PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Optimized transducer configuration for ultrasound waveform tomography in breast cancer detection

Nicolas Vinard, Naiara Korta Martiartu, Christian Boehm, Ivana Jovanović Balic, Andreas Fichtner

Nicolas Vinard, Naiara Korta Martiartu, Christian Boehm, Ivana Jovanović Balic, Andreas Fichtner, "Optimized transducer configuration for ultrasound waveform tomography in breast cancer detection," Proc. SPIE 10580, Medical Imaging 2018: Ultrasonic Imaging and Tomography, 105800I (6 March 2018); doi: 10.1117/12.2293600

SPIE.

Event: SPIE Medical Imaging, 2018, Houston, Texas, United States

Optimized transducer configuration for ultrasound waveform tomography in breast cancer detection

Nicolas Vinard[a,b], Naiara Korta Martiartu[b], Christian Boehm[b], Ivana Jovanović Balic[c], and Andreas Fichtner[b]

[a]Faculty of Civil Engineering and Geosciences, TU Delft, 2628 CN Delft, Netherlands
[b]Department of Earth Sciences, ETH Zurich, CH-8092 Zurich, Switzerland
[c]SonoView Acoustic Sensing Technologies, CH-2560 Nidau, Switzerland

ABSTRACT

Waveform inversion is a promising method for ultrasound computed tomography able to produce high-resolution images of human breast tissue. However, the computational complexity of waveform inversion remains a considerable challenge, and the costs per iteration are proportional to the number of emitting transducers. We propose a twofold strategy to accelerate the time-to-solution by identifying the optimal number and location of emitters using sequential optimal experimental design (SOED). SOED is a powerful tool to iteratively add the most informative transducer or remove redundant measurements, respectively. This approach simultaneously provides optimized transducer configurations and a cost-benefit curve that quantifies the information gain versus the computational cost.

First, we propose a method to identify the emitters that provide reconstructions with minimal expected uncertainties. Using a Bayesian approach, model uncertainties and resolution can be quantified with the trace of the posterior covariance. By linearizing the wave equation, we can compute the posterior covariance using the inverse of the Gauss-Newton approximation of the Hessian. Furthermore, this posterior is independent of the breast model and the experimental data, thus enabling pre-acquisition experimental optimization. Then, for the post-acquisition inversion, we present an approach to select a subsample of sources that accurately approximates the full gradient direction in each iteration. We control the convergence of the angular differences between consecutive gradient directions by randomly adding new emitters into the subsample.

We present synthetic studies in 2D and 3D that consider a ring-shaped and a semi-ellipsoidal scanning device, respectively. Numerical results suggest that the provided methods have the potential to identify redundancies from the corresponding cost-benefit curves. Furthermore, the gradient direction rapidly converges to the direction of the full gradient, which appears to be independent of the model and the emitter locations.

Keywords: sequential optimal experimental design, A-optimal design, ultrasound tomography, waveform inversion, resolution analysis, posterior covariance matrix, Gauss-Newton

1. INTRODUCTION

Ultrasound computed tomography (USCT) is an emerging technology for breast cancer screening with promising capabilities to differentiate benign from malign masses.[1,2] The acoustic properties of the breast tissue are usually reconstructed using ray-based tomography techniques which provide computationally efficient implementations.[3–5] However, due to the need to improve the diagnostic accuracy of USCT systems, waveform tomography has become an active field of research.[6–11]

Further author information: (Send correspondence to N.V.)
N.V.: E-mail: n.a.vinard@tudelft.nl, Telephone: +31 15 278 77 35
N.K.M.: E-mail: naiara.korta@erdw.ethz.ch, Telephone: +41 44 633 35 66
C.B.: E-mail: christian.boehm@erdw.ethz.ch, Telephone: +41 44 633 33 32
I.J.B.: E-mail: ivana.jovanovic@sono-view.com, Telephone: +41 76 463 18 92
A.F.: E-mail: andreas.fichtner@erdw.ethz.ch, Telephone: +41 44 632 25 97

Medical Imaging 2018: Ultrasonic Imaging and Tomography, edited by Neb Duric, Brett C. Byram, Proc. of SPIE Vol. 10580, 105800I · © 2018 SPIE
CCC code: 1605-7422/18/$18 · doi: 10.1117/12.2293600

In contrast to ray-based methods, waveform inversion is computationally significantly more challenging. Many studies[12–14] address this problem using 2D approximations of the acoustic wave equation and consider 2D layer-by-layer scanning systems. This approach considerably reduces the computational cost of simulating wave propagation. However, recent studies[11,15] indicate the benefits of solving the wave equation in 3D to avoid artifacts in the reconstructed images related to out-of-plane scattering and reflections.

The computational complexity of waveform inversion, which is particularly challenging for 3D simulations or for high frequencies, is proportional to the number of times that the wave equation needs to be solved. A natural solution, therefore, is to minimize the number of emitters, since this will reduce the recurring wave propagation simulations required in an iterative nonlinear optimization scheme. Recently, source encoding techniques have been introduced for USCT,[14,16] in which individual sources are treated as encoded simultaneous-source gathers in a stochastic reconstruction procedure.[17–19] Although this method drastically reduces the computing time per iteration, convergence is rather slow when it comes to mitigate the crosstalk noise from the model updates.[20] As an alternative approach, in geophysical studies, van Leeuwen et al.[20] proposed to approximate the full misfit and gradient using a different batch of sequential sources in each iteration. Even though it proves to be as efficient as source encoding, both methods miss to answer a fundamental question: How many emitters are actually necessary for an efficient 3D waveform tomography without incorporating redundancies into the measurements?

To answer this question, we apply sequential optimal experimental design (SOED)[21] to identify the redundant emitters from a preset transducer configuration. The method is a computationally attractive alternative to global optimization strategies,[22] and it has been successfully applied, for instance, in seismic exploration.[23,24] SOED strategies can be used for multiple purposes, depending on (1) the experimental design parameter that is intended to be optimized, and (2) the context in which the design parameter plays a substantial role. We refer the reader to Ref. 23 for an extensive overview in optimal experimental design in the context of seismic waveform inversion.

This paper is organized as follows. First, we describe the theory and general formulation for optimal experimental design when waveform tomography is used for the reconstruction. Here we introduce the Bayesian formulation for the inverse problem together with linearization assumptions that will allow us to quantify the expected posterior uncertainties in the parameters. In a second step, we propose pre- and post-acquisition applications to identify redundancies in the measurements and optimize the emitter configuration. In the pre-acquisition application we use properties of the posterior covariance matrix to quantify the quality of different transducer configurations, which leads to the so-called A-optimal design.[25] Then we develop a post-acquisition method to select subsamples of sources, from those previously identified as relevant, to accurately approximate the gradient direction in every iterative step of the waveform tomography approach.[20,26] Finally, in Section 3, we study the applicability of our methods to optimize a ring-shaped transducer array in a 2D numerical example, and we extend the results to 3D imaging systems where we consider a semi-ellipsoidal setup. Although we focus our work on reconstructing the speed of sound, extensions of the method to multiparametric inversions are straightforward.

