

Benchmark problems for transcranial ultrasound simulation: Intercomparison of compressional wave models

Jean-Francois Aubry, Oscar Bates, Christian Boehm, et al.

Citation: *The Journal of the Acoustical Society of America* **152**, 1003 (2022); doi: 10.1121/10.0013426

View online: <https://doi.org/10.1121/10.0013426>

View Table of Contents: <https://asa.scitation.org/toc/jas/152/2>

Published by the *Acoustical Society of America*

ARTICLES YOU MAY BE INTERESTED IN

[Measurement and simulation of steered acoustic fields generated by a multielement array for therapeutic ultrasound](#)

JASA Express Letters **1**, 012001 (2021); <https://doi.org/10.1121/10.0003210>

[Modeling ultrasonic wave fields using a Quasi-Monte Carlo method: Wave transmission through complicated interfaces](#)

The Journal of the Acoustical Society of America **152**, 994 (2022); <https://doi.org/10.1121/10.0013411>

[Nonlinear ultrasound simulation in an axisymmetric coordinate system using a k-space pseudospectral method](#)

The Journal of the Acoustical Society of America **148**, 2288 (2020); <https://doi.org/10.1121/10.0002177>

[An arbitrary high-order discontinuous Galerkin method with local time-stepping for linear acoustic wave propagation](#)

The Journal of the Acoustical Society of America **149**, 569 (2021); <https://doi.org/10.1121/10.0003340>

[Modelling and measurement of laser-generated focused ultrasound: Can interventional transducers achieve therapeutic effects?](#)

The Journal of the Acoustical Society of America **149**, 2732 (2021); <https://doi.org/10.1121/10.0004302>

[The mechanical potential of ultrasound on nervous tissue](#)



The Journal of the Acoustical Society of America **149**, R11 (2021); <https://doi.org/10.1121/10.0005066>

JASA
THE JOURNAL OF THE
ACOUSTICAL SOCIETY OF AMERICA

Special Issue:
Additive Manufacturing and Acoustics

Read Now!

Benchmark problems for transcranial ultrasound simulation: Intercomparison of compressional wave models^{a)}

Jean-Francois Aubry,¹ Oscar Bates,² Christian Boehm,³ Kim Butts Pauly,⁴ Douglas Christensen,⁵ Carlos Cueto,² Pierre G lat,⁶ Lluıs Guasch,⁷ Jiri Jaros,⁸  Yun Jing,⁹ Rebecca Jones,¹⁰ Ningrui Li,¹¹ Patrick Marty,³ Hazael Montanaro,^{12,b)} Esra Neufeld,^{12,c)} Samuel Pichardo,¹³ Gianmarco Pinton,¹⁰ Aki Pulkkinen,¹⁴ Antonio Stanziola,¹⁵ Axel Thielscher,^{16,d)} Bradley Treeby,^{15,e)}  and Elwin van 't Wout¹⁷

¹Physics for Medicine Paris, National Institute of Health and Medical Research (INSERM) U1273, ESPCI Paris, Paris Sciences and Lettres University, French National Centre for Scientific Research (CNRS) UMR 8063, Paris, France

²Department of Bioengineering, Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom

³Institute of Geophysics, Swiss Federal Institute of Technology (ETH) Z rich, Sonneggstrasse 5, 8092 Z rich, Switzerland

⁴Department of Radiology, Stanford University, Stanford, California 94305, USA

⁵Department of Biomedical Engineering and Department of Electrical and Computer Engineering, University of Utah, Salt Lake City, Utah 84112, USA

⁶Department of Surgical Biotechnology, Division of Surgery and Interventional Science, University College London, London NW3 2PF, United Kingdom

⁷Earth Science and Engineering Department, Imperial College London, London, United Kingdom

⁸Centre of Excellence IT4Innovations, Faculty of Information Technology, Brno University of Technology, Bozetechova 2, Brno 612 00, Czech Republic

⁹Graduate Program in Acoustics, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

¹⁰Joint Department of Biomedical Engineering, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA and North Carolina State University, Raleigh, North Carolina 27695, USA

¹¹Department of Electrical Engineering, Stanford University, Stanford, California 94305, USA

¹²Foundation for Research on Information Technologies in Society (IT²IS), Zurich, Switzerland

¹³Radiology and Clinical Neurosciences Departments, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

¹⁴Department of Applied Physics, University of Eastern Finland, 70211 Kuopio, Finland

¹⁵Department of Medical Physics and Biomedical Engineering, University College London, Gower Street, London WC1E 6BT, United Kingdom

¹⁶Technical University of Denmark, Kongens Lyngby, Denmark

¹⁷Institute for Mathematical and Computational Engineering, School of Engineering and Faculty of Mathematics, Pontificia Universidad Cat lica de Chile, Santiago, Chile

ABSTRACT:

Computational models of acoustic wave propagation are frequently used in transcranial ultrasound therapy, for example, to calculate the intracranial pressure field or to calculate phase delays to correct for skull distortions. To allow intercomparison between the different modeling tools and techniques used by the community, an international working group was convened to formulate a set of numerical benchmarks. Here, these benchmarks are presented, along with intercomparison results. Nine different benchmarks of increasing geometric complexity are defined. These include a single-layer planar bone immersed in water, a multi-layer bone, and a whole skull. Two transducer configurations are considered (a focused bowl and a plane piston operating at 500 kHz), giving a total of 18 permutations of the benchmarks. Eleven different modeling tools are used to compute the benchmark results. The models span a wide range of numerical techniques, including the finite-difference time-domain method, angular spectrum method, pseudospectral method, boundary-element method, and spectral-element method. Good agreement is found between the models, particularly for the position, size, and magnitude of the acoustic focus within the skull. When comparing results for each model with every other model in a cross-

^{a)}Invited.

^{b)}Also at: Laboratory for Acoustics/Noise Control, Empa, Swiss Federal Laboratories for Materials Science and Technology, D bendorf, Switzerland.

^{c)}Also at: Department of Information Technology and Electrical Engineering, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland.

^{d)}Also at: Danish Research Center for Magnetic Resonance, Copenhagen University Hospital Hvidovre, Hvidovre, Denmark.

^{e)}Electronic mail: b.treeby@ucl.ac.uk

comparison, the median values for each benchmark for the difference in focal pressure and position are less than 10% and 1 mm, respectively. The benchmark definitions, model results, and intercomparison codes are freely available to facilitate further comparisons.

© 2022 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1121/10.0013426>

(Received 9 February 2022; revised 18 July 2022; accepted 22 July 2022; published online 16 August 2022)

[Editor: James F. Lynch]

Pages: 1004–1019

I. INTRODUCTION

Ultrasound is increasingly used for therapeutic applications in the brain, including for tissue ablation,¹ opening of the blood-brain barrier,² and modulation of brain activity.³ One challenge is the non-invasive delivery of ultrasound through the skull bone, which can significantly distort and attenuate the transmitted waves.⁴ To account for this, computer simulations are now frequently used to make predictions of the intracranial pressure field⁵ and to correct for phase aberrations due to the skull.⁶ This is particularly important for transcranial ultrasound stimulation (TUS), as the low ultrasound intensities make it highly challenging to measure the delivered energy *in vivo*, e.g., using magnetic resonance (MR)-guided thermometry.⁷

At a high level, there are four main steps in the setup of an acoustic model for transcranial ultrasound: (1) defining the medium parameters, including the skull and soft-tissue geometry and the acoustic properties (using a medical image, for example); (2) defining the transducer characteristics, including the geometry, driving parameters, and relative position; (3) defining the numerical parameters for the model, including the grid resolution and boundary conditions; and (4) processing and interpreting the simulated results. One challenge for the community is that there is a large variation in these steps in the published literature, and there is currently little consensus on the best approach or the uncertainties associated with numerical modeling more generally.

As part of the International Transcranial Ultrasonic Stimulation Safety and Standards (ITRUSST) consortium, a working group focused on simulation and planning was convened. The primary goal was to perform a modeling intercomparison to systematically evaluate the steps involved in transcranial simulation, with a view to establishing best practice. A number of researchers active in the development of tools for transcranial ultrasound simulation were invited to take part. The first phase, which is reported here, was a model intercomparison using a series of numerical benchmarks relevant to transcranial ultrasound where the medium parameters and transducer characteristics were well defined. The primary research question was *do different modeling techniques and computer codes give the same answer when the inputs to the model are well specified?* This was taken as the first step to ensure that any differences in more complicated scenarios (e.g., where the skull properties are mapped from a medical image or the transducer properties are mapped from a hydrophone measurement) could be evaluated as systematically and independently as possible.

The working group met regularly throughout 2021. The list of benchmarks (discussed in Sec. II) was iteratively refined, including the source definitions, the medium geometry, the material properties, and the output domain size. File submission formats, mechanisms for data sharing, and comparison metrics (along with codes to compute them)⁸ were also defined. Benchmark submissions were non-blinded with multiple resubmissions allowed. The goal was not a competition to establish which model was the “best” by some definition. Rather, it was to establish consensus on different approaches to transcranial ultrasound modeling and how to implement these correctly using a range of modeling tools available to the community. In this spirit, work-in-progress results and comparison metrics were discussed at regular intervals. These discussions, along with the sharing of code, approaches, and processing steps, etc., ultimately allowed the benchmarks to be computed with a wide set of simulation tools with excellent agreement (see Sec. IV).

The primary goal of this phase of the intercomparison exercise was to establish a series of benchmarks relevant to transcranial ultrasound, along with consensus on the correct numerical solutions for these benchmarks. Consequently, simulations were typically performed with very high sampling to maximize accuracy. Because of this emphasis and the different computational resources available to each group, a comparison of the computational performance of the individual models was considered out-of-scope from the outset. However, it is still important to note that some models used in the intercomparison, in particular, those based on the angular spectrum method, have an efficiency/accuracy trade-off inherent in their formulation (see, e.g., Ref. 9). This should be considered when interpreting the intercomparison metrics presented in Sec. IV and the supplementary information.¹⁰

The final output from the intercomparison exercise is a set of nine well-defined numerical benchmarks relevant to transcranial ultrasound (with a total of 18 permutations of these benchmarks), along with publicly available simulation results for these benchmarks computed using 11 different modeling codes.^{8,11}

II. BENCHMARKS

A. Overview

The benchmarks were defined considering typical TUS scenarios, although they are also relevant to other therapeutic applications of transcranial ultrasound. Simulations were single frequency (time-harmonic) and performed assuming

linear wave propagation (previous studies have shown that nonlinear effects are negligible for typical TUS parameters)¹² Only compressional waves were considered for this stage of the intercomparison. In some circumstances, elastic wave effects will also play a role, particularly if the ultrasound waves are not close to normal incidence with respect to the skull bone,^{13,14} but these effects were not investigated here. All simulations were conducted in three dimensions.

B. Transducer characteristics

Two transducer definitions were used (see Fig. 1). The first was a spherically curved transducer with a 64 mm radius of curvature and a 64 mm aperture diameter. This is representative of the single-element transducers frequently used for TUS.¹⁵ The second was a plane piston transducer with a diameter of 20 mm. Piston transducers are often used in multi-element arrays. While the typical diameter of an element in a multi-element array is smaller than 20 mm, this diameter was used to provide identifiable beam characteristics within the simulation domain. For some numerical techniques, piston transducers are easier to model, particularly when aligned with the computational grid, which avoids staircasing artifacts.¹⁶ Both transducers were driven at

500 kHz with a constant surface velocity of 0.04 m/s.¹⁷ Assuming an acoustic impedance of 1.5 megarayls, this is equivalent to modeling the sources as a distribution of free-field monopole radiators with a source pressure of 60 kPa.

C. Material properties

The material properties used for the benchmarks are given in Table I (all materials are modeled as acoustic media, i.e., fluids). These are intended to be representative (rather than definitive) values and were taken from the range presented in the literature.^{18–25} For the simulations including absorption, the loss is defined to be non-dispersive, i.e., either frequency independent or, for power law models, dependent on frequency squared.

