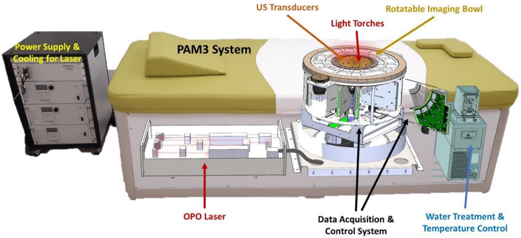
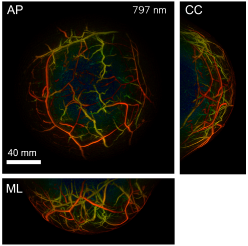
Centrum Wiskunde & Informatica

Photoacoustic & Ultrasonic Tomography for Breast Cancer Imaging

Felix Lucka (he/him/his) & the PAMMOTH team



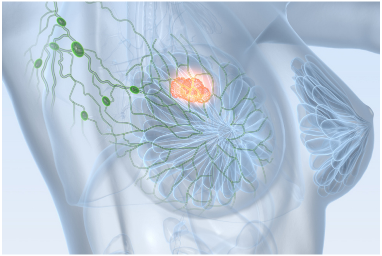


3rd IMA Conference on Inverse Problems from Theory to Application
3 May 2022

Motivation: Breast Cancer Imaging

Most common cause of cancer death in women worldwide.

- 25% of all cancer cases in women
- 15% of all cancer deaths in women

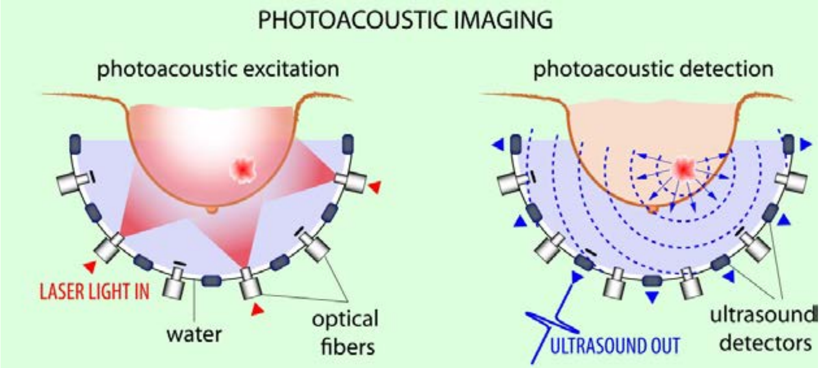


Despite advances in early detection and diagnosis:

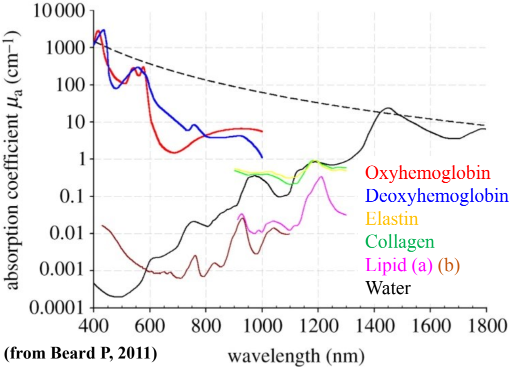
Urgent need for novel imaging techniques providing higher specificity, contrast and image resolution than X-ray mammography at lower costs than MRI.

Quantitative Photoacoustic Breast Imaging

- hybrid imaging: "light in, sound out"
- non-ionizing, near-infrared radiation
- quantitative images of optical properties
- novel diagnostic information



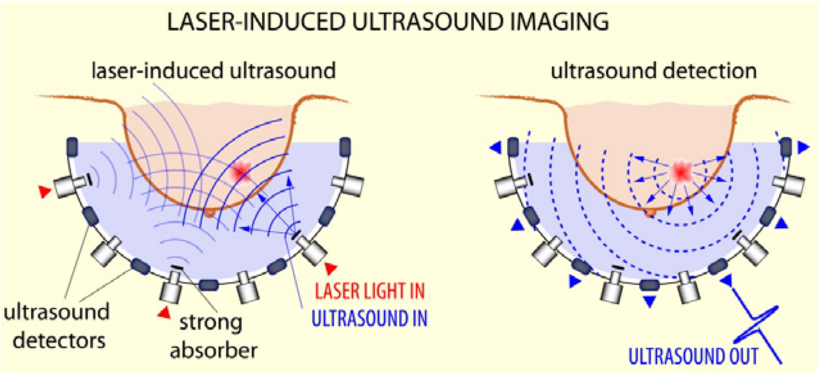
Photoacoustic Imaging: Spectral Properties



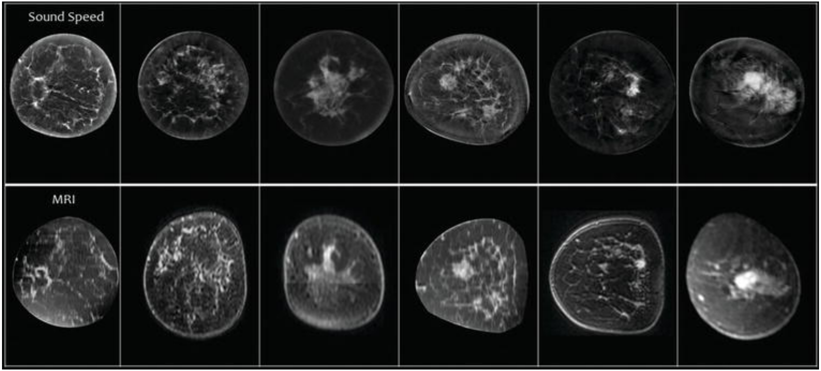
- different wavelengths allow quantitative spectroscopic examinations.
- gap between oxygenated and deoxygenated blood.