2. METHODS

2.1 Waveform inversion and optimal experimental design

Waveform inversion can be formulated as a nonlinear least-squares problem that infers tissue properties from measurements of ultrasonic waves

$$\min_{\mathbf{m}} \frac{1}{2}\|\mathbf{F}_{\mathbf{s}}(\mathbf{m}) - \mathbf{d}_{\text{obs}}\|^2_{\mathbf{\Gamma}^{-1}_{\text{noise}}} + \frac{1}{2}\|\mathbf{m} - \mathbf{m}_0\|^2_{\mathbf{\Gamma}^{-1}_{\text{prior}}}. \tag{1}$$

Here, $\mathbf{F}_{\mathbf{s}}$ represents the wave operator that maps model parameters $\mathbf{m}$, e.g., density or sound speed, to synthetic data that can be compared with observed data $\mathbf{d}_{\text{obs}}$. $\mathbf{F}_{\mathbf{s}}$ depends nonlinearly on $\mathbf{m}$, in general, and is parameterized by a set of design variables $\mathbf{s}$, e.g., the locations of the emitters. Furthermore, $\mathbf{m}_0$ denotes the prior information about the unknown model parameters, and $\mathbf{\Gamma}_{\text{noise}}$ and $\mathbf{\Gamma}_{\text{prior}}$ are the covariance matrices of the noise and the prior distribution, respectively, which we both assume to be Gaussian.

Generally speaking, the goal of optimal experimental design is to determine a set of design variables $\mathbf{s}$ that provide a good trade-off between the cost of acquisition or computation, on the one hand, and the expected

quality of the reconstruction, on the other hand.[25,27] As a concrete example, we consider the selection of emitter locations for USCT. To this end, let n be the number of all possible emitter locations and let $\mathbf{s} \in \{0,1\}^n$ denote binary decision variables that encode whether the locations are selected for the experiment or not. In addition, we impose an upper bound c on the number of selected emitters to control the cost. Finally, we introduce a scalar function Θ that quantifies the information content. Depending on the specific choice of Θ, the experimental design problem could either maximize the expected information that is gained from the experiment, or minimize the expected uncertainties in the reconstruction. Without loss of generality, we treat this as a minimization problem and obtain

$$\min_{\mathbf{s}} \ \Theta(\mathbf{s}) \quad \text{subject to} \quad \mathbf{s} \in \{0,1\}^n, \quad \|\mathbf{s}\|_1 \leq c. \tag{2}$$

A common way of quantifying the information quality of the problem setup is to consider a Bayesian formulation of the inverse problem,[28,29] and to extract properties from the posterior covariance matrix. For the above example in Eq. (1) , the posterior probability density function of the model parameters is given by[30]

$$\pi_{\text{post}}(\mathbf{s},\mathbf{m}) \propto \exp\left(-\frac{1}{2}\|\mathbf{F}_{\mathbf{s}}(\mathbf{m}) - \mathbf{d}_{\text{obs}}\|^2_{\mathbf{\Gamma}^{-1}_{\text{noise}}} - \frac{1}{2}\|\mathbf{m}-\mathbf{m}_0\|^2_{\mathbf{\Gamma}^{-1}_{\text{prior}}}\right). \tag{3}$$

The quality of a specific design can be linked to the expected uncertainties in the reconstruction, which can be extracted from the posterior distribution. One of the key challenges here is that the optimal set $\mathbf{s}$ needs to be determined prior to taking any measurements, so it needs to be independent from both model and data. For nonlinear inverse problems this poses two main challenges. First, evaluating the complete posterior covariance is computationally intractable and even using low-rank approximations and adjoint-based methods poses a significant challenge to approximate its properties. Second, the posterior covariance depends on the unknown true model which, of course, is data-dependent and would require proper sampling of the data space. To circumvent these issues, we consider the Born approximation and linearize the forward operator

$$\mathbf{F}_{\mathbf{s}}(\mathbf{m}) \approx \mathbf{F}_{\mathbf{s}}(\mathbf{m}_0) + \mathbf{F}'_{\mathbf{s}}(\mathbf{m}_0)\Delta\mathbf{m}, \quad \text{with} \quad \Delta\mathbf{m} = \mathbf{m} - \mathbf{m}_0. \tag{4}$$

This is a valid simplification, because the expected changes in the tissue properties are small and, thus, we can assume to have a prior $\mathbf{m}_0$ that is already in the vicinity of the global optimum. By substituting Eq. (4) into Eq. (3), we obtain a local approximation $\tilde{\pi}_{\text{post}}$ of the posterior probability density function

$$\tilde{\pi}_{\text{post}}(\mathbf{s},\Delta\mathbf{m}) \propto \exp\left(-\frac{1}{2}\|\mathbf{F}_{\mathbf{s}}(\mathbf{m}_0) + \mathbf{F}'_{\mathbf{s}}(\mathbf{m}_0)\Delta\mathbf{m} - \mathbf{d}_{\text{obs}}\|^2_{\mathbf{\Gamma}^{-1}_{\text{noise}}} - \frac{1}{2}\|\Delta\mathbf{m}\|^2_{\mathbf{\Gamma}^{-1}_{\text{prior}}}\right). \tag{5}$$

Both the regularization and the data misfit term now depend quadratically on $\Delta\mathbf{m}$, and the Hessian of the latter with respect to $\Delta\mathbf{m}$ is given by

$$\mathbf{H} = \mathbf{H}(\mathbf{s}) = \mathbf{F}'_{\mathbf{s}}(\mathbf{m}_0)^T \mathbf{\Gamma}^{-1}_{\text{noise}} \mathbf{F}'_{\mathbf{s}}(\mathbf{m}_0). \tag{6}$$

In particular, it does not depend on $\mathbf{m}$, but only on the prior $\mathbf{m}_0$. Furthermore, $\mathbf{H}$ is equivalent to the Gauss-Newton approximation of the Hessian of the nonlinear $\mathbf{F}_{\mathbf{s}}$ evaluated at the prior $\mathbf{m}_0$. It is important to note that $\mathbf{H}$ already contains information on inter-parameter trade-offs from first-order scattering in its off-diagonal terms. This is particularly interesting for multi-parameter inversions.

Returning to the posterior distribution defined in Eq. (5), we deduce that, after linearlization, the posterior for a fixed $\mathbf{s}$ is equivalent to a Gaussian with a covariance matrix $\mathbf{\Gamma}_{\text{post}}$ given by

$$\mathbf{\Gamma}_{\text{post}}(\mathbf{s}) = \left(\mathbf{H}(\mathbf{s}) + \mathbf{\Gamma}^{-1}_{\text{prior}}\right)^{-1}. \tag{7}$$

In particular, the posterior covariance only depends on the forward operator $\mathbf{F}_{\mathbf{s}}$ through the Gauss-Newton approximation, but neither on the model $\mathbf{m}$ nor the data $\mathbf{d}_{\text{obs}}$. Thus, the Gauss-Newton approximation renders optimal experimental design tractable provided that the space of design parameters can be handled efficiently, which is the focus of the next section.

We conclude this section with two specific examples of the design criterion Θ. *A-optimal*[25] design aims at minimizing the trace of the posterior covariance matrix, i.e.,

$$\Theta_1 = \mathrm{tr}(\mathbf{\Gamma}_{\mathrm{post}}(\mathbf{s})). \tag{8}$$

Hence, A-optimal design is equivalent to minimizing the average posterior variance, where the diagonal of $\mathbf{\Gamma}_{\mathrm{post}}(\mathbf{s})$ represents the expected spatial uncertainties in the solution of the inverse problem.[19,31] In particular, redundant information that is obtained, for example, by repeating the same experiment several times, can decrease the posterior variance. Alternatively, we can apply a criterion that accounts only for irredundant information. Here, we consider the negative trace of the Gauss-Newton approximation, and normalize it by the largest eigenvalue[22,32]

$$\Theta_2 = -\frac{\mathrm{tr}(\mathbf{H}(\mathbf{s}))}{\lambda_{\max}(\mathbf{H}(\mathbf{s}))}. \tag{9}$$

It is important to note that both criteria can be restricted to a subset of the computational domain. This is particularly helpful to focus the design on a specific region of interest. For instance, we can use prior knowledge about the surrounding water tank to target only the approximate area of the breast in USCT.