D. Simulation outputs

The simulation results were stored as two variables named *p_amp* and *p_phase*. These represent the amplitude and phase of the complex pressure field at 500 kHz over the specified comparison domain. For time domain models, these parameters can be extracted precisely by setting the time step to an integer number of points per period (PPP), recording the steady-state pressure field for an integer

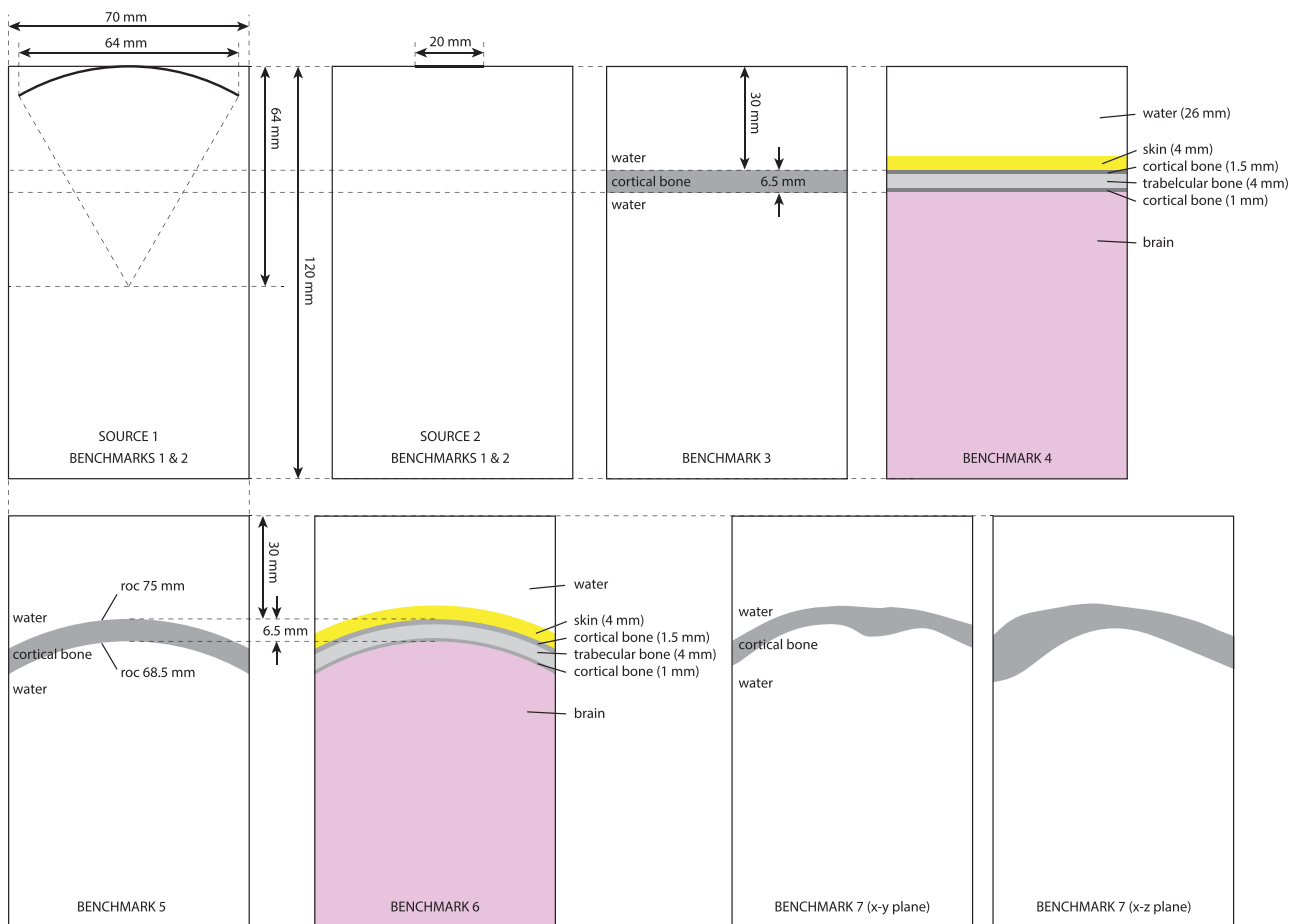


FIG. 1. (Color online) Transducer definitions and simulation layouts for benchmarks 1–7. Benchmarks 1–6 use a two-dimensional (2D) comparison domain of 120 mm (axial) by 70 mm (lateral) through the central *z* plane. Benchmark 7 uses a 3D comparison domain of 120 × 70 by 70 mm. The material properties used are given in Table I.

TABLE I. Compressional sound speed (c), mass density (ρ), and absorption coefficient (α) used in the benchmark simulations.

	c (m/s)	ρ (kg/m ³)	α (dB/cm at 500 kHz)
Water	1500	1000	0
Skin	1610	1090	0.2
Brain	1560	1040	0.3
Cortical bone	2800	1850	4
Trabecular bone	2300	1700	8

number of periods, and then extracting the amplitude and phase at the driving frequency using a Fourier transform. Note that the phase is optional and was not used for the comparisons presented in Sec. IV but was included for completeness. The results were saved either as MATLAB .mat files using the "-v7.3" flag where possible (this format can be easily opened as a HDF5 file outside MATLAB) or as HDF5 .h5 files with the variables saved as datasets in the root group.

Regardless of the sampling or mesh used for the simulations, the outputs stored in the comparison files were resampled onto a uniform Cartesian grid with 0.5 mm grid sampling. This corresponds to six points per wavelength (PPW) in water. For benchmarks 1–6, the comparison domain size was a 120 × 70 mm (axial × lateral) slice through the central z -coordinate, corresponding to a grid size of 241 × 141 grid points. For benchmark 7, the comparison domain was 120 × 70 × 70 mm (241 × 141 × 141 grid points). For benchmark 8, the comparison domain was 225 × 170 × 190 mm (451 × 341 × 381 grid points). For benchmark 9, the comparison domain was 212 × 224 × 184 mm (425 × 449 × 369 grid points).

For all benchmarks, the transducer was oriented such that the beam axis pointed in the x dimension, with the transducer positioned in the center of the y/z dimensions. Using 1-based indexing, the center of the source (rear of the bowl or center of the piston) relative to the output grid was positioned at [1, 71] for benchmarks 1–6, [1, 71, 71] for benchmark 7, [1, 171, 191] for benchmark 8, and [1, 225, 185] for benchmark 9. Note the comparisons for benchmarks 1–6 were made in two dimensions due to the axisymmetry of the geometry. All simulations were conducted in three dimensions.

TABLE II. Summary of benchmarks in phase 1 of the intercomparison. SC1 corresponds to the focused bowl transducer and SC2 to the plane piston transducer. Outputs are resampled to a regular Cartesian mesh with a grid spacing of 0.5 mm. Simulation layouts are shown in Figs. 1 and 2. gp = grid points.

Label	Description	Output grid size
PH1-BM1-SC1/2	Water (lossless)	120 × 70 mm (241 × 141 gp)
PH1-BM2-SC1/2	Water (artificial absorption of 1 dB/cm at 500 kHz)	120 × 70 mm (241 × 141 gp)
PH1-BM3-SC1/2	Flat, single-layer skull (cortical bone) in water	120 × 70 mm (241 × 141 gp)
PH1-BM4-SC1/2	Flat, skin, three-layered skull, and brain	120 × 70 mm (241 × 141 gp)
PH1-BM5-SC1/2	Curved, single-layer skull (cortical bone) in water	120 × 70 mm (241 × 141 gp)
PH1-BM6-SC1/2	Curved, skin, three-layered skull, and brain	120 × 70 mm (241 × 141 gp)
PH1-BM7-SC1/2	Truncated skull mesh in water, target in visual cortex	120 × 70 × 70 mm (241 × 141 × 141 gp)
PH1-BM8-SC1/2	Whole skull mesh, target in visual cortex	225 × 170 × 190 mm (451 × 341 × 381 gp)
PH1-BM9-SC1/2	Whole skull mesh, target in motor cortex	212 × 224 × 184 mm (425 × 449 × 369 gp)

E. Naming convention

The benchmarks were given unique identifiers in the following format: PH<NUM>-BM<NUM>-SC<NUM>. PH (phase) identifies the intercomparison phase (in this case 1). BM (benchmark) identifies the benchmark number within the phase. SC (source) identifies the source condition, where 1 is the bowl source and 2 is the plane piston source. A summary of the different benchmarks is given in Table II. File names for the intercomparison results follow the same convention with the model name appended (see Table IV): PH<NUM>-BM<NUM>-SC<NUM>_<MODELNAME>. The simulation outputs for each model for each benchmark are publicly available.¹¹

F. Benchmarks

A total of nine benchmarks relevant to transcranial ultrasound were devised. These are summarized in Table II. The benchmarks gradually increase in complexity, both adding additional tissue layers and increasing the geometric complexity of the skull. Benchmarks 1–7 are illustrated in Fig. 1, while benchmarks 8 and 9 are illustrated in Fig. 2.

Benchmark 1 considers the bowl and piston transducers in water (free-field) using the properties given in Table I. Benchmark 2 adds uniform artificial absorption of 1 dB/cm at 500 kHz. During the intercomparison exercise, these benchmarks served as a helpful reference to ensure the transducer properties, absorption units, and comparison domain were correctly specified. For these simulations, reference simulations were also computed using the fast near-field method as implemented in the FOCUS toolbox.^{26–28} Calculations using FOCUS were performed using 5000 integration points to give a high level of precision. Several models used the fields computed using FOCUS across a transverse y - z plane as the source definition (see Sec. III).

Benchmark 3 introduces a single flat 6.5 mm layer of cortical bone immersed in water, positioned 30 mm from the transducer as shown in Fig. 1. Benchmark 4 extends this to include a 4 mm skin layer, a three-layered skull (consisting of 1.5 mm cortical bone for the outer table, 4 mm trabecular bone, and 1 mm cortical bone for the inner table, giving the same overall skull thickness and position as benchmark 3)

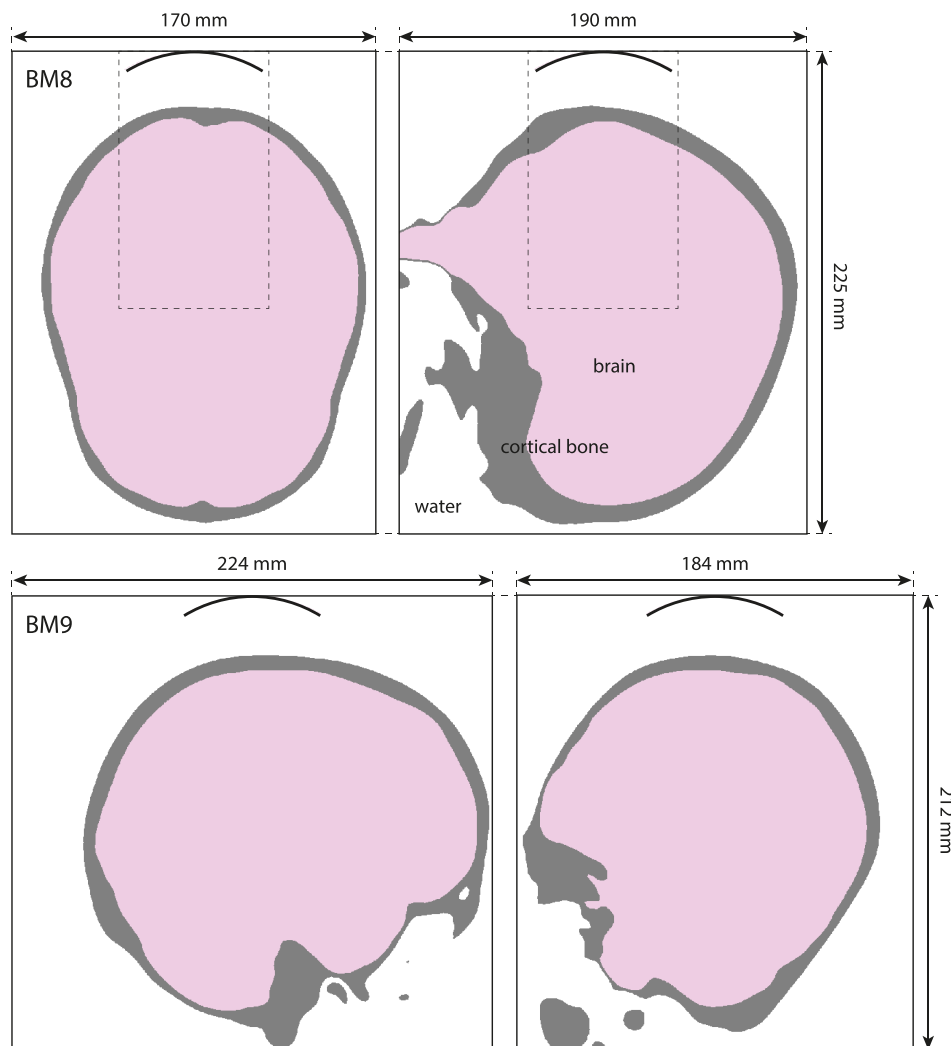


FIG. 2. (Color online) Simulation layouts for benchmarks 8 (top row) and 9 (bottom row) showing the central x - y and x - z slices. The position of the bowl transducer is shown for reference. Benchmark 7 (shown in Fig. 1) uses a subset of the skull mask and the same relative transducer position as benchmark 8, with a reduced comparison domain size as shown with the dashed line. The material properties used are given in Table I.

with water on the exterior and brain on the interior as shown in Fig. 1. The thickness values are based on average values for parietal bone²⁹ and scalp.³⁰

Benchmark 5 increases the geometric complexity of benchmark 3 by using a curved 6.5 mm layer of cortical bone immersed in water, with inner and outer radii of 68.5 and 75 mm, respectively. Note that the bone layer is spherically (not cylindrically) curved, meaning the curvature in the out-of-plane dimension is the same as that shown in Fig. 1. Benchmark 6 is a curved extension of benchmark 4, where the thickness values correspond to differences in the curvature radii.