Quantitative Ultrasonic Breast Imaging

- "sound in, sound out"
- different from conventional US but as safe
- quantitative images of acoustic properties
- novel diagnostic information



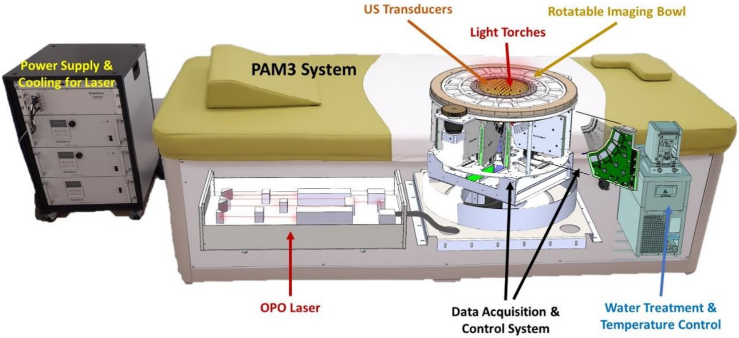
Speed of Sound vs MRI Images



Taken from:

Duric, Littrup, 2017. Breast Ultrasound Tomography, *IntechOpen.*

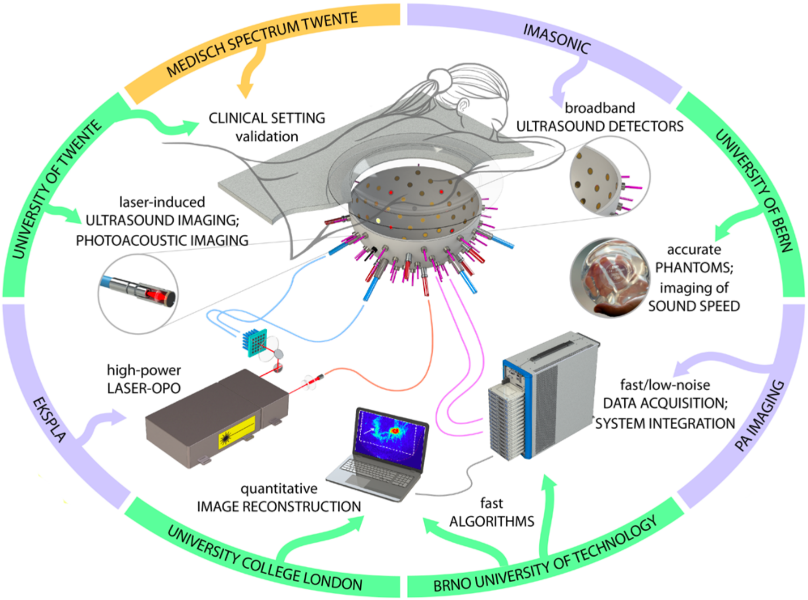
Photoacoustic and Ultrasonic Mammography Scanner



Aim: novel diagnostic information from high resolution maps of optical and acoustic properties

- 512 US transducers on rotatable half-sphere
- 40 optical fibers for photoacoustic excitation
- fully 3D, isotropic resolution ≤ 0.5mm

The PAMMOTH Team



Our Contributions

simulation studies for

- ultrasonic transducer specification
- light excitation design
- sensing pattern design
- measurement protocol design

reconstruction algorithm design:

- accuracy vs. computational time/resources/complexity
- scanner modelling
- assist high performance computing implementation

assist **phantom & calibration design**

process data, refine measurement procedures

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 Tomography (UST)

$$(c(x)^{-2}\partial_t^2 - \Delta)p_i^{US}(x,t) = s_i(x,t), \qquad f_i^{US} = M_i p_i^{US}, \qquad i = 1,\ldots,n_s$$

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.

Reconstruction of Initial Photoacoustic Pressure

$$(c(x)^{-2}\partial_t^2 - \Delta)p^{PA}(x,t) = 0, \quad p^{PA}(x, t=0) = p_0, \quad f^{PA} = Mp^{PA}$$

✓ linear inverse problem $f^{PA} = MAp_0$, moderately ill-posed.

✓ variational approach

$$\hat{p}_0 = \underset{p_0 \in \mathcal{C}}{\operatorname{argmin}} \left\| MAp_0 - f^{PA} \right\|_W^2 + \mathcal{R}(p_0)$$

✓ first order optimization with early stopping

! model acoustic properties, model discrepancies

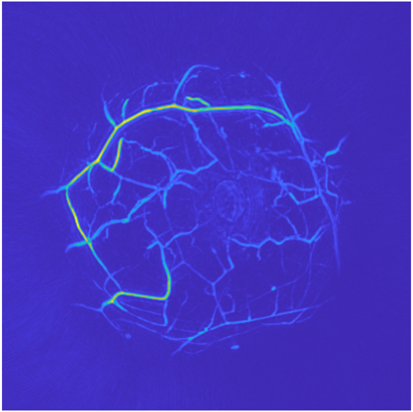
! model /calibrate sensors: sensitivity, impulse response, angular sensitivity, ...

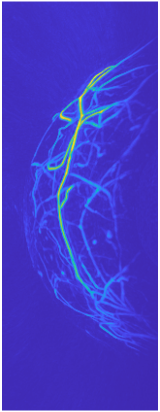
! parameter choices, image artifacts,...

