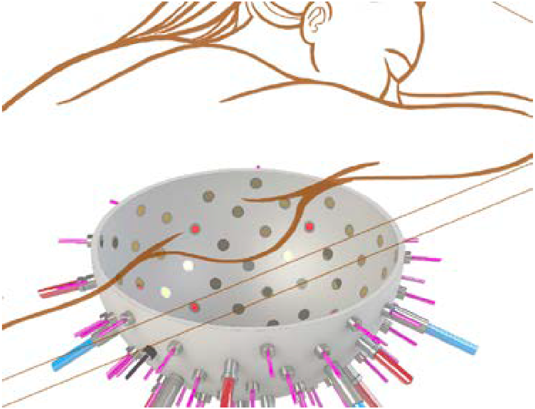
Centrum Wiskunde & Informatica

Time-Domain Full Waveform Inversion for High Resolution 3D Ultrasound Computed Tomography of the Breast

Felix Lucka

International Workshop on Medical Ultrasound Tomography
Detroit
15the October 2019

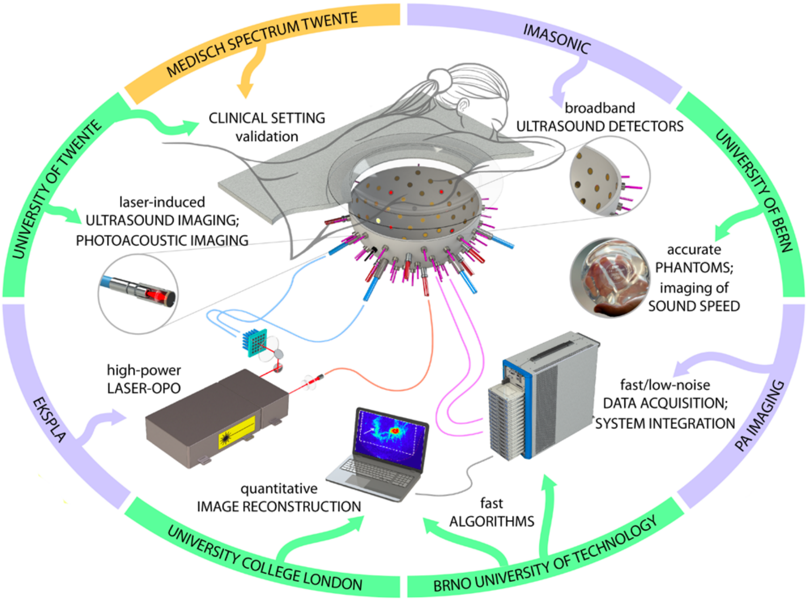
H2020 Project: Novel PAT+USCT Mammography Scanner



Diagnostic information from optical and acoustic properties

- 512 US transducers on rotatable half-sphere
- 40 optical fibers for photoacoustic excitation
- 40 inserts for laser-induced US (LIUS)

H2020 Project: Partners



USCT Reconstruction Approaches

$$(c(x)^{-2}\partial_t^2 - \Delta)p_i(x,t) = s_i(x,t), \qquad f_i = M_i p_i, \qquad i = 1, \dots, n_{src}$$

Travel time tomography (TTT): geometrical optics approximation.

- ✓ robust & computationally efficient
- **!** valid for high frequencies ($\rightarrow$ attenuation), low res, data size

Reverse time migration (RTM): forward wavefield correlated in time with backward wavefield (adjoint wave equation) via imaging condition.

- ✓ 2 wave simulations, better quality than TTT.
- **!** approximation, needs initial guess, quantitative errors

Full waveform inversion (FWI): fit full model to all data.

- ✓ high res from little data, include constraints, regularization
- **!** many wave simulations, non-convex PDE-constrained optimization.

time domain vs frequency domain methods

Time Domain Full Waveform Inversion

$$F(c)p_i := (c^{-2}\partial_t^2 - \Delta)p_i = s_i, \qquad f_i = M_i\, p_i, \quad i = 1, \ldots, n_{src}$$

$$\min_{c \in \mathcal{C}} \sum_i^{n_{src}} \mathcal{D}\left(f_i(c), f_i^\delta\right) \quad s.t. \quad f_i(c) = M_i F^{-1}(c) s_i$$

gradient for **first-order optimization** via **adjoint state method**:

$$\nabla_c \mathcal{D}\left(f(c), f^\delta\right) = 2 \int_0^T \frac{1}{c(x)^3} \left(\frac{\partial^2 p(x,t)}{\partial t^2}\right) q^*(x,t) \quad ,$$

where $(c^{-2}\partial_t^2 - \Delta)q^* = s^*$, $s^*(x,t)$ is time-reversed data discrepancy

$\rightarrow$ **two wave simulations for one gradient**

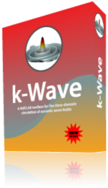
Acoustic Wave Propagation: Numerical Solution

- **Direct methods**, such as finite-difference, pseudospectral, finite/spectral element, discontinous Galerkin.
- **Integral wave equation methods**, e.g. boundary element
- **Asymptotic methods**, e.g., geometrical optics, Gaussian beams

Acoustic Wave Propagation: Numerical Solution

- Direct methods, such as finite-difference, **pseudospectral**, finite/spectral element, discontinous Galerkin.
- Integral wave equation methods, e.g. boundary element.
- Asymptotic methods, e.g., geometrical optics, Gaussian beams.