Benchmarks 7–9 increase the geometric complexity further by using a homogeneous skull mesh generated from the MNI152 T1_1 mm magnetic resonance imaging template brain.^{31,32} The template image was run through an adapted version of SimNIBS headreco.³³ Additional smoothing of the tissue surfaces while simultaneously preventing intersections between neighboring surfaces was performed using SimNIBS functions. Benchmarks 7 and 8 use a transducer position targeted at the foveal representation of the primary visual cortex, while benchmark 9 uses a transducer position targeted at the hand area of the primary motor cortex.

The skull mesh was stored as two `.stl` files representing the inner and outer surfaces of the skull bone. Position transforms were stored as three-dimensional (3D) affine transformations that position the transducer relative to the coordinates in the `.stl` files. Grid-based discretizations containing a binary skull mask were also generated using the `iso2mesh` MATLAB toolbox.^{34,35} These were generated on a regular Cartesian mesh at a range of resolutions after applying the appropriate inverse position transforms (to move the skull mesh relative to the transducer) and were truncated to the appropriate comparison domain (see Sec. IID). The skull files and position transforms are available alongside the simulation results.¹¹

G. Intercomparison metrics

A number of metrics were chosen to compare the simulated fields. Mathematical definitions for some metrics are given in Table III. Metrics based on the entire field were taken from the exit plane of the source, excluding the first grid point in the x -direction for the piston transducer and the first 19 grid points in the x -direction for the bowl transducer. The relative L^2 and L^∞ errors provide a useful (and strict) measure of the overall differences between simulations.

TABLE III. Difference metrics used for the intercomparison. Here, p_1 and p_2 are the amplitude of the pressure field over the 2D or 3D comparison domains for the reference field and comparison field, respectively (these are assumed to be positive). Sums and maximum values are assumed to be over all values in the comparison domain starting from the exit plane of the transducer. Focal values are taken from inside the brain (or post-skull) region only. pos max is used to denote the position of the maximum value in the comparison domain.

Metric	Definition
Relative L^2	$\sqrt{\frac{\sum (p_1 - p_2)^2}{\sum p_1^2}}$
Relative L^∞	$\frac{\max p_1 - p_2 }{\max(p_1)}$
Focal (peak) pressure	$\frac{ \max(p_1) - \max(p_2) }{\max(p_1)}$
Focal position	$\ \text{pos max}(p_1) - \text{pos max}(p_2)\ _2$

However, for more complex geometries, these become dominated by differences in the rapidly varying near-field region between the source and the skull. For this reason, differences in the focal characteristics were also computed. This included the magnitude and position of the peak pressure within the brain and differences in the full-width at half-maximum (FWHM) and -6 dB focal volume. The FWHM values were taken in each Cartesian direction present in the comparison domain (i.e., in the x and y dimensions for benchmarks 1–6 and in the x , y , and z dimensions for benchmarks 7–9). For benchmarks 3–9 for the piston transducer, the acoustic field gradually decays within the brain; thus, there is no natural focus in the axial direction. In this case, the axial focal position and lateral profiles and FWHM values were taken at $x = 60$ mm (corresponding to a grid index of 121). For benchmarks 7–9, differences in the -6 dB focal volume were also computed. The focal volume was calculated by thresholding the pressure field inside the brain to 50% of the maximum value and then counting the voxels in the largest connected component. Code to compute the intercomparison metrics is available on GitHub.⁸

III. MODELS

A. Overview

A total of 11 modeling tools were used for the intercomparison, in addition to the free-field reference values calculated using FOCUS discussed in Sec. II F. These are summarized in Table IV. A short description of each model is given in Secs. III B–III L, with additional details given in the supplementary material.¹⁰

B. BABELVISCOFDTD

BABELVISCOFDTD solves the viscoelastic wave equation expressed in stress tensors and displacement vectors, where the bone material is modeled as a viscoelastic isotropic medium.²³ The term “Babel” refers to the multiple computing backends (CUDA, OpenCL, Apple Metal, and X86–64) that are supported for calculations. Nodes of stress and displacement are placed in a staggered-grid arrangement.³⁶

TABLE IV. Summary of models used to calculate the benchmark results. Additional details are given in the supplementary material (Ref. 10). Authors correspond to the authors of the current manuscript directly contributing to the intercomparison exercise, not necessarily the authors of the model.

Label	Authors	Domain	Method
BABELVISCOFDTD	S.P.	Time	FDTD ^a
FULLWAVE	R.J., G.P.	Time	FDTD
GMFDTD	A.P.	Time	FDTD
HAS	N.L., K.B.P.	Frequency	HAS ^b
JWAVE	A.S.	Frequency	Fourier spectral method with iterative solver
KWAVE	B.T., J.J.	Time	Pseudospectral time domain
MSOUND	Y.J.	Frequency	Modified angular spectrum
OPTIMUS	P.G., E.v.W.	Frequency	BEM ^c
SALVUS	P.M., C.B.	Time	Spectral-element
SIM4LIFE	H.M., E.N.	Time	FDTD
STRIDE	C.C., O.B., L.G.	Time	FDTD

^aFinite-difference time-domain (FDTD).

^bHybrid angular spectrum (HAS).

^cBoundary-element method (BEM).

Calculations are solved using a 4th-order in space and 2nd-order in time FDTD scheme in Cartesian coordinates.^{37,38} Stress tensors and displacement vectors are solved a half time step separated from each other. Attenuation losses are modeled using a quality factor for narrowband conditions.^{38,39} Liquid-bone interfaces and heterogeneity of tissue material are modeled using averaging operators.⁴⁰ Optional reduction of staircasing artifacts can be enabled using a superposition operator.⁴¹ A perfectly matched layer (PML) condition for viscoelastic propagation is used to absorb waves at the boundaries.⁴²

All benchmarks were computed using a resolution of 12 grid PPW. Sources were modeled as stress nodes using the same staircase-free formulation and dispersion correction as in the k-Wave model (see Sec. III G). The time PPP for benchmarks 1 and 2 was 25, and for benchmarks 3–9, it was 48. Benchmarks 1–7 used a total grid size of $305 \times 305 \times 521$ grid points, including the PML. For benchmarks 8 and 9, the grid size was, respectively, $785 \times 705 \times 941$ and $761 \times 921 \times 889$. Simulation outputs were resampled to the comparison grid using a spline interpolation of order 3.

C. FULLWAVE

FULLWAVE2 3D solves the wave equation with quadratic nonlinearity and multiple relaxations using a staggered-grid FDTD approach with fourth-order accuracy in time and variable accuracy in space.^{43,44} This model uses a staggered-grid Cartesian mesh with a convolutional PML at the boundaries, utilizing high-order adaptive stencils that minimize dispersion and dissipation errors. The source and output can take the shape of any arbitrary geometry that can be defined on a Cartesian grid, with sources modeled either as free-

field particle displacement, velocity, or a monopole pressure source.

For the benchmark comparison, the bowl and piston geometries were modeled as monopole pressure sources on a Cartesian grid, emitting a continuous sinusoidal wave. All benchmarks were computed with 12 PPW and 60 PPP, giving a Courant–Fredrichs–Lewy condition (CFL)⁴⁵ of 0.2. This created a simulation grid 2 times the size of the comparison grid. To account for this, the simulations were run with a spatial step size of 2 voxels in each direction, down-sampling the output grid to the comparison grid size. The output over one steady-state cycle was then scaled based on the CFL and driving signal to account for the use of additive sources.

D. GMFDTD

The GMFDTD model simulates acoustic wave propagation based on coupling of the second-order acoustic and viscoelastic wave equations using a combined grid method and FDTD method. The model operates using a regular Cartesian mesh. For fluid simulations (as described in this work), GMFDTD solves the acoustic wave equation using a FDTD approach with fourth-order spatial and second-order time stencils. First-order absorbing boundary conditions are used on exterior boundaries of the simulation domain. A finite thickness absorbing layer was placed on the exterior boundaries to further reduce acoustic reflections. A heterogeneous Neumann boundary condition is used to model the sound sources making the source-medium interface work as an acoustically hard reflector for incoming sound waves. For a more thorough description of the model, see Ref. 46.

All simulations were computed using 12 PPW and 75 PPP. Grid sizes for the simulations were $576 \times 376 \times 376$ grid points for benchmarks 1–7, $996 \times 776 \times 856$ for benchmark 8, and $944 \times 992 \times 832$ for benchmark 9. The grid sizes include a 48 grid point absorbing layer surrounding the domain, which had attenuation linearly increasing from zero to 50 Np/m, corresponding to about 94% amplitude attenuation for a normally incident reflected wave. Simulations were computed for 3900 time steps for benchmarks 1–7, 15 075 time steps for benchmark 8, and 15 075 time steps for benchmark 9. The simulations produced a complex valued steady-state pressure field, which was resampled to the comparison grid using spline interpolation before computing the pressure amplitude and the phase angle.

E. HAS

The HAS method is a generalization of the angular spectrum method, enabling propagation of pressure fields in heterogeneous media.^{47,48} An initial pressure distribution is first defined on a plane perpendicular to the direction of propagation. To produce the full 3D steady-state pressure field, pressures on subsequent planes are calculated in the spatial-frequency domain by solving the Helmholtz equation using the angular spectrum method. Errors due to local variations in attenuation and acoustic velocity are corrected for

using a spatial step between each spatial-frequency step. Reflected pressures are saved, backpropagated, and summed with the incident pressure field, and this process is repeated until convergence to produce the final steady-state pressure field.

Initial pressure fields were computed using the fast near-field method as implemented in the FOCUS toolbox (see Sec. II F). Benchmarks 1–6 were computed using a grid size of $1001 \times 1001 \times 1001$ with 6 PPW in the transverse directions and 24 PPW in the axial direction. Benchmarks 7 and 8 were computed using a grid size of $1401 \times 1401 \times 501$ with an isotropic resolution of 12 PPW. Benchmark 9 was computed using a grid size of $1201 \times 1201 \times 501$ with an isotropic resolution of 12 PPW. Calculated pressure fields were resampled to the comparison grid using bilateral interpolation.

F. JWAVE

JWAVE simulates the solution of time-harmonic wave propagation problems by solving the heterogeneous Helmholtz equation in the complex domain, using a regular Fourier spectral collocation method and linear iterative solvers such as restarted generalized minimal residual method (GMRES).^{49,50} Absorbing boundary conditions are enforced using a PML,⁵¹ while the definition of the sources is done by projecting them on the discrete collocation grid by approximately convolving them with the band limited interpolant.⁵² The source field is modeled as a mass source. JWAVE is a PYTHON software written using JaxDF,⁵³ which in turn is based on JAX.⁵⁴ The code is just-in-time compiled for the hardware at hand [e.g., graphical processing units (GPUs) or tensor processing units (TPUs)] and allows for automatic differentiation to be applied with respect to any continuous parameter.

Benchmarks 1 and 2 were computed using 6 PPW, while benchmarks 3–7 were computed using 12 PPW. The PML size was fixed to 30 voxels. To reduce the computation time of the FFTs, the domain dimensions were padded to the nearest integers with prime factors smaller than 7. When required, the results were resampled to the intercomparison grid using Fourier interpolation. Benchmarks 8 and 9 were too large for the available computational resources, so results for these benchmarks were not computed.