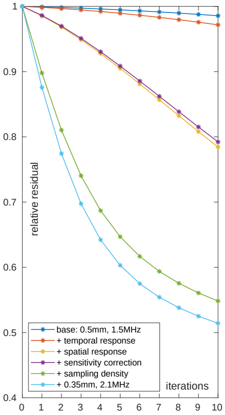
! numerical wave simulations: broadband up to $\geqslant 1.5$MHz, $\leqslant 0.5$mm

Healthy volunteer, 797nm, baseline

$$\hat{p}_0 = \text{argmin}_{p_0 \geqslant 0} \left\| MAp_0 - f^{PA} \right\|_W^2$$



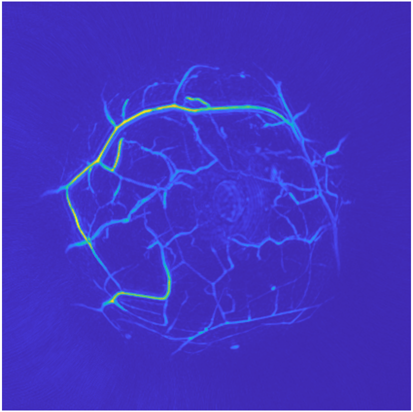


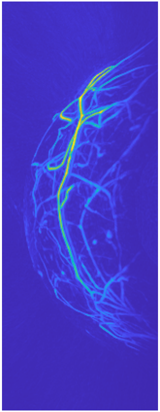


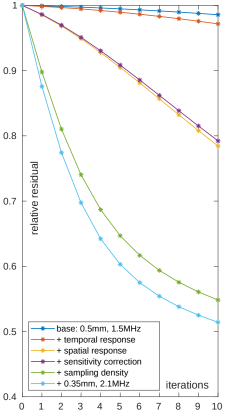
single wave simulation on 0.50mm: 2min 20; 10 iterations: 1h 20min

Healthy volunteer, 797nm, add impulse response

$$\hat{p}_0 = \mathrm{argmin}_{p_0 \geqslant 0} \left\| M A p_0 - f^{PA} \right\|_W^2$$



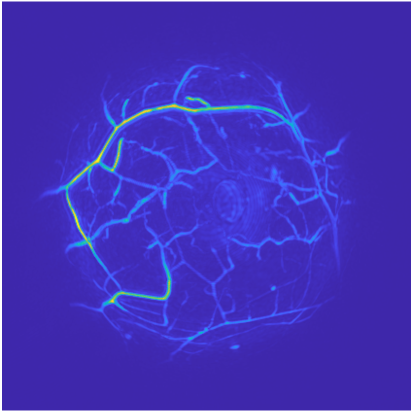


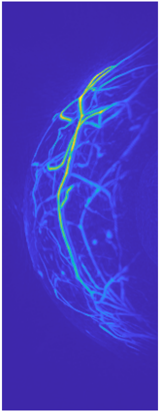


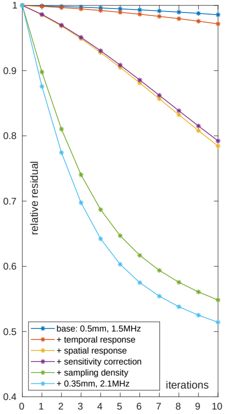
single wave simulation on 0.50mm: 2min 20; 10 iterations: 1h 20min

Healthy volunteer, 797nm, add spatial response

$$\hat{p}_0 = \operatorname{argmin}_{p_0 \geqslant 0} \left\| M A p_0 - f^{PA} \right\|_W^2$$



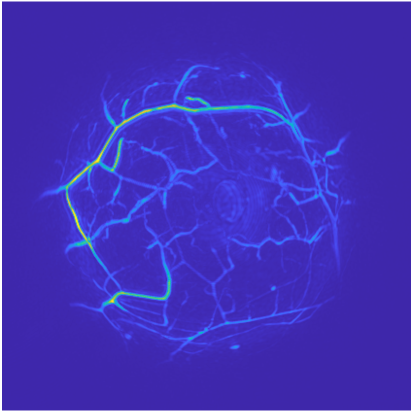
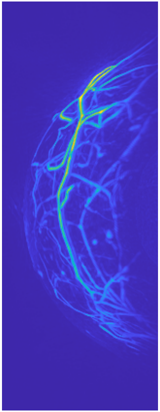
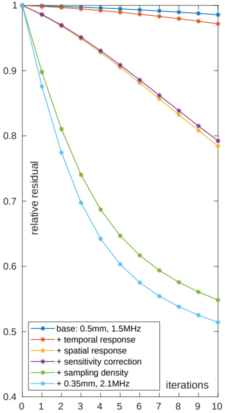




single wave simulation on 0.50mm: 2min 20; 10 iterations: 1h 20min

Healthy volunteer, 797nm, add sensitivity

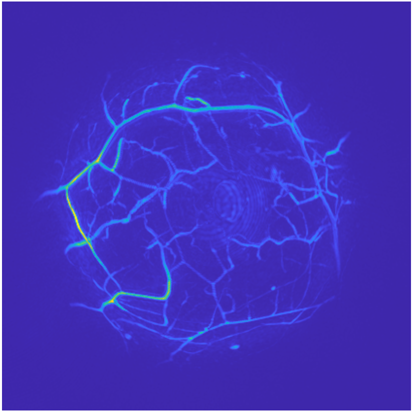
$$\hat{p}_0 = \operatorname{argmin}_{p_0 \geqslant 0} \left\| M A p_0 - f^{PA} \right\|_W^2$$

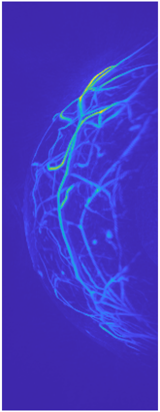




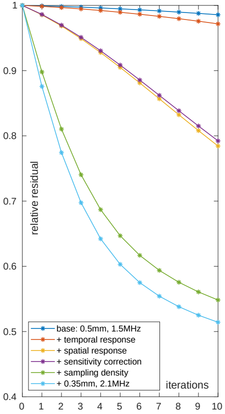
single wave simulation on 0.50mm: 2min 20; 10 iterations: 1h 20min

