

Ultrasonic imaging through aberrating layers using covariance matching

van der Meulen, Pim; Coutino, Mario; Bosch, Johannes G.; Kruizinga, Pieter; Leus, Geert

DOI

[10.1109/TCI.2023.3302233](https://doi.org/10.1109/TCI.2023.3302233)

Publication date

2023

Document Version

Final published version

Published in

IEEE Transactions on Computational Imaging

Citation (APA)

van der Meulen, P., Coutino, M., Bosch, J. G., Kruizinga, P., & Leus, G. (2023). Ultrasonic imaging through aberrating layers using covariance matching. *IEEE Transactions on Computational Imaging*, 9, 745-759. <https://doi.org/10.1109/TCI.2023.3302233>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.


Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

<https://www.openaccess.nl/en/you-share-we-take-care>

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

Ultrasonic Imaging Through Aberrating Layers Using Covariance Matching

Pim van der Meulen , Mario Coutiño , *Member, IEEE*, Johannes G. Bosch , *Member, IEEE*, Pieter Kruizinga , and Geert Leus , *Fellow, IEEE*

Abstract—We consider the scenario of finding the transfer function of an aberrating layer in front of a receiving ultrasound (US) array, assuming a separate non-aberrated transmit source. We propose a method for blindly estimating this transfer function without exact knowledge of the ultrasound sources or acoustic contrast image, and without directly measuring the transfer function using a separate controlled calibration experiment. Instead, the measurement data of many unknown random images is collected, such as from blood flow, and its second-order statistics are exploited. A measurement model is formulated that explicitly defines the layer's transfer function. A covariance domain problem is then defined to eliminate the image variable, and it is solved for the layer's transfer function using manifold-based optimization. The proposed approach and calibration algorithm are evaluated on a range of challenging and realistic simulations using the *k-Wave* toolbox. Our results show that, given a sufficiently efficient parameterization of the layer's transfer function, and by jointly estimating the transfer function at multiple frequencies, the proposed algorithm is able to obtain an accurate estimate. Subsequent simulated imaging experiments using the obtained transfer function also show increased imaging performance in various aberrating layers, including a skull layer.

Index Terms—Aberration correction, blind calibration, covariance matching, projection algorithms, ultrasonic imaging.

I. INTRODUCTION

ULTRASOUND imaging is one of the most widely used medical imaging techniques due to its affordability, portability, its capability of imaging at very high frame rates (more than 1,000 Hz), and its real-time feedback for a clinician. Additionally, it does not utilize ionizing radiation or strong magnetic fields, and as such does not create potential adverse health effects for patients. However, it suffers from an inherent trade-off

between resolution and penetration depth. Moreover, imaging through certain layers in the human body, such as human skull bone tissue or fat and muscle layers, cause aberrations in the ultrasonic pulse-echo wavefields, resulting in imaging artifacts and loss of resolution [1], [2]. These wavefield aberrations are a result of both non-linear distortions (due to absorption) as well as linear distortions (due to echo time-of-arrival shifts, scattering within the layer, etc.). In this article we intend to address the linear distortions by estimating the Green's functions through an aberrating layer in front of an ultrasound array, and assume that the source is capable of transmitting the same (aberrated or not) pulse to each pixel. This limitation needs to ensure the mathematical tractability of solving the proposed blind calibration problem (discussed in more detail in footnote 1 in Sec. II). One-way aberrations are relevant for e.g. photo-acoustic imaging, and have potential when combined with, e.g., a moderately invasive ultrasound transmitter probe inserted beyond the aberrating layer (which needs to be only of a small size as it does not necessarily require a full array, but only a single point-like transmitter). If, even after aberration, the same pulse is transmitted to each pixel, our proposed method would also be applicable, and be non-invasive.

Most imaging techniques rely on the Born approximation [3], [4], which approximates a pulse-echo scenario by modeling transmitted waves as travelling throughout the region of interest (ROI), such as human tissue, uninterruptedly (as if there was a homogeneous background medium). Small inhomogeneities in human tissue cause small reflections of this forward field, which are also assumed to propagate to a receiving sensor without being altered. Hence, multiple scattering is ignored, and the model consists of an independent wave model for the propagation path from a transmitter to a scatterer, and from a scatterer to each receiver. More importantly, the unknown ultrasound image becomes linear w.r.t. the ultrasonic measurements, simplifying the imaging problem greatly. One of the conditions underlying the Born approximation is that scatterers do not scatter very strongly, so that the transmit field is indeed nearly unaltered by scatterers, and so that second-order (and higher) scattering is negligible. The Born approximation is implicitly assumed by most ultrasound beamforming algorithms, and the resulting beamforming algorithms are successful in scenarios with no aberrating layers and where the image mostly consists of weakly scattering contrasts.

Traditional approaches, however, are not effective when there are strongly aberrating layers placed between the sensor array

Manuscript received 16 October 2022; revised 5 May 2023 and 16 July 2023; accepted 24 July 2023. Date of publication 4 August 2023; date of current version 18 August 2023. This work was supported by the Netherlands Organisation for Scientific Research (NWO) through the ASPIRE Project (project 14926 within the STW OTP programme). The Associate editor coordinating the review of this manuscript and approving it for publication was Dr. Patrick LaRiviere. (Corresponding author: Pim van der Meulen.)

Pim van der Meulen, Mario Coutiño, and Geert Leus are with the Circuits and Systems group, Delft University of Technology, 2628 Delft, The Netherlands (e-mail: pqvandermeulen@gmail.com; m.a.coutinominguez@tudelft.nl; g.j.t.leus@tudelft.nl).

Johannes G. Bosch is with the Department of Cardiology, Erasmus Medical Center Thorax Biomedical Engineering, 3015 Rotterdam, The Netherlands (e-mail: j.bosch@erasmusmc.nl).

Pieter Kruizinga is with the Department of Neuroscience, Erasmus Medical Center, 3015 Rotterdam, The Netherlands (e-mail: p.kruizinga@erasmusmc.nl). Digital Object Identifier 10.1109/TCI.2023.3302233

and the ROI, even if scattering outside these layers is still linear, since it is not known how echoes propagate through this layer to the array of receiving sensors. To image correctly, we would have to know how the acoustic field propagates through the aberrating layer, which would require one to e.g. remove part of the skull, and thus violates the purpose of non-invasive imaging. Hence, we propose a method for *blindly* estimating an aberrating layer's transfer function. That is, to characterize it without an a priori known calibration image. Instead, we propose to measure many images of which only the second-order statistics need to be known, a condition which is much less demanding than prior knowledge of the true contrast image. We then use the second-order statistics of the set of measured echo signals in combination with the assumed known covariance of the images, to estimate the layer's transfer function.

An interesting layer in the context of this article is the human skull. As mentioned earlier, it is known to cause strong aberrations, and prohibits imaging of the brain, as well as non-invasive ablation of (malignant) tissue using high intensity focused ultrasound (HIFU) [1]. This is highly relevant for recent advances in ultrasound such as functional ultrasound imaging (fUS) [5], [6], [7]. Functional US imaging allows for inferring brain activity based on changes in the local haemodynamics that cause measurable changes in the ultrasound Doppler signal. Compared to fMRI, fUS allows for a more affordable, portable, and precise way of imaging brain activity [8]. It helps us to better understand the human brain, and it is a promising tool for surgeons to more accurately distinguish between healthy and malignant tissue when removing a brain tumor, preventing a surgeon from accidentally damaging healthy tissue, which could have disastrous results for the patient [9]. All of these examples urge us to investigate more advanced imaging techniques to deal with the distortions caused by the aforementioned layers.

Most studies address this problem by applying a time shift to each measured signal in the array (i.e., it is assumed aberrations are modeled using a phase screen directly in front of the array), based on a variety of goodness of fit criteria. Some examples from many are the studies of [10], [11], [12], [13], [14], [15]. However, these approaches often need to estimate a different phase screen for each imaging direction (direction into the ROI w.r.t. to ultrasound array), since a single phase screen would not be able to model more complex layers for all directions simultaneously. Moreover, one could argue such models are not accurate for more complicated layers where e.g. internal multiple scattering takes place. In such cases, the echo from a reflector as measured on different array elements could be different not only by a time shift, but also by the shape of pulse-echo signal. In addition, this makes it hard to use the methods above, as they rely on delay-and-sum beamforming, which requires signals to be in-phase after aberration correction and delaying before summing. A more holistic approach would view the US field in a single sensor on the array-side of the layer to be a *linear combination* of the *entire* 'input' field on the other side (Fig. 1).

In [16], [17], the authors address this issue by modeling aberration as a phase screen positioned some distance away from the aperture. Thus, an incident wave is first aberrated by the phase screen, followed by propagation to the sensor array.

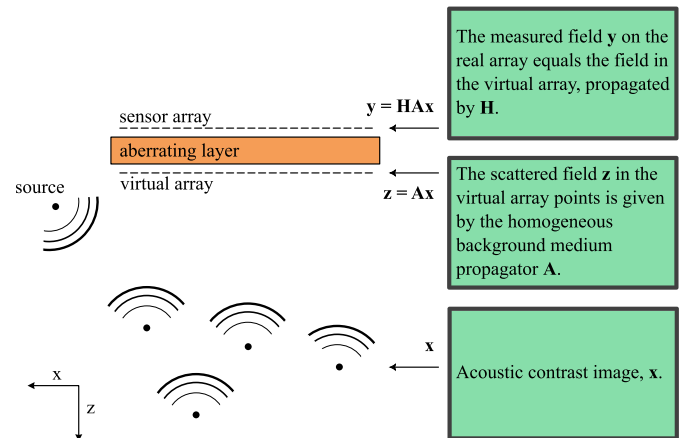


Fig. 1. Aberration layer distorts incident wavefields, leading to incorrect estimation of the acoustic contrast image. We are interested in estimating the transfer function from each point on the virtual array just before the aberration layer, to each point on the real array.

As a result, measurements on the array are linear combinations of the incident field just after the phase screen aberration occurred. Although such a model is appropriate for e.g. the human abdominal wall, we don't expect it to be capable of modeling more complicated layers such as the human skull, where it is not obvious how the porous bone would be modeled as a (collection of) phase screen(s).

Another large body of research is specifically focused on imaging through the human skull. In [18], time-reversal techniques are used to focus waves on transmit inside the brain. This is useful for acoustic therapy, but less so for directly imaging the brain. In [19], a similar goal is achieved by the design of a spatio-temporal filter. However, both of these techniques require one to first measure the transmitted ultrasound waves inside the brain to characterize the joint effect of the imaging setup and the skull, prohibiting non-invasive procedures. Recently, trans-cranial brain vasculature imaging was demonstrated using localization microscopy utilizing micro-bubbles inserted into the brain vasculature [20]. Their work exploits the use of micro-bubbles, which can be isolated in the back scattered echo signals (reflectors are spaced sufficiently apart in the brain vessels to do so), to analyze aberration profiles and model them using a phase-screen for regions in the ROI where bubbles were estimated to have roughly similar aberration profiles.

Another approach was proposed in [21], where two opposing ultrasound arrays are placed on the skull. The effect of the two skull layers on each side has to be estimated. To simplify matters, the authors assume that the skull wall on one side can be approximated as a thin phase layer, allowing for a more complex formulation and estimation of the effect of the other skull wall. However, a thin phase screen is arguably not an accurate approximation of the skull wall and consequently, it is possible that the estimated skull wall's transfer functions are sub-optimal. The technique proposed in our article only involves a single skull wall's transfer function, and it does not rely on approximations using thin phase screens.

A more recent approach ignores the Born approximation entirely, and tries to solve the acoustic wave equations using gradient descent optimization techniques [22], known as full waveform inversion (FWI). Their method relies on a sufficient initial estimate of the skull wall, and requires a measurement setup where sensors are distributed across the entire skull, instead of using only an array. Moreover, the technique is currently very computationally intensive, and requires many hours to obtain a single (3D) frame. This currently makes it entirely unsuitable for Doppler or fUS imaging, where hundreds to thousands of frames need to be imaged to obtain a single Doppler or fUS image.

Finally, we have already proposed an SVD-based technique to blindly estimate the effect of the aberrating layer from random images or source signals in the ROI. However, it relies on structural requirements of the aberration layer's transfer functions, and on the setup of the imaging scenario [23]. In the techniques we will discuss in this article, in contrast to existing studies, we will propose a more flexible and general method, where many kinds of layers can be characterized blindly, and without the need for two sensor arrays.