G. KWAVE

KWAVE solves three coupled equations equivalent to a generalized Westervelt equation, where spatial gradients are calculated using a Fourier collocation spectral method, and time integration is performed using a dispersion-corrected finite-difference scheme.^{55,56} Calculations are performed on a regular Cartesian mesh with a space and time staggered grid. A split-field PML is used to absorb the waves at the domain boundaries. Sources are modeled as free-field monopoles (injection of mass) using a stair-case-free formulation to represent the bowl and piston geometries⁵² and a dispersion-corrected time-stepping scheme.⁵⁷

Benchmarks 1–6 were computed using the axisymmetric version of k-Wave to provide a high-resolution reference simulation.⁵⁸ Benchmarks 1, 2, 3, and 5 used 60 PPW and 2400 PPP, while benchmarks 4 and 6 used 60 PPW and 6000 PPP. In both cases, the total grid size was 2700×864 grid points, including the PML, and the simulation time was 120 μ s, giving 144 000 and 360 000 time steps, respectively. Benchmarks 7–9 were computed using the 3D version of k-Wave optimized for high performance computing clusters.⁵⁹ Benchmark 7 used 30 PPW and 1200 PPP, with a grid size of $1296 \times 768 \times 768$ grid points and 72 000 time steps (120 μ s simulation time). Benchmarks 8 and 9 used 18 PPW and 360 PPP with 72 000 time steps (400 μ s simulation time). The grid sizes were $1458 \times 1080 \times 1200$ and $1350 \times 1440 \times 1152$, respectively. The simulation times were sufficient to reach steady state and were chosen via a convergence test. All simulations used a grid spacing that was an integer division of the comparison resolution (0.5 mm); thus, simulation outputs were resampled to the comparison grid using decimation.

H. MSOUND

MSOUND solves the Helmholtz equation with the absorption term for linear acoustics cases.⁶⁰ For layered media, the conventional angular spectrum approach coupled with the analytical plane wave transmission and reflection coefficients is used. For arbitrarily heterogeneous media, a split-step Fourier method with interpolation is used. Calculations are performed on a regular Cartesian mesh in space. A non-reflecting layer can be used to reduce the spatial aliasing error. Sources are modeled by assigning the complex pressure distribution on the initial plane. In these simulations, the initial plane pressure fields were obtained by FOCUS, as MSOUND currently only considers the pressure-release boundary condition ($p=0$) for the region outside the source. All benchmarks were computed using the function `Forward3D_fund`. Benchmarks 1 and 2 were computed using 6 PPW in all directions. Benchmarks 7–9 were computed using 12 PPW in all directions. Benchmarks 3 and 4 were computed using 6 PPW in the lateral directions and 48 PPW in the axial (propagation) direction. Benchmarks 5 and 6 were computed using 6 PPW in the lateral directions and 24 PPW in the axial direction. For benchmarks 3–9, simulation outputs were down-sampled to the comparison grid.

I. OPTIMUS

OPTIMUS is a full wave solver based on the BEM.^{61,88} The BEM employs the Green's function of the Helmholtz equation to reformulate the volumetric wave problem into a boundary integral equation at the interfaces of piecewise homogeneous domains embedded in free space.⁶² Benchmarks 3, 5, and 7 were modeled with the Poggio–Miller–Chew–Harrington–Wu–Tsai (PMCHWT) formulation,⁶³ benchmarks 4 and 6 were solved with a multi-trace formulation,⁶⁴ and a nested version of the PMCHWT formulation solves benchmarks 8 and 9. The numerical discretization leads to a dense system of linear equations, whose computational footprint is reduced through

hierarchical matrix compression.⁶⁵ The convergence of the iterative GMRES linear solver was improved with OSRC preconditioning.⁶⁶ All models were implemented in PYTHON, using version 3 of the open-source BEMPP library.⁶⁷ The triangular surface meshes were created with Gmsh⁶⁸ for benchmarks 3–6 and using Meshmixer⁶⁹ for benchmarks 7–9.

The size of the mesh elements was specified as 4.3 PPW (0.7 mm) in benchmark 3, 6 PPW (0.5 mm) in benchmarks 4 and 5, 4 PPW (0.75 mm) for benchmark 6, and 10 PPW (0.3 mm) for benchmark 7. A compromise in terms of memory requirements and accuracy of results had to be sought on benchmarks 8 and 9, and a value of 4 PPW (0.75 mm) was used on the skull mesh in the vicinity of the transducer with a value of 2.4 PPW (1.25 mm) elsewhere. The bowl and piston transducers were implemented using a Rayleigh integral formulation, consisting of a summation of evenly spaced monopole radiators positioned on their surface. The transducer surfaces were discretized using 23 and 6 monopole sources per wavelength in water for benchmarks 1–7 and benchmarks 8 and 9, respectively. In cases where the position of monopole sources coincided with a field evaluation point, NaN was assigned to the acoustic pressure. The acoustic field was evaluated from the surface potentials by interpolation for points on, or very close to, the material interface and with Green's functions for points in the material volume.

J. SALVUS

SALVUS solves the second-order linear wave equation in the time domain and can handle acoustic and elastic media.⁷⁰ It utilizes a matrix-free implementation of the continuous-Galerkin spectral-element method⁷¹ and an explicit second-order Newmark time-stepping scheme. The computational domain is discretized using unstructured conforming hexahedral meshes,⁷² which enable the exact representation of interfaces and discontinuities in the tissue parameters. Absorbing boundaries are imposed using the first-order Sommerfeld radiation condition in addition to sponge layers.⁷³ The transducers are modeled as a collection of monopole point sources distributed evenly over the surface of the transducer.

Spectral elements of order 4 were utilized for all simulations; this corresponds to 125 nodes per element. Due to the interfaces being represented precisely using hexahedral meshes generated within Coreform Cubit 2021.5,⁷⁴ utilizing 2–3 elements per wavelength for all benchmarks proved to be sufficient. The maximum pressure distributions were computed by propagating the wavefield in the time domain and then applying the on-the-fly temporal Fourier transform.⁷⁵ All simulation results were output on the same hexahedral discretizations used as inputs and were subsequently resampled onto the comparison grid using fourth-order Lagrange polynomials in the spectral-element basis.

K. SIM4LIFE

SIM4LIFE solves acoustic pressure wave equations (linear, or Westervelt–Lighthill, which considers dispersion and frequency mixing), using a multi-GPU-accelerated FDTD

method on adaptive, rectilinear meshes (to adapt grid-steps to the local wavelength and refine relevant geometric features) with cell-centered pressure degrees-of-freedom. Flux conserving virtual auxiliary points are used to improve accuracy at interfaces and boundaries, and PMLs—according to the stretched coordinate formulation⁷⁶—are used to avoid reflections at domain boundaries (for more details on the numerical methods, see Ref. 77). Results can be recorded as phasors (at the base frequency and, if relevant, higher harmonics) or transient 3 + 1D fields, and the solver has been verified and validated,⁷⁸ also for transcranial focused ultrasound modeling.^{79,80} The original hard sources (imposed pressure; sinusoidal with rise time or user-defined transient profiles) were extended for the purpose of this work by soft sources (cosine function to avoid slowly decaying low frequency components).

The present benchmarks were simulated using isotropic voxel meshes (24 voxels per wavelength, 0.125 mm resolution) over the prescribed simulation domain padded with 96 layers of inhomogeneous PML, with a time step chosen to satisfy the CFL stability criterion (0.026 μ s or 76.9 PPP with bone, 0.048 μ s or 41.7 PPP without). Fifty periods were simulated for benchmarks 1–7 (561 \times 561 \times 961 voxels), while 200 periods were simulated for benchmark 8 (1521 \times 1361 \times 1801) and benchmark 9 (1473 \times 1793 \times 1697). To facilitate comparison, voxeling was offset by half a cell compared to the defined transducer surface, such that transducer grid points correspond to voxel cell centers and material interfaces to voxel faces.

L. STRIDE

STRIDE solves the second-order, isotropic, linear acoustic wave equation using an FDTD approximation over a rectangular Cartesian grid,⁸¹ which is generated using the domain-specific language Devito.⁸² Spatial derivatives are calculated using a 10th-order finite-difference approximation, while time integration is performed using a 4th-order time-stepping scheme optimized for increased stability.⁸³ Acoustic waves at the boundaries are absorbed using either a sponge absorbing boundary⁸⁴ or a complex frequency-shifted PML.⁸⁵ Sources are introduced as free-field monopoles, which can be defined at locations both on and off the grid.⁸⁶

Benchmarks 1–7 were computed using 24 PPW and 120 PPP, resulting in a grid size of 1061 \times 661 \times 661, including absorbing boundaries. Benchmarks 8 and 9 were computed using 18 PPW and 90 PPP, with grid sizes of 1451 \times 1121 \times 1241 and 1373 \times 1445 \times 1205, respectively. A complex frequency-shifted PML was used as the absorbing boundary for all benchmarks. Computed results were resampled onto the comparison mesh using linear interpolation.

IV. BENCHMARK RESULTS

A. Field characteristics

Representative simulation results for all benchmarks are given in Figs. 3 and 4. These illustrate the pressure

amplitudes over the comparison domains given in Table II. The beam shapes for benchmarks 1 and 2 are characteristic of focused bowl and unfocused piston transducers. The introduction of a flat skull bone with a single layer (benchmark 3) or multiple layers (benchmark 4) causes a drop in the focal pressure. Hot-spots (localized regions of increased pressure) are introduced on the skull surface, and the reflected waves generate a complex interference pattern between the transducer and the skull. For the focused bowl transducer (PH1-BM3-SC1 and PH1-BM4-SC1), the reflected waves also generate a secondary focus near the rear surface of the transducer. When a curved skull is used (benchmarks 5 and 6), the hot-spots and secondary focus are reduced. For all benchmarks with the piston transducer, a distinct last-axial maximum is no longer present after the introduction of the skull. Instead, the spatial peak pressure is typically either inside or immediately adjacent to the skull bone, and the acoustic beam gradually diverges after the skull surface. The introduction of a more complex skull geometry in benchmarks 7–9 generates additional features in the pressure fields. For benchmarks 7 and 8, the internal occipital protuberance of the skull bone causes a noticeable deflection of the acoustic beam. The use of the whole skull for benchmarks 8 and 9 also introduces small amplitude reflections from the opposite skull surface (e.g., see PH1-BM8-SC2 in Fig. 4).

B. Difference metrics

Aggregated difference metrics are given in Figs. 5–7. These were calculated by comparing each model with every other model in a cross-comparison and then computing the metrics described in Sec. II G. The box plots (generated using `boxchart` in MATLAB) illustrate the minimum, maximum, median, and first and third quartiles, along with any outliers. The same metrics were also computed for each model and benchmark using `KWAVE` as a reference. This reference was used due to the very high spatial and temporal sampling possible for the `KWAVE` simulations, particularly for benchmarks 1–6, which allowed an axisymmetric formulation to be used. Field plots, axial and lateral profiles, difference plots, and summary tables against `FOCUS` (for benchmarks 1 and 2) and `KWAVE` (for benchmarks 1–9) for each model are given in the supplementary material.¹⁰ These outputs are grouped both by benchmark and by model for ease of reference. Note that the simulation results and the comparison codes are freely available;^{8,11} thus, it is straightforward to generate other comparisons as required or add new modeling results to the intercomparison.