Healthy volunteer, 797nm, add sampling density

$$\hat{p}_0 = \text{argmin}_{p_0 \geqslant 0} \left\| MAp_0 - f^{PA} \right\|_W^2$$



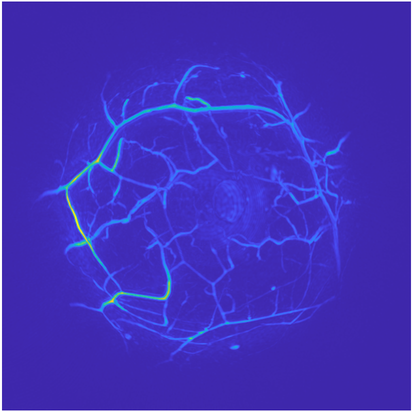


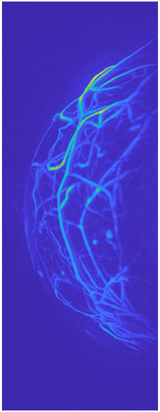


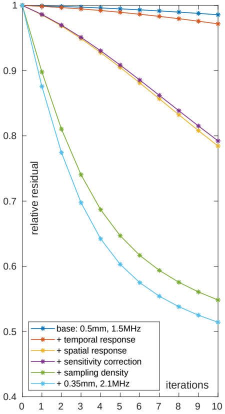
single wave simulation on 0.50mm: 2min 20; 10 iterations: 1h 20min

Healthy volunteer, 797nm, increase resolution/bandwidth

$$\hat{p}_0 = \text{argmin}_{p_0 \geqslant 0} \left\| MAp_0 - f^{PA} \right\|_W^2$$







single wave simulation on 0.35mm: 12min 10; 10 iterations: 4h 8min

UST Reconstruction Approaches

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

Travel time tomography: geometrical optics approximation.

- ✓ robust & computationally efficient
- ! valid for high frequencies (attenuation!), low res, lots of data

Javaherian, L, Cox, 2020. Refraction-corrected ray-based inversion for three-dimensional ultrasound tomography of the breast, *Inverse Problems.*

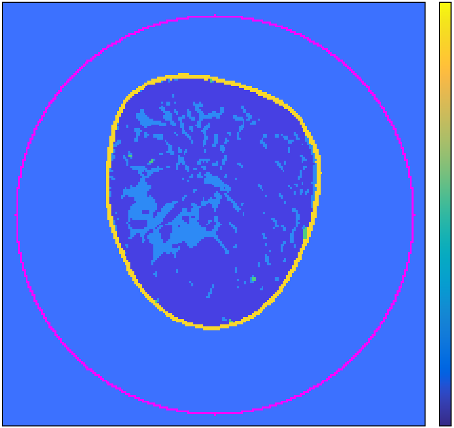
Full waveform inversion (FWI): fit full wave model to all data.

- ✓ high res from little data, transducer modelling, constraints
- ! many wave simulations, complex numerical optimization
- • low TRL but already used in 2D systems

Pérez-Liva, Herraiz, Udías, Miller, Cox, Treeby 2017. Time domain reconstruction of sound speed and attenuation in ultrasound computed tomography using full wave inversion, *JASA.*

FWI Illustration in 2D

SOS ground truth c^{true}

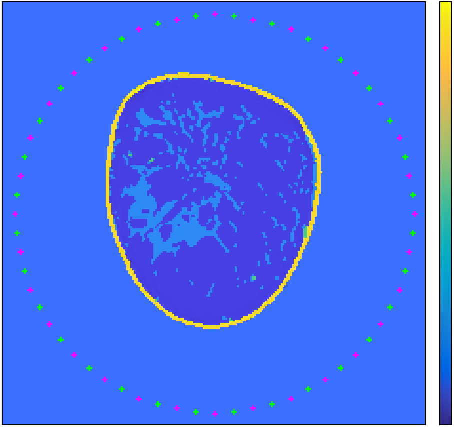


color range 1450 - 1670 m/s

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

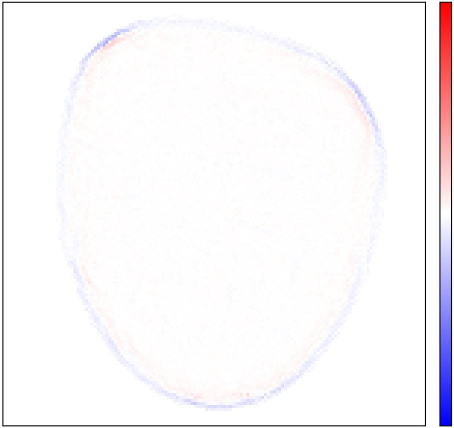
FWI Illustration in 2D: 32 Sensors, 32 Receivers

SOS reconstruction c^{rec}



color range 1450 - 1670 m/s

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



color range -50 - 50 m/s

3D Time Domain FWI for Breast UST

$$\min_{c \in \mathcal{C}} \sum_{i}^{n_{src}} \mathcal{D}\left(M_i A^{-1}(c) s_i, f_i^{\delta}\right)$$

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

Challenges and solutions for 3D:

- **!** 2 x n_{src} wave simulations per gradient
- **!** computationally & stochastically efficient gradient estimator
- **!** memory requirements of gradient computation
- **!** slow convergence and local minima
- **!** computational resources

3D Time Domain FWI for Breast UST

$$\min_{c\in\mathcal{C}} \sum_{i}^{n_{src}} \mathcal{D}\left(M_i A^{-1}(c) s_i, f_i^{\delta}\right)$$

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

Challenges and solutions for 3D:

! $2 \times n_{src}$ wave simulations per gradient
$\longrightarrow$ **stochastic quasi-newton optimization (SL-BFGS)**

! computationally & stochastically efficient gradient estimator

! memory requirements of gradient computation

! slow convergence and local minima

! computational resources

3D Time Domain FWI for Breast UST

$$\min_{c \in \mathcal{C}} \sum_{i}^{n_{src}} \mathcal{D}\left(M_i A^{-1}(c) s_i, f_i^{\delta}\right)$$

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

Challenges and solutions for 3D:

- ! 2 x n_{src} wave simulations per gradient
  $\longrightarrow$ **stochastic quasi-newton optimization (SL-BFGS)**
- ! computationally & stochastically efficient gradient estimator
  $\longrightarrow$ **source encoding for time-invariant systems**
- ! memory requirements of gradient computation
- ! slow convergence and local minima
- ! computational resources

3D Time Domain FWI for Breast UST

$$\min_{c \in \mathcal{C}} \sum_{i}^{n_{src}} \mathcal{D}\left(M_i A^{-1}(c) s_i, f_i^{\delta}\right)$$

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

Challenges and solutions for 3D:

- ! 2 x n_{src} wave simulations per gradient
  $\longrightarrow$ **stochastic quasi-newton optimization (SL-BFGS)**
- ! computationally & stochastically efficient gradient estimator
  $\longrightarrow$ **source encoding for time-invariant systems**
- ! memory requirements of gradient computation
  $\longrightarrow$ **time-reversal based gradient computation**
- ! slow convergence and local minima
- ! computational resources

3D Time Domain FWI for Breast UST

$$\min_{c \in \mathcal{C}} \sum_{i}^{n_{src}} \mathcal{D}\left(M_i A^{-1}(c) s_i, f_i^{\delta}\right)$$

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

Challenges and solutions for 3D:

- ! 2 x n_{src} wave simulations per gradient
  $\longrightarrow$ **stochastic quasi-newton optimization (SL-BFGS)**
- ! computationally & stochastically efficient gradient estimator
  $\longrightarrow$ **source encoding for time-invariant systems**
- ! memory requirements of gradient computation
  $\longrightarrow$ **time-reversal based gradient computation**
- ! slow convergence and local minima
  $\longrightarrow$ **coarse-to-fine multigrid schemes**
- ! computational resources

3D Time Domain FWI for Breast UST

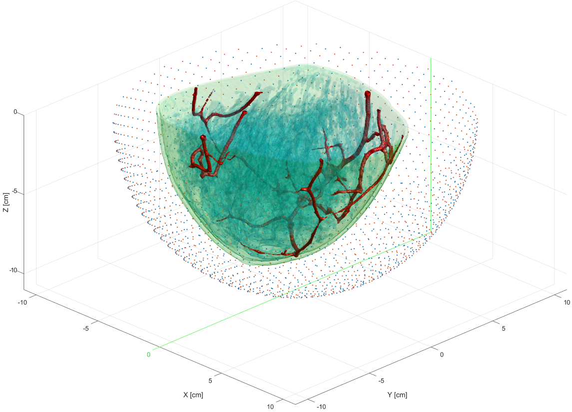
$$\min_{c \in \mathcal{C}} \sum_{i}^{n_{src}} \mathcal{D}\left(M_i A^{-1}(c) s_i, f_i^{\delta}\right)$$

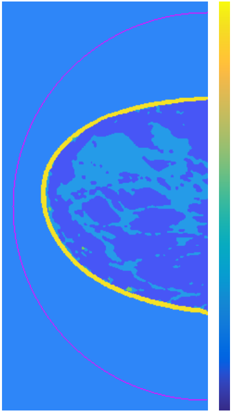
$$\nabla_c \mathcal{D}\left(f(c), f^{\delta}\right) = 2 \int_0^T \frac{1}{c(x)^3} \left(\frac{\partial^2 p(x,t)}{\partial t^2}\right) q^*(x,t)$$

Challenges and solutions for 3D:

- ! 2 x n_{src} wave simulations per gradient
  $\longrightarrow$ **stochastic quasi-newton optimization (SL-BFGS)**
- ! computationally & stochastically efficient gradient estimator
  $\longrightarrow$ **source encoding for time-invariant systems**
- ! memory requirements of gradient computation
  $\longrightarrow$ **time-reversal based gradient computation**
- ! slow convergence and local minima
  $\longrightarrow$ **coarse-to-fine multigrid schemes**
- ! computational resources
  $\longrightarrow$ **runs on single GPU, can utilize multiple GPUs**

3D FWI: Setup

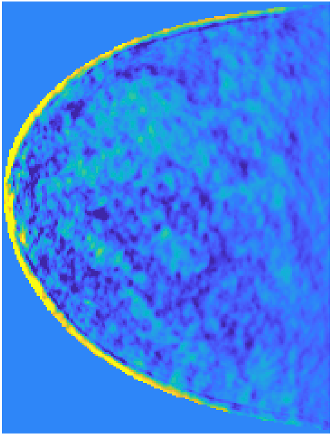




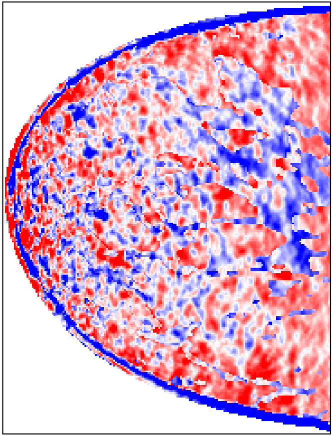
- color range 1435-1665 m/s
- 3D breast phantom at 0.5mm resolution, 1024 sources and receivers
- 442 $\times$ 442 $\times$ 222 voxel, 3912 time steps