The method we propose in this new contribution heavily relies on *model-based imaging* techniques. Such techniques take into account the entire wave propagation model when forming an image, and try to find an image that, based on a model that relates the image to measured echo signals, best explains the observed echoes. Such techniques have been used before (see e.g., [24], [25], to just name a few), but are especially relevant to scenarios where traditional imaging techniques like delay-and-sum cannot be utilized. We have particularly shown this in [26], [27], where we intentionally distort transmitted fields using an aberration layer, and show that this allows 3D imaging using just a single sensor. In [26], [27], we carefully characterize the aberrating layer in order to incorporate it into the wave propagation model, after which an image is formed from echo signals based on this model. Hence, even heavily distorted signals where the echo of a single reflector is different on each sensor, and strongly smeared across time could be resolved, as long as it is known *how* the distortion takes place.

In this contribution we will show how to explicitly incorporate an extra variable representing the layer's transfer function into our echo signal model, an approach which has partly been introduced before in [21], but has not seen much use in literature. Specifically, we use the concept of a transfer function from an imaginary (virtual) array just in front of the aberrating layer, to the true sensor after the layer (Fig. 1). Next, we will show how to get rid of the contrast image variable, which we assume is unknown. This is done by collecting many measurements, and looking at their second-order statistics, instead of the raw measurements directly. As a result, the contrast image variable is integrated out of the signal model equations, after which the layer's transfer function is the only unknown. Unfortunately, this is only mathematically feasible if it is assumed that waves travel through the aberrating layer once (as explained in the next section). This limits the applicability of our model to photo-acoustic imaging and the use of e.g. a moderately invasive point-like source inserted beyond the aberrating layer. If the

source aberration is such that the same (although aberrated) waveform is transmitted to each pixel, our method would also still be applicable (as this pulse waveform would be trivially incorporated into \mathbf{H}), and be fully non-invasive. Finally, we propose a method based on covariance matching [28], using a manifold optimization technique, to obtain an estimate of the layer's transfer function. Other work in ultrasound imaging exploiting second-order statistics can be found in [29], [30], where the cross-correlations of the measured signals are used. However, these papers are still focused on time delays or assume that the aberrating layer acts identically on each point of the layer. More general layer transfer functions with interaction between channels are not included.

Using covariance measurements to find the array gain and phase has been studied in signal processing literature before, most notably by Friedlander and Weiss [31]. Covariance matching for array calibration has also been applied before to antenna arrays for radio astronomy imaging [32], and to acoustic vector sensor arrays [33]. However, only phase and gain changes per antenna are considered (instead of linear combinations of inputs), similar to [11], [12], [15] in ultrasonic imaging, instead of more complicated transfer functions. They thus assume a specific structure for the layer's transfer function, and can thus be seen as a special case of our approach. The authors use alternating descent and semidefinite programming methods, which are not directly applicable to the mathematical problem we discuss in this article. As far as we know, there has been no prior work where the estimation of general transfer functions using covariance matching is considered.

Notation: We will represent a vector \mathbf{x} in lower-case bold, and a matrix \mathbf{X} in upper-case bold. The operator $\mathbb{E}\{\cdot\}$ represents the statistical expectation operator, $\|\mathbf{X}\|_F$ the Frobenius norm of \mathbf{X} , and $\|\mathbf{x}\|_2$ the l_2 -norm of \mathbf{x} . The notation \mathbf{X}^\dagger represents the Moore-Penrose pseudo-inverse of \mathbf{X} . The matrix transpose and conjugate transpose are denoted using $(\cdot)^T$ and $(\cdot)^H$, respectively. Finally, the \otimes symbol represents the Kronecker product.

II. SIGNAL MODEL AND PROBLEM FORMULATION

In the case of a homogeneous medium with an acoustic contrast image, and using the Born approximation, we can assume the following linear measurement model for an M element array, and a single temporal frequency l :

$$\mathbf{y}_l = \mathbf{A}_l \mathbf{x} + \mathbf{n}_l, \quad (1)$$

where $\mathbf{y}_l \in \mathbb{C}^M$ groups the array pressure measurements for frequency l , $\mathbf{x} \in \mathbb{R}^N$ represents the scattering coefficients for N pixels, and $\mathbf{A}_l \in \mathbb{C}^{M \times N}$ is a known model that relates the measurements in \mathbf{y}_l to the image \mathbf{x} , and thus contains the expected reflected signals at frequency l received by the array for all pixels of interest. Finally, \mathbf{n}_l represents additive measurement noise. The matrix \mathbf{A}_l is well known when no aberration layer is present, and the Born approximation holds. It is easily generated since the Green's functions for the pulse propagation paths (from transmitter, to each pixel, to each receiver) and the speed of sound are known. In that case, the received echo signal for each such path is straightforwardly found by delaying the

transmitted pulse based on the total path length. Under the Born approximation, the scattering intensity of each pixel depends only on the difference in speed of sound and density between the pixel and the background speed of sound (contrast image), and is thus real-valued and frequency-independent. With no aberration layer present, image reconstruction boils down to estimating \mathbf{x} from \mathbf{y}_l , knowing \mathbf{A}_l , using the entire available bandwidth instead of a single frequency.

In the case of an aberration layer, we assume there is a linear transfer function from each point on the virtual array in Fig. 1 to each element in the real array. That is, there exists a Green's function $h_{i,j}(t)$ that represents the impulse response for the linear time-invariant channel from point i on the virtual array to sensor j on the real array. The measurement on any given sensor j is then the superposition of all the convolved channel responses from each virtual sensor i to j . In the frequency domain, that means the problem can be approached on a per-frequency basis, and we have the following new measurement model:

$$\mathbf{y}_l = \mathbf{H}_l \mathbf{A}_l \mathbf{x} + \mathbf{n}_l, \quad (2)$$

where $\mathbf{A}_l \in \mathbb{C}^{M \times N}$ is now related to the virtual array, and $\mathbf{H}_l \in \mathbb{C}^{M \times M}$ contains the frequency domain transfer function coefficients for frequency l , from each point on the virtual array to each point on the true array (note that in (2) it is assumed that waves travel through the aberrating layer only *once*,¹ forcing us to use a transmitter that is not impeded by the aberrating layer). Since \mathbf{H}_l represents the transfer function of an aberrating layer, it is unknown, and it is not possible to estimate \mathbf{x} .

When considering multiple frequencies, the \mathbf{A}_l matrices of the corresponding L frequencies can be stacked vertically into

¹In (2) and (4), \mathbf{A} contains the pulse-echo signal received across the virtual array from each pixel. This implicitly relies on the known Green's functions in a homogeneous medium from transmitter to pixel to each point on the virtual array. Note that it is thus assumed that there is a separate transmitter from which a pulse propagates into the imaging medium, which reflects of contrast inhomogeneities in the ROI towards the virtual and true array. After all, if the array is used for transmission, we would not know the field that is transmitted into the ROI. In that case, (4) would contain the variable \mathbf{H} *twice* (as pulses would propagate through the aberration layer both on transmit and receive), and (4) would become $\mathbf{y} = \mathbf{H} \mathbf{A} \text{diag}(\mathbf{x}) \mathbf{A}^T \mathbf{H}^T \mathbf{s}$ (where $\mathbf{s} \in \mathbb{C}^{ML}$ would contain the transmitted pulses on the array, and \mathbf{A} now represents the one-way propagation in a linear medium from a pixel to a point on the virtual array). This would make it much harder to estimate \mathbf{H} from \mathbf{y} , since \mathbf{y} would be quadratic w.r.t. \mathbf{H} (and (9) would be bi-quartic w.r.t. \mathbf{H}), and the techniques described in the next section are not trivially applicable to this equation. Hence, it is assumed that a small transmitter is inserted into the medium. When the aberration layer consists of human tissue, this would mean that the calibration procedure would be much less invasive, since only a small device has to be inserted (as opposed to an opening large enough to fit an array). We also point out that, if the source is aberrated but still sends the same waveform to each pixel, the waveform is trivially incorporated by scaling each \mathbf{H}_l with the corresponding waveform DFT coefficients. These coefficients are then calibrated 'for free', and the calibration would be fully non-invasive. As a result, only *one* layer's transfer function has to be estimated, instead of two (like in [21]), and we do not have to estimate one of the layers using e.g. a thin phase mask (sometimes an insufficient approximation), thus retaining the validity of our model (4). Photoacoustic scenarios would also be covered under the formulation in (4), since in that case all pixels in the optic transmit view would be excited simultaneously, and start transmitting a pulse towards the array at the same time.

a bigger matrix $\mathbf{A} \in \mathbb{C}^{ML \times N}$,

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_L \end{bmatrix}. \quad (3)$$

Defining $\mathbf{y} \in \mathbb{C}^{ML}$ and $\mathbf{n} \in \mathbb{C}^{ML}$ in a similar way, we have a new multi-frequency measurement equation

$$\mathbf{y} = \mathbf{H} \mathbf{A} \mathbf{x} + \mathbf{n}. \quad (4)$$

Since convolution in the time domain becomes a product in the frequency domain, $\mathbf{H} \in \mathbb{C}^{ML \times ML}$ will have a block diagonal structure and mostly contains zeros,

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 & & \vdots \\ \vdots & & \ddots & \\ \mathbf{0} & \dots & & \mathbf{H}_L \end{bmatrix}. \quad (5)$$

Note that the measurement (2) is simply a special case of (4) with $L = 1$. Furthermore, we point out that the usage of \mathbf{H} generalizes all linear aberration layers. For example, if the aberrating layer is modelled as a phase screen, \mathbf{H} will be a diagonal matrix, and is thus a special case of the formulation (4). Another example would be laterally invariant layers (i.e., layers that vary in the z -dimension, but not in the x -dimension), which would have a strictly Toeplitz structure.

In a *blind* calibration scenario, we wish to estimate \mathbf{H} from (4), without knowing \mathbf{x} . Estimating both \mathbf{H} and \mathbf{x} jointly would result in a bi-linear measurement equation, which is typically challenging to solve. Instead, we consider removing \mathbf{x} from our measurement equation by using covariance measurements, where \mathbf{x} would be integrated out of the measurement equation, and only the covariance matrix of \mathbf{x} would have to be known.

Therefore, we will interpret \mathbf{x} as a realization of a random process for which multiple realizations are obtained. Consider for example a liquid medium with moving micro-bubbles or contrast agents, so that each pulse-echo measurement has a different contrast image \mathbf{x} due to the displacement of the bubbles between measurements. Instead of bubbles, one could consider the flow of blood inside a patient. If stationary reflections are removed prior to calibration (using e.g. a high-pass filter in each pixel across the frame-dimension), one is left with only the signals originating from moving particles inside the blood stream. Assuming the measurement noise \mathbf{n} is uncorrelated with the contrast image \mathbf{x} , and defining the covariance matrices, $\mathbf{C}_n = \mathbb{E}\{\mathbf{n}\mathbf{n}^H\}$, $\mathbf{C}_y = \mathbb{E}\{\mathbf{y}\mathbf{y}^H\}$, and $\mathbf{C}_x = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$, we have the following measurement equation in the covariance domain:

$$\mathbf{C}_y = \mathbb{E}\{\mathbf{H} \mathbf{A} \mathbf{x} \mathbf{x}^T \mathbf{A}^H \mathbf{H}^H\} + \mathbb{E}\{\mathbf{n}\mathbf{n}^H\} \quad (6)$$

$$= \mathbf{H} \mathbf{A} \mathbf{C}_x \mathbf{A}^H \mathbf{H}^H + \mathbf{C}_n. \quad (7)$$

Interestingly, the image variable \mathbf{x} has been integrated out of the original measurement (4), and we now only need to know the second-order statistics of \mathbf{x} . This is a much less demanding

requirement. For example, if the ROI primarily consists of blood flow, and many pulse-echo measurements of that ROI are taken with sufficient time in between, it can be safely assumed that the corresponding \mathbf{x} -vectors are uncorrelated, and thus $\mathbf{C}_x = \mathbf{I}$. Correlated measurements due to non-moving ‘background’ tissue can be easily removed from measurements using clutter filtering, as is common in US Doppler imaging. The covariance data of the measurement noise \mathbf{C}_n can be obtained by taking measurements with no ultrasound transmission, and \mathbf{C}_y can be estimated by taking P measurements, and then using the estimate

$$\hat{\mathbf{C}}_y = \frac{1}{P} \sum_{p=1}^P \mathbf{y}^{(p)} \mathbf{y}^{(p)\text{H}} - \mathbf{C}_n, \quad (8)$$

where the known noise covariance matrix has already been subtracted for convenience in our derivations. We are interested, then, in estimating \mathbf{H} from $\hat{\mathbf{C}}_y$, knowing \mathbf{C}_x and \mathbf{A} . This is a covariance matching problem [28], where we want to minimize the following cost function:

$$\hat{\mathbf{H}} = \arg \min_{\mathbf{H}} \|\hat{\mathbf{C}}_y - \mathbf{H} \mathbf{A} \mathbf{C}_x \mathbf{A}^{\text{H}} \mathbf{H}^{\text{H}}\|_F^2. \quad (9)$$

This is not a trivial problem to solve since \mathbf{H} is *quartic* w.r.t. $\hat{\mathbf{C}}_y$. In the next section we will provide methods for solving this optimization problem.