2.2 Sequential optimal experimental design for pre-acquisition emitter selection

Solving the design problem in Eq. (2) is computationally challenging, because the number of feasible emitter configurations grows exponentially with the number of emitters n. In addition, instead of fixing the costs c a priori, it is often more desirable to see the value of adding more emitters. The gain in information from additional measurements typically decreases after a certain point, where new data mostly add redundant information. Assuming that the costs are proportional to the number of experiments, this results in a cost-benefit curve that flattens out with increasing number of experiments. This curve enables us to determine a suitable number of emitters needed to obtain a good trade-off between cost and information, as well as their locations. The downside, however, is that problem (2) needs to be solved repeatedly for different values of c. This becomes intractable when the computational cost of the quality measure is high.

An alternative approach is SOED.[21,23] Starting from an experimental configuration that considers all possible locations, the method iteratively removes the emitter which least contributes to the quality measure. In this way, a locally optimal experimental design is found in each iteration. Although this greedy approach does not guarantee a globally optimal solution for a fixed number of transducers, previous studies have shown that this approach works very well in practice.[24]

For selecting the emitter locations, the design parameters $\mathbf{s}$ give a special structure to the inverse problem (1), which can be exploited in the computations. In particular, the misfit term is separable with respect to different events, i.e.,

$$\frac{1}{2}\sum_{i=1}^{n} \|\mathbf{F}_{s_i}(\mathbf{m}) - \mathbf{d}_{\mathrm{obs}}\|^2_{\mathbf{\Gamma}^{-1}_{\mathrm{noise}}}, \tag{10}$$

which allows for simulating the individual emitters independently. The same structure applies to the derivatives, which enables us to compute the contributions to the Gauss-Newton approximation individually for every emitter.

Furthermore, due to the sequential nature of the algorithm, we directly obtain a cost-benefit curve that indicates the quality measure as a function of the number of emitters. This enables us to (1) remove the emitters that provide the most redundant information about the region of interest and (2) determine the minimum amount of emitters required such that the model parameters can still be resolved accurately enough.

2.3 Post-acquisition emitter subsampling

The method presented in the previous section identifies a suitable set of emitter locations. To reconstruct the parameters $\mathbf{m}$ we return to the nonlinear inverse problem where the misfit term in the objective function is given by

$$J_{\mathcal{S}}(\mathbf{m}) := \frac{1}{2 \cdot |\mathcal{S}|} \sum_{s \in \mathcal{S}} \|\mathbf{F}_s(\mathbf{m}) - \mathbf{d}_{\text{obs}}\|^2_{\mathbf{\Gamma}^{-1}_{\text{noise}}}, \tag{11}$$

where $\mathcal{S}$ is now the set of selected sources, and $|\mathcal{S}|$ denotes the size of the set. Because $J_{\mathcal{S}}$ depends nonlinearly on $\mathbf{m}$, the inverse problem can only be solved iteratively, e.g., with gradient-based descent algorithms. Each iteration k requires the gradient of $J_{\mathcal{S}}$ with respect to the current model $\mathbf{m}^k$ to compute an update. Using the adjoint method[33] this requires two simulations of the wave equation per emitter.

Because the recurring simulations of the wave equation vastly dominate the computational cost of an iteration, we can reduce the computational complexity by approximating the misfit function and its gradient by selecting only a subset of emitters $\mathcal{S}^k \subset \mathcal{S}$ in every iteration. This approach is related to the Kacmarz method[34] and stochastic descent algorithms,[20] and it exploits redundancies in the gradient direction. The subset $\mathcal{S}^k$ can be chosen adaptively in every iteration, ensuring that all emitters will be considered during the inversion. Hence, this approach does not change the set of optimal models, but it reduces the computational cost to obtain a reconstruction.

In the following, we drop the superscript k and the dependence on $\mathbf{m}$ to simplify the notation. To select the emitters for computing the gradient, we can apply SOED to determine the subsets of transducers, which give the smallest angular difference to the comprehensive gradient that is built from all emitters in $\mathcal{S}$. Here, we only need to consider the data misfit term, because the regularization term is independent of $\mathcal{S}$. We start with an empty set $\mathcal{S}_0$ and iteratively add new emitters by choosing the one that gives the smallest angular difference to the comprehensive gradient $\nabla J_{\mathcal{S}}$. I.e., we sequentially select the emitter s_{i+1} that satisfies

$$s_{i+1} = \underset{s \in \mathcal{S} \setminus \mathcal{S}_i}{\text{argmax}} \left\{ \frac{\left(\nabla J_{\mathcal{S}_i \cup s}\right)^T \left(\nabla J_{\mathcal{S}}\right)}{\|\nabla J_{\mathcal{S}_i \cup s}\| \cdot \|\nabla J_{\mathcal{S}}\|} \right\}, \tag{12}$$

and update the set $\mathcal{S}_{i+1} = \mathcal{S}_i \cup s_{i+1}$.

However, this would require the gradient for all $s \in \mathcal{S}$, which would not save any cost. Instead we apply a simple heuristic and check the incremental change in the gradient direction from adding additional emitters. The emitters can either be selected randomly or from a deterministic pattern, and the selection can be stopped as soon as the incremental change,

$$\frac{\left(\nabla J_{\mathcal{S}_{i+1}}\right)^T \left(\nabla J_{\mathcal{S}_i}\right)}{\|\nabla J_{\mathcal{S}_{i+1}}\| \cdot \|\nabla J_{\mathcal{S}_i}\|,} \tag{13}$$

is smaller than a predefined threshold. This approach neither requires knowledge of the comprehensive data set $\mathcal{S}$, nor any additional simulations. Furthermore, it allows for seamless integration into existing inversion frameworks.

3. RESULTS

In this section, we present two use cases for selecting the emitters in USCT based on optimal experimental design. We consider a synthetic breast phantom with inclusions of varying size, as well as a more realistic breast phantom provided by the OA-breast database.[35] All numerical simulations use waveform synthetics created with Salvus[36] and Macesopt.[37]

3.1 Pre-acquisition emitter selection

As a first example, we consider the experimental setup illustrated in Fig. 1 which consists of a 2D ring-shaped device with a radius of 10 cm. Here, a total of 62 transducers are regularly placed - 1 cm apart from each other - from which we aim to select the optimal ones acting as emitters. These are considered as point sources and modelled using a Ricker wavelet with a center frequency of 200 kHz. Additionally, for each emitter, the ultrasound signals are recorded by the remaining 61 transducers.

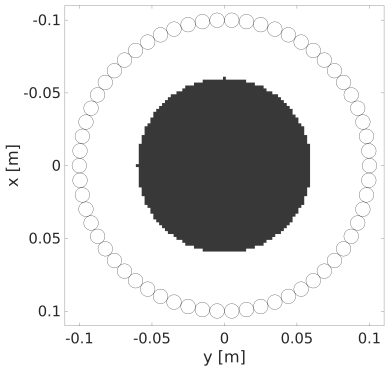


Figure 1: The ring-shaped device with 62 regularly placed transducers (circles). The white region delimits the region of interest, which is represented by a homogeneous sound speed model (water).