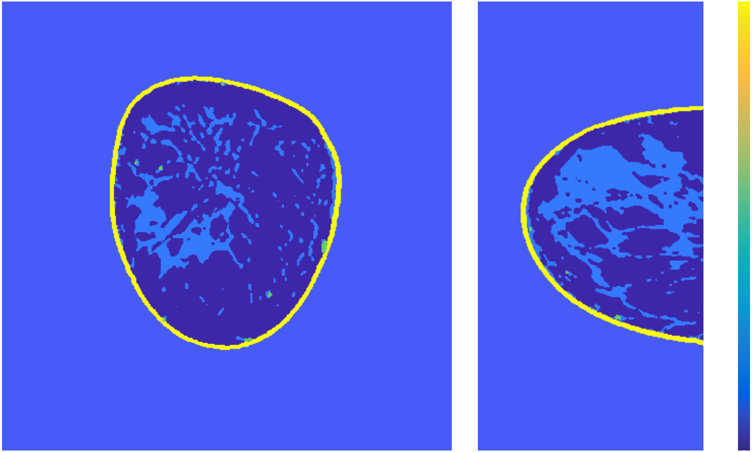
k-Wave: *k*-space pseudospectral method solving the underlying system of first order conservation laws.



- Compute spatial derivatives in Fourier space: 3D FFTs.
- Parallel/GPU computing leads to massive speed-ups.
- Modify finite temporal differences by *k*-space operator and use staggered grids for accuracy and robustness.
- Perfectly matched layer to simulate free-space propagation.

B. Treeby and B. Cox, 2010. k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields, *Journal of Biomedical Optics.*

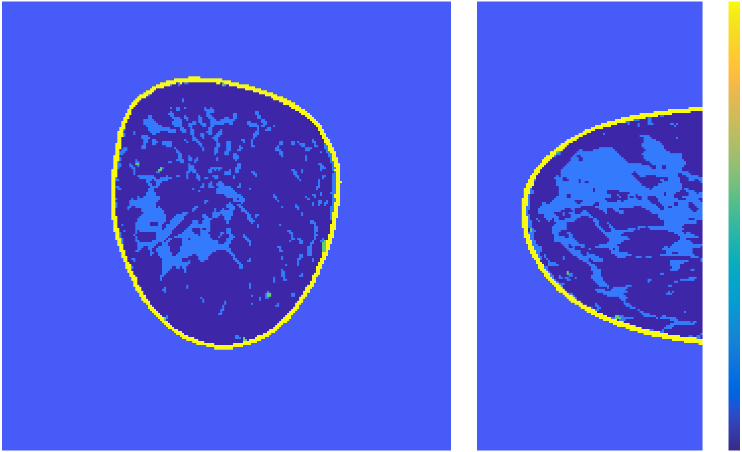
Numerical Phantoms



color range 1470 - 1650 m/s, resolution 0.5mm

Yang Lou et al. Generation of anatomically realistic numerical phantoms for photoacoustic and ultrasonic breast imaging, *JBO, 2017.*
https://anastasio.wustl.edu/downloadable-contents/oa-breast/

Numerical Phantoms

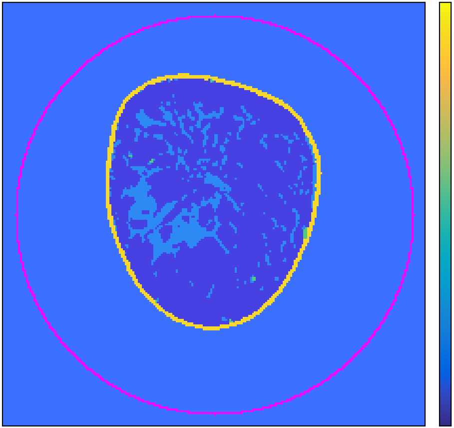


color range 1470 - 1650 m/s, resolution 1mm

Yang Lou et al. Generation of anatomically realistic numerical phantoms for photoacoustic and ultrasonic breast imaging, *JBO, 2017.*
https://anastasio.wustl.edu/downloadable-contents/oa-breast/

FWI Illustration in 2D

SOS ground truth c^{true}

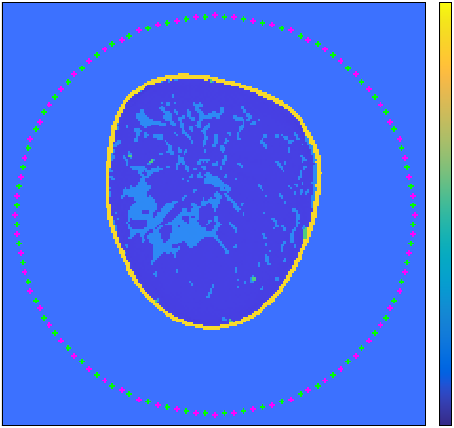


color range 1450 - 1670 m/s

- 1mm resolution
- 222^2 voxel
- 836 voxels on surface (pink)
- TTT would need 836^2 source-receiver combos for high res result