III. CALIBRATION ALGORITHMS

A. Alternating Minimization

Defining $\mathbf{B} \triangleq \mathbf{A} \mathbf{C}_x \mathbf{A}^{\text{H}}$, one solution to estimate \mathbf{H} from \mathbf{C}_y , assuming it’s perfectly known for now, is given by

$$\hat{\mathbf{H}} = \mathbf{C}_y^{\frac{1}{2}} \mathbf{B}^{-\frac{1}{2}}. \quad (10)$$

Here, we define the square root of a matrix \mathbf{C} as any matrix $\mathbf{C}^{\frac{1}{2}}$ such that $\mathbf{C}^{\frac{1}{2}} \mathbf{C}^{\frac{1}{2}} = \mathbf{C}$. However, there is a larger set of solutions, which can be described by

$$\hat{\mathbf{H}} = \mathbf{C}_y^{\frac{1}{2}} \mathbf{Q} \mathbf{B}^{-\frac{1}{2}}, \quad (11)$$

where \mathbf{Q} is any orthogonal matrix, resulting in a non-unique solution.

To further narrow down the solution space, we assume that \mathbf{H} can be parameterized by fewer parameters than the total number of entries in \mathbf{H} . These parameters will be represented by $\boldsymbol{\theta}$, and the related matrix is then denoted as $\mathbf{H}(\boldsymbol{\theta})$. Hence, to estimate \mathbf{H} , we would like to find the correct \mathbf{Q} and $\boldsymbol{\theta}$ by solving the following problem instead of (9):

$$\{\hat{\boldsymbol{\theta}}, \hat{\mathbf{Q}}\} = \arg \min_{\boldsymbol{\theta}, \mathbf{Q}} \|\mathbf{C}_y^{\frac{1}{2}} \mathbf{Q} \mathbf{B}^{-\frac{1}{2}} - \mathbf{H}(\boldsymbol{\theta})\|_F^2, \text{ s.t. } \mathbf{Q} \in \mathcal{Q}, \quad (12)$$

where \mathcal{Q} represents the set of orthogonal $ML \times ML$ matrices. In other words, we try to find a solution that will give a zero cost function for (9) (any orthogonal \mathbf{Q}), while also penalizing the distance to the set of parameterized matrices $\mathbf{H}(\boldsymbol{\theta})$. Using this formulation the cost function is no longer quartic w.r.t. \mathbf{H} . However, even if $\mathbf{H}(\boldsymbol{\theta})$ is linearly parameterized, the problem above is not a convex problem, due to the orthogonality constraint.

In order to arrive at a solution for this problem, we propose to alternately optimize for $\boldsymbol{\theta}$ and \mathbf{Q} , as outlined in Algorithm 1. When \mathbf{Q} is fixed, $\boldsymbol{\theta}$ can be found by solving a linear least squares problem. We assume that \mathbf{H} can be linearly parameterized by $\boldsymbol{\theta}$ with parameterization matrix² \mathbf{G} , such that $\text{vec}(\mathbf{H}(\boldsymbol{\theta})) = \mathbf{G}\boldsymbol{\theta}$. This way, solving for this step admits a unique global minimum which can be found either iteratively or analytically. When $\boldsymbol{\theta}$ is fixed, (12) becomes a weighted orthogonal Procrustes problem (WOPP). The latter is hard to solve, since it does not allow an analytical solution, the cost function contains multiple local minima, and iterative methods are not guaranteed to find the global minimizer (see [34] and the references therein). These two problems are solved alternately until the algorithm converges, or a maximum amount of iterations has been completed. In the next subsection we propose an improved version of this algorithm that addresses the problems mentioned above.

When jointly calibrating for multiple frequencies, the block-diagonal structure of \mathbf{H} in (5) is enforced by \mathbf{G} , while additionally imposing a parameterization for the blocks on the diagonal representing each \mathbf{H}_l . Due to the block-diagonal structure of \mathbf{H} , the number of unknowns increases linearly with each added frequency, whereas the number of new covariance measurements increases quadratically. This is of course limited by the echo signal bandwidth and the Nyquist rate in both space and time. Moreover, space-frequency samples in \mathbf{C}_y tend to decorrelate for large frequency and space differences between entries. The additional information in \mathbf{C}_y consists of the cross-correlations between temporal frequencies, which now also have to be matched by the covariance-matching algorithm. Hence, we expect the probability of finding a unique solution for \mathbf{H} to increase as the bandwidth L increases.

The above procedure, summarized in Algorithm 1, is an alternating projection algorithm, where $\boldsymbol{\theta}$ and \mathbf{Q} are alternately projected onto a linear subspace, and the Stiefel manifold, respectively. We have not been able to find a similar problem or algorithm in literature, and are not aware of any analyses concerning convergence and global and local minima. However, given the non-convex constraint imposed by the Stiefel manifold, it is likely that there are multiple local minima. Moreover, there are at least two global minima, because the solutions $\{\hat{\boldsymbol{\theta}}, \hat{\mathbf{Q}}\}$ and $\{-\hat{\boldsymbol{\theta}}, -\hat{\mathbf{Q}}\}$ should have the same cost function value in (12).

In [35], cyclical algorithms are proposed, which use a similar alternating minimization problem, although the authors in [35] are not necessarily looking for a unique solution, since any solution in the feasible set should suffice. In our work, however, we are looking for a specific solution: the one that is as close as possible to the true \mathbf{H} . The work does not analyse local and global minima. However, similar to [35], we can point out that (9) is zero if and only if (12) is also zero. Moreover, if the global minimum of (9) is small, then (12) should be small as well. It is then argued that the two cost functions are ‘nearly’ equivalent

²A basis \mathbf{G} can be found by simulating the layer of interest in an accurate simulator such as *k-Wave*, and studying the structure in \mathbf{H} . For the sake of simplicity, we used a 2D Fourier basis in our experiments, removing the high frequencies from the basis. Finding proper bases is a topic of future research.

Algorithm 1: Alternating Optimization Algorithm.

-
- 1: Input: N_{it} , $\mathbf{C}_y^{\frac{1}{2}}$, $\mathbf{B}^{-\frac{1}{2}}$, \mathbf{G} .
 - 2: Initialize: random orthogonal matrix $\hat{\mathbf{Q}}$, $n = 0$.
 - 3: **repeat**
 - 4: $\hat{\boldsymbol{\theta}} = \mathbf{G}^\dagger \text{vec} \left(\mathbf{C}_y^{\frac{1}{2}} \hat{\mathbf{Q}} \mathbf{B}^{-\frac{1}{2}} \right)$
 - 5: $\hat{\mathbf{H}} = \text{unvec}(\mathbf{G} \hat{\boldsymbol{\theta}})$
 - 6: Solve the WOPP $\hat{\mathbf{Q}} = \arg \min_{\mathbf{Q}} \|\mathbf{C}_y^{\frac{1}{2}} \mathbf{Q} \mathbf{B}^{-\frac{1}{2}} - \hat{\mathbf{H}}\|_F^2$
 - 7: $n = n + 1$
 - 8: **until** $n = N_{it}$
-

in the sense that if the global minimizer of (9) is sufficiently small, that the sequences minimizing both cost functions can be expected to lead to similar results. However, the cost functions are not equivalent, and will have different local minima and minimization sequences in general.

B. Alternating Optimization for Ill-Conditioned $\mathbf{B}^{\frac{1}{2}}$ and Low-Rank $\mathbf{C}_y^{\frac{1}{2}}$

Algorithm 1 has several drawbacks that we intend to address in this subsection. Most importantly, a WOPP has to be solved in Algorithm 1, which requires iterative methods with no guarantee of obtaining a global minimizer. The altered cost function proposed in this subsection will lead to a formulation where only an *unweighted* orthogonal Procrustes problem needs to be solved, for which an analytical solution is available. Another issue is that when \mathbf{H} is low-rank, $\mathbf{C}_y^{\frac{1}{2}}$ is low rank as well. Consequently, there will be no unique solution to the weighted orthogonal Procrustes problem in Algorithm 1. Moreover, the matrix \mathbf{A} is typically ill-conditioned in the case of (multi-frequency) imaging problems, making the inversion of $\mathbf{B}^{\frac{1}{2}}$ problematic, and also causing $\mathbf{C}_y^{\frac{1}{2}}$ to be ill-conditioned. The issue of ill-conditioning is the one we encountered in our experiments, and thus of most importance. For the sake of completeness, we also show how to deal with a low-rank \mathbf{C}_y .

To address these issues, we start by observing that instead of solving (9), we can also take the matrix square root of the terms in (9), similar to the work in [35], to obtain a solution using the following minimization problem:

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \|\mathbf{C}_y^{\frac{1}{2}} - \mathbf{H}(\boldsymbol{\theta}) \mathbf{A} \mathbf{C}_x^{\frac{1}{2}}\|_F^2 \quad (13)$$

$$= \arg \min_{\boldsymbol{\theta}} \|\mathbf{C}_y^{\frac{1}{2}} - \mathbf{H}(\boldsymbol{\theta}) \mathbf{B}^{\frac{1}{2}}\|_F^2. \quad (14)$$

Of course, this is only one of many solutions, and similar to the previous subsection, multiplication by an orthogonal matrix \mathbf{Q} of the left term will also provide a solution to the original cost function (9). Incorporating this term into the new cost function leads to:

$$\{\hat{\boldsymbol{\theta}}, \hat{\mathbf{Q}}\} = \arg \min_{\boldsymbol{\theta}, \mathbf{Q}} \|\mathbf{C}_y^{\frac{1}{2}} \mathbf{Q} - \mathbf{H}(\boldsymbol{\theta}) \mathbf{B}^{\frac{1}{2}}\|_F^2, \text{ s.t. } \mathbf{Q} \in \mathcal{Q}. \quad (15)$$

When $\boldsymbol{\theta}$ is fixed, solving for \mathbf{Q} now has become an OPP instead of a WOPP, and the global optimum for the OPP in the current iteration can be found analytically. The solution for $\boldsymbol{\theta}$ when \mathbf{Q}

Algorithm 2: Alternating Optimization for Ill-Conditioned $\mathbf{B}^{\frac{1}{2}}$ and Low-Rank $\mathbf{C}_y^{\frac{1}{2}}$.

-
- 1: Input: N_{it} , $\mathbf{B}^{\frac{1}{2}}$, \mathbf{G} , \mathbf{U} , $\boldsymbol{\Sigma}$.
 - 2: Initialize: random matrix $\mathbf{P} \in \mathcal{Q}_{k \times ML}$, $n = 0$.
 - 3: **repeat**
 - 4: $\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \|\mathbf{C}_y^{\frac{1}{2}} - \mathbf{H}(\boldsymbol{\theta}) \mathbf{B}^{\frac{1}{2}}\|_F^2 + \lambda \|\boldsymbol{\theta}\|_2^2$
 - 5: $\hat{\mathbf{H}} = \text{unvec}(\mathbf{G} \hat{\boldsymbol{\theta}})$
 - 6: Decompose: $\boldsymbol{\Sigma}^{\frac{H}{2}} \mathbf{U}^H \hat{\mathbf{H}} \mathbf{B}^{\frac{1}{2}} = \tilde{\mathbf{U}} \tilde{\boldsymbol{\Sigma}} \tilde{\mathbf{V}}^H$
 - 7: $\hat{\mathbf{P}} = \tilde{\mathbf{U}} \tilde{\mathbf{V}}^H$
 - 8: $n = n + 1$
 - 9: **until** $n = N_{it}$
-

is fixed is found by solving a (linear) least squares problem. In case of an ill-conditioned $\mathbf{B}^{\frac{1}{2}}$, one can use any regularization technique when solving for $\boldsymbol{\theta}$, such as Tikhonov regularization. As discussed in the previous subsection, the local minima and minimization sequences of (9) and (15) are not equivalent in general, although they should have similar behaviour for small enough values of the cost functions, and (9) can only be zero if (15) is also zero.

