comparison study between SWS and SoS for breast lesion differentiation is yet not available. Nevertheless, it was shown in several individual studies that SoS is able to discriminate between adipose and fibrous tissues, and has been applied in several clinical contexts such as breast density classification (Sak et al. 2015, Sanbria et al. 2018), diagnosis of liver steatosis (Imbault et al. 2017) and assessment of fatty musculoskeletal degeneration (Qu et al. 2017).

A family of methods called Ultrasound Computed Tomography (USCT) has been proposed to quantitatively measure SoS (Jeong et al. 2008, Nebeker & Nelson 2012, Duric et al. 2013). USCT reconstructs ultrasound propagation speed from transmission measurements based on an elevated number of transducer elements located around the inspected tissue, which is immersed in a degassed water tank. A tomographic reconstruction scheme uses time delays and amplitudes of the ultrasound signal transmitted along multiple angular directions. However, most body parts are not easily submersible and hence are not accessible by USCT. Furthermore, this method requires a bulky and burdensome setup and imaging work-flow, negating most advantages of typical

ultrasound systems, and thus constraining its clinical translation. Medical applicability of USCT is currently limited mainly to the breast. These systems allow for the automatic imaging of the whole breast in a 3D coordinate system, familiar to radiologists and can help reduce operator-dependency, thus having potential in screening and diagnosis.

For widespread clinical use of USCT capabilities, it would be desirable to use standard ultrasound probes for SoS imaging. Ultrasound is currently used in the clinics as a supplementary diagnostic modality, following an initial imaging modality – often the X-ray mammography. Diagnosis in B-mode US images is currently based on qualitative interpretation of lesion geometric features, with low specificity for lesion differentiation, and therefore any suspicious cases ending up with biopsies as histology being the gold-standard in diagnosis. Adding SoS capabilities to conventional ultrasound imaging could improve lesion differentiation, thus reducing unnecessary biopsies and providing early staging of cancers. Several approaches have been proposed to estimate SoS in pulse-echo mode using reflected signals. Iterative optimization of average SoS values based on the blurring effect of point-like structures in B-mode image reconstruction or based on registration of ultrasound images acquired from different directions were discussed in (Krücker et al. 2004, Shin et al. 2010, Cho et al. 2014), mostly aiming for estimating a single homogeneous-equivalent SoS value for correcting aberration artifacts. Another method (Byram et al. 2012) relies in the co-registration of different ultrasound views for isolated bright point reflectors, such as a wire target, however, the latter are typically not available in real tissue textures, which consist of diffuse speckle patterns. Other early attempts to measure SoS distributions used a beam tracking technique (Robinson et al. 1991, Kondo et al. 1990, Cespedes et al. 1992), in which two separated transducers generate intersecting beams, but this can provide a very coarse SoS resolution. More recently, a Fourier domain reconstruction has been proposed to link echo phase shifts from different beamforming directions to reconstruct local SoS variations, for reconstructing SoS images (Jaeger et al. 2015) and with preliminary *in-vivo* results showing promise in differentiating abdominal tissue layers (Jaeger & Frenz 2015b). This method, however, suffers from significant image reconstruction artifacts and loss of resolution in the axial direction for moderate noise levels, i.e 10% (20 dB) noise, or for incomplete echo shift datasets. As a result, it does not allow quantitative interpretation of the SoS images, which limits its clinical applicability. To address these challenges, spatial domain algorithms have been discussed as a possible extension in (Jaeger & Frenz 2015a); however, to the best of our knowledge, a specific method has not yet been proposed, nor quantitatively validated in the literature. Another recent work demonstrated quantitative SoS imaging based on the time-of-flight measurements to a passive reflector as the timing reference (Sanabria & Goksel 2016) and a spatial domain regularization; however, this requires two-sided access to the investigated tissue.

In this work, an SoS image reconstruction method is proposed by solving an inverse problem directly in the spatial domain. This allows spatial-domain regularization, where we utilize here an anisotropically-weighted total-variation method. This enables the

delineation of tumorous inclusions, and the accurate prediction of their speed-of-sound values. This reconstruction method utilizes a set of echo shift measurements in the observed domain, which can be obtained using image registration and displacement tracking techniques on different ultrasound transmitter excitations, such as plane-wave transmits in different angles. With a spatial domain reconstruction, we can also incorporate incomplete echo measurements and prior information about the inclusion geometry. The proposed method only requires single-sided access to the tissue and is applicable with a conventional ultrasound array transducer, being well-suited to image internal organs such as the liver and the kidney. Furthermore, this can also significantly simplify the clinical work-flow for organs in which double-sided access is still possible, in particular the breast, since no water bath or additional hardware elements are necessary.

2. Methods

Below, we first present the SoS problem formulation and its spatial discretization for differential time-of-flight measurements along arbitrary wave paths. Next our proposed solution to the regularized *inverse problem* is introduced. Finally, a practical beamformer implementation for producing different wave paths and differential measurements is given.

2.1. Spatial Domain Reconstruction (SDR)

In the following, the wavefront propagation is estimated/represented by rays, similarly to light propagation in optics. Without loss of generality, it is assumed that an ultrasound transducer emits and receives ultrasound waves along a total of P wave propagation paths p, corresponding to trajectories along which ultrasound waves propagate, as in Fig. 1(a). Accordingly, the time delay t_p for each path, also known as Time-of-Flight (ToF), is defined as the time interval between the time of transmission and of reception of the ultrasonic wave. To relate time and SoS quantities, the imaging region is first discretized into a predefined grid $[x, z]$ containing a finite number of cells C determining the resolution of the image. With a path length $l_{p,c}$ per ray path p and grid cell c from the transmitter Tx to an arbitrary scatterer S and back to the receiver Rx, the ToF values can be calculated in function of the slowness σ_c ($= v_c{}^{-1}$, the inverse of discretized SoS) as follows:

$$t_p = \int_C \sigma dl \approx \sum_{c=1}^{C} l_{p,c}\sigma_c. \tag{1}$$

However, in general, the position of inner tissue structures is not accurately known, and only relative time measurements τ_m along different paths can be obtained as:

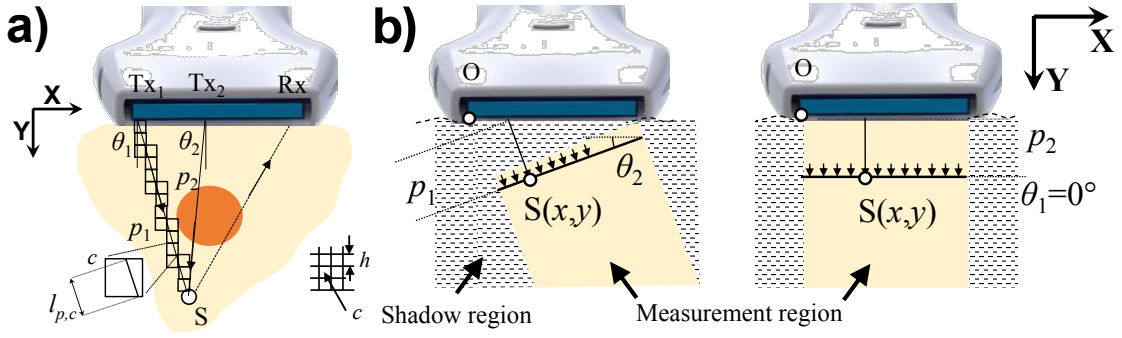$$\tau_m = \sum_{p=1}^{P} w_{m,p}t_p. \tag{2}$$

Figure 1. Speed-of-sound (SoS) reconstruction setup. (a) Rays p_1 and p_2 represent paths traversed by wavefronts arriving at point S for different plane-waves at angular directions θ_1 and θ_2. Discretization for proposed Spatial Domain Reconstruction (SDR) is also shown. (b) Wave path definition based on plane wave beamforming.

For each measurement $m=\{1,\ldots,M\}$, τ_m provides relative time delays between paths t_p using the weights $w_{m,p} = \{0, \pm 1\}$. These relative delays summarize the sound speed deviations accumulated along different wave propagation paths from the probe to the acoustic scatterer S. Substituting (1) in (2), the relative time measurement can be expressed as

$$\tau_m = \sum_{c=1}^{C}\left(\sum_{p=1}^{P} w_{m,p} l_{p,c}\right)\sigma_c = \sum_{c=1}^{C} L_{m,c}\sigma_c, \tag{3}$$

where the last equation merges the path lengths $l_{p,c}$ and relative measurement weights $w_{m,p}$ in a single linear combination of weighted path lengths $L_{m,c}$. The system (3) is a *forward problem*, which can be expressed in a matrix form as

$$\boldsymbol{\tau} = \mathbf{L}\boldsymbol{\sigma}, \tag{4}$$

where $\mathbf{L} \in \mathbb{R}^{M \times C}$ represents geometric information that depends on the transmitter-receiver setup of the transducer. The slowness values per cell c are assembled in $\boldsymbol{\sigma} \in \mathbb{R}^C$ and the time measurements τ_m in $\boldsymbol{\tau} \in \mathbb{R}^M$, respectively, and $\boldsymbol{\sigma}$ can be reconstructed using (4).

A complete set of path orientations $[-\pi, \pi]$ traversing each cell c is necessary for a complete computed tomographic reconstruction (Kak & Slaney 1988). However, for pulse-echo imaging with a conventional ultrasound probe with single-sided tissue access, each cell is only traversed by a limited set of angular directions θ_1, θ_2..., thus the reconstruction problem is *incomplete*. Moreover, relative path measurements lead to an *ill-conditioned* matrix $\mathbf{L}$. Therefore, we solve (4) as a regularized optimization problem:

$$\hat{\boldsymbol{\sigma}} = \arg\min_{\boldsymbol{\sigma}}||\boldsymbol{\tau} - \mathbf{L}\boldsymbol{\sigma}||_1 + \lambda||\mathbf{D}\boldsymbol{\sigma}||_1\,, \tag{5}$$

where ℓ_1 norm total variation (TV) ((Sidky et al. 2006)) is used for edge-preserving regularization. λ is the regularization constant and $\mathbf{D}$, in a simple form, can be the first-order discrete differentiation operator.