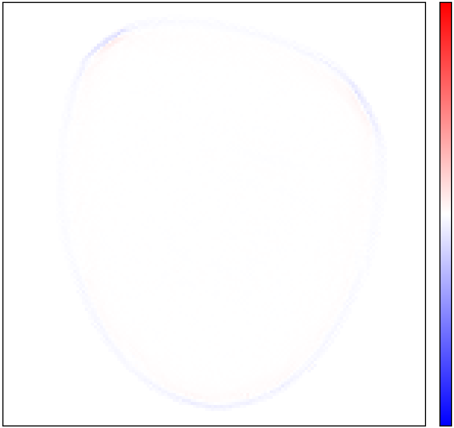
FWI Illustration in 2D: 64 Sensors, 64 Receivers

SOS reconstruction c^{rec}



color range 1450 - 1670 m/s

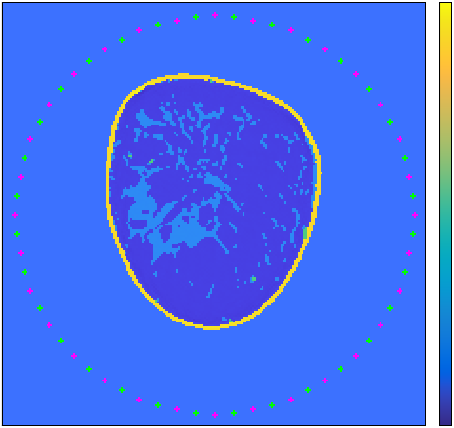
reconstruction error $c^{true} - c^{rec}$



color range -50 - 50 m/s

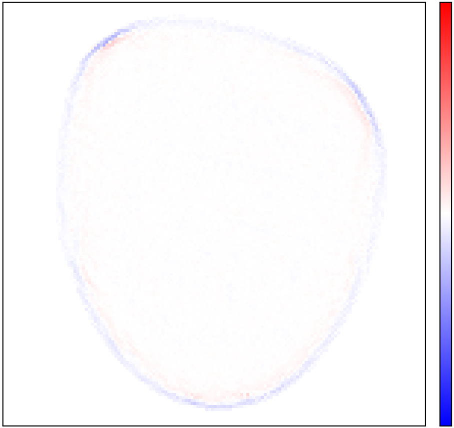
FWI Illustration in 2D: 32 Sensors, 32 Receivers

SOS reconstruction c^{rec}



color range 1450 - 1670 m/s

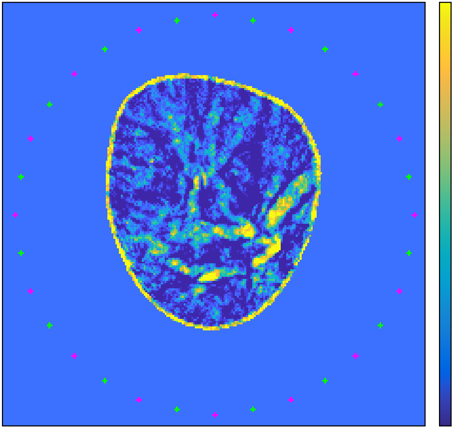
reconstruction error $c^{true} - c^{rec}$



color range -50 - 50 m/s

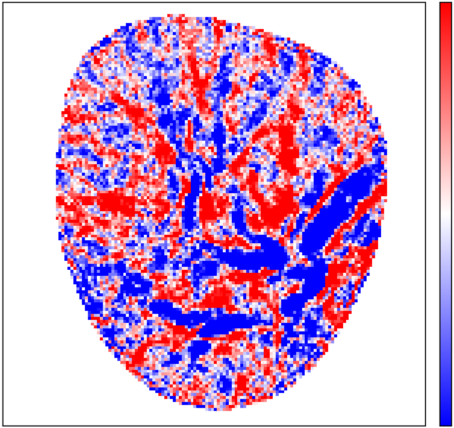
FWI Illustration in 2D: 16 Sensors, 16 Receivers

SOS reconstruction c^{rec}



color range 1450 - 1670 m/s

reconstruction error $c^{true} - c^{rec}$



color range -50 - 50 m/s

Challenges of High-Resolution FWI in 3D

$$\min_{c \in \mathcal{C}} \sum_{i}^{n_{src}} \mathcal{D}\left(f_i(c), f_i^\delta\right) \quad s.t. \quad f_i(c) = M_i F^{-1}(c) s_i$$

$$\nabla_c \mathcal{D}\left(f(c), f^\delta\right) = 2 \int_0^T \frac{1}{c(x)^3} \left(\frac{\partial^2 p(x,t)}{\partial t^2}\right) q^*(x,t)$$

PAMMOTH scanner example:

- 0.5mm res: comp grid $560 \times 560 \times 300$ voxel $= 94$M, ROI $= 7$M
- 1024 transducers, 4000 time samples (multiple sources);

Gradient computation:

- 1 wave sim: $\sim$30 min.
- ! **2 wave sim per source**, $n_{src} = 1024 \rightarrow 20$ days per gradient.

- ! **storage of forward field** in ROI: $\sim$ 200GB.

Challenges of High-Resolution FWI in 3D

$$\min_{c \in \mathcal{C}} \sum_{i}^{n_{src}} \mathcal{D}\left(f_i(c), f_i^{\delta}\right) \quad s.t. \quad f_i(c) = M_i F^{-1}(c) s_i$$

$$\nabla_c \mathcal{D}\left(f(c), f^{\delta}\right) = 2 \int_0^T \frac{1}{c(x)^3} \left(\frac{\partial^2 p(x,t)}{\partial t^2}\right) q^*(x,t)$$

PAMMOTH scanner example:

- 0.5mm res: comp grid $560 \times 560 \times 300$ voxel $= 94$M, ROI $= 7$M
- 1024 transducers, 4000 time samples (multiple sources);

Gradient computation:

- 1 wave sim: $\sim$30 min.
- ! **2 wave sim per source**, $n_{src} = 1024 \rightarrow 20$ days per gradient.
  stochastic gradient methods $\rightarrow$ 60 min per gradient
- ! **storage of forward field** in ROI: $\sim$ 200GB.
  time-reversal based gradient computation $\rightarrow 5 - 25$GB.