Next, we rewrite the problem to only solve for the column space of $\mathbf{C}_y^{\frac{1}{2}}$ as follows. First, we define the *economy size* SVD of \mathbf{C}_y : $\mathbf{C}_y = \mathbf{U} \boldsymbol{\Sigma} \mathbf{U}^H$, and use $\mathbf{C}_y^{\frac{1}{2}} = \mathbf{U} \boldsymbol{\Sigma}^{\frac{1}{2}} \mathbf{U}^H$. Here $\mathbf{U} \in \mathbb{C}^{ML \times k}$, where $k = \text{rank}(\mathbf{C}_y)$, and $\boldsymbol{\Sigma} \in \mathbb{C}^{k \times k}$. Next, we rewrite the cost function in (15) by left-multiplying the terms in (15) by \mathbf{U}^H , leading to the problem

$$\{\hat{\boldsymbol{\theta}}, \hat{\mathbf{Q}}\} = \arg \min_{\boldsymbol{\theta}, \mathbf{Q}} \|\boldsymbol{\Sigma}^{\frac{1}{2}} \mathbf{U}^H \mathbf{Q} - \mathbf{U}^H \mathbf{H}(\boldsymbol{\theta}) \mathbf{B}^{\frac{1}{2}}\|_F^2, \text{ s.t. } \mathbf{Q} \in \mathcal{Q}, \quad (16)$$

or equivalently,

$$\{\hat{\boldsymbol{\theta}}, \hat{\mathbf{P}}\} = \arg \min_{\boldsymbol{\theta}, \mathbf{P}} \|\boldsymbol{\Sigma}^{\frac{1}{2}} \mathbf{P} - \mathbf{U}^H \mathbf{H}(\boldsymbol{\theta}) \mathbf{B}^{\frac{1}{2}}\|_F^2, \text{ s.t. } \mathbf{P} \in \mathcal{Q}_{k \times ML}, \quad (17)$$

where $\mathbf{P} = \mathbf{U}^H \mathbf{Q}$, and $\mathcal{Q}_{k \times ML}$ represents the set of $k \times ML$ matrices with unit-norm orthogonal rows. The solution to (17) for \mathbf{Q} (an OPP) is given by [34]:

$$\hat{\mathbf{P}} = \tilde{\mathbf{U}} \tilde{\mathbf{V}}^H, \quad (18)$$

where $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{V}}$ collect the k most significant left and right singular vectors of $\mathbf{Z} = \boldsymbol{\Sigma}^{\frac{H}{2}} \mathbf{U}^H \mathbf{H}(\boldsymbol{\theta}) \mathbf{B}^{\frac{1}{2}} \in \mathbb{C}^{k \times ML}$, respectively.

Although we have changed the cost function, it is now more easily solved for \mathbf{Q} if $\boldsymbol{\theta}$ is fixed: instead of a WOPP, we only have to solve an OPP, for which a global minimum can be found analytically. Moreover, it avoids computing the inverse of $\mathbf{B}^{\frac{1}{2}}$, and ensures there is a unique solution to the OPP when \mathbf{C}_y is rank-deficient, by only solving for the column space of \mathbf{C}_y . We will use cost function (17) throughout this article to avoid the aforementioned issues. The updated algorithm is shown in Algorithm 2.

C. Manifold Gradient Descent

From Algorithm 1, line 4, we can directly substitute the solution for θ into (12), resulting in the following simpler cost function where \mathbf{Q} is the only optimization variable:

$$\begin{aligned}\hat{\mathbf{Q}} &= \arg \min_{\mathbf{Q} \in \mathcal{Q}} \left\| \text{vec} \left(\mathbf{C}_y^{\frac{1}{2}} \mathbf{Q} \mathbf{B}^{-\frac{1}{2}} \right) - \mathbf{G} \mathbf{G}^\dagger \text{vec} \left(\mathbf{C}_y^{\frac{1}{2}} \mathbf{Q} \mathbf{B}^{-\frac{1}{2}} \right) \right\|_2^2, \\ &= \arg \min_{\mathbf{Q} \in \mathcal{Q}} \|\mathbf{F} \text{vec}(\mathbf{Q})\|_2^2,\end{aligned}\quad (19)$$

where $\mathbf{F} = (\mathbf{I} - \mathbf{G} \mathbf{G}^\dagger) \left(\mathbf{B}^{-\frac{T}{2}} \otimes \mathbf{C}_y^{\frac{1}{2}} \right)$. Given a \mathbf{Q} , the operator \mathbf{F} first computes the corresponding estimate of \mathbf{H} , $(\mathbf{B}^{-\frac{T}{2}} \otimes \mathbf{C}_y^{\frac{1}{2}}) \text{vec}(\mathbf{Q})$, and then projects it onto the nullspace of \mathbf{G} . Toolboxes such as ManOpt [36] can be used to find a solution to (19), which uses a Riemannian based gradient descent method.

In a similar fashion, (17) can be rewritten as

$$\begin{aligned}\hat{\mathbf{P}} &= \arg \min_{\mathbf{P} \in \mathcal{Q}_{k \times ML}} \left\| \text{vec} \left(\Sigma^{\frac{1}{2}} \mathbf{P} \right) - \left(\mathbf{B}^{\frac{T}{2}} \otimes \mathbf{U}^H \right) \mathbf{G} \theta \right\|_2^2 \quad (20) \\ &= \arg \min_{\mathbf{P} \in \mathcal{Q}_{k \times ML}} \left\| \text{vec} \left(\Sigma^{\frac{1}{2}} \mathbf{P} \right) - \tilde{\mathbf{G}} \tilde{\mathbf{G}}^\dagger \text{vec} \left(\Sigma^{\frac{1}{2}} \mathbf{P} \right) \right\|_2^2,\end{aligned}\quad (21)$$

where $\tilde{\mathbf{G}} = \left(\mathbf{B}^{\frac{T}{2}} \otimes \mathbf{U}^H \right) \mathbf{G}$. Defining $\tilde{\mathbf{F}} = (\mathbf{I} - \mathbf{G} \mathbf{G}^\dagger) (\mathbf{I} \otimes \Sigma^{\frac{1}{2}})$, we end up with

$$\hat{\mathbf{P}} = \arg \min_{\mathbf{P} \in \mathcal{Q}_{k \times ML}} \|\tilde{\mathbf{F}} \text{vec}(\mathbf{P})\|_2^2. \quad (22)$$

Our initial investigations showed that, although both algorithms converge to the same solution, Algorithm 2 converges faster with faster iterations, so we do not consider these gradient descent methods in the Results section. On the other hand, they show what kind of problem is being solved in a more compact manner.

IV. NUMERICAL AND SIMULATION RESULTS

We will evaluate the proposed algorithm using various simulated aberration layers, in increasing complexity and difficulty. Our main approach will be as follows.

- 1) Obtain the true \mathbf{H} from a k -Wave simulation containing only the aberrating layer. This will act as a ground truth.
- 2) Generate calibration data synthetically from (4), using the true \mathbf{H} from step 1 and P random $\mathbf{x}^{(p)}$ -vectors, resulting in P \mathbf{y} -vectors.
- 3) Estimate the covariance matrix \mathbf{R}_y from the obtained \mathbf{y} -vectors, using (8).
- 4) Estimate \mathbf{H} from the estimated \mathbf{R}_y using Algorithm 2.
- 5) Evaluate imaging performance by simulating the echo RF signal of the aberration layer with a verification image (different from the data in step 2) in k -Wave, and attempt to image it with the calibrated model.

These steps are also presented in Fig. 2, and explained in more detail below.

1) *Ground Truth of \mathbf{H}* : first, we will design the layer of interest in the k -Wave toolbox [37]. To avoid lengthy simulation times, and to have a reference ground truth, we directly measure the \mathbf{H}_l matrices in k -Wave, from which we can form \mathbf{H} . To do

so, we transmit the excitation pulse (the same one that we would transmit from our transmit transducer) from one point on the virtual array, and measure the entire response on the receiving array. This single simulation is then repeated for each virtual array point. Thus, only the aberration layer, and no acoustic contrast image was present in these simulations. This \mathbf{H} will act as our ground truth throughout most of the experiments, and will also be used to efficiently generate many measurements.

2) *Generate Calibration Measurements From Many Random Contrast Images*: since \mathbf{H} is now known, we can generate synthetic calibration measurements using (4) for many random \mathbf{x} . To this end, we generate P realizations of a spatially white zero mean i.i.d. Gaussian image, with $\mathbf{C}_x = \mathbf{I}$. The contrast images are zero mean, because we assume that the speed of sound of the inhomogeneities (consider e.g. various moving particles in a blood stream) can be either higher or lower than the background medium speed of sound. This results in P \mathbf{y} -vectors, from which an estimated covariance matrix $\hat{\mathbf{R}}_y$ is obtained using (8). To every measurement we add a different realization of spatially and temporally white zero mean i.i.d. Gaussian measurement noise, using $\mathbf{C}_n = \sigma_n^2 \mathbf{I}$.

3) *Estimation of the Covariance Matrix \mathbf{R}_y* : from these synthesized measurements, we use the estimated covariance matrix as defined in (8), using $\mathbf{C}_n = \sigma_n^2 \mathbf{I}$, to estimate \mathbf{H} . We assume the noise covariance is known, since it could be measured, e.g., by measuring on the used ultrasound transducer without any echo signals.

4) *Calibration*: We use Algorithm 2 based on the alternative cost function described in Section III-B for all the results, using $k = LM$ (i.e., we don't use a low-rank approximation). We found it unnecessary to use regularization as described by Algorithm 2.

5) *Imaging Performance*: since the measured \mathbf{H} acts as a ground truth, we can compare our estimated \mathbf{H} with the true one. We will use the normalized correlation coefficient

$$\rho = \text{vec}(\mathbf{H})^H \text{vec}(\hat{\mathbf{H}}) / (\|\mathbf{H}\|_F \|\hat{\mathbf{H}}\|_F) \quad (23)$$

to do so. However, how well \mathbf{H} describes the true propagation throughout the aberrating layer depends on (a) whether the assumption of linear propagation holds, and (b) whether the virtual array samples the incident ultrasound field sufficiently well. Since we will generate measurement data for estimating \mathbf{C}_y using an assumed true \mathbf{H} as described above, these assumptions are automatically satisfied. However, we will verify if we can truly obtain better imaging performance, by performing a separate imaging experiment using the calibrated model (i.e., estimating \mathbf{x} from a single measurement \mathbf{y}). The measurement is obtained using a k -Wave simulation containing both the aberrating layer and the acoustic contrast image. Since the k -Wave toolbox iteratively evaluates the acoustic wave equations, and no model assumptions such as the Born approximation are made, the entire signal model and imaging performance is tested in a realistic way.