2.2. Multi-Angle Anisotropically-Weighted Total Variation

Considering limited-angle nature of our USCT problem, i.e., with naturally lower resolution in the axial direction, one can regularize gradients in each axes differently. This leads to the following Anisotropically-Weighted Total-Variation (AWTV) regularization (Sanabria & Goksel 2016):

$$||\mathbf{D}\boldsymbol{\sigma}||_{\text{AWTV}} = \sum_{i,j} \kappa|\sigma_{i+1,j} - \sigma_{i,j}| + (1-\kappa)|\sigma_{i,j+1} - \sigma_{i,j}|. \tag{6}$$

where i and j are discretization indices, in horizontal (x) and vertical (y) directions. Derivation of AWTV is further elaborated in Appendix A. Similarly to (Sanabria & Goksel 2016), $\kappa = 0.9$ was used, leading empirically to an optimum performance for the imaging setup discussed in this paper as well.

Extending on this idea, one can also use multiple gradient directions in the spatial regularization term, instead of two orthogonal gradient estimates, which is then referred to as Multi-Angle Anisotropically-Weighted Total-Variation (MA-AWTV)

$$||\mathbf{D}\boldsymbol{\sigma}||_{\text{MA-AWTV}} = \sum_{i,j} \sum_{\alpha=\{\alpha_1,\alpha_2,...,\alpha_N\}} \kappa_\alpha|D_\alpha\boldsymbol{\sigma}|, \tag{7}$$

where $D_\alpha\boldsymbol{\sigma}=\mathbf{D}\boldsymbol{\sigma} \cdot \mathbf{e}_\alpha$ is the directional derivative along the unit vector $\mathbf{e}_\alpha$ with angle α. Given the maximum imaging angle θ_{max}, the gradient directions α for a total of different N_α directions are chosen as $\{\alpha_k\} = \{\theta_k, 180° - \theta_k\}$ to ensure an optimum coverage of the available angular directions with

$$\theta_k = \left\{0, \theta_{max}\frac{1}{N_\alpha/2 - 1}, \theta_{max}\frac{2}{N_\alpha/2 - 1}, \dots, \theta_{max}\right\}. \tag{8}$$

The weights are calculated based on the setup geometry and hence available projection directions spatially at each cell as follows: For each wave path p crossing cell c, we calculate the inclination θ of the wave path. If $\theta < \theta_{max}$, we then increment the weight κ_α of the closest gradient direction $\{\alpha\}$ with respect to θ by the geometric overlap of path p with cell c. We then compute the mean $\{\kappa_\alpha\}$ over all cells, and use these set of weights for the subsequent SoS reconstruction.

The proposed reconstruction problem using MA-AWTV is given as

$$\hat{\boldsymbol{\sigma}}_{\text{MA-AWTV}} = \arg\min_{\boldsymbol{\sigma}}||\boldsymbol{\tau} - \mathbf{L}\boldsymbol{\sigma}||_1 + \lambda \sum_{i,j} \sum_{\alpha} \kappa_\alpha|D_\alpha\boldsymbol{\sigma}|, \tag{9}$$

where the weights are normalized such that $\sum_\alpha \kappa_\alpha=1$. Note that not including the directions outside the imaging angular span $[-\theta_{max}, \theta_{max}]$ effectively assigns them no regularization. This avoids over-constraining directions, for which no boundary information is available, which reduces image artifacts. For conventional medical ultrasound arrays, $\theta > 30°$ leads to strong grating lobes and cluttering and is typically not used (Hoskins et al. 2010). The regularization term λ is an optimization constant that balances the solution and regularization term. The quantitative speed of sound values are then calculated from the slowness values as SoS $=\hat{\sigma}^{-1}$.

2.3. Incorporating prior information

For cases when the position and geometry of inclusions are known *a priori* (for instance, from an ultrasound B-mode image), a constrained reconstruction can be defined to determine homogeneous SoS values robustly in the inclusions. For instance, a target region can be a suspect breast lesion, for which tissue malignancy (BIRADS classification) needs to be staged (Arribas et al. 2016). Proposed prior information helps to reduce the number of degrees of freedom in the reconstruction inverse problem. An average SoS value obtained with this approach in a suspicious lesion may facilitate tumor classification and differential diagnosis (Sanabria et al. 2018). Tumor heterogeneity via spatially-resolved images can be an additional diagnostic factor. Given (9), prior information is incorporated to (5) as follows: all R values $\boldsymbol{\sigma}_{\mathrm{ROI}} = \{\sigma_1, \sigma_2, \ldots \sigma_R\}$ within a target region-of-interest (ROI) are grouped as a single value σ_{ROI}. Accordingly, the corresponding columns of the geometric matrix $\mathbf{L}$ can be summed up $l_{p,1}\sigma_1 + l_{p,2}\sigma_2 + \ldots + l_{p,R}\sigma_{\mathrm{ROI}} = (\sum_{c=1}^{R} l_{p,c})\sigma_{\mathrm{ROI}} = l_{p,\mathrm{ROI}}\sigma_{\mathrm{ROI}}$. The columns of gradient matrix $\mathbf{D}$ corresponding to σ_{ROI} are also summed up, which leads to the neighborhood derivatives vanishing inside the grouped ROI, leaving only regularization along the circumference.

2.4. Multiple plane-wave transmits for a practical imaging scenario

An ultrasound probe consists of N transducer elements, each of which can be fired individually as transmitter Tx or used as receiver Rx. Recording each combination (which is called *multi-static* or *full-matrix* data) results in a total of $N \times N$ radio-frequency (RF) wave traces $A(t, \mathrm{Tx}, \mathrm{Rx})$ in function of echo time t. Assuming linearity in system electronics, scattering response, and acoustic propagation at low-power ultrasonic signals, different beamforming processes (*synthetic apertures*) can then be retrospectively constructed based on superposition from this data (Holmes et al. 2005). For instance, by delaying, scaling and adding up these traces, one can synthesize full-aperture plane, single-element circular, and collimated aperture wavefronts and focused beams, among others. After such beamforming, relative delay measurements τ_m can be taken between beamformed ray paths (Robinson et al. 1991, Cespedes et al. 1992, Byram et al. 2012, Jaeger et al. 2015).

For the sake of simplicity and without loss of generality, we hereafter focus on full-aperture plane-wave beamforming (Fig. 1(b)), which allows a direct comparison of our reconstructed method with another recent method (Jaeger et al. 2015). In this case, spatial diversity is obtained by controlling the angle of incidence θ of the plane wavefront. Consequently, the same tissue position (x, y) can be accessed from two different wave propagation paths (p_1, p_2) corresponding to two different angles (θ_1, θ_2). The weights $\mathbf{L}$ are calculated by subtracting the ray path lengths $l_{p,c}$ at steering angle θ_2 from the ones at reference angle θ_1. $l_{p,c}$ are calculated with ray tracing assuming straight ray paths. A line anti-aliasing algorithm (Wu 1991) is used to avoid interpolation errors. The relative time measurements τ_m at coordinate (x, y) can be obtained from the shift

of point observed in images seen from these two angles.

Clutter may complicate time measurements τ_m if the matched tissue scatterer S at (x, y) (Fig. 1(b)) is overlapped with echoes from other reflecting tissue structures. Therefore, the received echoes at each receiver element Rx are delayed and summed to achieve a focused beam, so that tissue reflectivity is maximized in a small region around the desired measurement point (x, y). Applying both plane wave Tx beamforming and focused Rx beamforming, tissue reflectivity (speckle) images $B(x, y, \theta)$ at different angles θ are generated as:

$$B(x, y, \theta) = \sum_{Tx=1}^{N} \sum_{Rx=1}^{N} A(t_{\text{Rx}} - t_{\text{Tx}}, \text{Tx}, \text{Rx}) \tag{10}$$

$$t_{\text{Rx}} = 1/v_B(x \sin(\theta) + y \cos(\theta) + \sqrt{(x - \xi \cdot Rx)^2 + y^2}) \tag{11}$$

$$t_{\text{Tx}} = 1/v_B \cdot \text{Tx} \cdot \xi \cdot \sin(\theta)/v_B \tag{12}$$

where t_{Tx} and t_{Rx} are respectively the transmit and receive delays, ξ is the transducer inter-element separation (array pitch) and v_B is the average speed of sound in the medium. A nominal value of $v_B = 1540$ m/s for soft tissue is used for beamforming.

Time measurements $\tau(x, y, \theta)$ are obtained by registering the two images $B_1 = B(x, y, \theta)$ and $B_2 = B(x, y, \theta + \Delta\theta)$. Speckle tracking methods are well-known in the literature, such as based on cross-correlation or optical-flow (Lubinski et al. 1999, Pan et al. 2009, Zahiri-Azar et al. 2010, Zahiri-Azar et al. 2011, Makhinya & Goksel 2015). In this work, a one-dimensional correlation search is performed in y, using a zero-normalized cross-correlation (ZNCC) (Pan et al. 2009) to find the vertical displacement Δy at each point (x, y) (Fig. 2), from which $\tau = \Delta y / v_B / \cos(\theta)$. The correlation coefficient q_{ZNCC}, which ranges between -1 and 1, with the largest positive values providing the best correlation, is used as a quality metric of the speckle tracking.

Due to the finite transducer aperture size, there are *shadow regions* in the imaging area where a plane wavefront may not arrive and no time measurements can be recorded (Fig. 1(b)). The proposed SDR method can handle these regions by simply excluding them from the input τ_m measurements, since τ_m are independent from the reconstruction grid σ_c. Time measurements collected at multiple different angles $\{\theta_1, \theta_2, \ldots, \theta_A\}$ can also be combined by stacking as input to (5) and solved together at once, in order to increase accuracy and robustness.

3. Experiments and Results

The reconstruction was implemented in Matlab® (The MathWorks, Inc.) with version r2017b (2018). The optimization problems (A.2) and (9) were solved with the CVX package (Grant & Boyd 2014, Boyd & Vandenberghe 2009), using the numerical solver Mosek v8.0, which is a high-performance optimizer for large-scale Second Order Cone Programming (SOCP) problems using an interior-point method. Cells with little information, i.e. those traversed by $< 1\%$ of the rays in comparison to the most
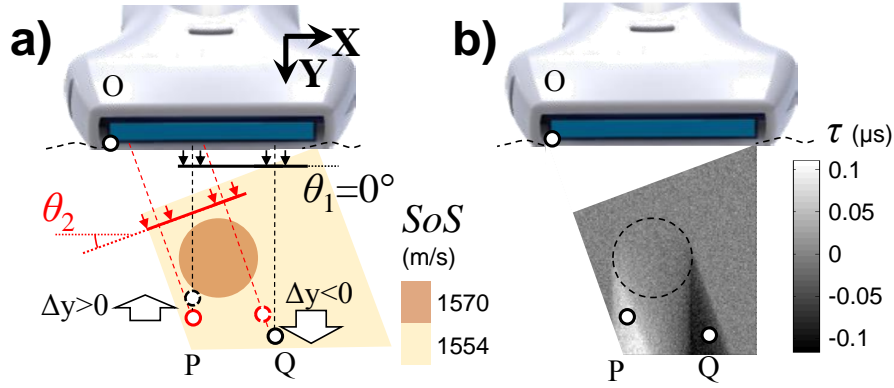
Figure 2. Plane-wave relative delay measurements τ_m at an angle $\theta_2=20°$ with respect to $\theta_1=0°$. a) Setup, b) Simulated τ_m at noise level of 10%.

traversed cell, were not included in the reconstruction problem. A typical MA-AWTV reconstruction in CVX presently runs in $<30\,$s on an Intel® Core™ i7-4770k CPU @3.5 GHz with 16 GB RAM for a single frame and for a domain of size 40x38 mm with 0.3mm reconstruction discretization using a pair of plane-wave angular projections with the same discretization resolution. For our synthetic tests, we first studied the effects of noise, regularization, and parametrization in reconstructions using ray-based synthetic data generation. This allows to study the effects of measurement errors using simulated noise on identical forward and inverse problems. In order to study any physical wave effects in reconstructions, we later used full-wave simulations of our physical acquisition setup, to study image reconstruction from subsequent computed displacements. Finally, we present reconstructed images from data collected with an ultrasound machine.