Yang Lou et al. Generation of anatomically realistic numerical phantoms for photoacoustic and ultrasonic breast imaging, *JBO, 2017.*

Starting point in 24h on desktop with single GPU



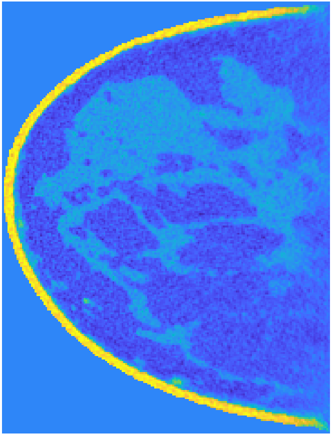
color range 1435 to 1665 m/s



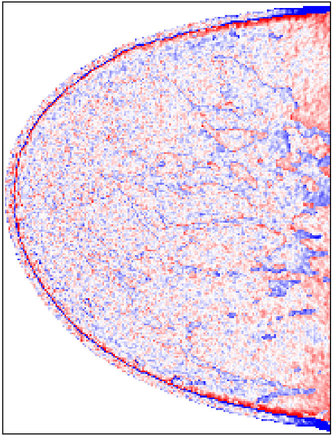
color range -50 to +50 m/s

- single grid
- SGD
- normal single source gradient estimator

3D FWI in 24h on desktop with single GPU



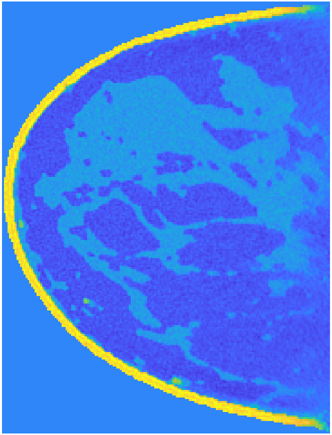
color range 1435 to 1665 m/s



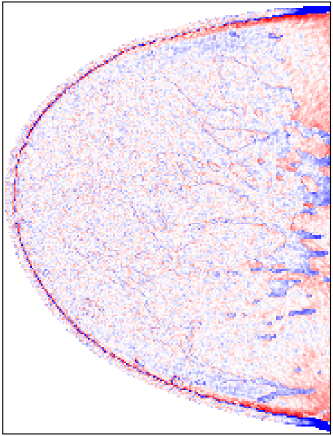
color range -50 to +50 m/s

- multi-grid with 3 level, coarsening factor 2
- SL-BFGS, slowness transform, prog. iter averaging
- time-reversal based source encoding gradient estimator

3D FWI in 24h on cluster with 4 GPU



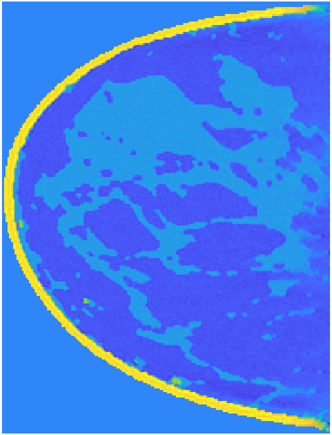
color range 1435 to 1665 m/s



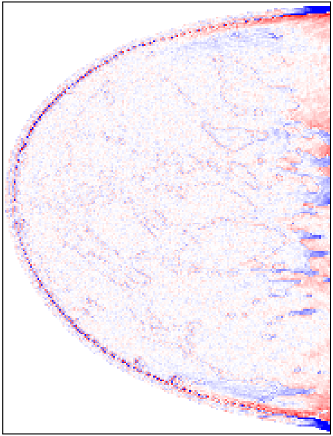
color range -50 to +50 m/s

- multi-grid with 3 level, coarsening factor 2
- SL-BFGS, slowness transform, prog. iter averaging
- time-reversal based source encoding gradient estimator

3D FWI in 24h on cluster with 16 GPU



color range 1435 to 1665 m/s



color range -50 to +50 m/s

- multi-grid with 3 level, coarsening factor 2
- SL-BFGS, slowness transform, prog. iter averaging
- time-reversal based source encoding gradient estimator

FWI for Experimental Data: Where Are We?

✓ data from phantom objects, volunteers & patients

✓ ray-based SOS reconstructions

✓ photoacoustic reconstructions $\rightarrow$ data pre-processing, scanner & transducer modelling, wave simulations

✓ modeling of US protocol, data read-in & pre-processing

! model calibration

! FWI of phantom objects, quantitative evaluation

! FWI of volunteer data

! clinical evaluation

Summary & Outlook

- need for novel breast imaging techniques
- photoacoustic (PAT) and ultrasound tomography (UST) give access to high-quality images of optical and acoustic tissue properties
- combined PAT+UST scanner designed & build
- evaluation on data from phantoms, volunteers & patients
- PAT reconstructions work well
- proof-of-concept studies of TD-FWI for high resolution 3D UST
- realization of FWI for experimental data

Thank you for your attention!