Best apodized phase screen approximation: We also compare the imaging performance of the proposed calibration method to that of the phase screen that comes closest to the true \mathbf{H} -matrix. Since phase screens only apply local phase changes, they are

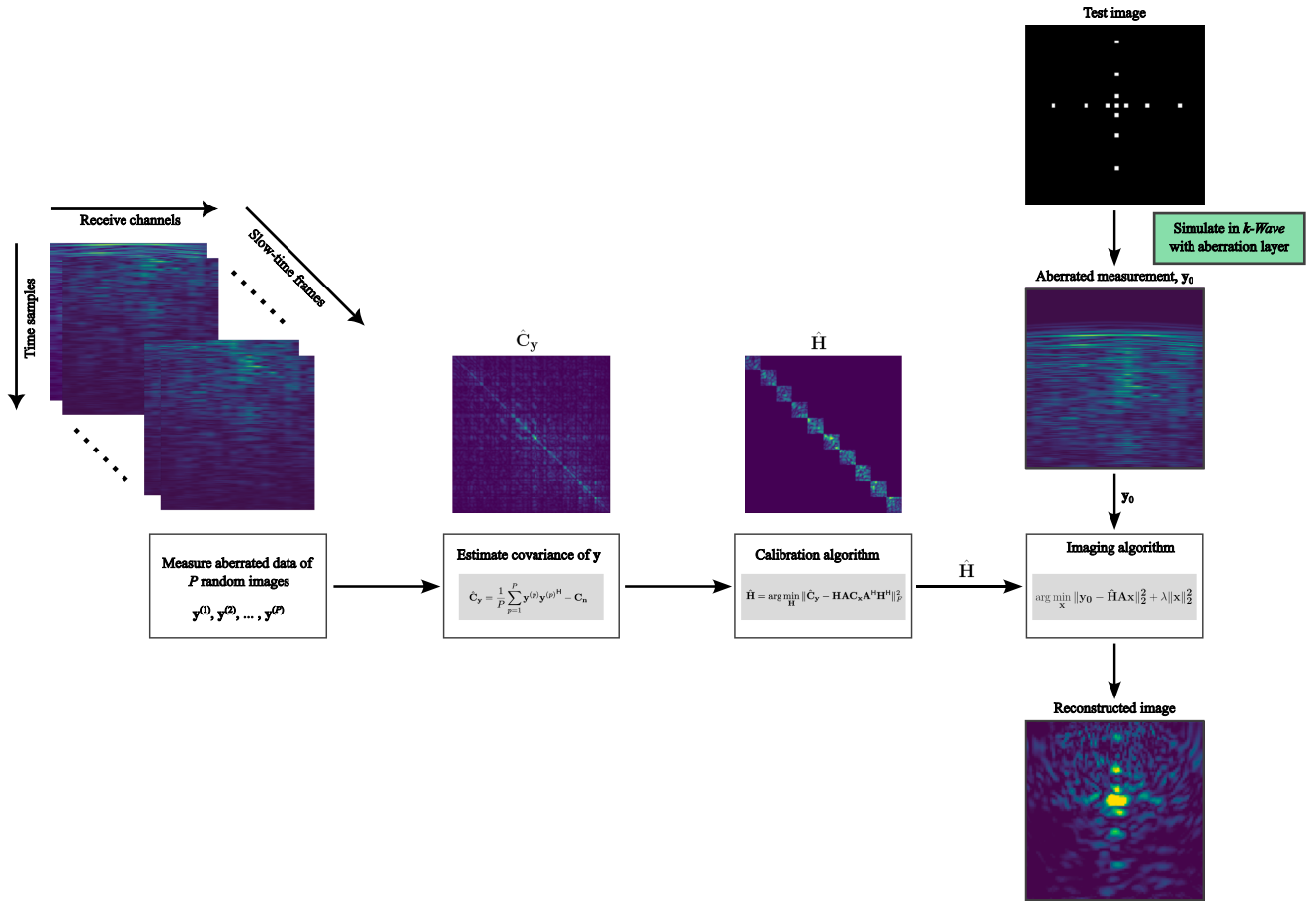


Fig. 2. Signal processing and imaging performance evaluation flow for the Results section. From left to right; first, images from random images are synthetically generated using (4). From these measurements, the covariance matrix of y can be estimated (second step). Next, we use our proposed calibration algorithm to obtain an estimate of \mathbf{H} (third step). Finally, we use an independent k -Wave simulation with a test image and aberration layer to generate a single test measurement. We use the estimated \mathbf{H} to estimate an image from this measurement and evaluate performance (right-hand side).

represented using purely diagonal \mathbf{H} -matrices. *The best possible apodized phase screen is then defined as taking only the diagonal values of the true \mathbf{H}* (both phase and amplitude). Consequently, any method that estimates a apodized phase screen to model aberrations, and takes it into account for imaging, cannot do better than this approximation. Here, the adjective ‘apodized’ is used since the amplitude is taken, in addition to the complex phase.

PSF analysis: for a quantitative measure of the imaging performance, we also generate the PSF for a single pixel in the ROI, and compute its full-width at half maximum (FWHM) measure, the integrated sidelobe level (ISL), and the distance of the PSF peak to the true pixel position. The FWHM is defined as the -3 dB cutoff point relative to the PSF peak. The ISL is obtained by averaging all the sidelobe levels, and are reported in dB using $10 \log_{10}$ (i.e., we don’t use the power of the image).

Parameterization of \mathbf{H} : Throughout this section, we use 2 different parameterizations of \mathbf{H} . For laterally invariant layers, such as in Section IV-A and IV-D, \mathbf{H} is easily parameterized using its Toeplitz structure. This results in $2M - 1$ parameters per frequency. Since there was no such obvious representation for the other, more complicated, aberration layers (we have tried

several popular transforms, such as the 2D DCT, wavelet, and wave atom bases), we decided to use a very general representation using the 2D FFT of the transfer function matrix \mathbf{H}_l of each frequency, removing only the higher spatial frequencies, resulting in approx. 350 (out of $M^2 = 400$) unknowns per frequency. The rationale behind this approach is that the 2D FFT is generally applicable, and we are thus testing our calibration algorithm using a sub-optimal parameterization, resulting in pessimistic results. Consequently, we expect improved results when using layer-specific \mathbf{H} -matrix parameterizations, as discussed in the Discussion section.

To choose these 350 coefficients, we took the 2D FFT of the \mathbf{H} for each temporal frequency, and selected the basis vectors of the largest coefficients. The selected vectors of all frequencies were then joined together, resulting in a single, basis that could be used regardless of which frequency is calibrated (and is not specifically tailored for each frequency specifically).

Other physical parameters: the matrix \mathbf{A} is generated assuming wave propagation from a known source location, to each pixel, towards the virtual array. For this, we assume a homogeneous background medium with a speed of sound of 1490 m/s, and a density of 1000 kg/m³ (similar to human

TABLE I
SIMULATION PARAMETERS FOR THE VARIOUS CALIBRATION EXPERIMENTS IN SECTION IV

Experiment	Aberrating layer	Parameterization of \mathbf{H}	Single/multi freq. calibration	# Calibration measurements	SNR
A	Laterally invariant	Toeplitz matrix	Calibration per frequency ($L = 1$)	$P = 500$	25 dB
B	Smoothly varying	2D FFT	Joint frequency calibration ($L = 36$)	$P = 10,000$	30 dB
C	Skull-like	2D FFT	Joint frequency calibration ($L = 46$)	$P = 10,000$	30 dB
D	Laterally invariant	Toeplitz matrix	Joint frequency calibration ($L = 96$)	$P = 1,000$	30 dB

tissue). For the sake of simplicity, all pixels act as independent sources, omni-directionally transmitting a Gaussian pulse centered around 5 MHz, unless otherwise specified. Note that this does not invalidate our model assumptions, since the Born approximation ignores multiple scattering, implying each acoustic contrast acts as an independent source. In the final experiment of Section IV-D, we will forego these assumptions, and use a full *k-Wave* simulation for each of the P calibration measurements using full contrast-based scattering.

All simulations were done using relatively small arrays, in order to keep most matrices and computations small. The parameter P was chosen relatively large but realistic for most experiments so that a good approximate of \mathbf{R}_y is obtained. We think this is a reasonable assumption due to the high frame-rates in US imaging. Finally, the choice of the number of frequencies L being jointly calibrated was chosen using trial-and-error. Generally speaking, experiments were done in the bandwidth 2–8 MHz, which we evenly divided into L equally-spaced frequency bins. Here, we tried to keep L as small as possible, and increased it until our algorithms were able to converge to a sensible solution, using trial-and-error.

Some of the important calibration data parameters have been summarized in Table I.

A. Triple-Layered Aberration Layer

First, we investigate the aberration layer in Fig. 3. It consists of a water-layer ‘sandwiched’ in between two homogeneous bone layers. It is important to realize that this layer is laterally *space-invariant*, so that the spatial transfer function also becomes space-invariant (i.e., it is only dependent on the lateral distance between the virtual and true array element). Consequently, each \mathbf{H}_l will have a Toeplitz structure, which was also confirmed by inspecting the \mathbf{H}_l matrices obtained from *k-Wave* simulations. Using this Toeplitz structure, \mathbf{G} only requires $2M - 1$ basis vectors per frequency, resulting in less unknowns. We use random scatterers in a ROI of 10 mm width by 8 mm depth for calibration. The calibration is performed for each temporal frequency independently ($L = 1$), and using $P = 500$ frames. Both the true and virtual array have $M = 20$ point-like sensors, placed at $175 \mu\text{m}$ intervals.

In Fig. 3, we show the calibration performance for various SNRs and temporal frequencies. Since the signal power is frequency dependent, we adjust the noise power accordingly to obtain the same SNR for each frequency. In the case of joint multi-frequency calibration in subsequent subsections we don’t do this, so that the SNR will vary per frequency. The transfer functions in \mathbf{H} are Toeplitz for all frequencies, and thus the calibration performance depends only on SNR. Each calibration finished within less than a second per frequency.

Finally, the bottom half of Fig. 3 shows the imaging results using our calibration method for a single point and a small collection of points. After estimating \mathbf{H} for all temporal frequencies, all frequencies are combined when estimating the contrast image \mathbf{x} . All measurement data \mathbf{y} is obtained from *k-Wave* simulation experiments. All images are reconstructed using the LSQR algorithm [38] to solve (2) for \mathbf{x} , which is regularized by limiting the amount of LSQR iterations. As can be seen from the single point, the uncalibrated PSF deteriorates in the axial direction due to the multiple reflections in the aberration layer, which are mistaken as echoes from separate sources when no calibration is utilized. This becomes even more clear in Fig. 3, bottom right, where the calibrated model is able to resolve most scatterers, in contrast to the uncalibrated model. The calibrated model gives much better reconstruction results, on par with the perfectly calibrated model. The best phase screen approximation is also able to correct for the depth offset delay of the aberrating layer, but is not able to account for multiple bouncing inside the layer. Consequently, it looks similar to the uncalibrated image, shifted in depth, exhibiting the same lack of resolution. These observations are also visible in the FWHM and ISL results in Table II.

We also attempted to obtain a good image by only tweaking the speed of sound in \mathbf{A} (and $\mathbf{H} = \mathbf{I}$). However, we were not able to obtain a reasonable image reconstruction using this approach, since the multitude of layers reverberate, which is probably the main limiting factor for imaging through this layer.

B. Joint Multi-Frequency Calibration: Smoothly Varying Layer

In the previous subsection we showed that calibration is possible using a single frequency if \mathbf{H}_l has a Toeplitz structure, i.e., when the aberration layer is space-invariant in the x -dimension. However, we have encountered significant difficulty calibrating for a single frequency for non-Toeplitz matrices. In such cases, we typically are able to find many solutions with a near-zero cost function.

In this subsection, we resolve this issue by using the multi-frequency approach to still obtain a good estimate of \mathbf{H} . We simulate a homogeneous, smoothly shaped layer as shown in Fig. 4. For the parameterization θ , we use the fact that the measured \mathbf{H}_l -matrices in *k-Wave* have a clear and smooth structure (Fig. 4, top right). Although probably not optimal, for now the parameterization matrix \mathbf{G} will consist of a 2D FFT representation, omitting some of the higher spatial frequencies, as described in the introduction to this section. The random images are located in a ROI of 10 mm in width by 5 mm in depth. Both the virtual and true array contain $M = 20$ elements, spaced at $100 \mu\text{m}$ intervals.