Stochastic Gradient Optimization

$$\mathcal{J} := n_{src}^{-1} \sum_{i}^{n_{src}} \mathcal{D}_i(c) := n_{src}^{-1} \sum_{i}^{n_{src}} \mathcal{D}\left(M_i F^{-1}(c) s_i, f_i^{\delta}\right)$$

approx $\nabla \mathcal{J}$ by $|\mathcal{S}|^{-1} \sum_{j \in \mathcal{S}} \nabla \mathcal{D}_j(c)$, $\mathcal{S} \subset \{1, \ldots, n_{src}\}$ predetermined.
$\rightarrow$ **incremental gradient, ordered sub-set methods**

Stochastic Gradient Optimization

$$\mathcal{J} := n_{src}^{-1} \sum_{i}^{n_{src}} \mathcal{D}_i(c) := n_{src}^{-1} \sum_{i}^{n_{src}} \mathcal{D}\left(M_i F^{-1}(c) s_i, f_i^{\delta}\right)$$

approx $\nabla \mathcal{J}$ by $|\mathcal{S}|^{-1} \sum_{j \in \mathcal{S}} \nabla \mathcal{D}_j(c)$, $\mathcal{S} \subset \{1, \ldots, n_{src}\}$ predetermined.
$\rightarrow$ **incremental gradient, ordered sub-set methods**

Instance of **finite sum minimization** similar to **training in machine learning**. Use **stochastic gradient descent (SGD)**:

- momentum, gradient/iterate averaging (SAV, SAGA), variance reduction (SVRG), choice of step size, mini-batch size
- include non-smooth regularizers (SPDHG, SADMM)
- quasi-Newton-type methods, e.g., **stochastic L-BFGS**

Bottou, Curtis, Nocedal. Optimization Methods for Large-Scale Machine Learning, *arXiv:1606.04838.*

Fabien-Ouellet, Gloaguen, Giroux, 2017. A stochastic L-BFGS approach for full-waveform inversion, *SEG.*

Gradient Estimates: Sub-Sampling vs Source Encoding

Computationally & stochastically efficient gradient estimator?

Gradient Estimates: Sub-Sampling vs Source Encoding

Computationally & stochastically efficient gradient estimator?

Source Encoding for linear PDE constraints:

$$\text{Let} \quad \hat{s} := \sum_{i}^{n_{srt}} w_i s_i, \quad \hat{f}^{\delta} := \sum_{i}^{n_{srt}} w_i f_i^{\delta}, \quad \text{with} \quad \mathbb{E}\left[w\right] = 0,\ \mathbb{C}\text{ov}[w] = I,$$

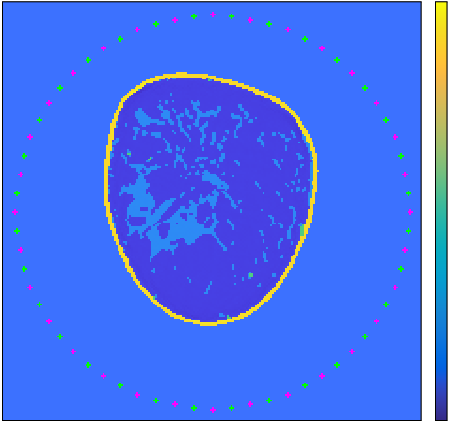
$$\text{then} \quad \mathbb{E}\left[\nabla \left\| MF^{-1}(c)\hat{s} - \hat{f}^{\delta} \right\|_2^2\right] = \nabla \sum_{i}^{n_{src}} \left\| MF^{-1}(c)s_i - f_i^{\delta} \right\|_2^2$$

- related to covariance trace estimators
- Rademacher distribution ($w_i = \pm 1$ with equal prob)
- add time-shifting for time-invariant PDEs $\rightarrow$ variance control
- can be turned into scanning strategy

Haber, Chung, Herrmann, 2012. An effective method for parameter estimation with PDE constraints with multiple right-hand sides, *SIAM J. Optim.*

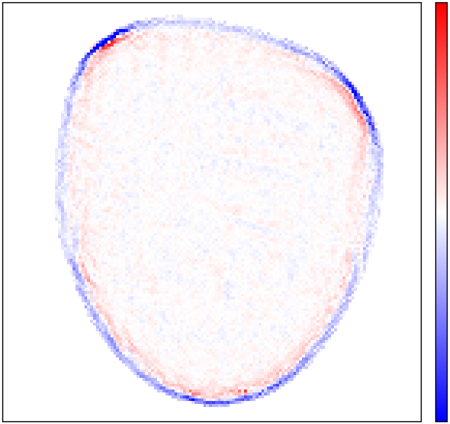
Stochastic Optimization Illustration

SOS reconstruction c^{rec} L-BFGS



color range 1450 to 1670 m/s

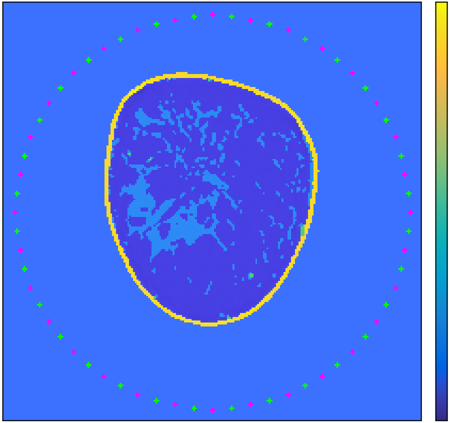
reconstruction error $c^{true} - c^{rec}$