L, Pérez-Liva, Treeby, Cox, 2021. High Resolution 3D Ultrasonic Breast Imaging by Time-Domain Full Waveform Inversion, *Inverse Problems 38(2).*

PHOTONICS[21]
PHOTONICS PUBLIC PRIVATE PARTNERSHIP
EPSRC
Engineering and Physical Sciences Research Council
NWO
NVIDIA

Challenges of High-Resolution FWI in 3D

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

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

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

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

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**

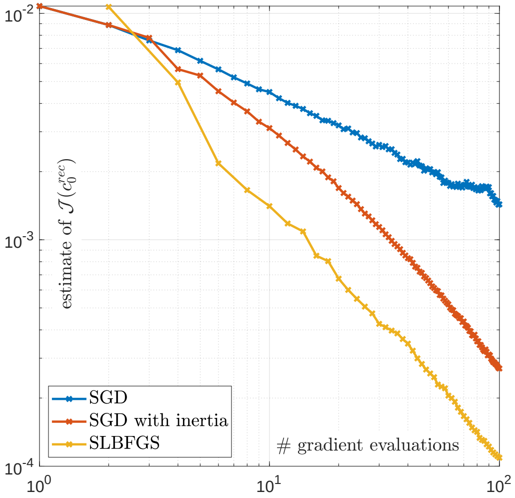
Instance of **finite sum minimization** similar to **training in machine learning**. Use **stochastic gradient descent (SGD)**:

- 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.*

Stochastic Gradient Optimization



Gradient Estimates: Sub-Sampling vs Source Encoding

Computationally & stochastically efficient gradient estimator?

Source Encoding for linear PDE constraints:

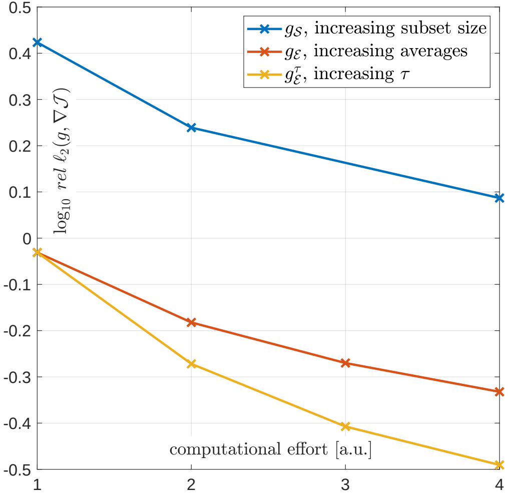
$$\text{Let} \quad \hat{s} := \sum_{i}^{n_{srt}} w_i s_i, \quad \hat{f}^{\delta} := \sum_{i}^{n_{srt}} w_i f_i^{\delta}, \quad \text{with} \quad \mathbb{E}\left[w\right] = 0, \ \mathbb{C}\text{ov}[w] = I,$$

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

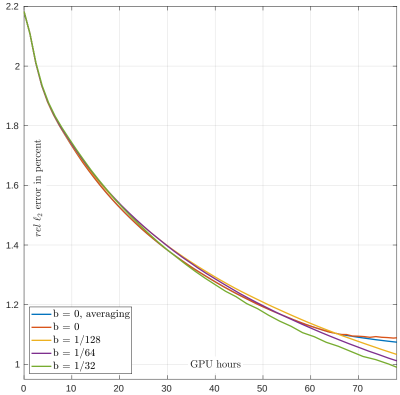
- related to covariance trace estimators
- Rademacher distribution ($w_i = \pm 1$ with equal prob)
- add time-shifting for time-invariant PDEs $\rightarrow$ variance control
- can be turned into scanning strategy

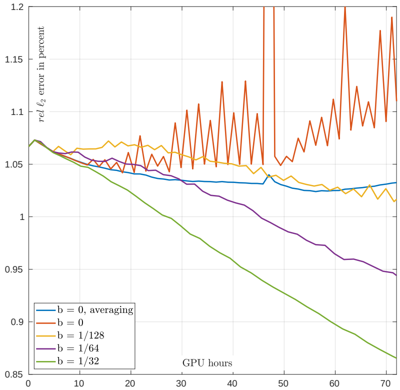
Haber, Chung, Herrmann, 2012. An effective method for parameter estimation with PDE constraints with multiple right-hand sides, *SIAM J. Optim.*

Stochastic Gradient Estimates



Delayed Source Encoding





Time-Reversal Gradient Computations

Avoid storage of forward fields!

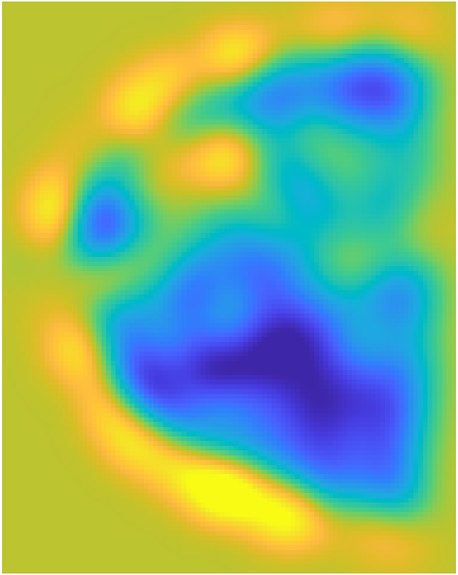
$$(c(x)^{-2}\partial_t^2 - \Delta)p(x,t) = s(x,t), \qquad \textit{in } \mathbb{R}^d \times [0,T]$$
$$\nabla_c \mathcal{D} = 2\int_0^T \frac{1}{c(x)^3}\left(\frac{\partial^2 p(x,t)}{\partial t^2}\right) q^*(x,t)$$

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

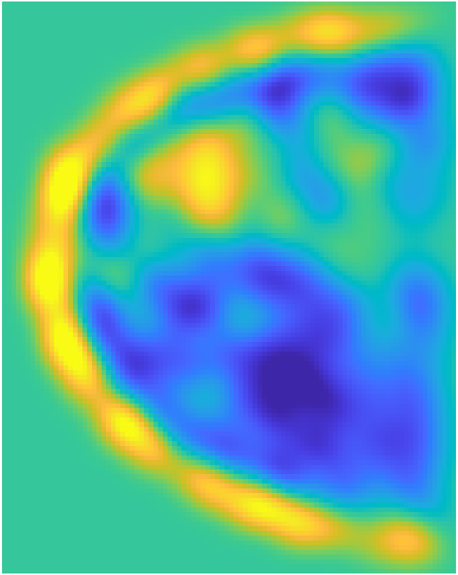
- easy due to regular grids in space and time
- coarsening by 2: (in principle) **speed up of 16**
- most basic multi-grid usage for now: initialization



level 6: upsampled from 5.66mm.