Figure 5 gives a summary of the L^∞ and L^2 intercomparison metrics computed across the comparison domains outlined in Table II. Results are presented for each benchmark (summarizing the cross-comparison results across all codes) and for each code (summarizing the cross-comparison results across all benchmarks). For benchmarks 1 and 2 (water and water with artificial absorption), the level of agreement is very high. For the bowl transducer, seven

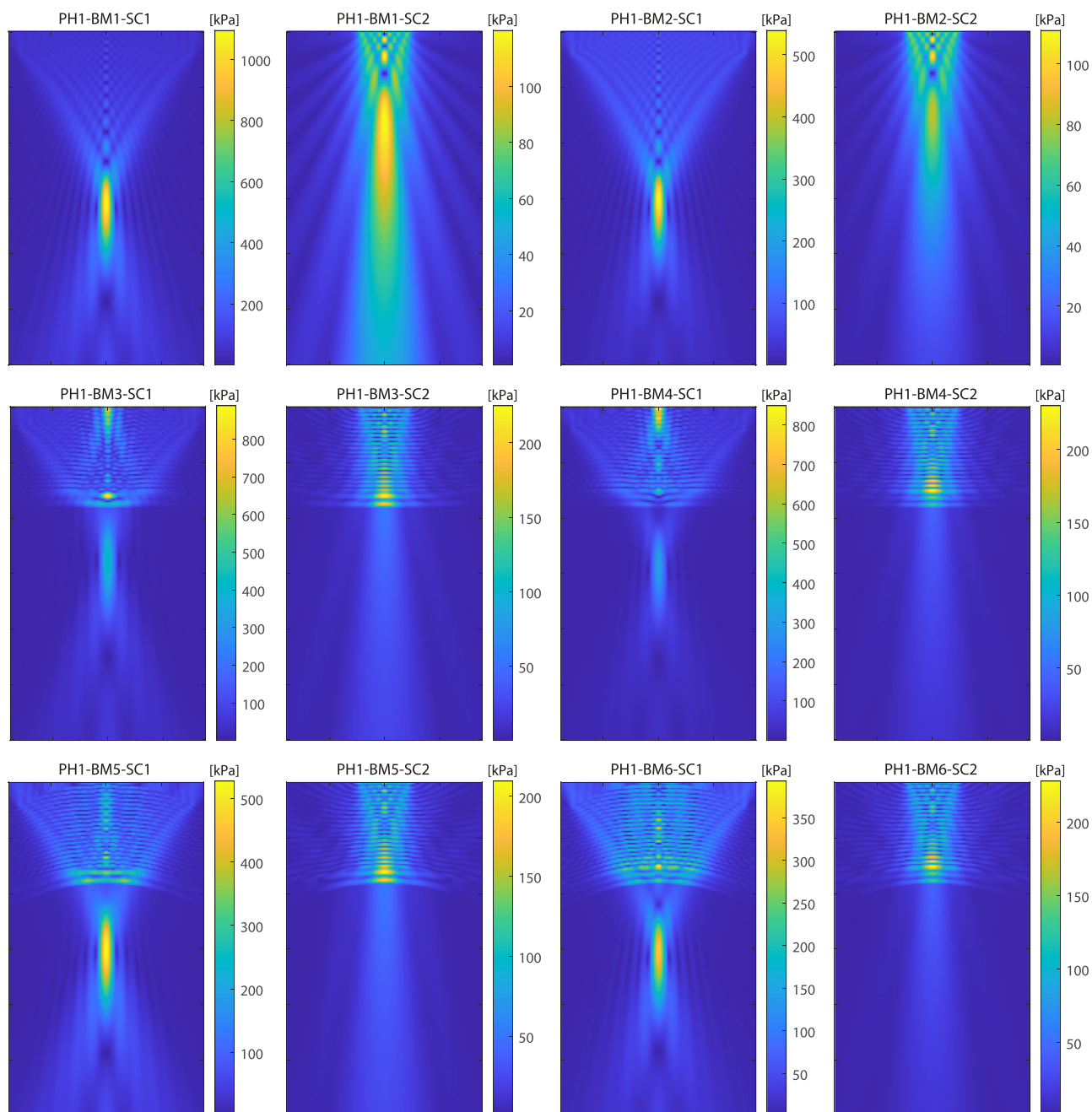


FIG. 3. (Color online) Pressure amplitudes computed using *kwave* for benchmarks 1–6 showing *x-y* slices through the central *z* plane for a comparison domain of 120 mm (axial) by 70 mm (lateral).

models have L^∞ values of less than 1% when compared to *FOCUS*, and all values are less than 10% (see supplementary material). For the piston transducer, the simulations are slightly less accurate. Four models have L^∞ values of less than 1% when compared to *FOCUS*, and the maximum L^∞ value against *FOCUS* is 15%. Examining the difference plots (see supplementary material),¹⁰ the largest differences are in the complex near-field pattern close to the transducer surface, where the pressure varies rapidly.

For benchmarks 3–9, the L^∞ and L^2 metrics both increase noticeably, with median values for the cross-comparison between 10% and 100% [Fig. 5(a)]. There is still close agreement between some models, for example, three

models have median L^∞ values less than 10% across all benchmarks when compared to *kwave* [see Fig. 5(b)]. However, in general, the differences are larger than those found for benchmarks 1 and 2. Examining the difference plots (see supplementary material),¹⁰ the largest variations are in the region between the transducer and skull bone. These arise due to a combination of errors in modeling the near-field of the transducer, even in free-field (described above), along with errors in modeling the reflection from the bone and soft-tissue surfaces (e.g., due to errors in the positions of the interfaces and amplitude and phase errors in the reflected waves).

Overall, the L^∞ and L^2 intercomparison metrics demonstrate that, on a pixel-by-pixel basis, there are often large

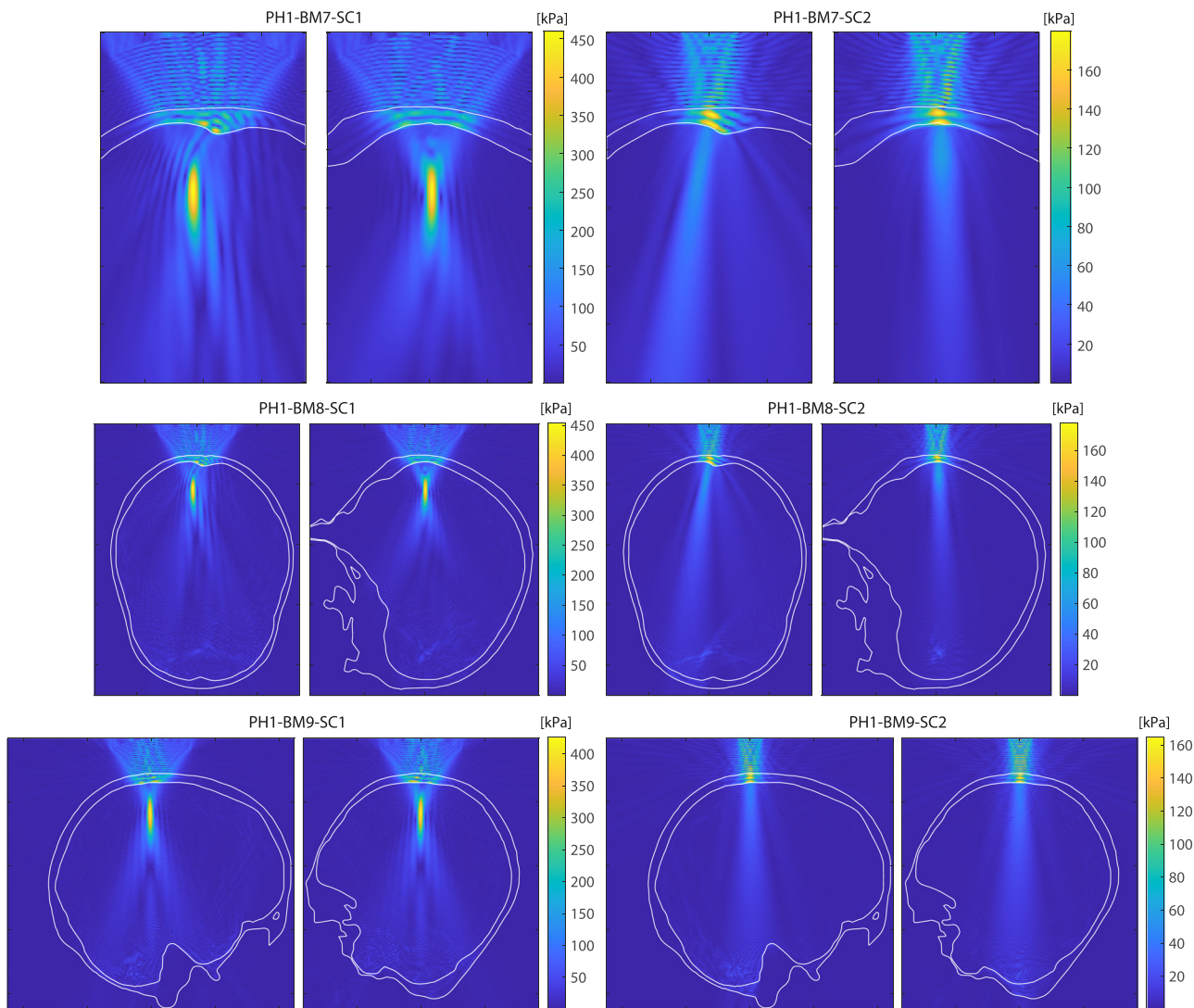


FIG. 4. (Color online) Pressure amplitudes computed using *kwave* for benchmarks 7–9 showing *x*-*y* (left) and *x*-*z* (right) slices through the location of the peak pressure. The approximate location of the skull is shown with the white overlay. The size of the comparison domain for each benchmark is given in Table II.

variations between the model outputs. This is true despite there being no uncertainty in the material parameters and transducer characteristics. This highlights the inherent uncertainties when using computational models for transcranial ultrasound simulation, which must be considered when interpreting model results.

Figures 6 and 7 give a summary of the intercomparison metrics for the focal position, size, and pressure. Despite the variations in the full-field error norms discussed above, there is very close agreement in the focal metrics. When compared by benchmark, the median values for the difference in focal pressure are all less than 10% [see Fig. 6(a)]. Similarly, when compared by code, 10 of 11 models have median differences less than 10%. Differences of this level are on par with experimental repeatability and reproducibility measurements conducted using similar ultrasound transducers and a range of hydrophones.⁸⁷ Compared to *kwave*, seven models have maximum differences in the focal pressure across all benchmarks of less than 10%, and five

models have median differences across all benchmarks on the order of 1% or less [see Fig. 6(b)]. Considering the focal position, all values including outliers are within 2.5 mm [see Fig. 6(a)], with median values for all benchmarks of 1 mm or less. Compared to *kwave*, the median values for all models are less than 0.5 mm, with seven models having a median value of 0 mm [see Fig. 6(b)].

Figure 7 gives a summary of the intercomparison metrics for focal size. Note, as mentioned in Sec. II G, the axial focal size for the piston transducer (SC2) for benchmarks 3–9 is not calculated as there is no focus after propagation through the skull. For reference, in water (BM1), the axial and lateral focal size for the focused bowl transducer is 26.2 and 4.1 mm, respectively, and the lateral focal size for the piston transducer at $x = 60$ mm is 13.2 mm. For all benchmarks, the median differences in the axial focal size for the focused bowl transducer are less than 0.6 mm [Fig. 7(a)], although there are a small number of outliers with differences up to 2.3 mm. The median differences in the lateral