color range -10 to 10 m/s

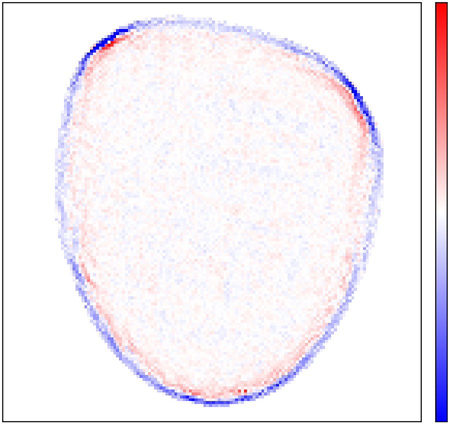
Stochastic Optimization Illustration

SOS reconstruction c^{rec} SL-BFGS



color range 1450 to 1670 m/s

reconstruction error $c^{true} - c^{rec}$



color range -10 to 10 m/s

Time-Reversal Gradient Computations

Avoid storage of forward fields!

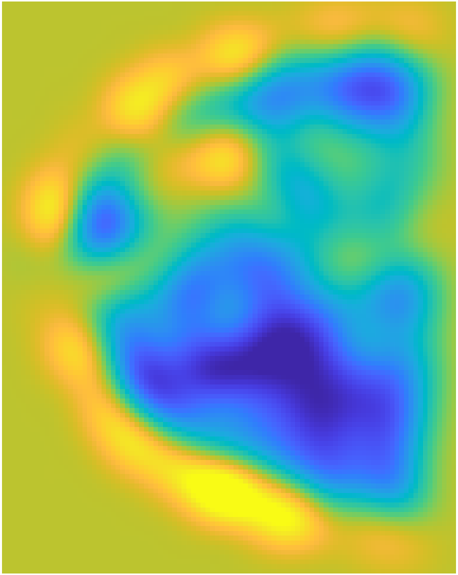
$$(c(x)^{-2}\partial_t^2 - \Delta)p(x,t) = s(x,t), \qquad in\ \mathbb{R}^d \times [0,T]$$
$$\nabla_c \mathcal{D} = 2\int_0^T \frac{1}{c(x)^3}\left(\frac{\partial^2 p(x,t)}{\partial t^2}\right) q^*(x,t)$$

Idea: ROI Ω, $\operatorname{supp}(s) \in \Omega^c \times [0,T]$. As $p(x,0) = p(x,T) = \partial_t p(x,0) = \partial_t p(x,T) = 0$ in Ω, $p(x,t)$ can be reconstructed from $p(x,t)$ on $\partial\Omega \times [0,T]$ by **time-reversal (TR)**.

- store fwd fields on ROI boundary during forward wave simulation
- interleave backward (adjoint) simulation with TR of boundary data
- 3 instead of 2 wave simulations (unless 2 GPUs used).
- code up efficiently
- multi-layer boundary increases accuarcy for pseudospectral method

Multigrid Schemes

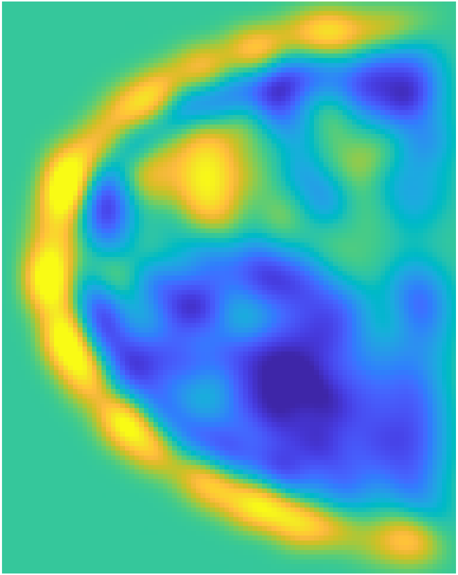
- easy due to regular grids in space and time
- coarsening by 2: (in principle) **speed up of 16**
- most basic multi-grid usage for now: initialization



level 6: upsampled from 5.66mm.

Multigrid Schemes

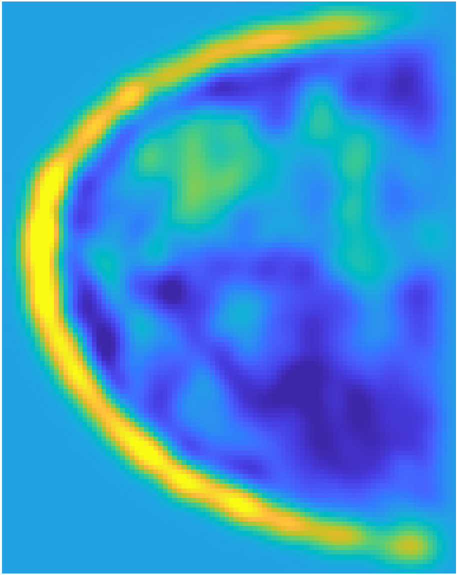
- easy due to regular grids in space and time
- coarsening by 2: (in principle) **speed up of 16**
- most basic multi-grid usage for now: initialization



level 5: upsampled from 4mm.