3.1. Synthetic Dataset via Ray-based Wavefront Modeling

Ten different examples of artificial inclusions were defined in a homogeneous tissue with a depth and width of 40 mm, as shown in Fig. 3(a). The colorbar that is used for the quantitative reconstruction outputs is shown on the left corner of the figures. Background SoS was set to 1554 m/s and the inclusions are either 1538 m/s or 1570 m/s, which correspond to 1% SoS contrast relative to the background, which is in the range of pathophysiological changes in the breast tumor (Sanabria et al. 2018). P8-P10 simulate smoothly varying SoS distributions. P10 was generated by multiplying P7 with a raised-cosine window with 0.5 roll-off. P9 and P10 simulate vertical and horizontal gradients in the background SoS, with a cosine reaised between 1540 m/s and 1554 m/s with 1.0 roll-off.

    Relative time measurements τ_m were simulated at each cell c by comparing plane wavefronts at angles $\theta_1=20°$ and $-20°$ with respect to $\theta_2=0°$. To isolate image degradation effects due to the incomplete reconstruction problem, τ_m values were directly generated using ray tracing for calculating $\mathbf{L}$ and the same geometric weights $\mathbf{L}$ were used in the inverse problem. To study sensitivity to errors in time-delay estimation,

time-delay errors were synthesized and added to the τ_m matrix, following a Gaussian distribution with zero mean and a standard deviation expressed as percentage with $\{1,2,5,10,20,50,56,63\}\%$ of the peak τ_m value for each simulated example, and hereafter referred to as noise level. Frequency Domain Reconstruction (FDR) (Jaeger et al. 2015) performed evaluations using a similar numerical simulation scheme, therefore allowing us for a direct comparison herein.

An example of τ_m is plotted in Fig. 2 for tissue model P8 and a noise level of 10%. The inclusion SoS is higher than the background. This leads to values $\tau_m > 0$ below the left edge of the inclusion region (P), where the path θ_2 through the background is slower than the reference path $\theta_1 = 0°$, which traverses the inclusion. Conversely, $\tau_m < 0$ are observed below the right corner of the inclusion region (Q), where only the path θ_2 traverses the inclusion and is thus faster than path θ_1.

3.2. Frequency- vs. Spatial-Domain Reconstruction

In Fig. 3 our proposed method Spatial Domain Reconstruction (SDR) is compared to FDR. For SDR, the proposed MA-AWTV was used with three gradient directions. The regularization parameters were adjusted via grid search separately for both FDR and SDR cases to report the best possible image quality with each. For FDR, this led to a regularization parameter $\gamma = 60$ for noiseless case, $\gamma = 1000$ for noise levels between 1% and 10%, and $\gamma = 3000$ for noise levels between 20 and 63%. For SDR, the same regularization constant $\lambda = 0.025$ was used at all noise levels. Missing τ_m readings for regions outside the overlap of angled plane-wave traversal were ignored implicitly by SDR, and were set to zero in FDR. Results at four measurement noise levels ($\{0,1,10,50\}\%$) are shown qualitatively here in Fig. 3, whereas all results are evaluated and compared quantitatively later in figures 6 and 7.

Fig. 3 shows that the FDR images present low contrast already at moderate noise levels of 10%, e.g. the inclusion geometries in P2 and P4 cannot be resolved. Image distortions such as resolution loss and strong streaking artifacts common with limited-angle computed tomography algorithms and were previously reported in (Jaeger et al. 2015) are also observed in the examples herein. Since ray paths orthogonal to the imaging direction are missing, the image resolution is reduced in the vertical direction y. Such strong streaking artifacts with FDR occlude the inclusions and also degrades the vertical resolution (e.g., smearing the two circles in P6 into one blob) already at noise level of 1%.

Given the same τ_m data, the proposed SDR method provides a visible image quality improvement with respect to FDR. The SDR method achieves piecewise delineation of homogeneous inclusions, effectively filtering out noise from the images. Both reconstructions at zero noise level and up to a noise level of 10% show consistent results, with minor image degradation. For instance, the two vertical inclusions in P6 are separable up to a noise level of 5% in Fig. 3. Any remaining artifacts include a loss of vertical resolution for very flat inclusions (P4) due to limited angular information,
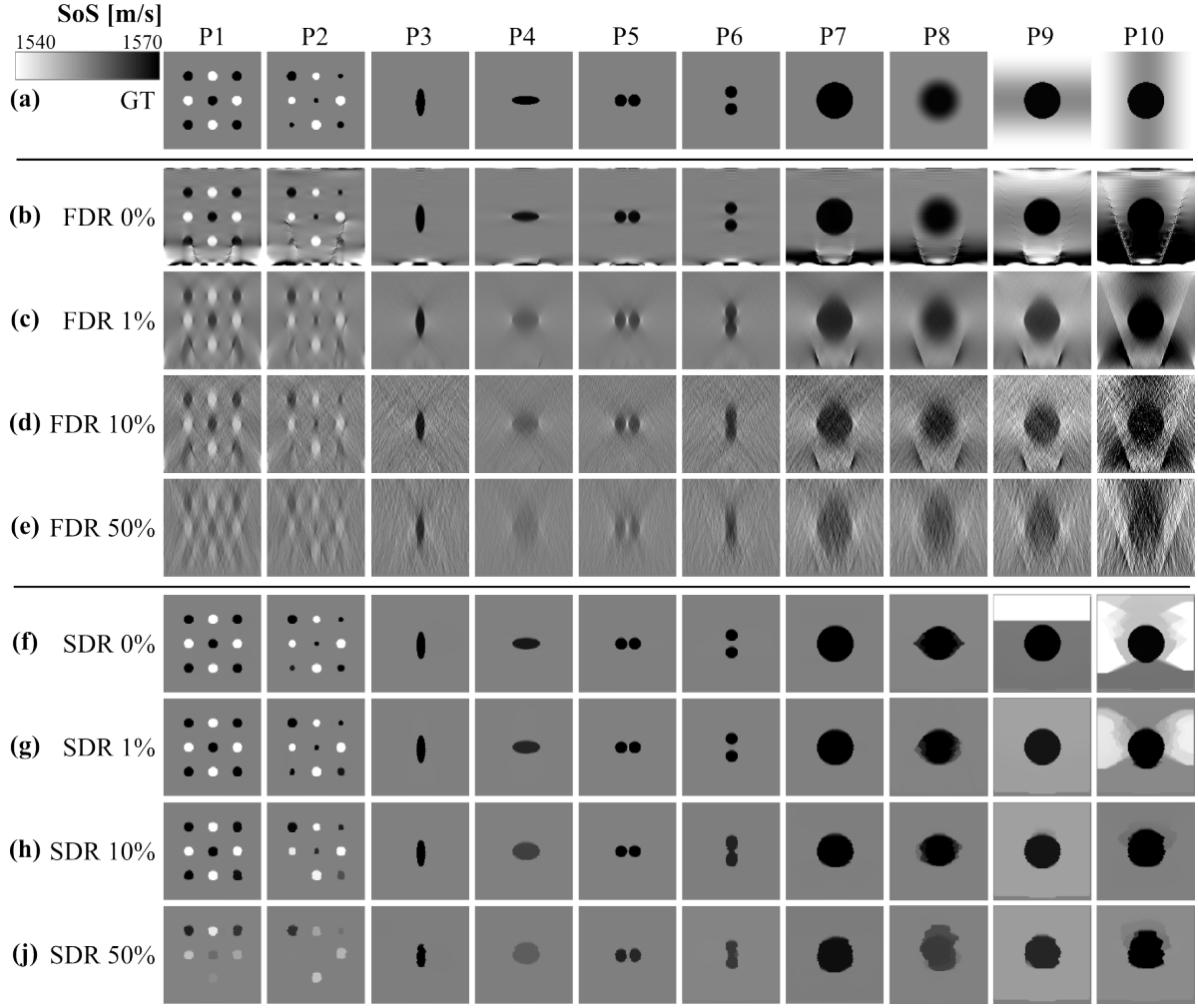
Figure 3. (a) Reconstructions of 10 different examples of artificial inclusions in homogeneous and non-homogeneous tissue samples, with the colorbar shown on top-left. SoS reconstruction using the baseline method FDR (b)-(e) and our proposed SDR with MA-AWTV (f)-(i) for different noise levels. Time measurements were simulated at angles $\theta_1{=}20°$ and $-20°$ with respect to the reference angle $\theta_2{=}0°$.

and the disappearance of one of the inclusions of P2 at a noise level of 10%, which corresponds to a corner region where only a few τ_m measurements are available. For the simulations of vertical and horizontal background (P9-P10), both methods could not yield an accurate reconstruction, this is due to the fact that our regularization assumes a piece-wise constant data.