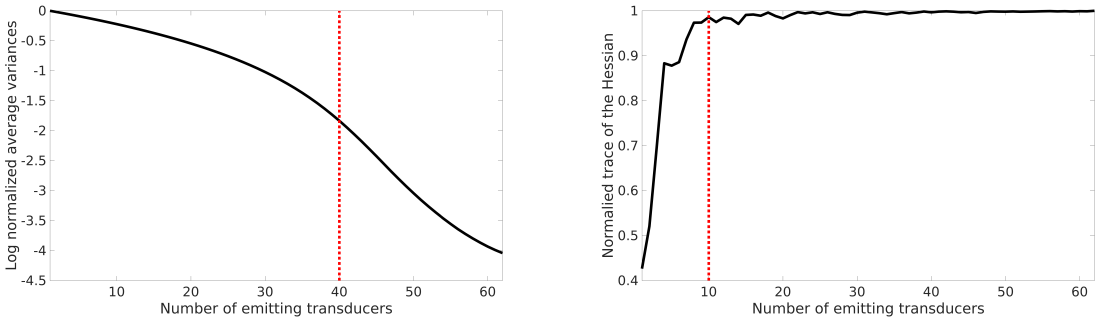
We apply A-optimal SOED to identify the optimal emitting transducers: starting from a setup that considers all transducers as emitters, we iteratively remove the less sensitive emitters according to the quality measure described in Eq. (8). To compute the trace of the covariance matrix, Θ_1, we generate the observed data using a homogeneous water model and we choose the prior $\mathbf{m}_0$ as the smoothed version of a circular breast phantom without any inclusions. Both models are discretized on a rectilinear mesh with a grid size of 2 mm.

We select the optimal number of emitters by analyzing the cost-benefit curves shown in Fig. 2. Here we represent the quantities in Eqs. (8) and (9) with respect to the emitting transducers selected by the A-optimal SOED algorithm in each iteration:

1. Fig. 2a shows the logarithm of the normalized trace of the posterior covariance matrix, which corresponds to the average of the expected variances in the reconstructed parameters. The posterior variances tend to decrease - increasing the certainty about the parameters - when more emitters are added into the experiment. This means that the measured ultrasound signals provide new information about the region of interest. However, as soon as the new emitters only add redundant information into the set of measurements, the curve flattens out and the variances remain stable. We select 40 emitters (red dashed line) as representative of the best cost-benefit ratio for our experiment, just before the sign of the curvature of the cost-benefit curve changes. In this way, we ensure that the posterior variances are improved by two orders of magnitude relative to the case of having no emitter at all, i.e., relative to our prior information about the model.

2. Fig. 2b shows a transformation of the previous curve. Here, we represent the trace of the Gauss-Newton approximation normalized by the largest eigenvalue, Eq. (9), with respect to the sequentially optimized emitters. This quantity is sensitive to the relative eigenvalue magnitudes of the Hessian,[22,32] which means that it can only be maximized by flattening the eigenvalue spectrum, i.e., when the design provides eigenvalues with a more equally distributed information about the region of interest. The link to the uncertainties in the parameters is rather straightforward if we note that, by definition, the inverse of the eigenvalues of the Hessian, are indeed the eigenvalues of the posterior covariance matrix. Thus, small eigenvalues, for instance, will contribute to higher posterior variances.

   The cost-benefit curve shows a pronounced rise at the beginning and rapidly converges to a maximum value where the relative eigenvalue magnitudes remain unchanged. The maximum value represents emitter configurations providing experimental data with the most balanced information about the region of interest. The best cost-benefit ratio is then produced by the number of emitters corresponding to the maximum curvature,[24] which in this case is 10 (red dashed line). Although adding more emitters can increase our certainty about the solution of the inverse problem - as we have observed in the previous curve -, it does not necessarily translate to a noticeable improvement in the reconstructed model, at least in the absence of modelling and measurement errors.



(a) (b)

Figure 2: Cost-benefit curves obtained from A-optimal SOED problem: (a) the normalized log of Θ_1 and (b) the normalized absolute value of Θ_2 with respect to the sequentially selected emitters. Red dashed lines indicate the optimal number of emitters identified in both cases.

The meaning of the cost-benefit curves can be intuitively understood by computing the reconstructions for different emitter configurations. Figs. 3b - 3d compare the sound speed reconstructions for the Derenzo breast phantom[22] using 5, 10 and 40 optimally selected emitters, respectively. While the experiment with only 5 emitters exhibits some artifacts in the reconstructed breast image, there are no significant differences between those obtained from 10 and 40 emitters. This is consistent with our previous observations. However, we can understand the actual differences between Figs. 3c and 3d by analyzing the expected posterior variances in both cases. Fig. 4a shows the absolute values of the comprehensive posterior variances when all transducers are considered as emitters, and in Figs. 4b - 4c we display the expected variances relative to the comprehensive one for experiments using 10 and 40 emitters, respectively. Here, regions with values close to 1 indicate that the experiment has retrieved the maximum possible information about that region, while values close to 0 represent regions in which uncertainties can be reduced by adding more emitters. As we can observe, even though one can hardly observe differences between both reconstruction by the naked eye, our confidence about the solution of the inverse problem is considerably increased when the experiment uses 40 emitters. Thus, we may interpret that by adding more than 10 emitters we are contributing to reduce the parameter trade-offs.

Finally, we test the emitter configuration found by A-optimal SOED in a more realistic breast phantom.[35] Fig. 5 shows the true, initial and reconstructed model when the experiment is performed using 10 optimally selected emitters. The quality of the reconstruction verifies that the features concluded about the experimental setup are actually model independent.

So far, we have identified the optimal number of emitters as 10 or 40. Although both configurations leads to reconstructed images with similar quality, in an actual experiment we suggest to collect the ultrasound signals using the number of emitters satisfying the lowest posterior variances - 40 in our case. Then, we can apply the method described in Sec. 2.3 to accelerate the iterative reconstruction by exploiting the redundancies in the gradient direction. In the following subsection we first present the results for the 2D experiment described above and then, we extend the results to the 3D waveform tomography.

3.2 Post-acquisition emitter subsampling

3.2.1 2D Waveform Tomography

Continuing with the previous example, we consider the 40 optimized emitters to study redundancies in the gradient direction. This numerical example is situated in the post-acquisition context, i.e., for when we have collected the experimental data for a particular breast. Hence, our goal is to accelerate the reconstruction algorithm. As shown before, not all the emitters incorporate new information about the region of interest into the set of measurements. This is why a subsample of emitters can already provide a good approximation of the gradient direction needed for the iterative reconstruction procedure. By selecting the optimal number of emitters in each iteration, one could avoid redundant and recurring wave propagation simulations, and thus reduce considerably the computational cost of waveform tomography.