Multigrid Schemes

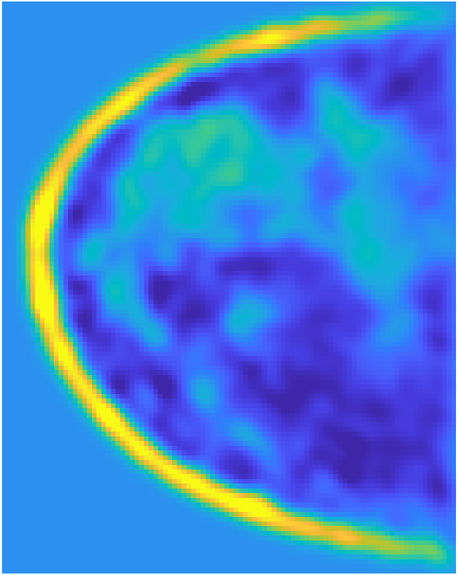
- easy due to regular grids in space and time
- coarsening by 2: (in principle) **speed up of 16**
- most basic multi-grid usage for now: initialization



level 4: upsampled from 2.83mm.

Multigrid Schemes

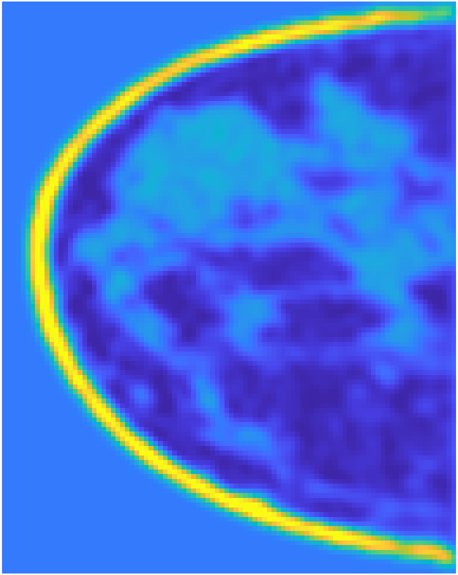
- easy due to regular grids in space and time
- coarsening by 2: (in principle) **speed up of 16**
- most basic multi-grid usage for now: initialization



level 3: upsampled from 2mm.

Multigrid Schemes

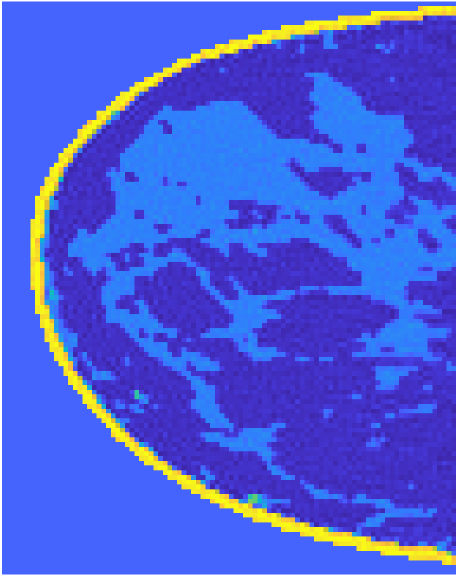
- easy due to regular grids in space and time
- coarsening by 2: (in principle) **speed up of 16**
- most basic multi-grid usage for now: initialization



level 2: upsampled from 1.41mm.

Multigrid Schemes

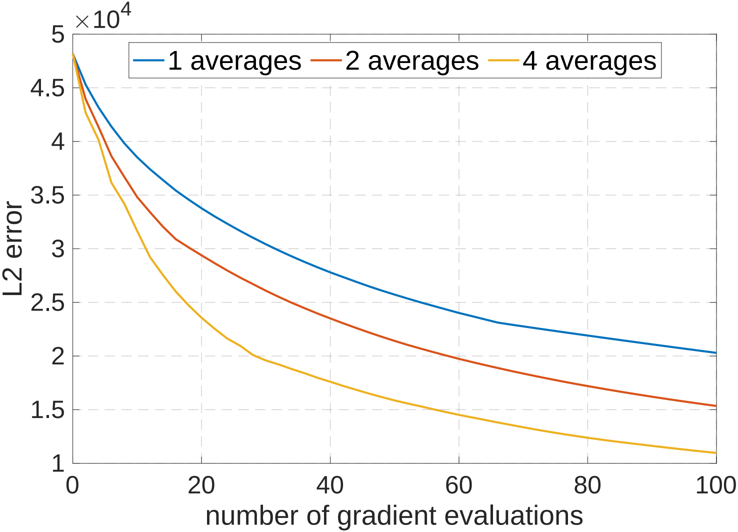
- easy due to regular grids in space and time
- coarsening by 2: (in principle) **speed up of 16**
- most basic multi-grid usage for now: initialization