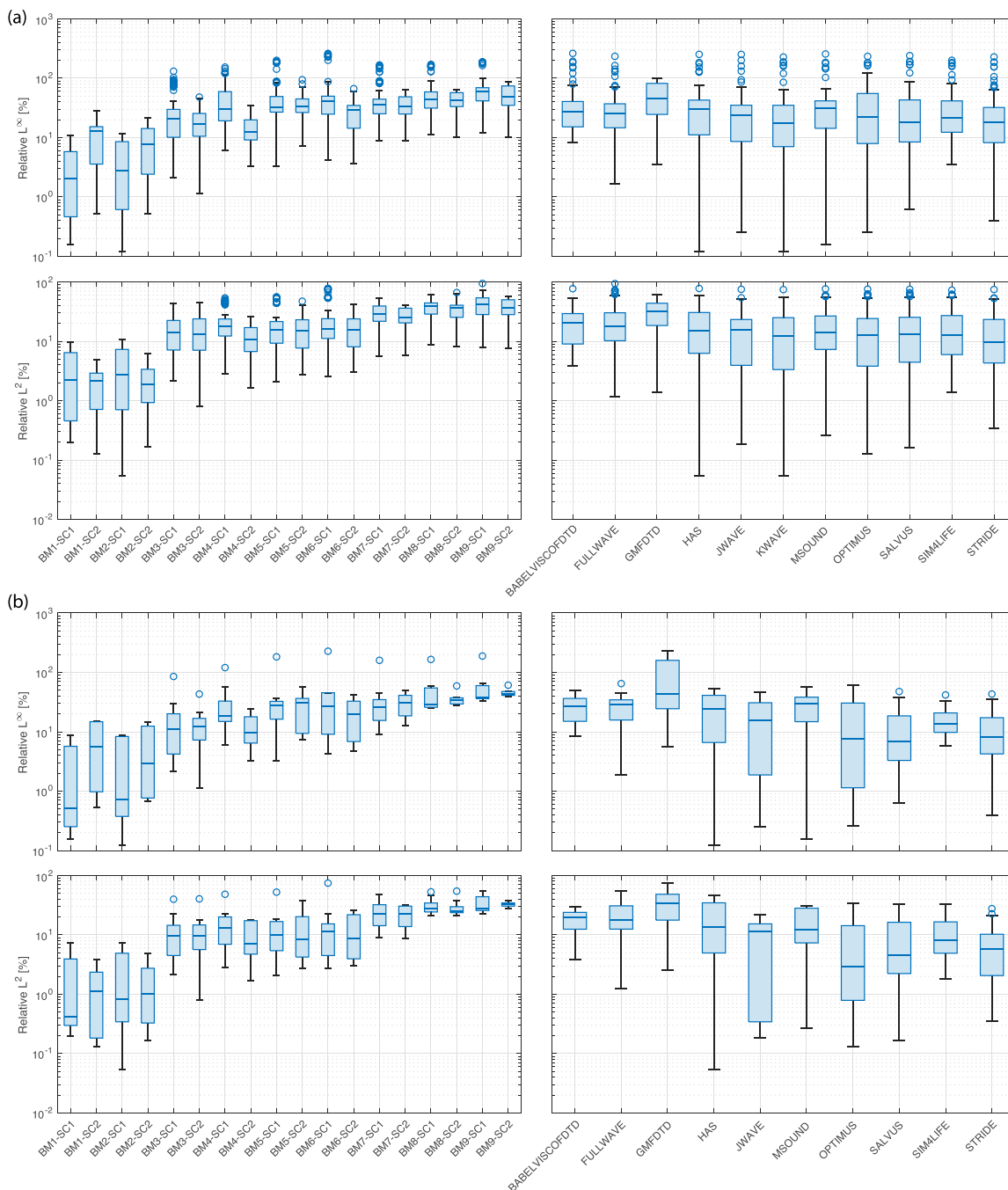


FIG. 5. (Color online) Summary of relative L^∞ and L^2 difference metrics computed across the entire field taken from the exit plane of the transducer. (a) Cross-comparison (all codes compared with all codes). (b) Comparison with *KWAVE*.

focal size for the focused bowl transducer for all benchmarks are 0.2 mm or less. Variations in the lateral focal size for the piston transducer are generally larger, noting the lateral focal size is also larger for this transducer. Similar results are evident for the comparison against *KWAVE* [Fig. 7(b)].

Overall, there is very close agreement for all benchmarks in the characteristics of the focal pressure field after propagation through the skull bone. Larger differences are evident in the full-field metrics, dominated by differences in the field between the transducer and the skull. The most

relevant metrics to compare, along with acceptable limits on the differences between models, depend strongly on the intended application of the computational results. For example, calculating phase delays, calculating the approximate position and size of the acoustic focus in the brain, and calculating the pressure in the skin and skull to subsequently estimate skull heating may each have different constraints and accuracy requirements. An analysis of these factors is beyond the scope of the current work. However, it is hoped that the benchmarks and computational results presented here may help to facilitate such investigations in the future.

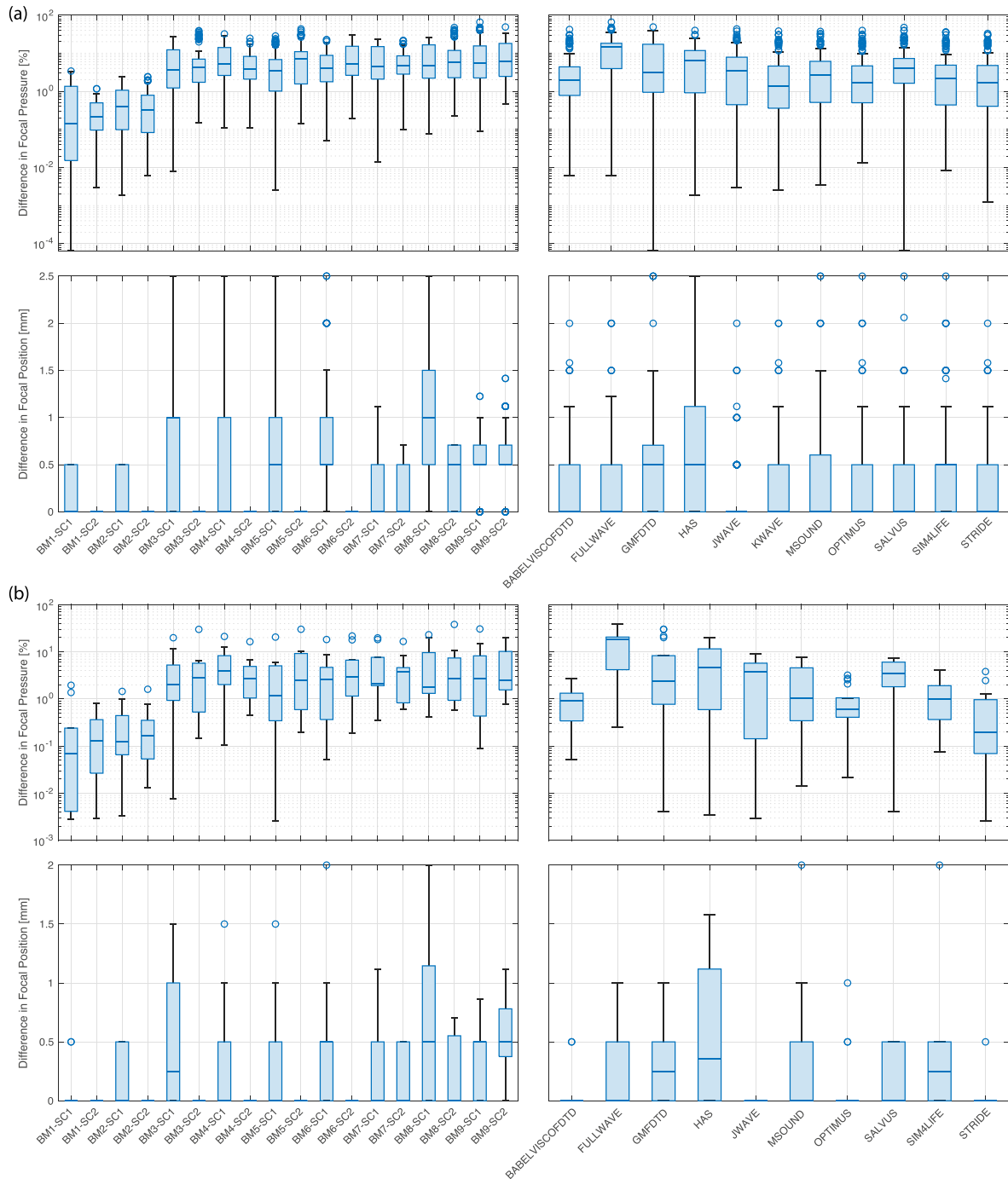


FIG. 6. (Color online) Summary of focal (peak) pressure and focal position metrics. (a) Cross-comparison (all codes compared with all codes). (b) Comparison with KWAVE.

V. SUMMARY

A series of numerical benchmarks relevant to transcranial ultrasound simulation are presented, along with inter-comparison results for 11 modeling tools used in the community. The intercomparison results show close agreement between the models, particularly for the position, size, and magnitude of the acoustic focus after propagating through the skull. When comparing each model with every other model in a cross-comparison, the median values for the difference in focal pressure and focal position are less

than 10% and 1 mm for all benchmarks. The differences in focal pressure are comparable to variations in experimental measurements,⁸⁷ and the median differences in the axial and lateral focal position (0.6 and 0.2 mm) for the focused transducer are small compared to the corresponding size of the -6 dB focal volume (26.2 and 4.1 mm). These results build confidence in the ability of the described computational models to produce consistent results when simulating wave propagation through skull layers at 500 kHz. The benchmark definitions and associated data files, simulation results, and

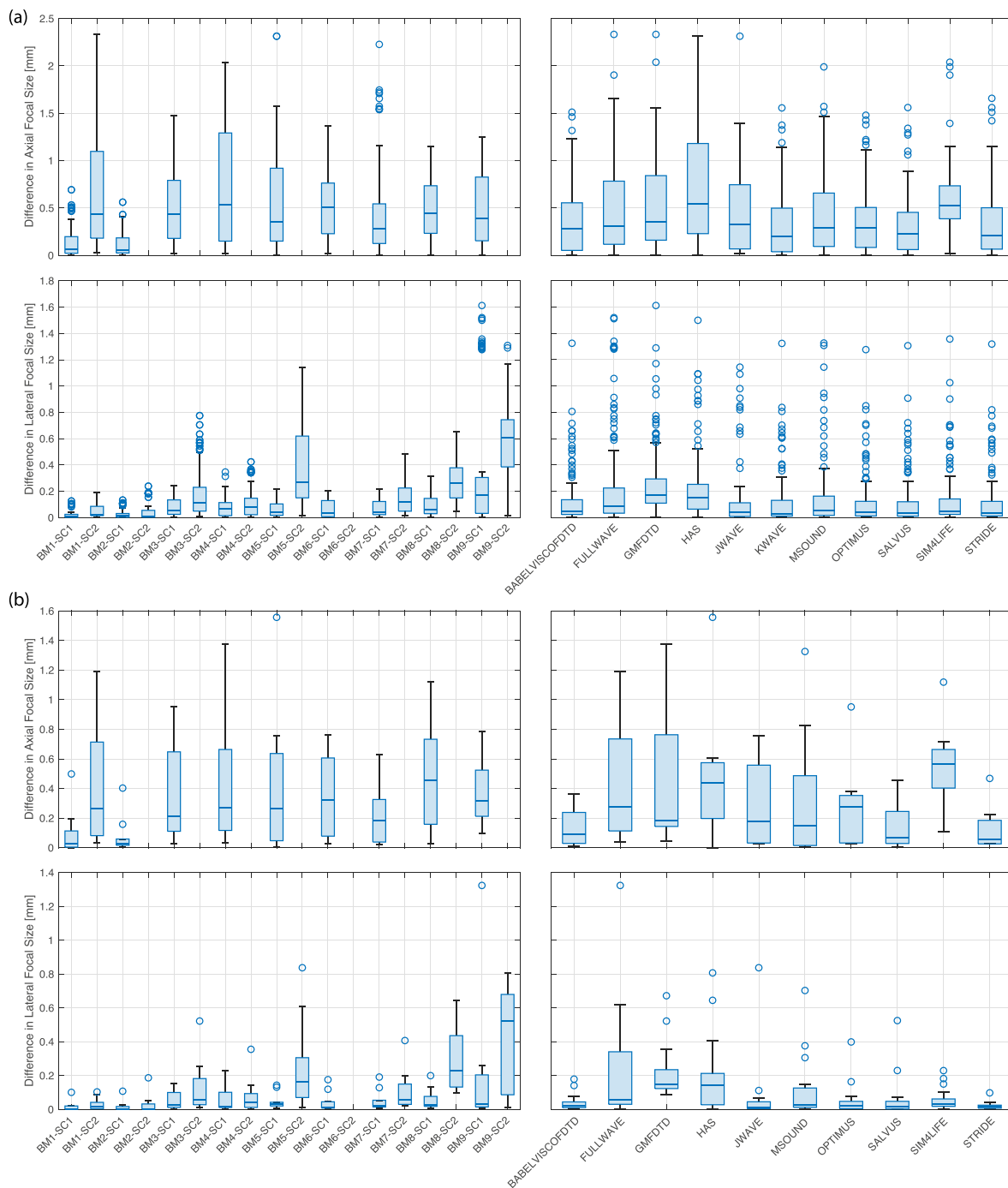


FIG. 7. (Color online) Summary of axial and lateral focal size metrics. Note that axial focal size was not computed for benchmarks 3–9 when using the plane piston source (SC2), as the field in this case did not have an axial maximum in the post-skull region. (a) Cross-comparison (all codes compared with all codes). (b) Comparison with KWAVE.

codes to compute the intercomparison metrics are all freely available.^{8,11} This allows the results to be replicated or further analysis to be conducted. Additional model results can also be easily added to the intercomparison, for example, to validate newly developed solvers. More generally, the intercomparison exercise provides a framework for creating benchmarks and performing model cross-comparisons. Further phases of the intercomparison exercise are currently under discussion, including benchmarks for elastic wave

models and model comparisons when using material parameters derived from CT images.