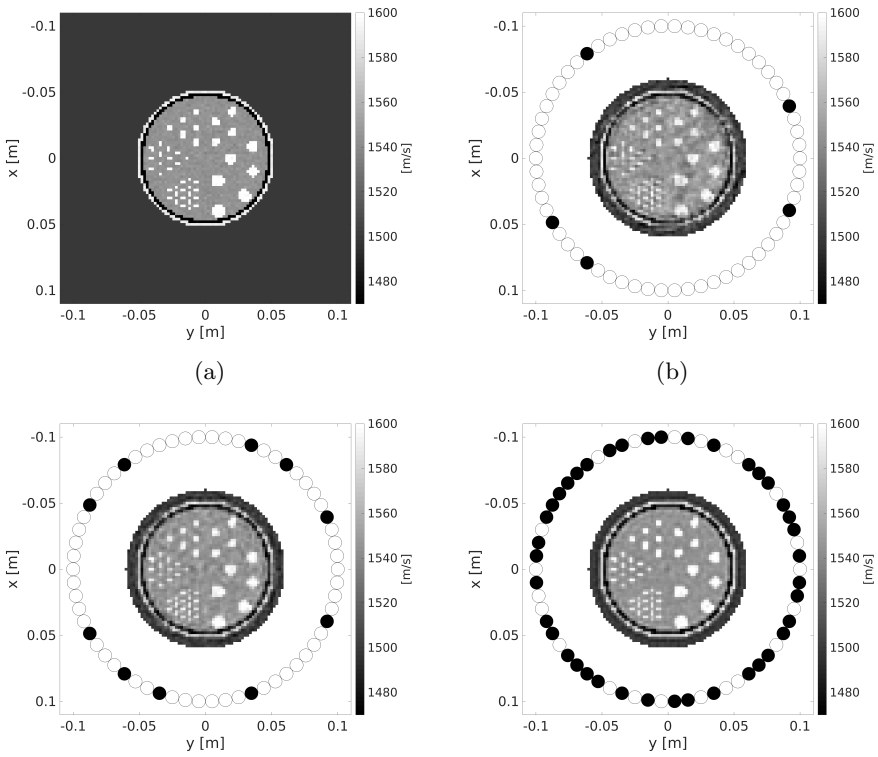


Figure 3: (a) True model (Derenzo). (b)-(d) Sound speed reconstructions using 5, 10 and 40 number of transducers as emitters, respectively. A-optimal SOED method was used to optimize the transducer configuration. Black and white circles represent the activated and deactivated emitting transducers, respectively.

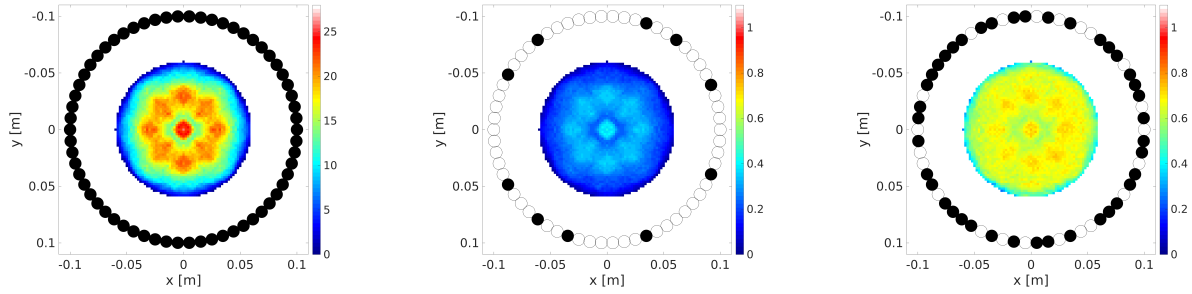
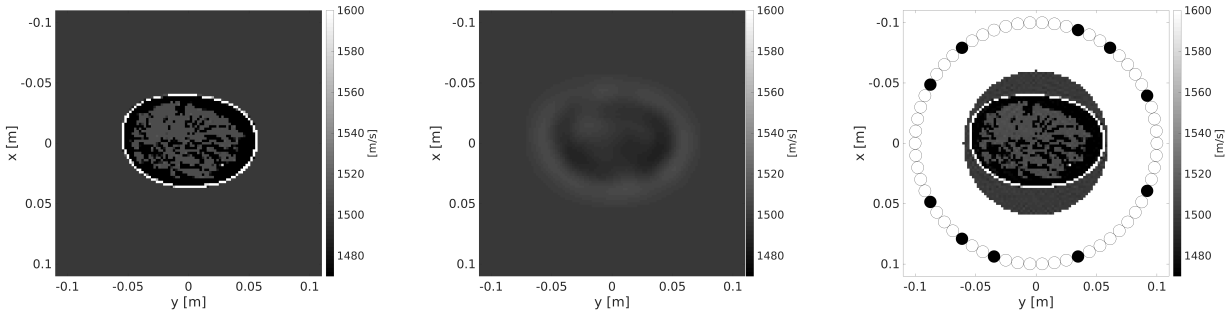


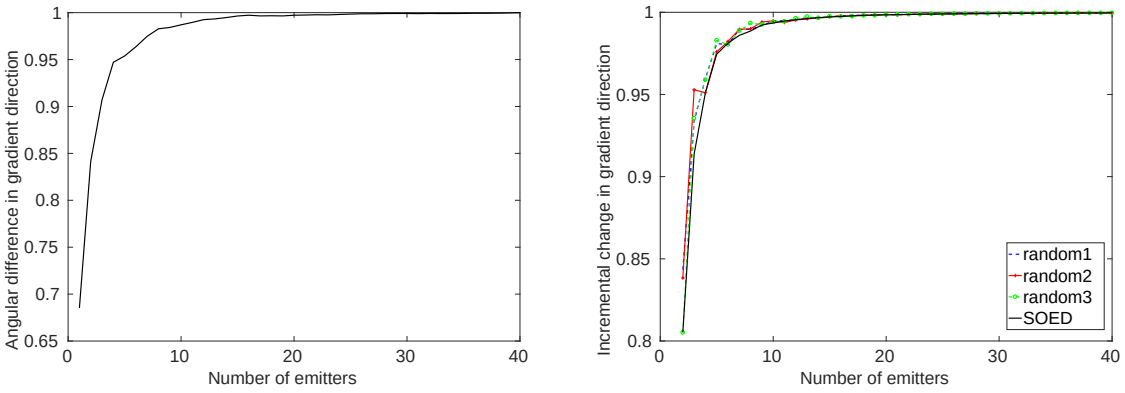
Figure 4: (a) Expected posterior variances when all transducers are acting as emitters. (b) - (c) Comprehensive posterior variances divided by the variances expected using 10 and 40 emitters, respectively.

We identify redundancies by applying the SOED method with the quality measure defined in Eq. (12). Here, we sequentially find subsamples of emitters that provide gradient directions with the smallest angular difference with respect to the comprehensive one computed from all emitters. Fig. 6a shows the cosine of the angular differences with respect to the number of emitters in the subsample. The curve convergences quickly towards its maximum value, which represents the minimum angular distance. We observe that approximately 10 emitters can already provide an accurate approximation of the comprehensive gradient direction. Note that this is consistent with the optimal number of emitters identified in the previous example, c.f. 2b.