3.3. Regularization Norm

We show in Fig. 4, SDR results with different data and regularization norms: ℓ_2-norm regularization, Total Variation (TV), Anisotropically-Weighted Total Variation (AWTV) with ℓ_1- and ℓ_2-norm in data term, and Multi-angle AWTV. For ℓ_2-norm regularization, $\lambda{=}0.0063$ was found to be optimal. For TV, AWTV and MA-AWTV,

the same regularization constant λ=0.025 as above was used. TV (A.1) provides piece-wise homogeneous inclusions effectively eliminating streaking artifacts; however, due to the limited angle information, a loss of resolution in the vertical direction y is still observed. AWTV with κ=0.9 (A.2) effectively regularizes the reconstructions with the available angular information while preserving the inclusion shape. AWTV with ℓ_1-norm in the data term (Fig. 4(e)) yields a minor improvement in reconstruction quality compared to AWTV with ℓ_2-norm data term (Fig. 4(d)). For remaining comparisons, ℓ_1-norm data term was used. The use of multiple regularization directions with MA-AWTV (9) improves the reconstruction for edges that are not axis-aligned. The proposed SDR is superior to FDR regardless of the data and regularization norm. Although FDR formulation is somewhat similar to SDR with ℓ_2-norm, the latter enables implicit handling of inaccurate and missing (out of plane-wave overlap) displacement readings, enabling better reconstructions seen herein. Of different SDR regularizations, throughout this paper we used our proposed MA-AWTV as the choice for further evaluations.

3.4. Sensitivity to Parameterization

SDR is observed to be robust against variations in the regularization constant λ, as shown in Fig. 5. A single choice of $\lambda^\star$=0.025 was sufficient for good performance for all inclusion geometries and noise levels. In addition, the reconstruction results
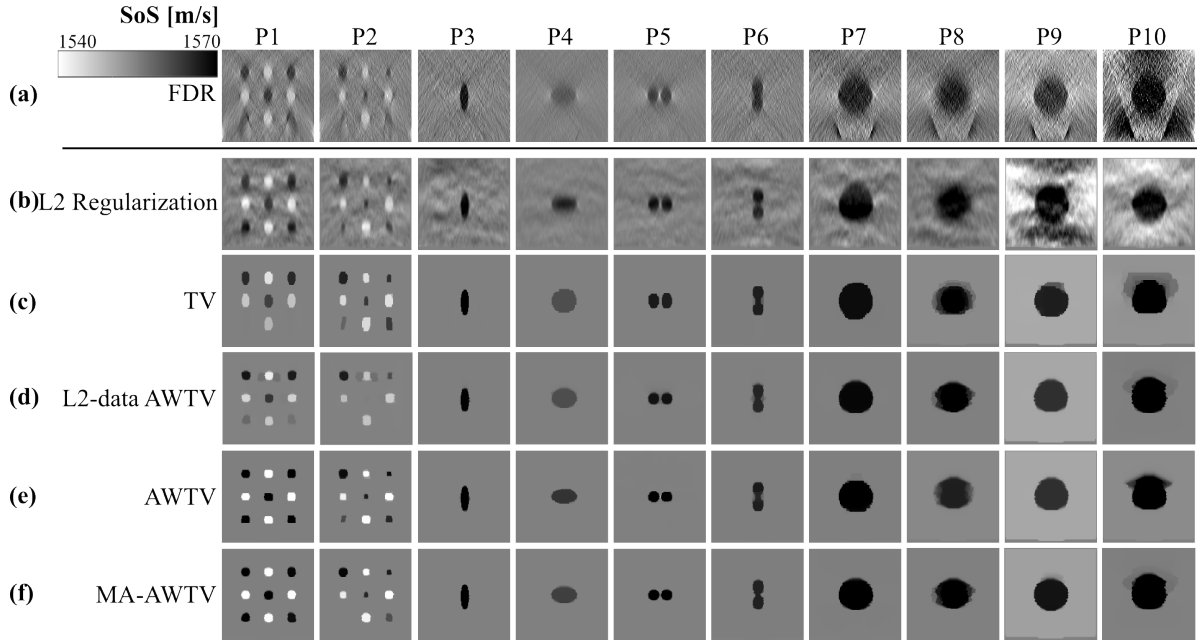

Figure 4. SoS reconstruction for artificial inclusions in Fig. 3 (a) and noise level of 10%, by using regularization alternatives. (a) FDR as baseline, (b) SDR with ℓ_2-norm regularization, (c) SDR with TV, (d) SDR with AWTV ($\kappa = 10$) with ℓ_2-norm in the data term, (e) SDR with AWTV ($\kappa = 10$) with ℓ_1-norm in the data term, (f) SDR with MA-AWTV (the same as Fig. 3(h).

were visually stable over an order of magnitude, i.e. $\lambda = [0.0062, 0.05]$. Too small λ (e.g., $\lambda = 0.00062$) introduce noise in the images, whereas too large λ lead to over-regularization, which smears the inclusion geometries.

3.5. Quantitative Metrics for Evaluation

We used the following metrics (Varghese & Ophir 1998) to quantitatively evaluate and compare reconstruction methods:

a) Background noise $\sigma_{\text{bg}}^{\star}$ calculated as the SoS standard-deviation in the background region

b) Root-Mean-Square Error as

$$\text{RMSE} = \sqrt{\frac{\sum_{c=1}^{C}(v_c^{\star} - v_c)^2}{C}}, \tag{13}$$

where $v_i^{\star}$ and v_i are the estimated and ground-truth SoS values at image cell c

c) Contrast Ratio in [%] as

$$\text{CR} = 100 \frac{|\mu_{\text{inc}}^{\star} - \mu_{\text{bg}}^{\star}|}{\mu_{\text{bg}}^{\star}} \tag{14}$$

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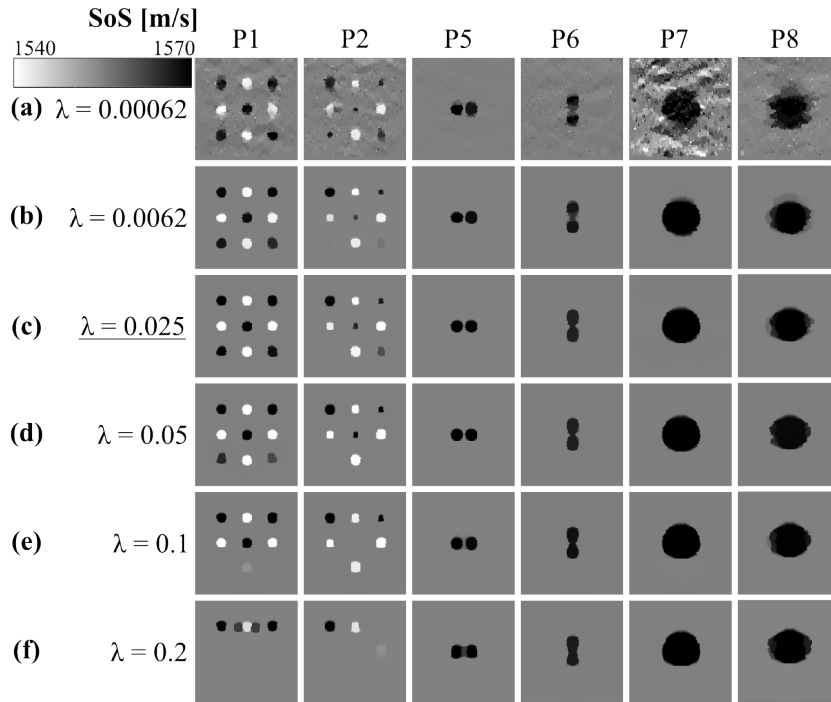

Figure 5. SoS reconstruction using our proposed SDR for different regularization constants λ at a noise level of 10%.

d) Dice coefficient of inclusion as

$$\frac{2|c_{\text{inc}}^{\star} \cap c_{\text{inc}}|}{|c_{\text{inc}}^{\star}| \cup |c_{\text{inc}}|} \tag{15}$$

where $c_{\text{inc}}^{\star}$ and c_{inc} are the regions of the reconstructed and ground-truth inclusions, respectively. 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.

RMSE provides a global quantitative assessment of image reconstruction quality, while $\sigma_{\text{bg}}^{\star}$ provides the noise floor. CR has a quantitative diagnostic meaning, since it describes the SoS contrast fidelity with which inclusions can be reconstructed. For instance, for an inclusion with a SoS contrast of 1%, a CR = 0.5% indicates that the inclusion is reconstructed with only half of its actual physical SoS contrast. Therefore, in a diagnostic setting, a low CR would increase the number of false negatives and decrease *Negative Predictive Value (NPV)*. The Dice coefficient enables a quantitative assessment of geometric fidelity for inclusion delineation in a synthetic setting, where ground-truth contrast is assumed to be known for a reconstruction-independent delineation threshold.

These metrics were used to evaluate our proposed method SDR and the baseline FDR for different noise levels and the geometries shown in Fig. 3. For each evaluation metric, the average over all ten simulation examples are shown in Fig. 6. All metrics show superior performance with SDR. For 1% noise level (SNR=40 dB) the CR is improved from 0.5% with FDR to 0.99% with SDR, in excellent agreement with respect to the ground truth at 1%. For moderate noise levels (10%, SNR=20 dB) SDR still achieves a sufficient contrast of 0.90%, while the contrast of FDR drops to 0.37%. For very large noise levels (for instance, 50%, SNR=6 dB), the CR of SDR drops to 0.67%. If prior information about the inclusion geometry is available, the quantitative prediction is significantly improved, as seen with CR=0.90% at the same noise level, as in Fig. 6(b). RMSE metrics for smooth inclusion geometries are plotted in Fig. 7. The sharp P7 and smooth P8 inclusions show comparable RMSE. TV may filter out SoS variations in the reconstructions, which is observed by P9 and P10 showing relatively higher RMSE. In all cases, SDR resulted in RMSE values lower than FDR, while also presenting a better agreement with the ground truth.