level 1: resolution 1mm

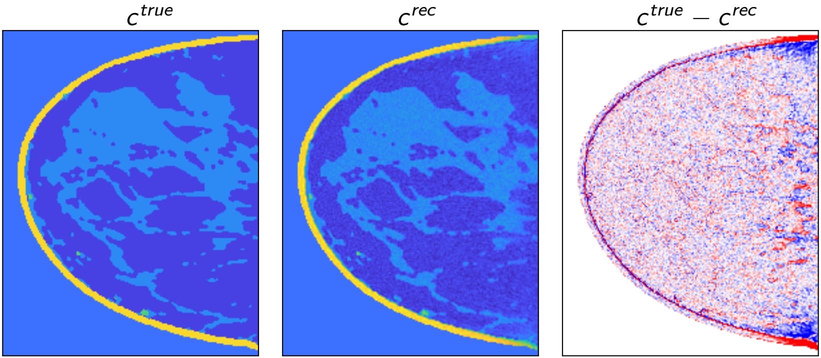
Utilizing Multiple GPUs

- average independent gradient estimates to reduce variance
- not be the best way to use multiple GPUs



Putting it all together

- 3D breast phantom at 0.5mm resolution, 1024 sources and receivers
- $442 \times 442 \times 232$ voxel, 3912 time steps
- multi-grid with 8 levels, coarsening factor $\sqrt{2}$.
- SL-BFGS (40 iter, 2d 4h on highest level), source encoding, 2 GPUs



color range 1450 to 1670 m/s

color range -50 to 50 m/s

Summary & Outlook

Summary:

- proof-of-concept studies of TD-FWI for high resolution 3D USCT
- stochastic L-BFGS with source encoding
- time reversal based gradient computation
- multi-grid initialization

Outlook:

- multi-GPU CUDA code
- realistic source/receiver modeling
- extension to acoustic attenuation, density, etc.
- **validation on experimental data!**

L, Pérez-Liva, Treeby, Cox, 2019. Time-Domain Full Waveform Inversion for High Resolution 3D Ultrasound Computed Tomography of the Breast, *in preparation.*

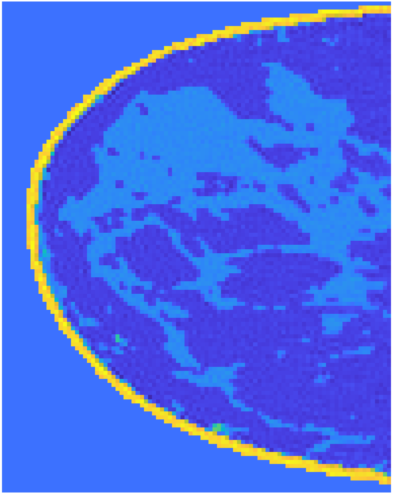
Thank you for your attention!

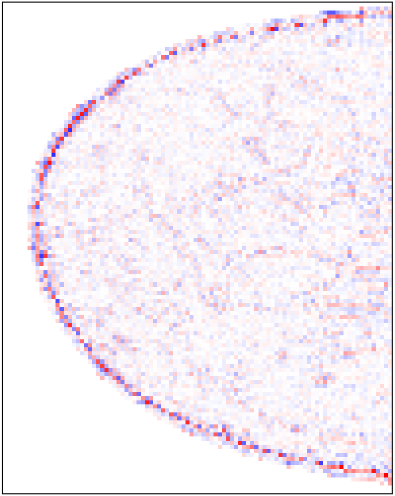
L, Pérez-Liva, Treeby, Cox, 2019. Time-Domain Full Waveform Inversion for High Resolution 3D Ultrasound Computed Tomography of the Breast, *in preparation.*