(a) (b) (c)

Figure 5: (a) True model, (b) initial model and (c) sound speed reconstruction using 10 optimized transducers as emitters. Realistic phantom acquired from Lou et al.[35]



(a) (b)

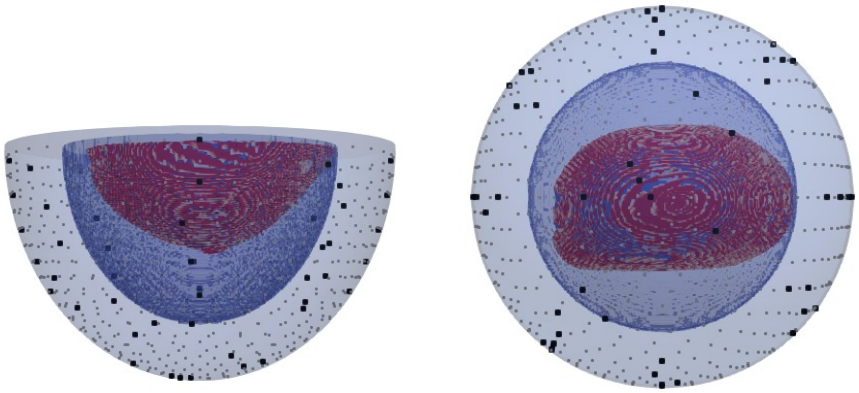
Figure 6: (a) Angular difference in the gradient direction, Eq. (12), with respect to the number of emitting transducers. (b) Incremental change in the consecutive gradient directions, Eq. (13), when the emitters are optimally and randomly selected.

Fig. 6b explores the possibility of deriving the same conclusions without the need of computing all the gradients beforehand - note that the quality measure used in SOED, Eq. (12), requires the computation of the comprehensive gradient direction. Here we represent the incremental change between two consecutive gradient directions, Eq. (13), with respect to the number of emitting transducers. We compare two cases and either apply SOED to determine the sequence in which the emitters are added, or randomly select the emitter locations. Both results show a similar convergence as Fig. 6a and the interesting point is, that the random source selection yields almost the same results as when the sources are selected using SOED method. In order to understand if the latter is a generalizable feature, in the next example we extend our study to the 3D waveform tomography.

3.2.2 3D Waveform Tomography

In this example, we consider a semi-ellipsoidal scanning system with semi-axes of 10 cm and 12 cm, as shown in Fig. 7. Here a total of 516 transducers are regularly placed on its surface, from which we only select fifty of them, uniformly distributed, to act as emitters with a center frequency of 150 kHz. The remaining transducers are set to receiving mode. We select 50 emitters just for illustrative purposes. A more accurate selection can be performed, for instance, by applying the pre-acquisition optimization described in the previous subsection.

Here, we consider the three numerical breast phantoms displayed in Figs. 8. These realistic models are available in the OA-Breast Database,[35] each of them having a different breast density level: a scattered fibroglandular (FG), a heterogeneously dense (HD) and an extremely dense (ED) level. We added six artificial tumors of varying sound speed, size and location into the otherwise healthy phantoms. The initial models required to compute the gradients are smoothed versions of the true models. For the HD model, in particular, we applied two



(a) (b)

Figure 7: (a) Lateral and (b) top view of the scanning device. The inner semi-ellipsoid specifies the region of interest considered for the SOED problem. Black and grey squares represent the sources and receivers, respectively.

different smoothing levels, hereinafter designated as HD and HD2. The synthetic measurements corresponding to the initial models are computed on a mesh with approximately 250'000 elements. To reduce the effects of the inverse crime affecting the synthetic computations, we compute the observed measurements on a finer mesh. A similar approach to mitigate the inverse crime has been applied in Ref. 28.

(a) HD (b) ED (c) FG

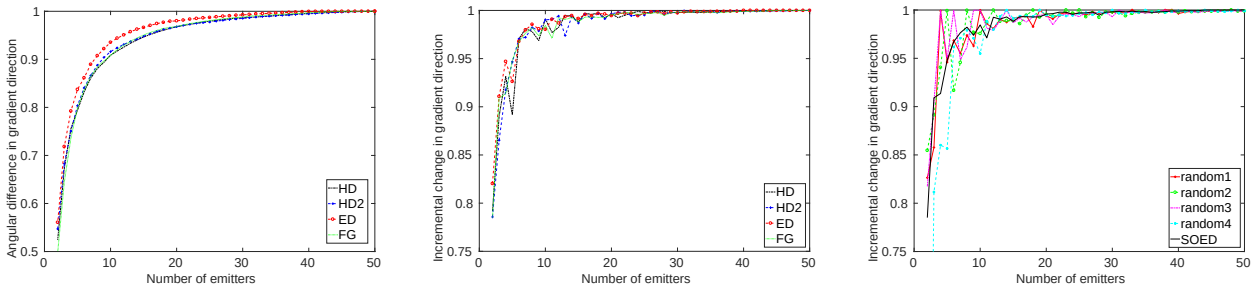
Figure 8: True models: heterogeneously dense (HD), extremely dense (ED) and scattered fibroglandular model (FG). Phantoms provided by Lou et al.[35]

Fig. 9a shows the SOED results of the emitter selection for the post-acquisition inversion. As we can observe, the gradient direction rapidly converges with increasing number of emitters, which seems to be model independent. In this example, we observe that approximately one third of the total amount of emitters can already provide a very good approximation of the desired gradient direction.

The interpretation of the curve becomes more intuitive if the images of the gradients are analyzed. Fig. 10 compares the gradients, normalized by the number of emitters, obtained from different subsamples of emitters for the case of the ED model. Each column displays a different slice of the 3D gradient and in the last row, we show the true model as reference. Clear improvements can be observed in the first three rows from the top - up

to twenty emitters - after which it gets harder to distinguish differences between the gradients. This agrees with the previous observation in Fig. 9a.

Finally, we represent the incremental changes between two consecutive gradient directions in Figs. 9b,c, for the optimal and random selection of the emitters, respectively. The latter is only applied with the gradients resulting from the FG model. All results show a fast convergence, although the curves obtained from the random selection show a more oscillatory behaviour. As in the previous example, the results suggest that the gradient direction can actually be adequately approximated through a randomly selected subsample of emitters, which avoids the need to compute the comprehensive gradient direction. Moreover, it also allows to design a useful post-acquisition optimization approach: few sources can be randomly incorporated at a time to check the desired convergence of the incremental change in the gradient direction.



(a) (b) (c)

Figure 9: (a) Angular difference in the gradient direction, Eq. (12), with respect to the number of emitting transducers. (b) - (c) Incremental changes in the consecutive gradient directions, Eq. (13), when the emitters are optimally and randomly selected, respectively. Random selection was only applied to the FG phantom.

4. SUMMARY AND CONCLUSIONS

Building upon a previous study,[22] in which the authors apply a global optimal experimental design algorithm to position transducers in the USCT scanning device using time-of-flight inversion, this work extended the method to waveform inversion for breast imaging. With the aim of alleviating the computational burden of waveform inversion, we proposed methods to reduce the number of redundant wave propagation simulations. For this purpose, we applied sequential optimal experimental design (SOED) to optimally select the number and locations of the emitters from a preset transducer array. We considered two possible scenarios: (1) pre-acquisition, i.e., before the experiment is conducted, and (2) post-acquisition, once the experimental data is collected for a particular breast.

Waveform tomography is a non-linear inverse problem, and therefore, the quantities used to measure the quality of an experimental configuration are dependent on both, the unknown true model and the experimental data. For the post-acquisition optimization the specific nature of the experimental data is included in the approach of the problem itself. However, in the pre-acquisition scenario linearization assumptions are required. Assuming small variations in the acoustic properties of the breast tissue, we used the Gauss-Newton approximation to quantify the expected posterior uncertainties in the solution to the inverse problem. The resulting emitter configuration may not be the optimal one for any particular breast, but it is optimal in the sense that it gives the smallest variances averaged over all possible breast models. Numerical examples using different breast phantoms have demonstrated this.