3.6. Multi-angle Reconstruction

Reconstruction quality can also be improved by combining measurements from several plane wave angles, instead of using only two angles as was in Fig. 3. SDR is better suited than FDR for multi-angle reconstruction, since SDR can inherently handle and seamlessly combine missing displacement readings from different angles. This is seen with increasing number of τ_m observations in Fig. 8, where displacements with respect to $\theta_2=0°$ for 2 angles at $\theta_1=\{\pm10\}°$ were compared to 6 angles at $\theta_1=\{\pm10, \pm20, \pm30\}°$ and 12 angles at $\theta_1=\{\pm5, \pm10, \pm15, \pm20, \pm25, \pm30\}°$. A noise level of 50% was simulated to represent a scenario of poor speckle tracking performance. While FDR does not show
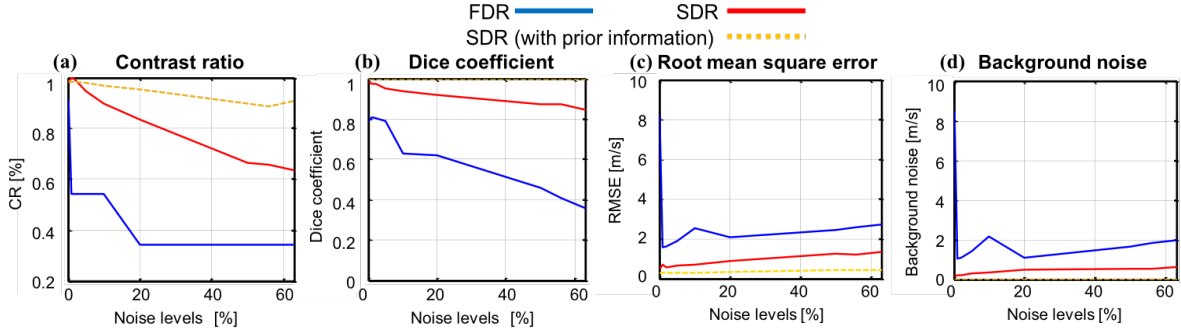
Figure 6. Quantitative metrics evaluated for baseline FDR and our proposed method SDR at different noise levels, where average metrics from all simulation geometries in Fig. 3 are plotted: (a) Contrast Ratio (CR); (b) Dice coefficient of inclusion; (c) root mean-square error (RMSE); and (d) background noise σ_{bg}^*.
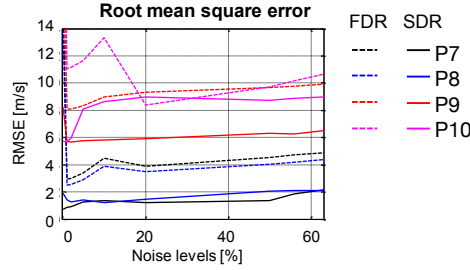

Figure 7. Quantitative RMSE metrics for smooth digital phantoms P7-P10 evaluated for FDR and our proposed SDR at different noise levels. Plots show the average across 5 noise realizations.

much visual improvement, SDR results significantly improve with increasing number of plane-wave angles. For instance, smaller inclusions are resolved in examples P1 and P2, and the vertical inclusions in P6 are separated. This is also confirmed by the quantitative metrics (Fig. 9). By using 6 instead of 2 angles, the CR at noise level 50% improves from 0.67% to 0.86% and the Dice coefficient from 0.87 to 0.94. For 12 angles, the CR is 0.87% and the Dice coefficient is 0.95, which shows an asymptotic improvement up to a saturation point. Angular steps smaller than 5° did not bring any significant improvement. If prior information is available, the CR improves to 93%.

3.7. Numerical Wave Simulations

The synthetic dataset experiments and studies above are limited due to their assumption of ideal data created geometrically. As a more realistic simulation, we performed numerical simulations of wave propagation. A tissue domain of size 45 mm ×40 mm was simulated, where an inclusion of 5 mm radius was located at the depth of 20 mm. The inclusion had a speed-of-sound of 1548 m/s and the background 1515 m/s, which were simulating a tumor with 2.2% speed-of-sound contrast. The sampling frequency of the transducer was set to 5 MHz.
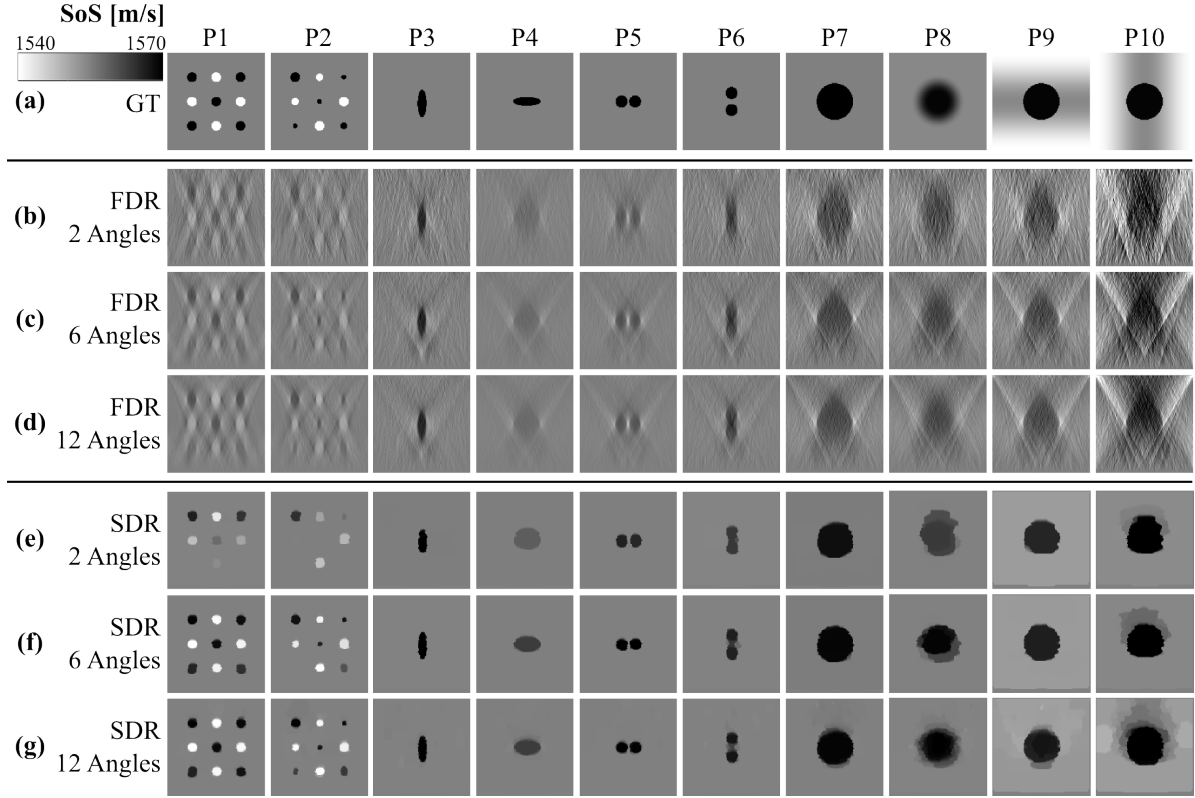
Figure 8. SoS reconstruction at 50% noise level, using FDR (b-d) and our proposed method SDR with MA-AWTV (e-g). Reconstructions with time measurements at multiple path angles: (b,e) 2 angles at $\theta_1=\{\pm20\}°$, (c,f) 6 angles at $\theta_1=\{\pm10,\pm20,\pm30\}°$ and (d,g) 12 angles at $\theta_1=\{\pm5,\pm10,\pm15,\pm20,\pm25,\pm30\}°$ with respect to the reference angle $\theta_2=0°$.
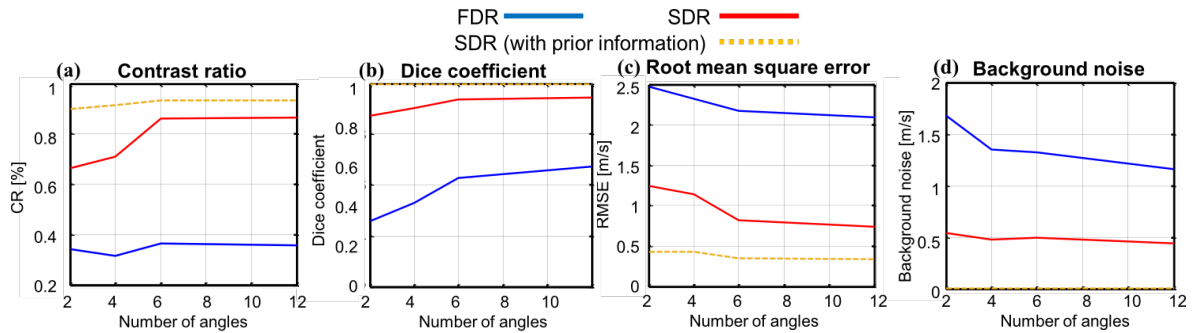

Figure 9. Quantitative metrics for baseline FDR and our proposed method SDR with time measurements at multiple path angles. (a) shows Contrast Ratio (CR), (b) Dice coefficient of inclusion, (c) Root mean square error (RMSE), and (d) Background noise σ_{bg}^*. The average of all simulation geometries in Fig. 8 is plotted.

For the simulation we used finite-difference time-domain (FDTD) method (Sanabria et al. 2015) in 2D. The model implements elastic wave propagation in terms of particle velocity v_p and stress tensor σ_{pq}, which leads to a system of equations $\partial_t v_p = \rho^{-1} \partial_q \sigma_{pq}$ and $\partial_t \sigma_{pq} = C_{pqrs} \partial_s v_r$ in function of the density ρ, the elasticity tensor C_{pqrs}, and the corresponding stiffness properties. A pixel size of $h_{\text{FDTD}} = 5\,\mu m$ and a time step $h_t = 1.5\,\text{ns}$ were used, which allowed to accommodate the bandwidth and dispersion effects, providing accurate time measurements down to 5 ns for known reflector positions (Sanabria & Goksel 2016). Multi-static data was simulated by running N=128 simulations for each individual Tx excitation. Due to the high time precision required, the simulations are computationally expensive, requiring 2.5 h per Tx element with a Tesla K40c GPU, while the synthetic tests above can be simulated in a few seconds. The results were binned to a pixel of 0.05 mm and a sampling frequency of 24 MHz. A convolution-based scattering model (Bamber & Dickinson 1980, Mattausch & Goksel 2018) with 120 scatterers-per-mm^2 was next simulated and used to modulate the Tx-Rx channel data via fast convolution in the frequency domain. This allowed for having high-frequency scattering effects that may go beyond the FDTD bandwidth.

Plane-wave data was then synthesized from the simulated multi-static data between $\theta = -15°$ and $15°$ with $0.5°$ steps and these images were beamformed with conventional delay-and-sum algorithm. The displacement tracking algorithm described in Section 2.4 was applied between consecutive pairs of beamformed images. A signal correlation block of $1 \times 1\,mm^2$ (52×3 samples) was used, with a search range of $0.125\,\mu s$ (5 samples). The calculated delays were cumulated for time measurements τ_m with respect to $0°$, i.e. for images at angles $\{-15, -13, .., 15\}°$ with $2°$ increments. As regularization constants, $\gamma = 5000$ for FDR and $\lambda = 0.05$ for SDR were used.