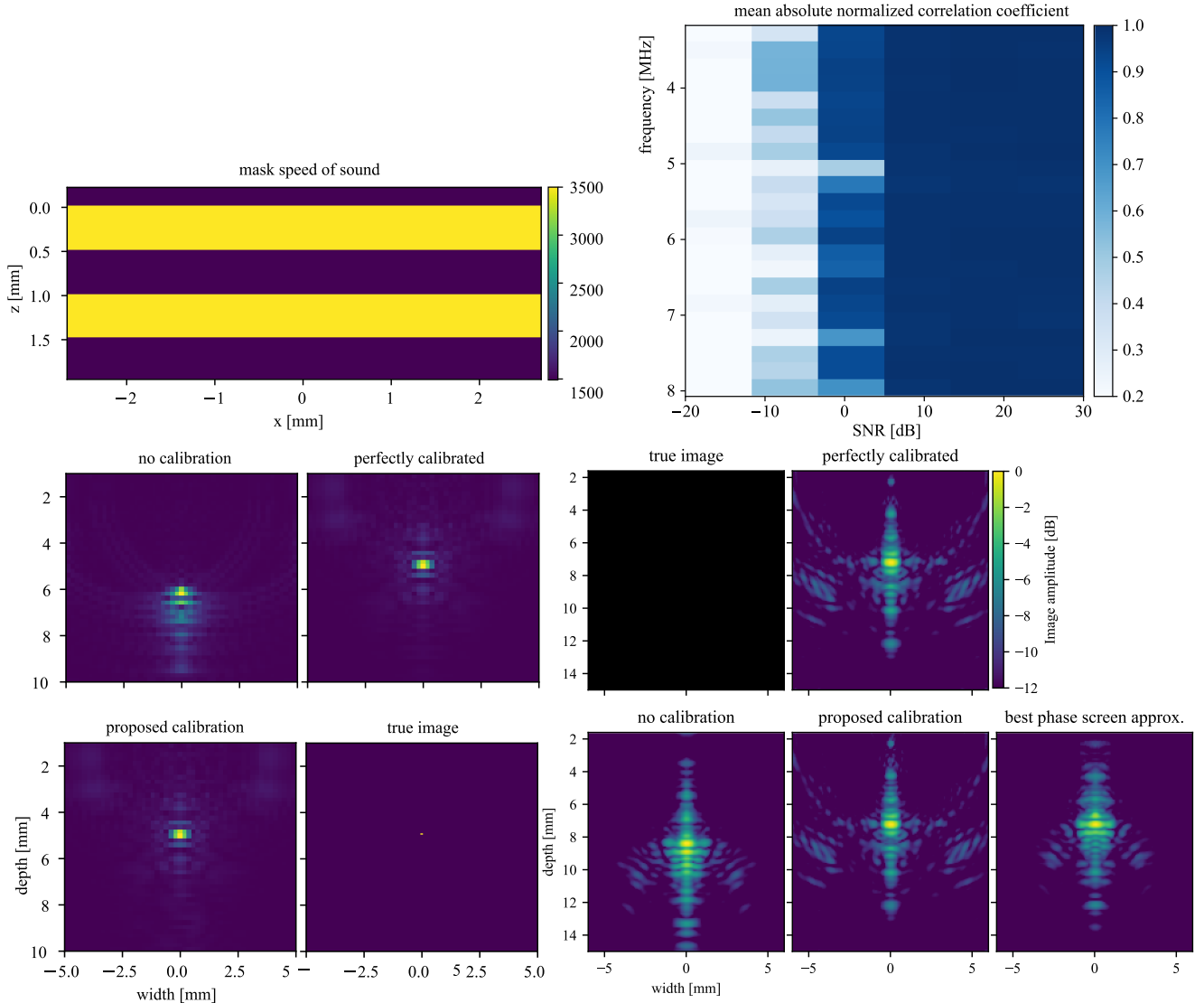


Fig. 3. Result for the simulations in Section IV-A. *Top left*: Laterally space-invariant aberration layer. *Top right*: Calibration performance per frequency and SNR. *Bottom left*: Imaging PSF performance comparison with/without calibration, for $P = 500$, and an SNR of 25 dB. *Bottom right*: Test image reconstruction performance comparison with/without calibration for reflectors at various inter-distances, for $P = 500$, and an SNR of 25 dB.

TABLE II
FWHM AND ISL FIGURES FOR ALL METHODS FOR THE RESULT OF IMAGING THROUGH A SKULL-LIKE LAYER IN SECTION IV-A

Method	FWHM width	FWHM depth	ISL width	ISL depth	Position error
No calibration	0.35 mm	0.28 mm	-10.7	-9.0 dB	0.28 mm
Best phase screen	0.48 mm	0.30 mm	-10.5 dB	-12.4 dB	0.03 mm
Proposed	0.38 mm	0.15 mm	-12.6 dB	-13.2 dB	0.03 mm
Perfect calibration	0.38 mm	0.15 mm	-12.7 dB	-13.7 dB	0.05 mm

We calibrate jointly for equidistantly spaced temporal frequencies between 1 and 7.5 MHz using $L = 36$. We form \mathbf{G} as mentioned before, using the 355 largest 2D FFT coefficients (out of 400). Using $P = 10,000$ and an SNR of 30 dB, we obtain an absolute normalized correlation coefficient of 0.94 after 30,000 iterations. When using an SNR of 20 dB, we obtain an absolute normalized correlation coefficient of 0.91. The used parameterization is very general and could possibly be used for a variety of aberration layers. Tighter parameterizations could

lead to improved estimation results and are a topic of future research, although it is encouraging that the proposed method is able to correctly calibrate using many unknowns.

In the bottom left of Fig. 4 we show the corresponding cost function. Interestingly, although we used the modified cost function of Section III-B, both the alternative cost function and the original covariance matching error decrease in a similar manner, suggesting that cost function (17) is a good surrogate for the original covariance matching cost function (9). Moreover, we

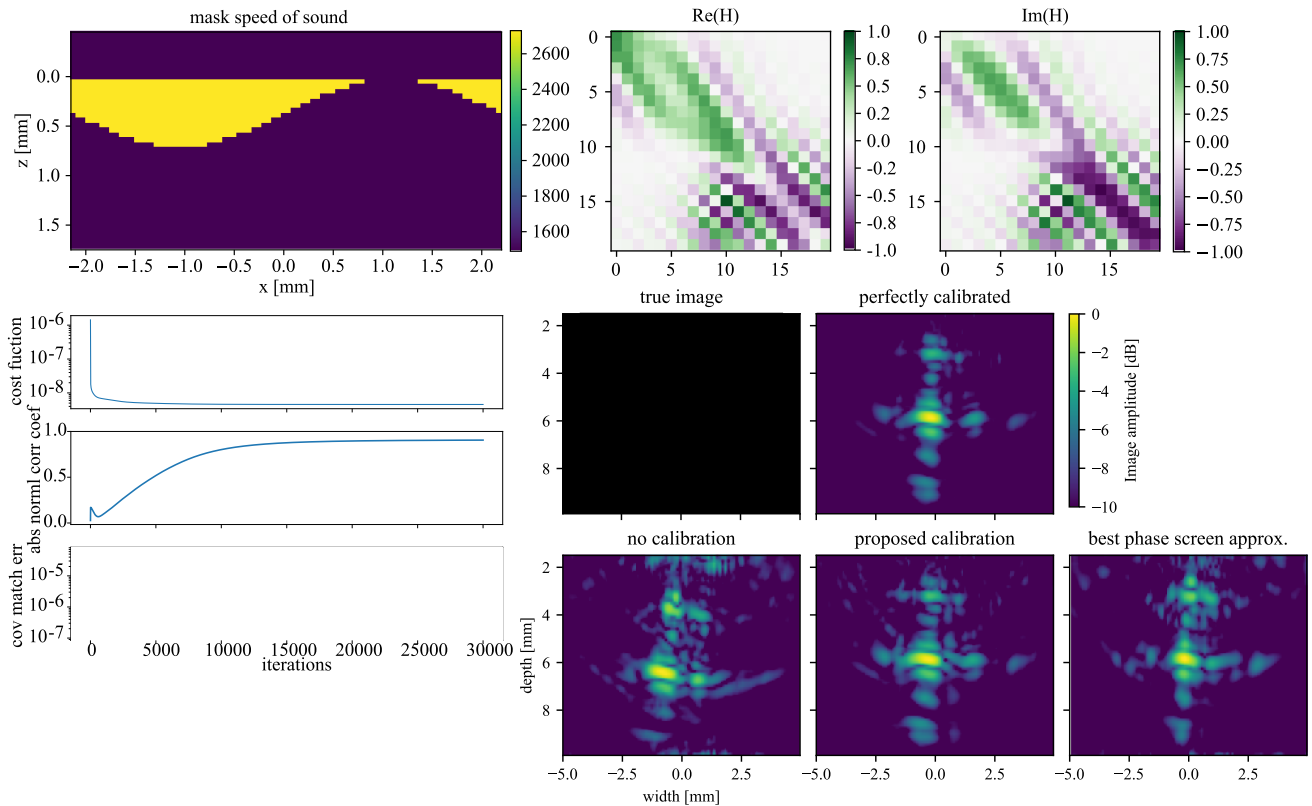


Fig. 4. *Top left*: Simulated smooth aberration layer used for the simulations in Section IV-B. *Top right*: \mathbf{H} -matrix for the aberration layer, for a frequency of 4.4 MHz. *Bottom left*: Various cost functions per iteration for the results of Section IV-B: proposed cost function (17), absolute normalized correlation coefficient (23), covariance matching error (9). *Bottom right*: Test image reconstruction performance comparison with/without calibration for reflectors at various inter-distances, for $P = 500$, and an SNR of 25 dB.

observe that even when the optimized cost function is decreasing slowly (after approx. 3,000 iterations), the actual error between the true solution and the estimated transfer function matrix is still increasing (here expressed using the normalized correlation coefficient). We hypothesize this is due to the fact that $\mathbf{A}\mathbf{A}^T$ has eigenvalues that are relatively low w.r.t. to the highest eigenvalues. Consequently, a small change in \mathbf{R}_y can result in a large change in \mathbf{H} . The calibration took approximately 20 minutes (using our non-optimized code), which we expect to strongly improve by better \mathbf{H} -matrix parameterization (as discussed in the Discussion section), and by optimizing code efficiency.

In the bottom right of Fig. 4, we show an image reconstruction for a phantom generated using k-wave. All image reconstructions are obtained using the LSQR algorithm with limited iterations. A number of observations can be made. Firstly, none of the reconstructions are of particularly high quality. This can be attributed to the fact that the \mathbf{H}_i matrices for this aberration layer have relatively *low singular values*. That is, one is not able to reconstruct (using e.g. a least-squares approach) the ‘input’ field to the layer from the ‘output’ field, since \mathbf{H} is low rank or ill-conditioned. Thus, given the general model equation $\mathbf{y} = \mathbf{H}\mathbf{A}\mathbf{x}$, information about \mathbf{x} is lost due to the conditioning and/or rank of \mathbf{H} . Secondly, we observe that knowledge of \mathbf{H} strongly improves the image reconstruction, making it apparent that a cross-like image is present. Finally, both the image result using the estimated \mathbf{H} , as well as imaging

using a phase screen approximation, come close to the perfect calibration case, with slightly worse background artefacts due to sidelobes etc. Note that sidelobes and other imaging artifacts do not have the same constitution as in imaging scenarios without an aberrating layer, where traditional imaging techniques, and their corresponding analysis, are applicable. In Table III, we see that calibration using a full \mathbf{H} matrix mostly improves positioning errors.

C. Joint Multi-Frequency Calibration: Skull Layer

As a more clinically relevant example, we calibrate a skull-like layer, consisting of a porous bone layer sandwiched between two homogeneous bone layers (Fig. 5). We exploit the fact that each \mathbf{H}_i matrix can be well represented using the 2D FFT basis for \mathbf{G} , omitting the highest frequencies, so that we have to estimate 342 variables (out of 400) per frequency. The random images are located in a ROI of 15 mm in width and 20 mm in depth. The true and virtual arrays contain $M = 20$ point-like sensors, spaced at 350 μm intervals.

We calibrate for $L = 46$ equidistantly spaced temporal frequencies between 1.8 and 6.0 MHz. Using $P = 10,000$ and an SNR of 30 dB, we obtain an estimate of \mathbf{H} with an absolute normalized correlation coefficient of 0.75. Fig. 6 shows the various cost functions for 30,000 iterations. The calibration took approximately 100 minutes (using our non-optimized code).