Multigrid Schemes

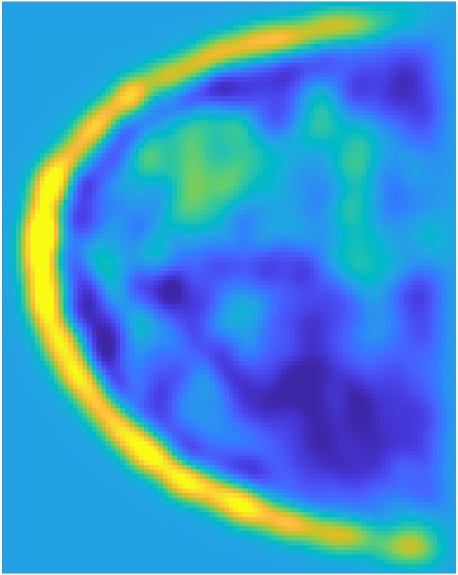
- easy due to regular grids in space and time
- coarsening by 2: (in principle) **speed up of 16**
- most basic multi-grid usage for now: initialization



level 5: upsampled from 4mm.

Multigrid Schemes

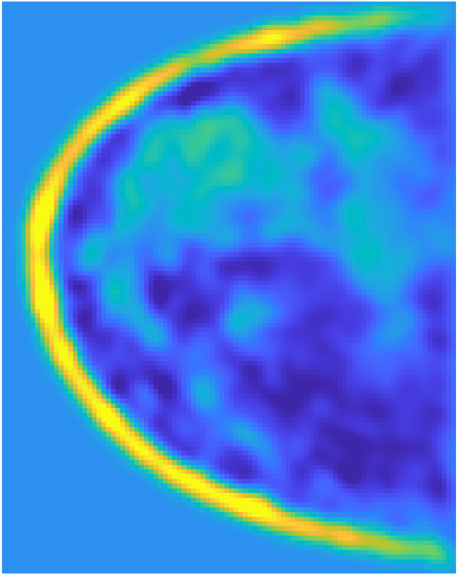
- easy due to regular grids in space and time
- coarsening by 2: (in principle) **speed up of 16**
- most basic multi-grid usage for now: initialization



level 4: upsampled from 2.83mm.

Multigrid Schemes

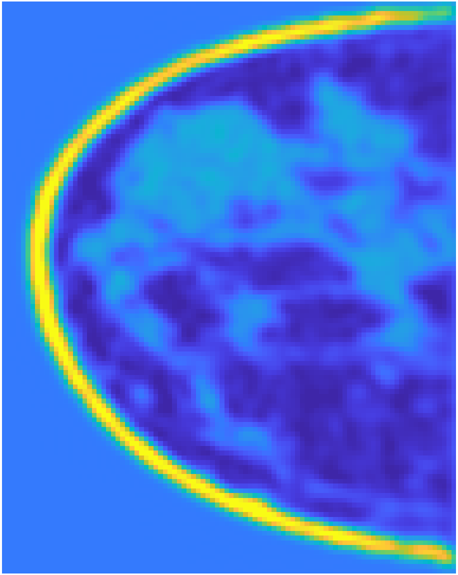
- easy due to regular grids in space and time
- coarsening by 2: (in principle) **speed up of 16**
- most basic multi-grid usage for now: initialization



level 3: upsampled from 2mm.

Multigrid Schemes

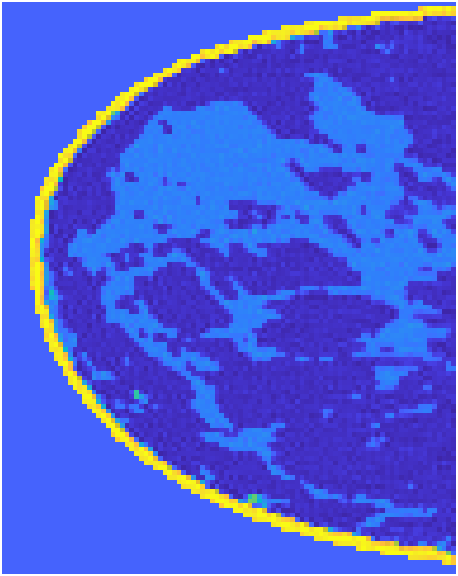
- easy due to regular grids in space and time
- coarsening by 2: (in principle) **speed up of 16**
- most basic multi-grid usage for now: initialization



level 2: upsampled from 1.41mm.

Multigrid Schemes

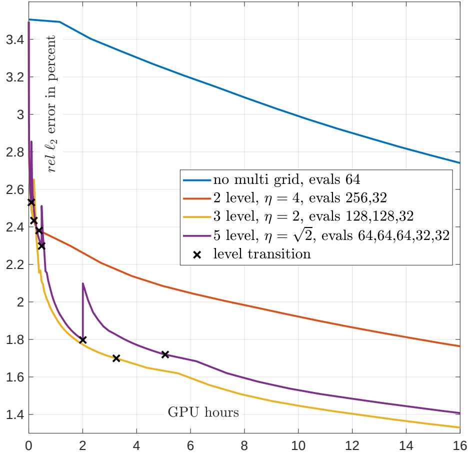
- easy due to regular grids in space and time
- coarsening by 2: (in principle) **speed up of 16**
- most basic multi-grid usage for now: initialization



level 1: resolution 1mm

Multigrid Schemes

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



Utilizing Multiple GPUs

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