ACKNOWLEDGMENTS

The authors would like to thank the International Transcranial Ultrasonic Stimulation Safety and Standards (ITRUSST) consortium for providing the motivation and framework to conduct this work and Robert McGough for

helpful discussions regarding the use of FOCUS as a reference simulation. O.B. was supported by the Engineering and Physical Sciences Research Council (EPSRC) Centre for Doctoral Training in Neurotechnology, Grant No. EP/L016737/1. K.B.P. and N.L. acknowledge the support of National Institutes of Health (NIH) Grant Nos. R01 CA227687, NIH T32 EB009653, and NSF DGE 1656518. D.C. acknowledges support from the Focused Ultrasound Foundation and NIH Grant Nos. R01 EB013433, R01 CA172787, R01 EB028316, and R37 CA224141. C.C. acknowledges the support of EPSRC Grant No. EP/T51780X/1. P.G. and E.v.W. acknowledge the support of EPSRC Grant No. EP/P012434/1 and use of the UCL Myriad High Performance Computing Facility (Myriad@UCL) and associated support services. J.J. was supported by Brno University of Technology under Project No. FIT-S-20-6309. Y.J. acknowledges the support of NIH Grant No. R01EB025205. P.M. and C.B. acknowledge support from the Swiss National Supercomputing Centre (CSCS) under Project Nos. s1040 and sm59. S.P. acknowledges the support of the Natural Sciences and Engineering Research Council of Canada and the Canada Foundation for Innovation. A.P. would like to acknowledge Academy of Finland Project Nos. 320166, 336119, and 336799. A.S. and B.T. were supported by EPSRC Grant No. EP/S026371/1. A.T. was supported by Lundbeck Foundation Grant No. R313-2019-622.

¹W. J. Elias, N. Lipsman, W. G. Ondo, P. Ghanouni, Y. G. Kim, W. Lee, M. Schwartz, K. Hynynen, A. M. Lozano, B. B. Shah, D. Huss, R. F. Dallapiazza, R. Gwinn, J. Witt, S. Ro, H. M. Eisenberg, P. S. Fishman, D. Gandhi, C. H. Halpern, R. Chuang, K. Butts Pauly, T. S. Tierney, M. T. Hayes, G. Rees Cosgrove, T. Yamaguchi, K. Abe, T. Taira, and J. W. Chang, "A randomized trial of focused ultrasound thalamotomy for essential tremor," *N. Engl. J. Med.* **375**(8), 730–739 (2016).

²A. Abrahao, Y. Meng, M. Llinas, Y. Huang, C. Hamani, T. Mainprize, I. Aubert, C. Heyn, S. E. Black, K. Hynynen, N. Lipsman, and L. Zinman, "First-in-human trial of blood–brain barrier opening in amyotrophic lateral sclerosis using MR-guided focused ultrasound," *Nat. Commun.* **10**(1), 4373 (2019).

³W. Legon, T. F. Sato, A. Opitz, J. Mueller, A. Barbour, A. Williams, and W. J. Tyler, "Transcranial focused ultrasound modulates the activity of primary somatosensory cortex in humans," *Nat. Neurosci.* **17**(2), 322–329 (2014).

⁴K. Hynynen and F. A. Jolesz, "Demonstration of potential noninvasive ultrasound brain therapy through an intact skull," *Ultrasound Med. Biol.* **24**(2), 275–283 (1998).

⁵G. Bouchoux, K. B. Bader, J. J. Korfhagen, J. L. Raymond, R. Shivashankar, T. A. Abruzzo, and C. K. Holland, "Experimental validation of a finite-difference model for the prediction of transcranial ultrasound fields based on CT images," *Phys. Med. Biol.* **57**(23), 8005–8022 (2012).

⁶F. Marquet, M. Pernot, J.-F. Aubry, G. Montaldo, L. Marsac, M. Tanter, and M. Fink, "Non-invasive transcranial ultrasound therapy based on a 3D CT scan: Protocol validation and *in vitro* results," *Phys. Med. Biol.* **54**(9), 2597–2613 (2009).

⁷R. F. Dallapiazza, K. F. Timbie, S. Holmberg, J. Gatesman, M. B. Lopes, R. J. Price, G. W. Miller, and W. J. Elias, "Noninvasive neuromodulation and thalamic mapping with low-intensity focused ultrasound," *J. Neurosurg.* **128**(3), 875–884 (2018).

⁸B. Treeby, "Benchmark problems for transcranial ultrasound simulation: Intercomparison library (version 1.0) [computer code]," <https://github.com/ucl-bug/transcranial-ultrasound-benchmarks> (Last viewed August 11, 2022).

⁹J. Gu and Y. Jing, "A modified mixed domain method for modeling acoustic wave propagation in strongly heterogeneous media," *J. Acoust. Soc. Am.* **147**(6), 4055–4068 (2020).

¹⁰See supplementary material at <https://www.scitation.org/doi/suppl/10.1121/10.0013426> for a table form of the model summaries given in Sec. III and summaries of the complete set of intercomparison results. SuppPub1.xlsx gives an alternate table form of the model summaries given in Sec. III. SuppPub2.zip provides summaries of the comparison results (including metrics, field plots, axial profiles, and difference plots) for each model compared against FOCUS (for benchmarks 1 and 2) and KWAVE (for benchmarks 1–9). The .zip file contains separate pdf files for each model and for each benchmark, as well as a summary of the cross-comparison metrics. The raw data files and MATLAB codes to process the results are also freely available (Refs. 8 and 11).

¹¹J.-F. Aubry, O. Bates, C. Boehm, K. Butts Pauly, D. Christensen, C. Cueto, P. Gelat, L. Guasch, J. Jaros, Y. Jing, R. Jones, N. Li, P. Marty, H. Montanaro, E. Neufeld, S. Pichardo, G. Pinton, A. Pulkkinen, A. Stanzola, A. Thielscher, B. Treeby, and E. van 't Wout, "Benchmark problems for transcranial ultrasound simulation: Datasets for intercomparison of compressional wave models (version 1.0) [dataset]," Zenodo, <https://doi.org/10.5281/zenodo.6020543> (2022).

¹²J. K. Mueller, L. Ai, P. Bansal, and W. Legon, "Numerical evaluation of the skull for human neuromodulation with transcranial focused ultrasound," *J. Neural Eng.* **14**(6), 066012 (2017).

¹³G. T. Clement, P. J. White, and K. Hynynen, "Enhanced ultrasound transmission through the human skull using shear mode conversion," *J. Acoust. Soc. Am.* **115**(3), 1356–1364 (2004).

¹⁴X.-D. Wang, W.-J. Lin, C. Su, and X.-M. Wang, "Influence of mode conversions in the skull on transcranial focused ultrasound and temperature fields utilizing the wave field separation method: A numerical study," *Chin. Phys. B* **27**(2), 024302 (2018).

¹⁵Y. Younan, T. Deffieux, B. Larrat, M. Fink, M. Tanter, and J.-F. Aubry, "Influence of the pressure field distribution in transcranial ultrasonic neurostimulation," *Med. Phys.* **40**(8), 082902 (2013).

¹⁶J. L. Robertson, B. T. Cox, J. Jaros, and B. E. Treeby, "Accurate simulation of transcranial ultrasound propagation for ultrasonic neuromodulation and stimulation," *J. Acoust. Soc. Am.* **141**(3), 1726–1738 (2017).

¹⁷H. O'Neil, "Theory of focusing radiators," *J. Acoust. Soc. Am.* **21**(5), 516–526 (1949).

¹⁸F. J. Fry and J. E. Barger, "Acoustical properties of the human skull," *J. Acoust. Soc. Am.* **63**(5), 1576–1590 (1978).

¹⁹T. D. Mast, "Empirical relationships between acoustic parameters in human soft tissues," *Acoust. Res. Lett. Online* **1**(2), 37–42 (2000).

²⁰G. Clement and K. Hynynen, "Correlation of ultrasound phase with physical skull properties," *Ultrasound Med. Biol.* **28**(5), 617–624 (2002).

²¹S. Pichardo, V. W. Sin, and K. Hynynen, "Multi-frequency characterization of the speed of sound and attenuation coefficient for longitudinal transmission of freshly excised human skulls," *Phys. Med. Biol.* **56**(1), 219–250 (2011).

²²G. Pinton, J.-F. Aubry, E. Bossy, M. Muller, M. Pernot, and M. Tanter, "Attenuation, scattering, and absorption of ultrasound in the skull bone," *Med. Phys.* **39**(1), 299–307 (2011).

²³S. Pichardo, C. Moreno-Hernández, R. A. Drainville, V. Sin, L. Curiel, and K. Hynynen, "A viscoelastic model for the prediction of transcranial ultrasound propagation: Application for the estimation of shear acoustic properties in the human skull," *Phys. Med. Biol.* **62**(17), 6938–6962 (2017).

²⁴P. J. White, G. T. Clement, and K. Hynynen, "Longitudinal and shear mode ultrasound propagation in human skull bone," *Ultrasound Med. Biol.* **32**(7), 1085–1096 (2006).

²⁵T. D. Webb, S. A. Leung, P. Ghanouni, J. J. Dahl, N. J. Pelc, and K. B. Pauly, "Acoustic attenuation: Multifrequency measurement and relationship to CT and MR imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **68**(5), 1532–1545 (2021).

²⁶R. J. McGough, T. V. Samulski, and J. F. Kelly, "An efficient grid sectoring method for calculations of the near-field pressure generated by a circular piston," *J. Acoust. Soc. Am.* **115**(5), 1942–1954 (2004).

²⁷D. Chen and R. J. McGough, "A 2D fast near-field method for calculating near-field pressures generated by apodized rectangular pistons," *J. Acoust. Soc. Am.* **124**(3), 1526–1537 (2008).

²⁸J. F. Kelly and R. J. McGough, "Transient fields generated by spherical shells in viscous media," *AIP Conf. Proc.* **1113**, 210–214 (2009).

²⁹S. L. Alexander, K. Rafaels, C. A. Gunnarsson, and T. Weerasooriya, "Structural analysis of the frontal and parietal bones of the human skull," *J. Mech. Behav. Biomed. Mater.* **90**, 689–701 (2019).