It is important to stress that one must be careful about the interpretation of the results - especially when it is about the specific number of emitters resulting from the optimization. In this theoretical study we only use synthetic data with neither measurement nor modelling errors. However, we think that the following conclusions are extendable to real applications:

1. We assume that the scanning device is already accessible such that we can select emitting transducers from a predefined set of locations. Therefore, we know the specific nature of the ultrasound signals produced

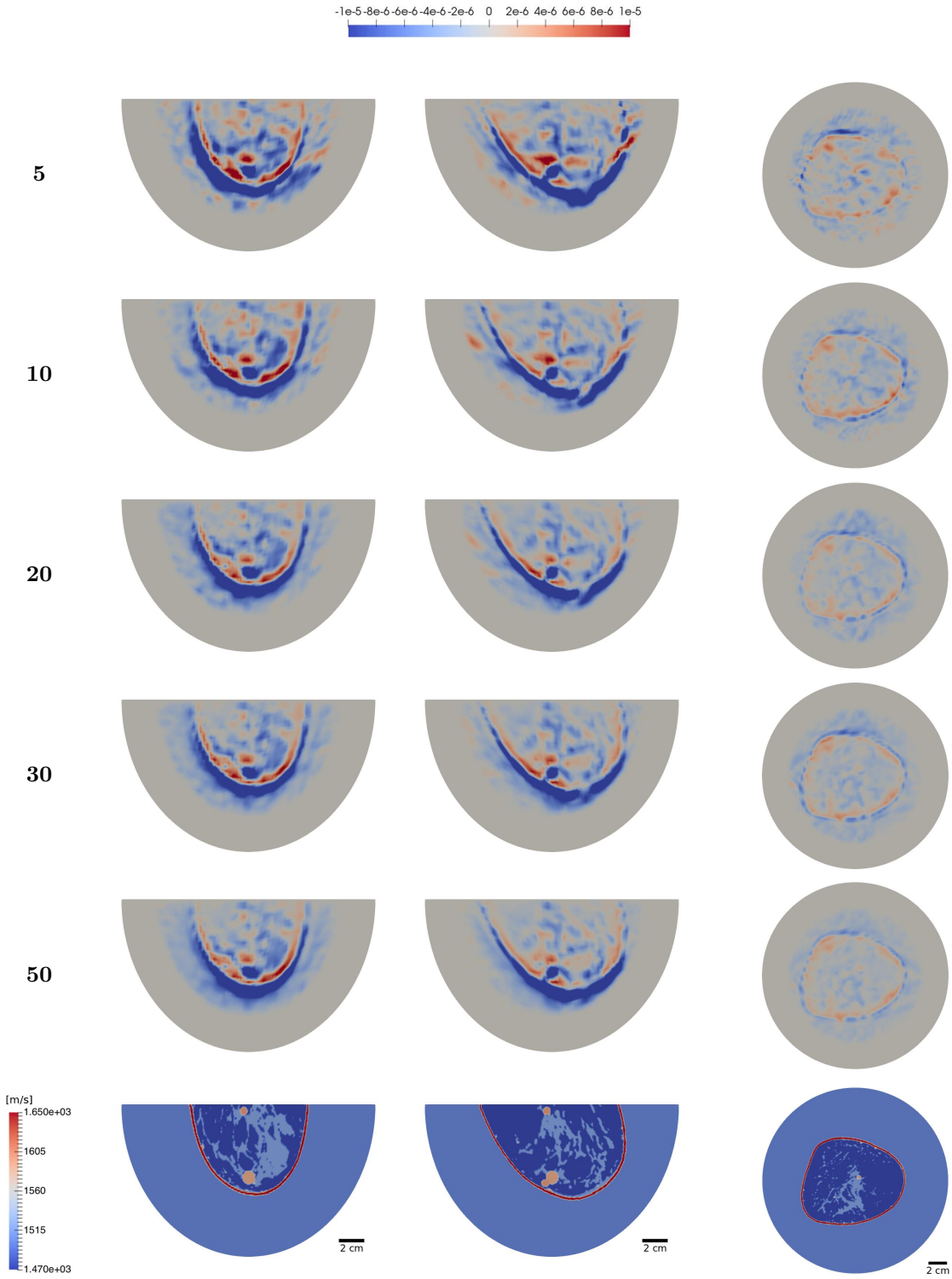


Figure 10: Gradients for ED phantom, normalized by the number of emitters. The first column indicates the number of sources and the last row shows the true model as reference.

by the experimental device, as well as the calibration data recorded in water. This information can be included in the pre-acquisition optimization approach.

2. We showed that acquiring more experimental data does not necessarily involve new information about the breast tissue. Redundant measurements exist and those can be identified from the cost-benefit curves provided by the SOED algorithm.

3. Although adding more measurements to the experiment may not imply noticeable improvements in the reconstructed model, they can still contribute to increase our certainty about it. We demonstrated that this kind of redundancies are useful to speed-up the post-acquisition reconstruction.

The latter led us to present a method to approximate the gradient direction. We observed that a subsample of emitters can already provide an accurate approximation of the comprehensive gradient direction, which helps us to prevent unnecessary and recurring wave propagation simulations. More interestingly, we observed that the fast convergence towards the comprehensive gradient direction is independent of both, the model and the selected emitter locations in the subsample. Thus, in every iterative step of the inversion, one could randomly select a subsample of emitters for the model update. In particular, we could add few emitters at a time and stop when the incremental change in the gradient direction is smaller than a suitably predefined threshold.

Finally, it is important to mention that we considered low frequencies to make the study computationally feasible. We expect the optimal number of emitters to depend on the frequencies that are used. However, the criteria presented in this study automatically adapt to the data to select the optimal number of emitters.

ACKNOWLEDGMENTS

The research leading to this study has received funding from the Swiss Commission for Technology and Innovation under grant number 17962.1 PFLS-LS. Furthermore, we gratefully acknowledge support by the Swiss National Supercomputing Centre (CSCS) under project grant d72.

REFERENCES

[1] Greenleaf, J. F., Johnson, S. A., and Bahn, R. C., "Quantitative cross-sectional imaging of ultrasound parameters," in [*1977 Ultrasonics Symposium*], 989–995 (Oct 1977).

[2] André, M., Wiskin, J., and Borup, D., "Clinical results with ultrasound computed tomography of the breast," in [*Quantitative Ultrasound in Soft Tissues*], Mamou, J. and Oelze, M. L., eds., 395–432, Springer Netherlands, Dordrecht (2013).

[3] Gemmeke, H. and Ruiter, N., "3D ultrasound computer tomography for medical imaging," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **580**(2), 1057 – 1065 (2007). Imaging 2006.

[4] Li, C., Duric, N., Littrup, P., and Huang, L., "In vivo breast sound-speed imaging with ultrasound tomography," *Ultrasound in medicine & biology* **35**(10), 1615–1628 (2009).

[5] Hormati, A., Jovanovic, I., Roy, O., and Vetterli, M., "Robust ultrasound travel-time tomography using the bent ray model," *Proc. SPIE* **7629**, 76290I–76290I–12 (2010).

[6] Pratt, R. G., Huang, L., Duric, N., and Littrup, P., "Sound-speed and attenuation imaging of breast tissue using waveform tomography of transmission ultrasound data," in [*Medical imaging*], 65104S–65104S, International Society for Optics and Photonics (2007).