Fig. 10(a) shows wave propagation snapshots for a plane wave with $\theta = 0°$, where all 128 elements are excited, before it reaches the inclusion and at the center of the inclusion. As one can see, the wave diffracts once it reaches the inclusion due to the speed-of sound difference and a phase difference is observed. The B-mode image of the domain is shown in Fig. 10(b). Since the same speckle density was used in inclusion and background, the inclusion appears as isoechoic. Two bright lines are observed on top and below the inclusion, corresponding to reflected ultrasound waves. These are generated at the inclusion boundaries due to the acoustic impedance change. Fig. 10(c) shows the input time measurements τ_m at $\theta = 15°$. A similar triangular pattern in Fig. 2 can be observed, which indicates an increased speed-of-sound in the inclusion region. Fig 10(d) shows results for the baseline FDR algorithm, in comparison with the proposed SDR reconstruction, as in Fig 10(e). *Lobular artifacts* are observed at the horizontal boundaries of the inclusion, which are associated to mode interference for rays that are incident (quasi)parallel to the inclusion edges, for which, according to Snell's law, strong refractions occur. Speckle decoherence is observed at the boundaries of the plane wavefront and in the near field region, which lead to shadow regions in the speckle pattern (Fig. 10(c)). Overall, these effects have a much larger impact on FDR (Fig. 10(d)), for which the inclusion position is distorted vertically by 2.4 mm, and SoS

deviations (11.3 m/s, 0.72%) and strong artifacts are observed. The SDR reconstruction shows a satisfactory inclusion delineation, with SoS values close to the baseline (error 0.3 m/s, 0.02%). Given the large area of missing displacement data, only a minor artifact of SoS reduction is observed in SDR, beneath the stiff inclusion at a depth of 30mm.

The full-wave simulations show that, even if the exact speckle positions are not known a priori (due to the unknown SoS distribution), relative displacements between the same speckle patterns observed from different angular directions allow still to reconstruct SoS images, with the uncertainties in the assumed geometric matrix **L** not having a large impact in the reconstruction. The approximation of (4) is shown to be valid since SoS variations in biological tissues are in general small, thus preserving speckle coherence.

3.8. In-Vivo Data

To illustrate the clinical applicability of our method, we herein present first preliminary *in-vivo* results. The data was recorded from a female patient (of 80-90 year-old age group) with a breast lesion, which the biopsy later revealed as Invasive Ductal Carcinoma (IDC). This preliminary data was acquired at the University Hospital of Zurich, as approved by the institutional review board and local ethics committee.

Conventional B-mode ultrasound images and multi-static (full-matrix) data were acquired using a SonixTouch ultrasound machine (Ultrasonix Medical Corporation, Richmond, BC, Canada). We used a standard commercial ultrasound probe (L14-5, Ultrasonix). The probe was operated at 5 MHz and is a linear array of $N{=}128$ transducer elements with a pitch of 0.3 mm between elements, an element elevation of 7 mm, and a total aperture of 38 mm. Ultrasound echoes were recorded as a function of time. The probe provides two-dimensional ultrasound imaging in a plane perpendicular to the transducer elements, with the width of the image corresponding to the linear array direction. Multi-static data was generated by sequentially transmitting an ultrasound pulse by a single transducer element and simultaneously recording the echoes received by all receiver elements. A multi-channel data acquisition board (SonixDAQ, Ultrasonix)
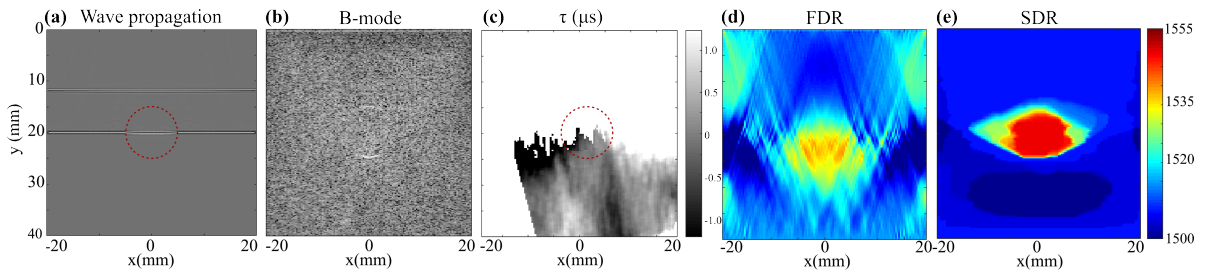

Figure 10. FDTD simulation of a domain with a circular inclusion at depth 20 mm. (a) Wave propagation at different depths. (b) B-mode image of the domain. (c) Displacement tracking for plane wave with $\theta = 15°$. (d) FDR baseline reconstruction. (e) Our proposed SDR reconstruction. An angular range $[-15° : 15°]$ is used in both cases.

helps collect raw pre-beamformed data in parallel with a sampling frequency of 40 MHz and 12 bits per sample. The process is repeated for all transmitter elements, so that a total of N^2 time traces were recorded in $< 0.1\,s$ (about 100 MB).

Plane-wave data was synthesized from the multi-static data between $\theta=-22.5°$ and $22.5°$ with $0.5°$ steps and the acquired images were beamformed with conventional delay-and-sum algorithm. The displacement tracking algorithm described in Section 2.4 was applied between consecutive pairs of beamformed images. This small angle step of $0.5°$ between the image pairs prevented any speckle decorrelation. A signal correlation block of $1 \times 1\,mm^2$ (52×3 samples) was used, with a search range of $0.125\,\mu s$ (5 samples). The calculated delays were accumulated for time measurements τ_m with respect to $0°$, i.e. for images at angles $\{-22.5, -17.5, .., 22.5\}°$ with $5°$ increments. As regularization constants, $\gamma=5000$ for FDR and $\lambda=0.07$ for SDR were chosen, for leading to visually best results. No prior information about the inclusion geometry was used.

Conventional ultrasound B-mode image is seen in Fig. 11(a). The B-mode image consists of a hypo-echoic region followed by some acoustic shadowing. An accurate delineation of the lesion was not possible from this B-mode image, but only its lateral extent could be estimated by the clinical expert. Fig. 11(b) shows the tracked displacements τ_m at $\theta = 12.5°$. In accordance with Fig. 2, positive and negative τ_m are respectively observed below the left and right sides of the inclusion region, which indicates an increased SoS in the inclusion region. Fig 11(c)-(d) show reconstruction results for the baseline FDR algorithm in comparison with our proposed SDR. Reconstruction color ranges were scaled individually to fit all SoS values reconstructed by each method.

The time measurements τ_m in Fig. 11 show large regions where no reliable displacement information could be extracted reliably (i.e. correlation coefficient $q_{\mathrm{ZNCC}} < 0.5$). These are mostly the areas close to the transducer where near-field effects dominate and a plane wavefront has not yet established as in Fig. 2. Moreover, in low-echogenicity regions around the breast lesion, the signal falls below the noise floor of the system, hindering an effective displacement estimation.
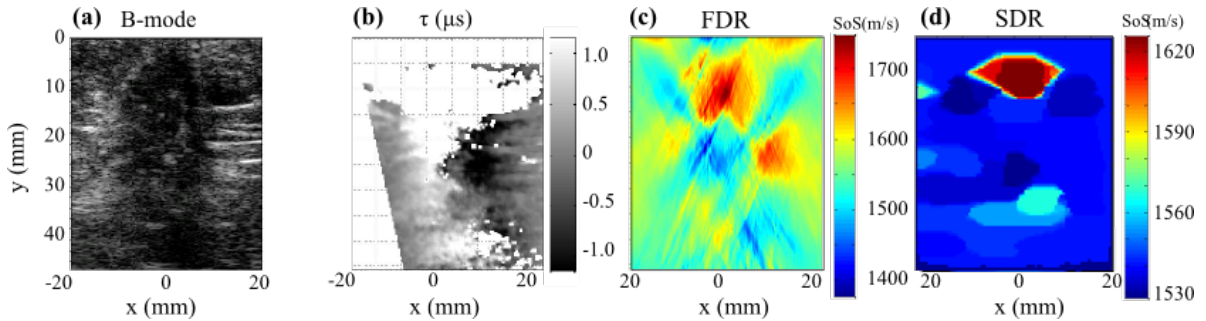

Figure 11. *In vivo* SoS imaging of an 88 year old woman with breast cancer : (a) B-mode image; (b) displacement tracking for plane wave with $\theta_1 = 12.5°$; (c) reconstruction with FDR; and (d) reconstruction with proposed SDR. An angular range of $(-22.5°, 22.5°)$ was used.

The baseline method FDR requires a complete set of $\boldsymbol{\tau}$ measurements for the entire tissue region to perform the Fourier Transform, whereas our proposed SDR method can effectively handle any missing tracking results, while still leveraging information from other tracked locations. The combination of a large part of measurement data missing, together with the noise of the speckle measurements, leads to a poor reconstruction with FDR, in which the tumorous inclusions cannot be delineated. In particular, streaking artifacts are observed. In Fig. 11(c), a SoS reduction is observed behind the tumorous inclusion (x = 0, y = 25). Moreover, a SoS elevation is observed at (x = 15, y = 25), where the corresponding B-mode image shows no tissue distortion. In contrast, our SDR algorithm delineates an inclusion geometry at (x = 0, y = 10) above the mid-range SoS value (1580 m/s), as a single focal region as confirmed by biopsy and with a lateral extent in agreement with that observed in B-mode. The sharp delineation of the lesion is remarkable, considering the lack of time measurement information in and around the tumor region. Similarly to the numerical example in Fig. 10, a potential artifact of reduced SoS is observed beneath the tumor around a depth of 25mm. Nevertheless, this cannot be confirmed as it can also be an actual heterogeneity in the breast tissue, and will be the focus of future evaluations.

3.9. Accommodating Missing Information

For testing the handling of dropout regions above, we evaluated the in-vivo data, where we first segmented the breast tumor geometry for carcinoma case in Fig. 11(d) as shown below in Fig. 12(a). We then simulated relative delay measurements τ_m and estimated speed-of-sound values while ignoring (b) and using (c) the dropout region mask in Fig. 11(b). For both cases, SDR outputs are very close to ground-truth, with a SoS error in the inclusion of 5 m/s (0.3%), where as FDR yields non-aligned reconstructions, with 5.1 mm vertical displacement for c), and strong deviations between both cases (20 m/s error, 1.3%). Overall, the time delay information relevant to the reconstruction is located below the tumor region, so drop out regions within the tumor or above have a minor impact in the reconstruction. Additional tests (not shown) were performed for synthetic geometries and different drop out region sizes, with 30% image dropouts being handled effectively by SDR with minor distortion. For instance, in comparison to Fig. 10(e), where artifacts of SoS reduction are observed in SDR reconstruction due to large regions of missing displacements, in Fig. 12(c) the missing displacement information covers a smaller region in the near field and these artifacts are no more observed in the SDR results, while being still present in FDR.