TABLE III
FWHM AND ISL FIGURES FOR ALL METHODS FOR THE RESULT OF IMAGING THROUGH A SKULL-LIKE LAYER IN SECTION IV-B

Method	FWHM width	FWHM depth	ISL width	ISL depth	Position error
No calibration	0.22 mm	0.27 mm	-8.9 dB	-16.9 dB	0.20 mm
Best phase screen	0.23 mm	0.60 mm	-10.4 dB	-17.6 dB	0.25 mm
Proposed	0.23 mm	0.33 mm	-13.5 dB	-14.7 dB	0.0 mm
Perfect calibration	0.23 mm	0.33 mm	-14.1 dB	-15.4 dB	0.0 mm

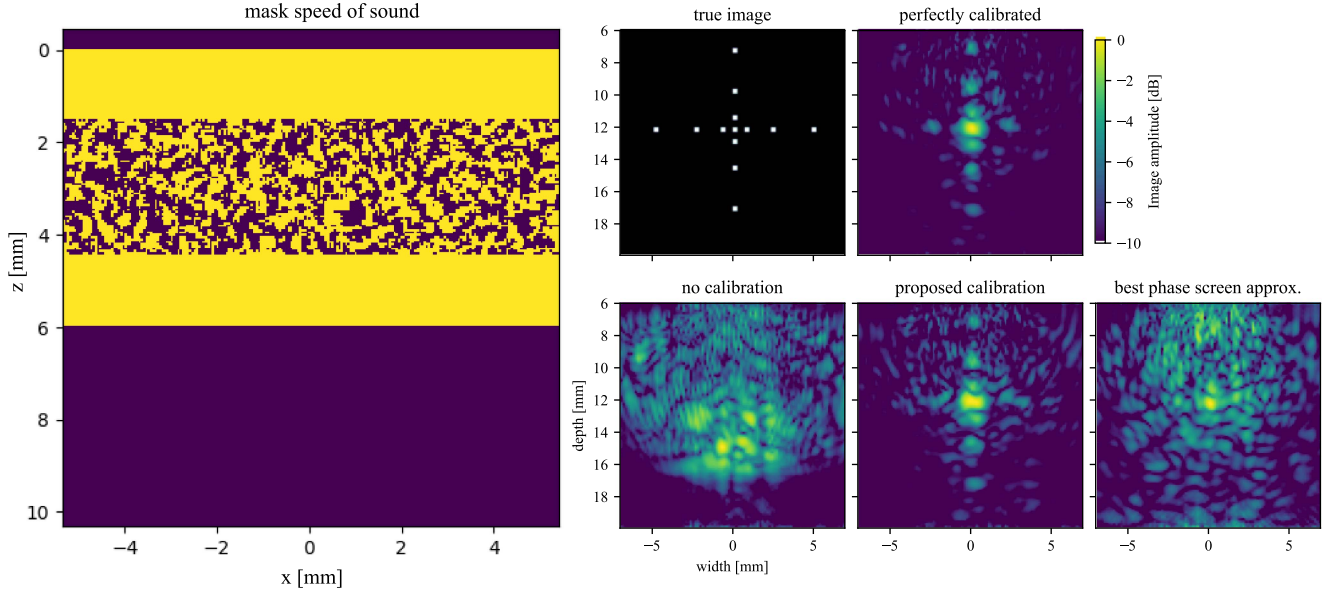


Fig. 5. *Left*: Results for the simulations in Section IV-C. Simulated skull-like aberration layer in k-wave, used for multi-frequency calibration experiments. *Right*: Image reconstruction example of simulated k-wave data of a set of point sources with varying inter-distances.

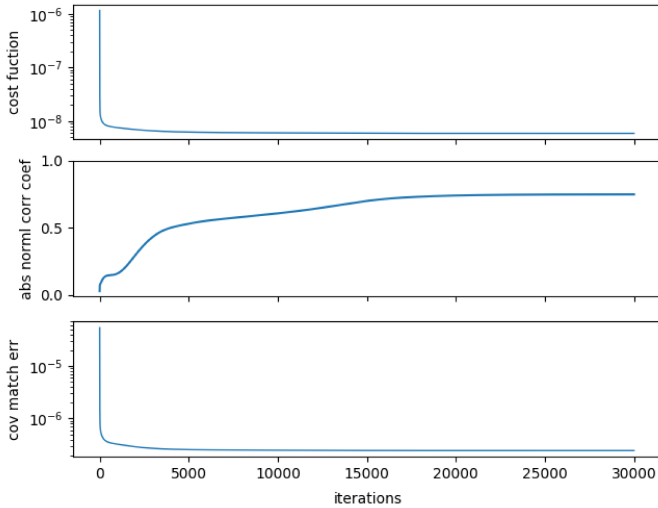


Fig. 6. Various cost functions per iteration for the results of Section IV-C: proposed cost function (17), absolute normalized correlation coefficient (23), covariance matching error (9).

In Fig. 5 we show the image reconstruction from RF data obtained from a *k-Wave* simulation. Image reconstructions are obtained using the LSQR algorithm to solve (2) with limited iterations. In this case, an enormous reconstruction improvement is obtained, as compared to imaging without any model

calibration. Although the phase screen approximation is able to create a strong peak at the correct position of the center reflector, most resolution is lost. This makes sense as waves propagate and scatter inside the aberrating layer, for which local delays would form a bad approximation. Seemingly, the phase screen is only able to correct for the bulk delay of the phase screen, but cannot account for other aberrating effects in the skull layer. As in the previous subsections, we tried to find the best image reconstruction for the calibration-less case by simply tweaking the speed of sound, but were not able to find a satisfactory result. The one in Fig. 5 was the best one obtained using a lower speed of sound than the background medium. Table IV shows the quantized PSF results. Since without calibration no sensible image can be obtained, the FWHM and ISL levels were not computed.

D. Blind Calibration From Blood Flow in a Vasculature Network

Next, we consider a more realistic scenario where images are not purely random. Instead, a vasculature tree with flowing particles is used, to emulate a more realistic and practical imaging scenario. Additionally, we will use true contrast reflections by simulating contrast-based reflections in *k-wave* instead of making each reflector a source as in the earlier results. This scenario is thus more realistic, and we can expect more model errors since the Born approximation may not strictly hold.

TABLE IV
FWHM AND ISL FIGURES FOR ALL METHODS FOR THE RESULT OF IMAGING THROUGH A SKULL-LIKE LAYER IN SECTION IV-C

Method	FWHM width	FWHM depth	ISL width	ISL depth	Position error
No calibration	n/a	n/a	n/a	n/a	n/a
Best phase screen	3,84 mm	0,92 mm	-6,4 dB	-5,9 dB	0,0 mm
Proposed	3,42 mm	0,93 mm	-7,6 dB	-6,2 dB	0,0 mm
Perfect calibration	3,52 mm	0,93 mm	-8,0 dB	-6,2 dB	0,0 mm

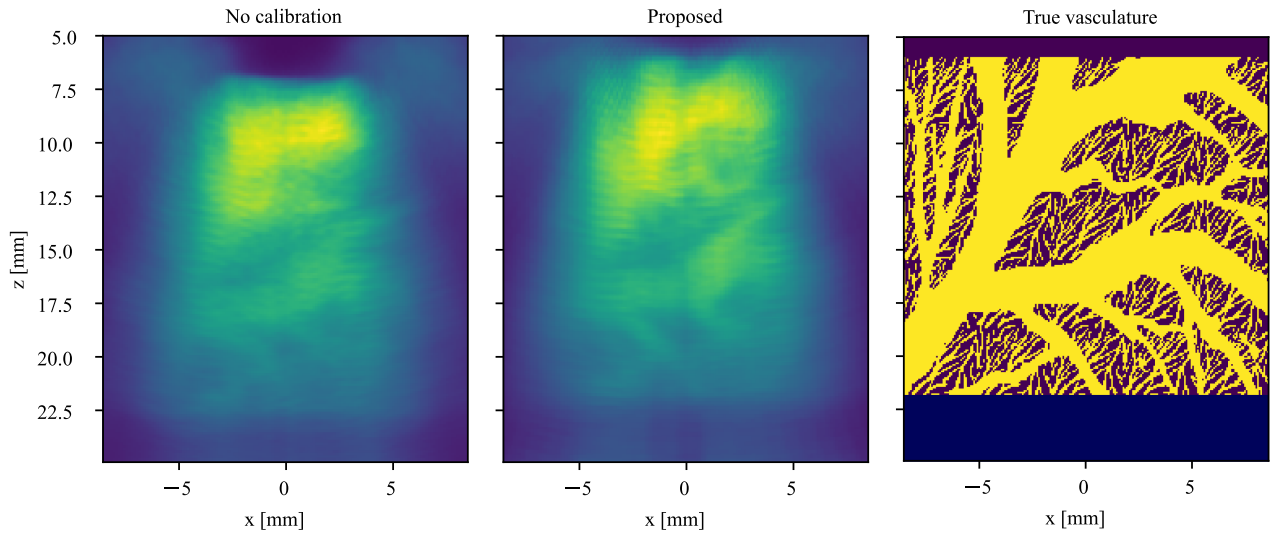


Fig. 7. Power Doppler images for $P = 1,000$ frames, for the results of Section IV-D. Each frame is simulated in k -Wave, and calibration is done using flow data from the same vasculature as visualized in the rightmost panel. The power Doppler images are normalized to have maximum of 1, and the images have been scaled equally in depth to be able to show image details at larger depths.

We use random i.i.d. normally distributed scatterers, where the mean speed of sound and density is that of the background medium (water). This field of random reflectors is shifted towards the array with a fixed distance every measurement ($500 \mu\text{m}$ steps). The random reflectors only occur inside the vasculature tree, and are set to the background medium properties when outside the vasculature. When computing the measurement covariance matrix $\hat{\mathbf{C}}_y$ from such a set of images, we found it to be nearly identical to the measurement covariance matrix from completely random images (where $\mathbf{C}_x = \mathbf{I}$), and hence we will assume \mathbf{I} as the a priori known image covariance matrix. *Note that we simulate each of the $P = 1,000$ frames in k -Wave, instead of using \mathbf{H} to generate many measurements.*

As an aberration layer, we use the same triple layer and Toeplitz basis as used in Section IV-A. The vasculature is positioned in a ROI from 6 to 22 mm depth, and 20 mm in width. The measurement array consists of $M = 30$ sensors, each one $250 \mu\text{m}$ wide, with $50 \mu\text{m}$ between them. The virtual array consists of $M = 30$ point-like sensors, at $300 \mu\text{m}$ intervals. We calibrate for equidistantly spaced temporal frequencies between 2.6 and 5.6 MHz using $L = 96$. We use only $P = 1,000$ measurements, and an SNR of 30 dB. Since our proposed algorithm relies on the forward field being known (see the footnote in Section II), we assume a point-like source transmits a spherical wave into the imaging medium without being aberrated.

In order to remove stationary signals, such as the transmit wave that propagates directly towards the sensor array without

scattering of pixels, we form a matrix of all the measured frames. Similar to SVD-based power Doppler [39], the 10 highest singular values are then set to zero. After this procedure, the filtered measurements $\mathbf{y}^{(p)}$ can be retrieved from the SVD-filtered matrix.

To evaluate the performance, we consider the estimated \mathbf{H} matrix to beamform each frame used in the calibration based on the LSQR algorithm with a large damping coefficient as regularization. We then apply a high-pass filter per pixel across all frames, followed by summing the absolute value of all frames to obtain a single power Doppler image. The result is shown in Fig. 7. Since the reconstructed images were suffering from decreasing image intensity for increasing depth and increasing lateral distance from the center (as was the case for the preceding image reconstructions in this section), the compounded power Doppler image has been scaled in those directions for visualization purposes. The power Doppler image using the \mathbf{H} matrix as estimated by the proposed algorithm shows a better resolved image compared to the case of no calibration. The improvement seems to be especially good in the ROI closer to the array, which can be explained due to the fact that columns of \mathbf{A} are more uncorrelated in that region, and thus more resolution is lost there if \mathbf{H} is unknown. As a simple calibration procedure, the speed of sound was somewhat tweaked to compensate for the delays of the aberration layer, but we found that using a speed of sound close to the true speed of sound gives the best result when not using our proposed calibration method. Hence, we expect

that the multiple reflections of the sub-layers cause the loss in (axial) resolution. Using our calibration algorithm, however, these are taken into account, leading to the improved image in Fig. 7. Importantly, this result suggests that power Doppler imaging (and thus, fUS imaging) through a layer is possible using the vasculature blood flow itself as random images. This is an especially encouraging result for fUS imaging of the brain from cerebral blood flow. The calibration took approximately 24 hours (using our non-optimized code).

V. CONCLUSION & DISCUSSION

We proposed a method for estimating the transfer function of an aberrating layer in ultrasound imaging scenarios by formulating a measurement model in the covariance domain. The ultrasound images do not need to be known, and instead only the second-order statistics of subsequent images are required, a condition that is considerably less demanding. The proposed measurement model explicitly exhibits the aberration transfer function, by observing that propagation through an aberrating layer is still a linear propagation process. By collecting a large number of measurements, the covariance matrix of the raw RF measurements can be estimated, from which the aberration transfer function is estimated. The differences between measurements is caused by the movement of e.g. blood flow or injected micro-bubbles. If enough time is taken between each frame/measurement, the positions of blood particles or micro-bubbles can be modelled as a random process, with no correlation between measurement frames and pixels. We can thus analyse data in the covariance domain where the ultrasound contrast image variable is replaced by the contrast image covariance matrix.

The approach results in a non-convex minimization problem with orthogonality constraints, for which we proposed algorithms to find an estimated transfer function. The proposed method is especially applicable for medical imaging scenarios with sufficient blood flow across the image, such as the human brain. It is much less invasive since only a small transmitter would need to be inserted, whereas the much larger sensor array can be positioned non-invasively outside the ROI and the aberration layer. Making the method suitable for pulse-echo imaging by transmitting and receiving using the same sensor array seems challenging as one of the main challenges observed in this article was the fact that the transfer function variables (in the case of non-aberrated transmit sensors) are quartic w.r.t. the observed RF data, and pulse-echo imaging would make it bi-quartic (octic), adding another layer of mathematical complexity to the problem.