[7] Wiskin, J., Borup, D., Johnson, S., Berggren, M., Abbott, T., and Hanover, R., "Full-wave, non-linear, inverse scattering," in [*Acoustical Imaging*], 183–193, Springer Netherlands, Dordrecht (2007).

[8] Roy, O., Jovanovic, I., Hormati, A., Parhizkar, R., and Vetterli, M., "Sound speed estimation using wave-based ultrasound tomography: theory and GPU implementation," *Proc. SPIE* **7629**, 76290J–76290J–12 (2010).

[9] Anis, F., Lou, Y., Conjusteau, A., Su, R., Oruganti, T., Ermilov, S. A., Oraevsky, A. A., and Anastasio, M. A., "Investigation of the adjoint-state method for ultrasound computed tomography: a numerical and experimental study," *Proc. SPIE* **8943**, 894337–894337–6 (2014).

[10] Perez-Liva, M., Herraiz, J. L., Udías, J. M., Cox, B. T., and Treeby, B. E., "Full-wave attenuation reconstruction in the time domain for ultrasound computed tomography," in [*Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*], 710–713, IEEE (2016).

[11] Sandhu, G. Y., West, E., Li, C., Roy, O., and Duric, N., "3D frequency-domain ultrasound waveform tomography breast imaging," *Proc. SPIE* **10139**, 1013909–1013909–14 (2017).

[12] Wiskin, J., Borup, D., Johnson, S., and Berggren, M., "Non-linear inverse scattering: High resolution quantitative breast tissue tomography," *The Journal of the Acoustical Society of America* **131**(5), 3802–3813 (2012).

[13] Sandhu, G., Li, C., Roy, O., Schmidt, S., and Duric, N., "Frequency domain ultrasound waveform tomography: breast imaging using a ring transducer," *Physics in medicine and biology* **60**(14), 5381 (2015).

[14] Matthews, T. P., Wang, K., Li, C., Duric, N., and Anastasio, M. A., "Regularized dual averaging image reconstruction for full-wave ultrasound computed tomography," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **64**(5), 811–825 (2017).

[15] Goncharsky, A., Romanov, S. Y., and Seryozhnikov, S. Y., "A computer simulation study of soft tissue characterization using low-frequancy ultrasonic tomography," *Ultrasonics* **67**, 136–150 (2016).

[16] Wang, K., Matthews, T., ans C. Li, F. A., Duric, N., and Anastasio, M., "Waveform inversion with source encoding for breast sound speed reconstruction in ultrasound computed tomography," *IEEE Trans. Ultrason. Ferroelectr. Freq. Contr.* **62** (2014).

[17] Krebs, J. R., Anderson, J. E., Hinkley, D., Neelamani, R., Lee, S., Baumstein, A., and Lacasse, M.-D., "Fast full-wavefield seismic inversion using encoded sources," *Geophysics* **74**(6), WCC177–WCC188 (2009).

[18] Ben-Hadj-Ali, H., Operto, S., and Virieux, J., "An efficient frequency-domain full waveform inversion method using simultaneous encoded sources," *Geophysics* **76**(4), R109–R124 (2011).

[19] Crestel, B., Alexanderian, A., Stadler, G., and Ghattas, O., "A-optimal encoding weights for nonlinear inverse problems, with application to the helmholtz inverse problem," *Inverse Problems* **33**(7), 074008 (2017).

[20] van Leeuwen, T. and Herrmann, F. J., "Fast waveform inversion without source-encoding," *Geophysical Prospecting* **61**, 10–19 (2013).

[21] Guest, T. and Curtis, A., "Iteratively constructive sequential design of experiments and surveys with nonlinear parameter-data relationships," *Journal of Geophysical Research: Solid Earth* **114**(B4) (2009). B04307.

[22] Korta Martiartu, N., Boehm, C., Vinard, N., Jovanović Balic, I., and Fichtner, A., "Optimal experimental design to position transducers in ultrasound breast imaging," *Proc. SPIE* **10139**, 10139 – 10139 – 15 (2017).

[23] Maurer, H., Curtis, A., and Boerner, D. E., "Recent advances in optimized survey design," *Geophysics* **75**(5), 75A177–75A194 (2010).

[24] Maurer, H., Nuber, A., Korta Martiartu, N., Reiser, F., Boehm, C., Manukyan, E., Schmelzbach, C., and Fichtner, A., "Optimized experimental design in the context of seismic full waveform inversion and seismic waveform imaging," in [*Advances in Geophysics*], Nielsen, L., ed., **58**, 1 – 45, Elsevier (2017).

[25] Uciński, D., [*Optimal measurement methods for distributed parameter system identification*], CRC Press, Boca Raton (2005).

[26] Boehm, C., Korta Martiartu, N., Vinard, N., Jovanović Balic, I., and Fichtner, A., "Time-domain spectral-element ultrasound waveform tomography using a stochastic quasi-Newton method," to appear in Proc. SPIE (2018).

[27] Haber, E., Horesh, L., and Tenorio, L., "Numerical methods for experimental design of large-scale linear ill-posed inverse problems," *Inverse Problems* **24**(5) (2008).

[28] Bui-Thanh, T., Ghattas, O., Martin, J., and Stadler, G., "A computational framework for infinite-dimensional Bayesian inverse problems part I: The linearized case, with application to global seismic inversion," *SIAM Journal on Scientific Computing* **35**(6), A2494–A2523 (2013).

[29] Alexanderian, A., Petra, N., Stadler, G., and Ghattas, O., "A fast and scalable method for A-optimal design of experiments for infinite-dimensional bayesian nonlinear inverse problems," *SIAM Journal on Scientific Computing* **38**(1), A243–A272 (2016).

[30] Martin, J., Wilcox, L. C., Burstedde, C., and Ghattas, O., "A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion," *SIAM Journal on Scientific Computing* **34**(3), A1460–A1487 (2012).

[31] Haber, E., Magnant, Z., Lucero, C., and Tenorio, L., "Numerical methods for A-optimal designs with a sparsity constraint for ill-posed inverse problems," *Computational Optimization and Applications* **52**, 293–314 (May 2012).

[32] Curtis, A., "Optimal design of focused experiments and surveys," *Geophysical Journal International* **139**(1), 205–215 (1999).

[33] Fichtner, A., [*Full Seismic Waveform Modelling and Inversion*], Springer, Heidelberg (2010).

[34] Burger, M. and Kaltenbacher, B., "Regularizing Newton–Kaczmarz methods for nonlinear ill-posed problems," *SIAM Journal on Numerical Analysis* **44**(1), 153–182 (2006).

[35] Lou, Y., Zhou, W., Matthews, T. P., Appleton, C. M., and Anastasio, M. A., "Generation of anatomically realistic numerical phantoms for photoacoustic and ultrasonic breast imaging," *Journal of Biomedical Optics* **22**(4), 041015–041015 (2017).

[36] Afanasiev, M., Boehm, C., van Driel, M., Krischer, L., May, D. A., Rietmann, M., and Fichtner, A., "Salvus: An open-source high-performance package for full waveform modelling and inversion from laboratory to global scales," *Geophysical Journal International* (2018). submitted, manuscript ID GJI-S-17-1139.

[37] Boehm, C., *Efficient Inversion Methods for Constrained Parameter Identification in Full-Waveform Seismic Tomography*, PhD thesis, Technische Universität München (2015).