4. Discussion and Conclusions

We propose a spatial-domain reconstruction method to calculate SoS distributions in an arbitrary grid from the time measurements recorded by a conventional linear-array transducer, given defined ultrasonic propagation paths. We assume in our imaging
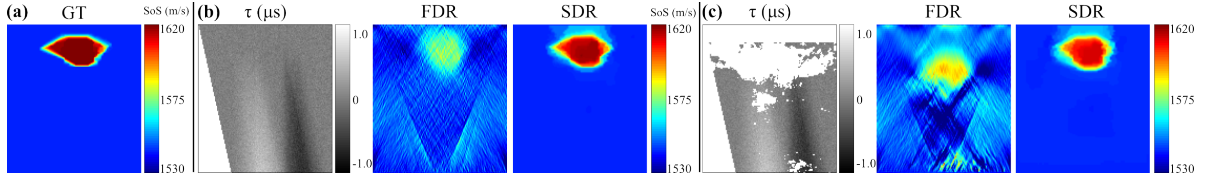
Figure 12. Simulation of in vivo speed-of-sound imaging output for testing the handling of drop-out regions. (a) Segmentation of breast tumor; (b) simulated full displacement tracking for plane wave with $\theta = 12.5°$ (at noise level of 10%), and reconstructions with FDR and SDR; and (c) results with simulated displacement tracking in which the experimentally erroneous info was masked out. An angular range of $(-22.5°, 22.5°)$ was used in both cases.

model that speed of sound variations in the tissue are small, within a few percent. Accordingly, ultrasound wavefronts are considered to propagate as straight ray paths of minimal cross-section, thus neglecting ray trajectory variations due to physical effects such as refraction, diffraction and interference effects. Our results of reconstructions on full-wave numerical simulation data indeed show that this simplified propagation model can produce satisfactory reconstructions. Speed-of-sound variation in function of ultrasound pulse frequency is known to be small in human tissues (Levy et al. 2007), and at low transmit powers, amplitude-dependent nonlinearities are minimal; thus wave dispersion was considered herein to be negligible, therefore calculating a single group velocity for speed-of-sound.

Quantitative speed-of-sound images are obtained, which can provide a biomarker for tissue diagnosis and differentiation. Our numerical experiments illustrate that the proposed SDR method allows for accurate reconstruction of sound speed maps with high spatial resolution and contrast, and low RMSE using a minimum of only two plane-wave transmit angles. The regularization parameter λ is used to regularize the inverse-problem using the geometric information based on the transmitter-receiver setup of the transducer and local ray-path information. Due to limited angle tomographic nature, especially at higher noise levels, the vertical resolution of our reconstructed images decrease. With our pulse-echo approach using a conventional transducer, increased number of angles for reconstruction are possible using higher steering angles, given that transducer elements have sufficiently wide main lobe response. Higher steering angles, however, limit the overlapping field-of-view (FoV) where scattering shifts and thus reconstructions can be performed. Thus, to keep the FoV sufficiently large at higher steering, larger transducer arrays (apertures) are in turn necessitated. One key advantage of our method with respect to the previous state of the art is that it can exclude regions of missing information in the reconstruction, therefore satisfactorily accommodating varying overlapping FoVs between different angular pairs. Besides these limitations, our solution strategy can easily be extended to other transducer geometries by taking this into account in our forward problem (**L**) formulation in Eq. (4).

Speed-of-sound values are known to vary among patients and with pathology, and different authors reported varying SoS values for malignant tissue (Li et al. 2009, Duric

et al. 2010). Moreover, breast density differences may already present SoS variations in healthy patients between approximately 1400 m/s (fatty breast) and 1550 m/s (extremely dense breast) (Sanbria et al. 2018, Sak et al. 2015). The average SoS value observed for the *in-vivo* results in Fig. 11(d) is 1618 m/s ($\pm$12 m/s standard deviation), corresponding to a 4.8% SoS increase with respect to the background SoS in that image. This finding corroborates with prior speed-of-sound observations with full-angle USCT of the breast, e.g. in (Andre et al. 2008). Although this is a promising in-vivo result, further clinical evidence is required to confirm any observations and to assert medically-relevant conclusions.

The inverse problem can be solved, even if time measurements are only available for partial regions of the tissue domain. Therefore, the reliability of the measurements can be directly incorporated, for instance, in terms of a correlation coefficient. Prior inclusion delineations and arbitrary complex path geometries can be easily incorporated with simple relations (scaling and additions). The latter can allow in the future to define arbitrary measurement configurations beyond plane wave beamforming and correct for physical effects such as wave refraction.

Our proposed method uses measurements of observed echo phase delays between different lines of propagation, which thus uses pulse-echo measurements from single-sided tissue access, in contrast to transmission USCT, e. g. , with circular arrays (Duric et al. 2010, Koch et al. 2015) and to those techniques that require a passive reflector on the other side of the investigated tissue (Krueger et al. 1998, Huang & Pai-Chi 2005, Sanabria & Goksel 2016). In contrast to other methods which use mis-registration of ultrasound images when scanning the tissue from different angles (Krücker et al. 2004, Shin et al. 2010, Cho et al. 2014), our proposed method utilizes the echo shift measurements, which is much more sensitive to sound-speed variations than use of the envelope(Jaeger et al. 2015). In comparison to beam tracking technique (Robinson et al. 1991, Kondo et al. 1990, Cespedes et al. 1992), in which two separated transducers generate intersecting beams, our method can provide finer SoS resolution. We have observed a loss of coherence and regions of missing information in the near field, which the SDR reconstruction handles satisfactorily. An optimization and comparison of different speckle tracking approaches, which was out of scope in this work, would be still desirable in the future to improve displacement tracking.

The proposed method contributes to a group of recent developments, which can be implemented as an add-on to conventional ultrasound equipment, particularly for focal disease detection and diagnosis. Apart from breast scanning, a broad range of applications can be envisaged for the musculoskeletal system, liver, kidney, spleen, prostate, brain, and thyroid. In general, the proposed method can be used at any organ site where conventional ultrasound imaging is used today. Future extensions are envisaged to obtain other acoustic parameters, such as the acoustic attenuation, to increase the reliability of the diagnosis across complex tumor populations. With computational optimizations and use of GPUs, our ultimate goal is to display speed-of-sound as a complementary modality overlaid on or shown side-by-side with standard

B-mode ultrasound. We have herein focused on widely available pulse-echo ultrasound linear probes, however, the proposed method is applicable to other probe geometries (convex, three-dimensional) and ultrasound imaging setups, for instance, Automatic Breast Ultrasound (ABUS)(Wenkel et al. 2008).

Acknowledgments

Appendix A. Anisotropically-Weighted Total Variation

Considering limited-angle nature of our USCT problem, one can weight individual directional derivatives differently, leading to the general form

$$||\mathbf{D}\boldsymbol{\sigma}||_1 = \frac{1}{2}\sum_{i,j}\left(|\sigma_{i+1,j} - \sigma_{i,j}| + |\sigma_{i,j+1} - \sigma_{i,j}|\right), \qquad (A.1)$$

where i and j are discretization indices, in horizontal (x) and vertical (y) directions. Taking the reconstruction outcome for different cost-function norm into account, ℓ_1-norm was observed to improve image quality compared to ℓ_2 norm, potentially reducing detrimental effects from outliers and allowing for an effective scaling between the two terms (Boyd & Vandenberghe 2009).

We next extend the TV regularization to incorporate the available angular information for each ray path. This has been shown recently to improve reconstructions of an incomplete USCT problem (Sanabria & Goksel 2016). For defined wave paths, the angular tomographic disparity $[-\theta, \theta]$ available for each cell can be different at different locations of the imaging field-of-view, due to the limited imaging aperture. Nevertheless, such angular tomographic coverage is known geometrically for each cell, and can be introduced accordingly in a spatial reconstruction. Hence, one can weight the SoS (slowness) gradient contributions in different angular directions according to the available ray information in these directions. The resulting regularization, combined with the TV approach, is called "Anisotropically Weighted Spatial Regularization" (AWTV).

In the simpler AWTV implementation (Sanabria & Goksel 2016), a constant κ is used as weight to balance horizontal and vertical gradients according to the available ray information in each direction:

$$||\mathbf{D}\boldsymbol{\sigma}||_{\mathrm{AWTV}} = \sum_{i,j}\kappa|\sigma_{i+1,j} - \sigma_{i,j}| + (1-\kappa)|\sigma_{i,j+1} - \sigma_{i,j}|. \qquad (A.2)$$

In ultrasound B-mode imaging, resolution is typically higher in the axial direction of beam propagation. In computed tomography, however, reconstruction resolution depends not only on the separability of backscattered echoes but also on the angular coverage of projections. In limited-angle tomographic reconstruction problems such as

the one presented here, missing projection directions prevent resolving gradients in the orthogonal direction, therefore, a larger regularization is applied in these directions to correctly delineate spatial transitions. With a conventional ultrasound transducer, as in Fig. 1(a), ultrasound beams naturally propagate in the vertical direction Y or with a moderate inclination θ with respect to Y, while ultrasound beams in the horizontal direction X cannot be synthesized. Therefore, we set a larger regularization for X gradients than for Y gradients, in order to accommodate a higher constraint in the directions where boundaries can be resolved. In particular, for the imaging setup discussed in this paper, an optimum performance is attained with $\kappa = 0.9$.

References

Andre, M. P., Barker, C., Sekhon, N., Wiskin, J., Borup, D. & Callahan, K. (2008). Pre-clinical experience with full-wave inverse-scattering for breast imaging, *Acoustical Imaging*, Vol. 29, pp. 73–80.

Arribas, E. M., Whitman, G. J. & De Bruhl, N. (2016). Screening breast ultrasound: Where are we today?, *Curr Breast Cancer Rep* **8**: 221–229.

Bamber, J. C. & Dickinson, R. J. (1980). Ultrasonic B-scanning: A computer simulation, *Physics in Medicine and Biology* **25**(3): 463–479.

Bamber, J. C. & Hill, C. R. (1981). Acoustic properties of normal and cancerous human liver - I. dependence on pathological condition, *Ultrasound in Medicine and Biology* **7**(2): 121–133.

Boyd, S. & Vandenberghe, L. (2009). *Convex Optimization*, Cambridge University Press.

Byram, B. C., Trahey, G. E. & Jensen, J. A. (2012). A method for direct localized sound-speed estimates using registered virtual detectors, *Ultrasonic Imaging* **34**(3): 159–180.

Cespedes, I., Ophir, J. & Huang, Y. (1992). On the feasibility of pulse-echo speed of sound estimation in small regions: Simulation studies, *Ultrasound in Medicine and Biology* **18**(3): 283–291.