A second critique on our method is that it does not model non-linear processes such as absorption, which may be more significant in certain aberrating layers than in soft tissue. Consequently, since our model is a purely linear one, it may not be able to capture these non-linear propagation effects sufficiently, and instead tries to capture it in our linear model, leading to incorrect solutions.

Although prior knowledge of the image covariance matrix instead of the actual images themselves relaxes the calibration

problem considerably, there can still be a model mismatch if there are relatively large areas in the ROI with no varying pixels (e.g. due to a lack of blood flow), or pixels with a relatively different variance compared to other pixels. In that case, using the identity matrix as the image covariance matrix might lead to faulty transfer function estimates. If an acceptable initial image can be obtained pre-calibration, or using an initial estimate of \mathbf{H} , such empty regions can be identified, and omitted from \mathbf{A} to obtain a more accurate model for \mathbf{C}_y .

We tested our proposed technique on a range of increasingly challenging calibration scenarios. We used *k-Wave* simulations to generate raw RF measurements to test our algorithm for a varying range of aberration layers. In each case, our technique was able to find a sufficiently good estimate of the transfer function of the aberration layer. Consequently, for each tested layer, image reconstructions were much better using our calibrated model than an uncalibrated model, even enabling imaging through a simulated skull layer.

There are currently no proofs or conditions for the existence of a unique solution to the mathematical calibration problem as formulated in this work. Consequently, we had to find whether this is the case for each scenario empirically. Furthermore, a parameterization of the transfer function matrix is required, which is dependent on the aberration layer characteristics. Nevertheless, we were able to successfully calibrate various layers using general FFT-based parameterizations, for various layers, including a skull layer with porous bone causing strong wave distortions. Due to practical considerations, some of the parameterizations were obtained by selecting the largest Fourier coefficients of \mathbf{H} beforehand, taken across all temporal frequencies (Section IV). Although this may seem like an inappropriate use of prior knowledge, we would argue that such a Fourier parameterization is sufficiently general (in all cases this method de facto only resulted in removal of the highest frequency components) and layer-agnostic (Fourier parameterizations are not expected to be the best parameterization for any specific aberration layer). Hence, we anticipate that the proposed method will yield even better results when using more intelligent layer- and frequency-specific parameterizations. For tightly parameterized transfer functions like those for laterally invariant layers, there seems to be a unique solution when calibrating for each single frequency independently, which our algorithm is able to find. For more challenging layers, a joint multi-frequency approach is required.

An important topic of future research constitutes the parameterization of \mathbf{H} . We expect that more efficient representations will make it easier to find a good solution using lower values of L , thereby decreasing the computational demands, and resulting in faster and better calibration results. We envision layer-specific parameterizations can be acquired by measuring the \mathbf{H} -matrices of e.g. many skull layers by varying wall thickness and spongy bone patterns. One could then use dimension reduction techniques such as PCA to represent these matrices using substantially fewer parameters. Even better representations could be obtained by taking into account the relation between \mathbf{H}_l -matrices across frequencies, and make each \mathbf{H}_l frequency-specific. An alternative approach would start with calibrating for the lower frequencies, since the corresponding \mathbf{H}_l -matrices are typically

easily represented using few coefficients, followed by estimation of the higher frequency matrices, taking into account the relation or correlation between the lower and higher frequency matrices.

We finally tested our techniques on a simulated blood flow phantom where the displacement of blood acts as random images. Assuming all images are uncorrelated, we were able to estimate the layer transfer functions, and we subsequently successfully computed a Doppler image of the same measurement data used for calibration. This is an encouraging result for reduced invasive functional ultrasound and Doppler imaging through the human skull, where the blood flow itself would act as random contrast images.

REFERENCES

- [1] J. G. Lynn, R. L. Zwemer, A. J. Chick, and A. E. Miller, "A new method for the generation and use of focused ultrasound in experimental biology," *J. Gen. Physiol.*, vol. 26, no. 2, pp. 179–193, 1942.
- [2] L. M. Hinkelman, T. D. Mast, L. A. Metlay, and R. C. Waag, "The effect of abdominal wall morphology on ultrasonic pulse distortion. Part I. Measurements," *J. Acoust. Soc. Amer.*, vol. 104, no. 6, pp. 3635–3649, 1998.
- [3] J. T. Fokkema and P. M. v.d. Berg, *Seismic Applications of Acoustic Reciprocity*. Amsterdam, The Netherlands: Elsevier, 1993.
- [4] M. Verweij, B. Treeby, K. V. Dongen, and L. Demi, "Simulation of ultrasound fields," in *Comprehensive Biomedical Physics*, Amsterdam, The Netherlands: Elsevier, 2014, pp. 465–499.
- [5] J. Bercoff et al., "Ultrafast compound Doppler imaging: Providing full blood flow characterization," *IEEE Trans. Ultrason., Ferroelectr. Freq. Control*, vol. 58, no. 1, pp. 134–147, Jan. 2011.
- [6] E. Macé, G. Montaldo, I. Cohen, M. Baulac, M. Fink, and M. Tanter, "Functional ultrasound imaging of the brain," *Nature Methods*, vol. 8, no. 8, pp. 662–664, 2011.
- [7] R. Rau et al., "3D functional ultrasound imaging of pigeons," *Neuroimage*, vol. 183, pp. 469–477, 2018.
- [8] B.-F. Osmanski, S. Pezet, A. Ricobaraza, Z. Lenkei, and M. Tanter, "Functional ultrasound imaging of intrinsic connectivity in the living rat brain with high spatiotemporal resolution," *Nature Commun.*, vol. 5, no. 1, pp. 1–14, 2014.
- [9] S. Soloukey et al., "Functional ultrasound (FUS) during awake brain surgery: The clinical potential of intra-operative functional and vascular brain mapping," *Front. Neurosci.*, vol. 13, 2020, Art. no. 1384.
- [10] M. Hirama and T. Sato, "Imaging through an inhomogeneous layer by least-mean-square error fitting," *J. Acoust. Soc. Amer.*, vol. 75, no. 4, pp. 1142–1147, 1984.
- [11] M. O'donnell and S. Flax, "Phase-aberration correction using signals from point reflectors and diffuse scatterers: Measurements," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 35, no. 6, pp. 768–774, Nov. 1988.
- [12] L. Nock, G. E. Trahey, and S. W. Smith, "Phase aberration correction in medical ultrasound using speckle brightness as a quality factor," *J. Acoust. Soc. Amer.*, vol. 85, no. 5, pp. 1819–1833, 1989.
- [13] D. Rachlin, "Direct estimation of aberrating delays in pulse-echo imaging systems," *J. Acoust. Soc. Amer.*, vol. 88, no. 1, pp. 191–198, 1990.
- [14] M. Karaman, A. Atalar, H. Koymen, and M. O'Donnell, "Experimental analysis of a computationally efficient phase aberration correction technique," in *Proc. IEEE Ultrasonics Symp. Proc.*, 1992, pp. 619–622.
- [15] M. Karaman, A. Atalar, H. Koymen, and M. O'Donnell, "A phase aberration correction method for ultrasound imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 40, no. 4, pp. 275–282, Jul. 1993.
- [16] D.-L. Liu and R. C. Waag, "Correction of ultrasonic waveform distortion using backpropagation and a reference waveform method for time-shift compensation," *J. Acoust. Soc. Amer.*, vol. 96, no. 2, pp. 649–660, 1994.
- [17] S.-E. Måsøy, T. F. Johansen, and B. Angelsen, "Correction of ultrasonic wave aberration with a time delay and amplitude filter," *J. Acoust. Soc. Amer.*, vol. 113, no. 4, pp. 2009–2020, 2003.
- [18] M. Fink et al., "Time-reversed acoustics," *Rep. Prog. Phys.*, vol. 63, no. 12, 2000, Art. no. 1933.
- [19] J.-F. Aubry, M. Tanter, J. Gerber, J.-L. Thomas, and M. Fink, "Optimal focusing by spatio-temporal inverse filter. II. Experiments. Application to focusing through absorbing and reverberating media," *J. Acoust. Soc. Amer.*, vol. 110, no. 1, pp. 48–58, 2001.
- [20] C. Demeñé et al., "Transcranial ultrafast ultrasound localization microscopy of brain vasculature in patients," *Nature Biomed. Eng.*, vol. 5, no. 3, pp. 219–228, 2021.
- [21] F. Vignon, J. Aubry, M. Tanter, A. Margoum, and M. Fink, "Adaptive focusing for transcranial ultrasound imaging using dual arrays," *J. Acoust. Soc. Amer.*, vol. 120, no. 5, pp. 2737–2745, 2006.
- [22] L. Guasch, O. C. Agudo, M.-X. Tang, P. Nachev, and M. Warner, "Full-waveform inversion imaging of the human brain," *NPJ Digit. Med.*, vol. 3, no. 1, pp. 1–12, 2020.
- [23] P. v. d. Meulen, M. Coutino, P. Kruizinga, J. G. Bosch, and G. Leus, "Blind calibration for arrays with an aberration layer in ultrasound imaging," in *Proc. IEEE 28th Eur. Signal Process. Conf.*, 2021, pp. 1269–1273.
- [24] R. Stoughton and S. Strait, "Source imaging with minimum mean-squared error," *J. Acoust. Soc. Amer.*, vol. 94, no. 2, pp. 827–834, 1993.
- [25] R. Lavarello, F. Kamalabadi, and W. D. O'Brien, "A regularized inverse approach to ultrasonic pulse-echo imaging," *IEEE Trans. Med. Imag.*, vol. 25, no. 6, pp. 712–722, Jun. 2006.
- [26] P. Kruizinga et al., "Compressive 3D ultrasound imaging using a single sensor," *Sci. Adv.*, vol. 3, no. 12, 2017, Art. no. e1701423.
- [27] J. Janjic et al., "Structured ultrasound microscopy," *Appl. Phys. Lett.*, vol. 112, no. 25, 2018, Art. no. 251901.
- [28] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digit. Signal Process.*, vol. 8, no. 3, pp. 185–210, 1998.
- [29] S.-E. Måsøy, B. Angelsen, and T. Varslot, "Estimation of ultrasound wave aberration with signals from random scatterers," *J. Acoust. Soc. Amer.*, vol. 115, no. 6, pp. 2998–3009, 2004.
- [30] T. Varslot, H. Krogstad, E. Mo, and B. A. Angelsen, "Eigenfunction analysis of stochastic backscatter for characterization of acoustic aberration in medical ultrasound imaging," *J. Acoust. Soc. Amer.*, vol. 115, no. 6, pp. 3068–3076, 2004.
- [31] A. J. Weiss and B. Friedlander, "Eigenstructure methods for direction finding with sensor gain and phase uncertainties," *Circuits, Syst. Signal Process.*, vol. 9, no. 3, pp. 271–300, 1990.
- [32] S. J. Wijnholds and A.-J. V.D. Veen, "Multisource self-calibration for sensor arrays," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3512–3522, Sep. 2009.
- [33] K. N. Ramamohan, S. P. Chepuri, D. F. Comesana, and G. Leus, "Blind calibration of sparse arrays for doa estimation with analog and one-bit measurements," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2019, pp. 4185–4189.
- [34] T. Viklands, "Algorithms for the weighted orthogonal procrustes problem and other least squares problems," Ph.D. dissertation, Department of Computing Science, Umea University, Umea, Sweden, 2006.
- [35] P. Stoica, H. He, and J. Li, "New algorithms for designing unimodular sequences with good correlation properties," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1415–1425, Apr. 2009.
- [36] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, "Manopt, a matlab toolbox for optimization on manifolds," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1455–1459, 2014.
- [37] B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *Proc. SPIE*, vol. 15, no. 2, 2010, Art. no. 021314.
- [38] C. C. Paige and M. A. Saunders, "LSQR: An algorithm for sparse linear equations and sparse least squares," *ACM Trans. Math. Softw.*, vol. 8, no. 1, pp. 43–71, 1982.
- [39] C. Demeñé et al., "Spatiotemporal clutter filtering of ultrafast ultrasound data highly increases Doppler and ultrasound sensitivity," *IEEE Trans. Med. Imag.*, vol. 34, no. 11, pp. 2271–2285, Nov. 2015.