PHOTONICS PUBLIC PRIVATE PARTNERSHIP

Engineering and Physical Sciences Research Council

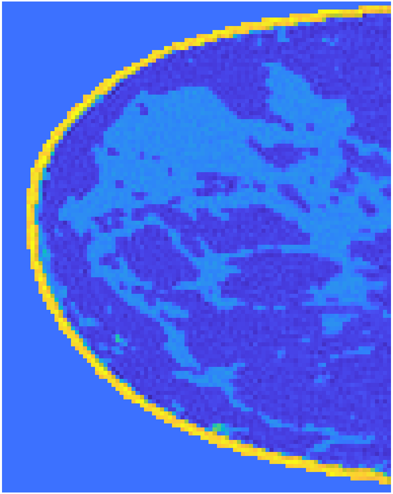
Influence of Noise

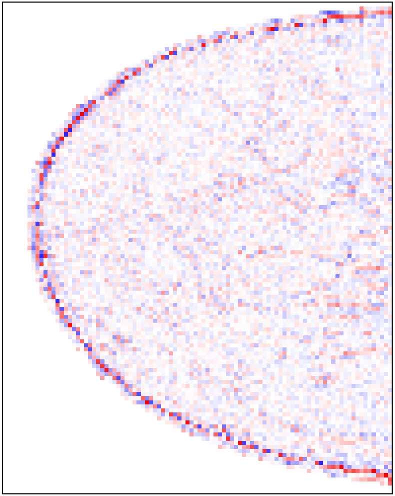




no noise

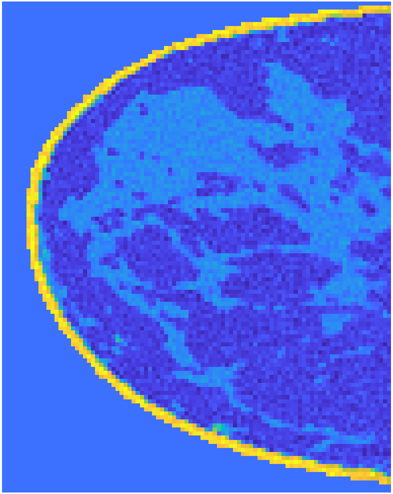
Influence of Noise

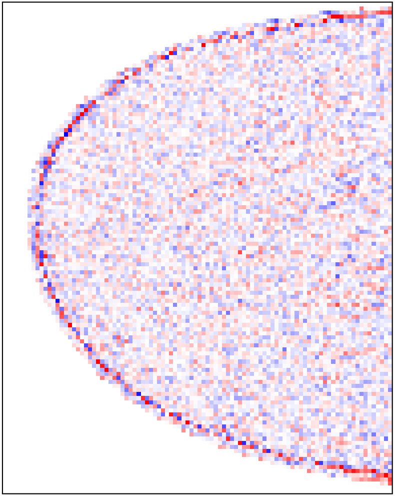




SNR 30dB

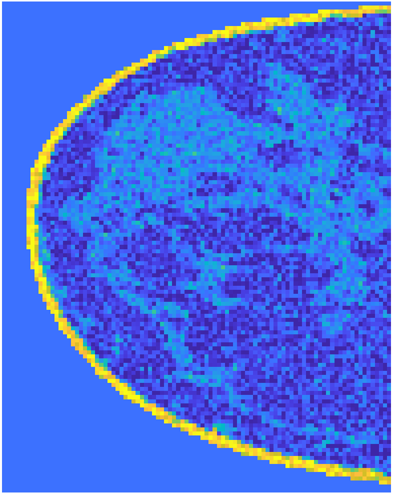
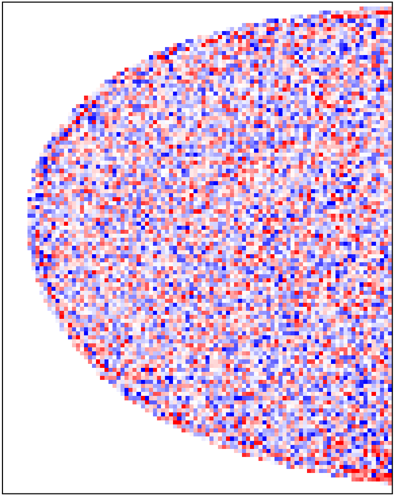
Influence of Noise





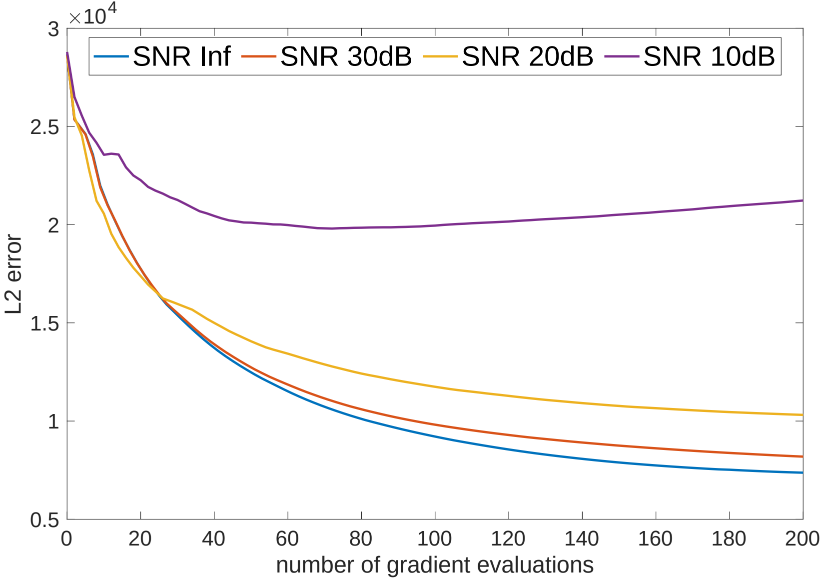
SNR 20dB

Influence of Noise

SNR 10dB

Influence of Noise



Mathematical Modelling (simplified)

Quantitative Photoacoustic Tomography (QPAT)

radiative transfer equation (RTE) + acoustic wave equation

$$(v \cdot \nabla + \mu_a(x) + \mu_s(x))\, \phi(x,v) = q(x,v) + \mu_s(x) \int \Theta(v,v')\phi(x,v')dv',$$

$$p^{PA}(x,t=0) = p_0 := \Gamma(x)\mu_a(x) \int \phi(x,v)dv, \qquad \partial_t p^{PA}(x,t=0) = 0$$

$$(c(x)^{-2}\partial_t^2 - \Delta)p^{PA}(x,t) = 0, \qquad f^{PA} = Mp^{PA}$$

Ultrasound Computed Tomography (USCT)

$$(c(x)^{-2}\partial_t^2 - \Delta)p^{US}(x,t) = s(x,t), \qquad f^{US} = Mp^{US}$$

Step-by-step inversion

1. $f^{US} \to c$: acoustic parameter identification from boundary data.
2. $f^{PA} \to p_0$: acoustic initial value problem with boundary data.
3. $p_0 \to \mu_a$: optical parameter identification from internal data.