Cho, S., Kang, J., Kang, J., Lee, Y. & Yoo, Y. (2014). Phantom and in-vivo evaluation of sound speed estimation methods: preliminary results, *Proc. IEEE International Ultrasonics Symposium*, pp. 1678–81.

Cosgrove, D., Piscaglia, F., Bamber, J., Bojunga, J., Correas, J. M. & Gilja, O. H. (2013). EFSUMB guidelines and recommendations on the clinical use of ultrasound elastography. part 2: Clinical applications, *Ultraschall in Medicine* **34**: 238–253.

Duric, N., Boyd, N., Littrup, P., Sak, M., Myc, L., Li, C., West, E., Minkin, S., Martin, L., Yaffe, M., Schmidt, S., Faiz, M., Shen, J., Melnichouk, O., Li, Q. & Albrecht, T. (2013). Breast density measurements with ultrasound tomography: A comparison with film and digital mammography, *Medical Physics* **40**(1).

Duric, N., Littrup, P., Chandiwala-Mody, P., Li, C., Schmidt, S., Myc, L. & Rama, O. (2010). In-vivo imaging results with ultrasound tomography: Report on an ongoing study at the karmanos cancer institute, *Proceedings of SPIE Vol. 7629*.

Eskandari, H., Salcudean, S., Rohling, R. & Ohayon, J. (2008). Viscoelastic characterization of soft tissue from dynamic finite element models, *Physics in Medicine and Biology* **53**(22): 6569–6590.

Glozman, T. & Azhari, H. (2010). A method for characterization of tissue elastic properties combining ultrasonic computed tomography with elastography, *Journal of Ultrasound in Medicine* **29**(3): 387– 398.

Goksel, O., Eskandari, H. & Salcudean, S. E. (2013). Mesh adaptation for improving elasticity reconstruction using the FEM inverse problem, *IEEE Transactions on Medical Imaging* **32**(2): 408–418.

Grant, M. & Boyd, S. (2014). CVX: Matlab software for disciplined convex programming, version 2.1, http://cvxr.com/cvx.

Holmes, C., Drinkwater, B. W. & Wilco, P. D. (2005). Post-processing of the full matrix of ultrasonic transmit-receive array data for non-destructive evaluation, *NDT & E International* **38**: 701–711.

Hoskins, P., Martin, K. & Thrush, A. (2010). *Diagnostic Ultrasound: Physics and Equipment*, Cambridge University Press.

Huang, S. W. & Pai-Chi, L. (2005). Ultrasonic computed tomography reconstruction of the attenuation coefficient using a linear array, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **52**(2): 2011–2022.

Imbault, M., Faccinetto, A., Osmanski, B. F., Tissier, A., Deffieux, T., Gennisson, J. L., Vilgrain, V. & Tanter, M. (2017). Robust sound speed estimation for ultrasound-based hepatic steatosis assessment, *Physics in Medicine and Biology* **62**(9): 3582–3598.

Jaeger, M. & Frenz, M. (2015a). Quantitative imaging of speed of sound in echo ultrasonography, *Proceedings of IEEE International Ultrasonics Symposium*.

Jaeger, M. & Frenz, M. (2015b). Towards clinical computed ultrasound tomography in echo-mode: Dynamic range artefact reduction, *Ultrasonics* **62**(1): 299–304.

Jaeger, M., Held, G., Peeters, S., Preisser, S., Grnig, M. & Frenz, M. (2015). Computed ultrasound tomography in echo mode for imaging speed of sound using pulse-echo sonography: proof of principle, *Ultrasound in Medicine and Biology* **41**(1): 235–250.

Jeong, J. W., Shin, D. C., Do, S. H., Blanco, C., Klipfel, N. E., Holmes, D. R., Hovanessian-Larsen, L. J. & Marmarelis, V. Z. (2008). Differentiation of cancerous lesions in excised human breast specimens using multiband attenuation profiles from ultrasonic transmission tomography, *Journal of Ultrasound in Medicine* **27**(3): 435–451.

Kak, A. C. & Slaney, M. (1988). *Principles of Computerized Tomographic Imaging*, IEEE Press.

Koch, A., Stiller, F., Lerch, R. & Ermert, H. (2015). An ultrasound tomography system with polyvinyl alcohol (PVA) moldings for coupling: in vivo results for 3-D pulse-echo imaging of the female breast, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **62**(2): 266–279.

Kondo, M., Takamizawa, K., Hirama, M., Okazaki, K., Linuma, K. & Takehara, Y. (1990). An evaluation of an in-vivo local sound speed estimation technique by the crossed beam method, *Ultrasound in Medicine and Biology* **16**(1): 65–72.

Krücker, J. F., Fowlkes, J. B. & Carson, P. L. (2004). Sound speed estimation using automatic ultrasound image registration, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **51**(9): 1095–1106.

Krueger, M., Burow, V., Hiltawsky, K. M. & Ermert, H. (1998). Limited-angle ultrasonic transmission tomography of the compressed female breast, *Proceedings of IEEE International Ultrasonics Symposium*.

Levy, Y., Agnon, Y. & Azhari, H. (2007). Ultrasonic speed of sound dispersion imaging, *Ultrasound in Medicine and Biology* **33**(5): 762–767.

Li, C., Duric, N., Littrup, P. & Huang, L. (2009). In vivo breast sound-speed imaging with ultrasound tomography, *Ultrasound in Medicine and Biology* **35**(10): 1615–1628.

Lubinski, M. A., Emelianov, S. Y. & ODonnell, M. (1999). Speckle tracking methods for ultrasonic elasticity imaging using short-time correlation, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **46**(1): 82–96.

Makhinya, M. & Goksel, O. (2015). Motion tracking in 2D ultrasound using vessel models and robust optic-flow, *Proceedings of the MICCAI 2015 Workshop*.

Mattausch, O. & Goksel, O. (2018). Image-based reconstruction of tissue scatterers using beam steering for ultrasound simulation, *IEEE Transactions on Medical Imaging* **37**(3): 767–780.

Nebeker, J. & Nelson, T. R. (2012). Imaging of sound speed using reflection ultrasound tomography, *Journal of Ultrasound in Medicine* **31**(9): 1389–1404.

Ophir, J., Cespedes, I., Ponnekanti, H., Yazdi, Y. & Li, X. (1991). Elastography: A quantitative method for imaging the elasticity of biological tissues, *Ultrasonic Imaging* **13**: 111–134.

Pan, B., Qian, K., Xie, H. & Asundi, A. (2009). Two-dimensional digital image correlation for in-plane

displacement and strain measurement: a review, *Measurement Science and Technology* **20**(6).

Qu, X., Azuma, T., Lin, H., Takeuchi, H., Itani, K., Tamano, S., Takagi, S. & Sakuma, I. (2017). Limb muscle sound speed estimation by ultrasound computed tomography excluding receivers in bone shadow, *Proceedings of SPIE Medical Imaging*, p. 101391B.

Robinson, D. E., Ophir, J., Wilson, L. S. & Chen, C. F. (1991). Pulse-echo ultrasound speed measurements: progress and prospects, *Ultrasound in Medicine and Biology* **17**(6): 633–646.

Sak, M., Duric, N., Littrup, P., Bey-Knight, L., Krycia, M., Sherman, M. E., Boyd, N. & Gierach, G. L. (2015). Comparison of breast density measurements made with ultrasound tomography and mammography, *Proceedings of SPIE Medical Imaging*, pp. 94190R–1/R–8.

Sanabria, S., Furrer, R., Neuenschwander, J., Niemz, P. & Schütz, P. (2015). Analytical modeling, finite-difference simulation and experimental validation of air-coupled ultrasound beam refraction and damping through timber laminates, with application to non-destructive testing, *Ultrasonics* **63**: 65–85.

Sanabria, S. & Goksel, O. (2016). Hand-held sound-speed imaging based on ultrasound reflector delineation, *Proceedings of MICCAI*.

Sanabria, S., Martini, K., Rominger, M., Dedes, K., Vorburger, D., Frauenfelder, T. & Goksel, O. (2018). Breast cancer assessment with hand-held ultrasound based speed of sound: preliminary results, *European Conference on Radiology*.

Sanbria, S., Goksel, O., Martini, K., Forte, S., Frauenfelder, T., Kubik-Huch, R. A. & Rominger, M. B. (2018). Breast density assessment with handheld ultrasound: A novel biomarker to assess breast cancer risk and to tailor screening, *European Radiology* **28**(8): 3165–3175.

Sarvazyan, A., Hall, T., Urban, M., Fatemi, M., Aglyamov, S. & Garra, B. (2011). An overview of elastography - an emerging branch of medical imaging, *Current Medical Imaging Reviews* **7**(4): 255–282.

Shin, H. C., Prager, R., Gomersall, H., Kingsbury, N., Treece, G. & Gee, A. (2010). Estimation of average speed of sound using deconvolution of medical ultrasound data, *Ultrasound in Medicine and Biology* **36**(4): 623–636.

Sidky, E. Y., Kao, C. M. & Pan, X. (2006). Accurate image reconstruction from few-views and limited-angle data in divergent-beam ct, *Journal of X-ray Science and Technology* pp. 119–139.

Varghese, T. & Ophir, J. (1998). An analysis of elastographic contrast-to-noise ratio, *Ultrasound in Medicine and Biology* **24**(6): 915–924.

Wenkel, E., Heckmann, M., Heinrich, M., Schwab, S. A., Uder, M., Schulz-Wendtland, R., Bautz, W. A. & Janka, R. (2008). Automated breast ultrasound: Lesion detection and BI-RADS(TM) classification - a pilot study, *ROFO* **180**(9): 804–808.

Wu, X. (1991). An efficient antialising technique, *Computer Graphics* **25**(4): 143–152.

Zahiri-Azar, R., Goksel, O. & Salcudean, S. E. (2010). Sub-sample displacement estimation from digitized ultrasound RF signals using multi-dimensional polynomial fitting of the cross-correlation function, *IEEE Trans Ultrasonics, Ferroelectrics, and Frequency Control* **57**(11): 2403–20.

Zahiri-Azar, R., Goksel, O. & Salcudean, S. E. (2011). Comparison between 2-D cross correlation with 2-D sub-sampling and 2-D tracking using beam steering, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **58**(8): 1534–7.

Zografos, G., Liakou, P., Koulocheri, D., Liovarou, I., Sofras, M., Hadjiagapis, S., Orme, M. & Marmarelis, V. (2015). Differentiation of BIRADS-4 small breast lesions via multimodal ultrasound tomography, *European Radiology* **25**(2): 410–418.