- ³⁰H. Hori, G. Moretti, A. Rebor, and F. Crovato, "The thickness of human scalp: Normal and bald," *J. Invest. Dermatol.* **58**(6), 396–399 (1972).
- ³¹V. S. Fonov, A. C. Evans, R. C. McKinstry, C. Almlı, and D. Collins, "Unbiased nonlinear average age-appropriate brain templates from birth to adulthood," *NeuroImage* **47**(Suppl. 1), S39–S41 (2009).
- ³²V. Fonov, A. C. Evans, K. Botteron, C. R. Almlı, R. C. McKinstry, D. L. Collins, and Brain Development Cooperative Group, "Unbiased average age-appropriate atlases for pediatric studies," *NeuroImage* **54**(1), 313–327 (2011).
- ³³J. D. Nielsen, K. H. Madsen, O. Puonti, H. R. Siebner, C. Bauer, C. G. Madsen, G. B. Saturnino, and A. Thielscher, "Automatic skull segmentation from MR images for realistic volume conductor models of the head: Assessment of the state-of-the-art," *NeuroImage* **174**, 587–598 (2018).
- ³⁴Q. Fang, "iso2mesh: a 3D surface and volumetric mesh generator for MATLAB/Octave," <http://iso2mesh.sourceforge.net> (Last viewed August 11, 2022).
- ³⁵Q. Fang and D. A. Boas, "Tetrahedral mesh generation from volumetric binary and grayscale images," in *Proceedings of the 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, Boston, MA (June 28–July 1, 2009), pp. 1142–1145.
- ³⁶J. Virieux, "P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method," *Geophysics* **51**(4), 889–901 (1986).
- ³⁷A. R. Levander, "Fourth-order finite-difference P-SV seismograms," *Geophysics* **53**(11), 1425–1436 (1988).
- ³⁸T. Bohlen, "Parallel 3-D viscoelastic finite difference seismic modelling," *Comput. Geosci.* **28**(8), 887–899 (2002).
- ³⁹J. O. Blanch, J. O. Robertsson, and W. W. Symes, "Modeling of a constant Q: Methodology and algorithm for an efficient and optimally inexpensive viscoelastic technique," *Geophysics* **60**(1), 176–184 (1995).
- ⁴⁰P. Moczo, J. Kristek, V. Vavrycuk, R. J. Archuleta, and L. Halada, "3D heterogeneous staggered-grid finite-difference modeling of seismic motion with volume harmonic and arithmetic averaging of elastic moduli and densities," *Bull. Seismol. Soc. Am.* **92**(8), 3042–3066 (2002).
- ⁴¹R. A. Drainville, L. Curiel, and S. Pichardo, "Superposition method for modelling boundaries between media in viscoelastic finite difference time domain simulations," *J. Acoust. Soc. Am.* **146**(6), 4382–4401 (2019).
- ⁴²F. Collino and C. Tsogka, "Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media," *Geophysics* **66**(1), 294–307 (2001).
- ⁴³G. F. Pinton, J. Dahl, S. Rosenzweig, and G. E. Trahey, "A heterogeneous nonlinear attenuating full-wave model of ultrasound," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **56**(3), 474–488 (2009).
- ⁴⁴G. Pinton, "A fullwave model of the nonlinear wave equation with multiple relaxations and relaxing perfectly matched layers for high-order numerical finite-difference solutions," [arXiv:2106.11476](https://arxiv.org/abs/2106.11476) (2021).
- ⁴⁵R. Courant, K. Friedrichs, and H. Lewy, "Über die partiellen differenzgleichungen der mathematischen physik" ("On the partial differential equations of mathematical physics"), *Math. Ann.* **100**(1), 32–74 (1928).
- ⁴⁶A. Pulkkinen, B. Werner, E. Martin, and K. Hynynen, "Numerical simulations of clinical focused ultrasound functional neurosurgery," *Phys. Med. Biol.* **59**(7), 1679–1700 (2014).
- ⁴⁷U. Vyas and D. Christensen, "Ultrasound beam simulations in inhomogeneous tissue geometries using the hybrid angular spectrum method," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **59**(6), 1093–1100 (2012).
- ⁴⁸S. Almqvist, D. L. Parker, and D. A. Christensen, "Rapid full-wave phase aberration correction method for transcranial high-intensity focused ultrasound therapies," *J. Ther. Ultrasound* **4**(1), 30 (2016).
- ⁴⁹A. Stanzola, S. R. Arridge, B. T. Cox, and B. E. Treeby, "j-Wave: An open-source differentiable wave simulator," [arXiv:2207.01499](https://arxiv.org/abs/2207.01499) (2022).
- ⁵⁰Y. Saad and M. H. Schultz, "GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems," *SIAM J. Sci. Stat. Comput.* **7**(3), 856–869 (1986).
- ⁵¹A. Bermúdez, L. Hervella-Nieto, A. Prieto, and R. Rodríguez, "An optimal perfectly matched layer with unbounded absorbing function for time-harmonic acoustic scattering problems," *J. Comput. Phys.* **223**(2), 469–488 (2007).
- ⁵²E. S. Wise, B. Cox, J. Jaros, and B. E. Treeby, "Representing arbitrary acoustic source and sensor distributions in Fourier collocation methods," *J. Acoust. Soc. Am.* **146**(1), 278–288 (2019).
- ⁵³A. Stanzola, S. R. Arridge, B. T. Cox, and B. E. Treeby, "A research framework for writing differentiable PDE discretizations in JAX," [arXiv:2111.05218](https://arxiv.org/abs/2111.05218) (2021).
- ⁵⁴J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, "JAX: Composable transformations of Python+NumPy programs," [http://github.com/google/jax](https://github.com/google/jax) (Last viewed August 11, 2022).
- ⁵⁵B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *J. Biomed. Opt.* **15**(2), 021314 (2010).
- ⁵⁶B. E. Treeby, J. Jaros, A. P. Rendell, and B. Cox, "Modeling nonlinear ultrasound propagation in heterogeneous media with power law absorption using a k-space pseudospectral method," *J. Acoust. Soc. Am.* **131**(6), 4324–4336 (2012).
- ⁵⁷B. Cox and B. Treeby, "Accurate time-varying sources in k-space pseudospectral time domain acoustic simulations," in *Proceedings of the 2018 IEEE International Ultrasonics Symposium (IUS)*, Kobe, Japan (October 22–25, 2018), pp. 1–4.
- ⁵⁸B. E. Treeby, E. S. Wise, F. Kuklis, J. Jaros, and B. Cox, "Nonlinear ultrasound simulation in an axisymmetric coordinate system using a k-space pseudospectral method," *J. Acoust. Soc. Am.* **148**(4), 2288–2300 (2020).
- ⁵⁹J. Jaros, A. P. Rendell, and B. E. Treeby, "Full-wave nonlinear ultrasound simulation on distributed clusters with applications in high-intensity focused ultrasound," *Int. J. High Perform. Comput. Appl.* **30**(2), 137–155 (2016).
- ⁶⁰J. Gu and Y. Jing, "MSOUND: An open source toolbox for modeling acoustic wave propagation in heterogeneous media," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **68**(5), 1476–1486 (2021).
- ⁶¹E. van 't Wout, P. Gélât, T. Betcke, and S. Arridge, "A fast boundary element method for the scattering analysis of high-intensity focused ultrasound," *J. Acoust. Soc. Am.* **138**(5), 2726–2737 (2015).
- ⁶²E. van 't Wout, S. R. Haqshenas, P. Gélât, T. Betcke, and N. Saffari, "Boundary integral formulations for acoustic modelling of high-contrast media," *Comput. Math. Appl.* **105**, 136–149 (2022).
- ⁶³S. R. Haqshenas, P. Gélât, E. van 't Wout, T. Betcke, and N. Saffari, "A fast full-wave solver for calculating ultrasound propagation in the body," *Ultrasonics* **110**, 106240 (2021).
- ⁶⁴E. van 't Wout, S. R. Haqshenas, P. Gélât, T. Betcke, and N. Saffari, "Benchmarking preconditioned boundary integral formulations for acoustics," *Int. J. Numer. Methods Eng.* **122**(20), 5873–5897 (2021).
- ⁶⁵T. Betcke, E. van 't Wout, and P. Gélât, "Computationally efficient boundary element methods for high-frequency Helmholtz problems in unbounded domains," in *Modern Solvers for Helmholtz Problems*, edited by D. Lahaye, J. Tang, and K. Vuik (Birkhäuser, Cham, Switzerland, 2017), pp. 215–243.
- ⁶⁶E. van 't Wout, S. R. Haqshenas, P. Gélât, T. Betcke, and N. Saffari, "Frequency-robust preconditioning of boundary integral equations for acoustic transmission," *J. Comput. Phys.* **462**, 111229 (2022).
- ⁶⁷W. Śmigaj, T. Betcke, S. Arridge, J. Phillips, and M. Schweiger, "Solving boundary integral problems with BEM," *ACM Trans. Math. Softw.* **41**, 1–40 (2015).
- ⁶⁸C. Geuzaine and J.-F. Remacle, "Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities," *Int. J. Numer. Methods Eng.* **79**(11), 1309–1331 (2009).
- ⁶⁹R. Schmidt and K. Singh, "Meshmixer: An interface for rapid mesh composition," in *Proceedings of ACM SIGGRAPH 2010 Talks*, Los Angeles, CA (July 26–30, 2010).
- ⁷⁰M. Afanasiev, C. Boehm, M. van Driel, L. Krischer, M. Rietmann, D. A. May, M. G. Knepley, and A. Fichtner, "Modular and flexible spectral-element waveform modelling in two and three dimensions," *Geophys. J. Int.* **216**(3), 1675–1692 (2019).
- ⁷¹A. Ferroni, P. F. Antonietti, I. Mazzeri, and A. Quarteroni, "Dispersion-dissipation analysis of 3-D continuous and discontinuous spectral element methods for the elastodynamics equation," *Geophys. J. Int.* **211**(3), 1554–1574 (2017).
- ⁷²V. Hapla, M. G. Knepley, M. Afanasiev, C. Boehm, M. van Driel, L. Krischer, and A. Fichtner, "Fully parallel mesh I/O using PETSc DMplex with an application to waveform modeling," *SIAM J. Sci. Comput.* **43**(2), C127–C153 (2021).
- ⁷³R. Kosloff and D. Kosloff, "Absorbing boundaries for wave propagation problems," *J. Comput. Phys.* **63**(2), 363–376 (1986).
- ⁷⁴Coreform LLC, "Coreform Cubit (version 2021.5) [computer software]," [http://coreform.com](https://coreform.com) (Last viewed August 11, 2022).
- ⁷⁵P. A. Witte, M. Louboutin, F. Luporini, G. J. Gorman, and F. J. Hermann, "Compressive least-squares migration with on-the-fly Fourier transforms," *Geophysics* **84**(5), R655–R672 (2019).

- ⁷⁶W. C. Chew and W. H. Weedon, "A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates," *Microw. Opt. Technol. Lett.* **7**(13), 599–604 (1994).
- ⁷⁷A. Kyriakou, E. Neufeld, B. Werner, G. Székely, and N. Kuster, "Full-wave acoustic and thermal modeling of transcranial ultrasound propagation and investigation of skull-induced aberration correction techniques: A feasibility study," *J. Ther. Ultrasound* **3**(1), 11 (2015).
- ⁷⁸E. Neufeld, A. Kyriacou, W. Kainz, and N. Kuster, "Approach to validate simulation-based distribution predictions combining the gamma-method and uncertainty assessment: Application to focused ultrasound," *J. Verif. Valid. Uncertain Quantif.* **1**(3), 031006 (2016).
- ⁷⁹C. Pasquinelli, H. Montanaro, H. J. Lee, L. G. Hanson, H. Kim, N. Kuster, H. R. Siebner, E. Neufeld, and A. Thielscher, "Transducer modeling for accurate acoustic simulations of transcranial focused ultrasound stimulation," *J. Neural Eng.* **17**(4), 046010 (2020).
- ⁸⁰H. Montanaro, C. Pasquinelli, H. J. Lee, H. Kim, H. R. Siebner, N. Kuster, A. Thielscher, and E. Neufeld, "The impact of CT image parameters and skull heterogeneity modeling on the accuracy of transcranial focused ultrasound simulations," *J. Neural Eng.* **18**(4), 046041 (2021).
- ⁸¹C. Cueto, O. Bates, G. Strong, J. Cudeiro, F. Luporini, Ö. C. Agudo, G. Gorman, L. Guasch, and M.-X. Tang, "Stride: A flexible software platform for high-performance ultrasound computed tomography," *Comput. Methods Programs Biomed.* **221**, 106855 (2022).
- ⁸²M. Louboutin, M. Lange, F. Luporini, N. Kukreja, P. A. Witte, F. J. Herrmann, P. Velesko, and G. J. Gorman, "Devito (v3. 1.0): An embedded domain-specific language for finite differences and geophysical exploration," *Geosci. Model Dev.* **12**(3), 1165–1187 (2019).
- ⁸³L. Amundsen and Ø. Pedersen, "Time step n -tupling for wave equations," *Geophysics* **82**(6), T249–T254 (2017).
- ⁸⁴G. Yao, N. V. Da Silva, and D. Wu, "An effective absorbing layer for the boundary condition in acoustic seismic wave simulation," *J. Geophys. Eng.* **15**(2), 495–511 (2018).
- ⁸⁵Y. Gao, J. Zhang, and Z. Yao, "Unsplit complex frequency shifted perfectly matched layer for second-order wave equation using auxiliary differential equations," *J. Acoust. Soc. Am.* **138**(6), EL551–EL557 (2015).
- ⁸⁶G. J. Hicks, "Arbitrary source and receiver positioning in finite-difference schemes using Kaiser windowed sinc functions," *Geophysics* **67**(1), 156–165 (2002).
- ⁸⁷E. Martin and B. Treeby, "Investigation of the repeatability and reproducibility of hydrophone measurements of medical ultrasound fields," *J. Acoust. Soc. Am.* **145**(3), 1270–1282 (2019).
- ⁸⁸P. Gélat, S. R. Haqshenas, and E. van 't Wout, "OptimUS: A Python library for solving 3D acoustic wave propagation," <https://github.com/optimuslib> (Last viewed August 11, 2022).