

PAPER • OPEN ACCESS

## Accelerated high-resolution photoacoustic tomography via compressed sensing

To cite this article: Simon Arridge *et al* 2016 *Phys. Med. Biol.* **61** 8908

View the [article online](#) for updates and enhancements.

### You may also like

- [Enhancement of photoacoustic tomography in the tissue with speed-of-sound variance using ultrasound computed tomography](#)  
Ren-Xiang Cheng, , Chao Tao et al.
- [Photo-acoustic tomographic image reconstruction from reduced data using physically inspired regularization](#)  
N.A. Rejesh, S.K. Kalva, M. Pramanik et al.
- [Another decade of photoacoustic imaging](#)  
Dhiman Das, Arunima Sharma, Praveenbalaji Rajendran et al.

**JOIN US | ESTRO 2024**

**In-Booth Talks, Demos,  
& Lunch Symposium**

[Browse talk schedule >](#)



# Accelerated high-resolution photoacoustic tomography via compressed sensing

Simon Arridge<sup>1</sup>, Paul Beard<sup>2</sup>, Marta Betcke<sup>1</sup>, Ben Cox<sup>2</sup>,  
Nam Huynh<sup>2</sup>, Felix Lucka<sup>1</sup>, Olumide Ogunlade<sup>2</sup>  
and Edward Zhang<sup>2</sup>

<sup>1</sup> Department of Computer Science, University College London, WC1E 6BT London, UK

<sup>2</sup> Department of Medical Physics and Bioengineering, University College London, WC1E 6BT London, UK

E-mail: [f.lucka@ucl.ac.uk](mailto:f.lucka@ucl.ac.uk)

Received 30 April 2016, revised 28 September 2016

Accepted for publication 4 October 2016

Published 2 December 2016



## Abstract

Current 3D photoacoustic tomography (PAT) systems offer either high image quality or high frame rates but are not able to deliver high spatial and temporal resolution simultaneously, which limits their ability to image dynamic processes in living tissue (4D PAT). A particular example is the planar Fabry–Pérot (FP) photoacoustic scanner, which yields high-resolution 3D images but takes several minutes to sequentially map the incident photoacoustic field on the 2D sensor plane, point-by-point. However, as the spatio-temporal complexity of many absorbing tissue structures is rather low, the data recorded in such a conventional, regularly sampled fashion is often highly redundant. We demonstrate that combining model-based, variational image reconstruction methods using spatial sparsity constraints with the development of novel PAT acquisition systems capable of sub-sampling the acoustic wave field can dramatically increase the acquisition speed while maintaining a good spatial resolution: first, we describe and model two general spatial sub-sampling schemes. Then, we discuss how to implement them using the FP interferometer and demonstrate the potential of these novel compressed sensing PAT devices through simulated data from a realistic numerical phantom and through measured data from a dynamic experimental phantom as well as from *in vivo* experiments. Our results show that images with good spatial resolution and contrast can be obtained from highly sub-sampled PAT data if variational image reconstruction techniques that describe the



Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

tissues structures with suitable sparsity-constraints are used. In particular, we examine the use of total variation (TV) regularization enhanced by Bregman iterations. These novel reconstruction strategies offer new opportunities to dramatically increase the acquisition speed of photoacoustic scanners that employ point-by-point sequential scanning as well as reducing the channel count of parallelized schemes that use detector arrays.

Keywords: photoacoustic tomography, compressed sensing, variational image reconstruction, sparsity, Bregman iteration, Fabry–Pérot scanner

 Online supplementary data available from [stacks.iop.org/PMB/61/8908/mmedia](http://stacks.iop.org/PMB/61/8908/mmedia)

(Some figures may appear in colour only in the online journal)

## 1. Introduction

*Photoacoustic tomography (PAT)* is an emerging biomedical, ‘imaging from coupled physics’-technique (Arridge and Scherzer 2012) based on laser-generated ultrasound (US). It allows the rich contrast afforded by optical absorption to be imaged with the high spatial resolution of ultrasound. Furthermore, the wavelength dependence of the optical absorption can, in principle, be utilized to provide spectroscopic (chemical) information on the absorbing molecules (chromophores). While PAT’s potential for an increasing variety of clinical applications is currently being explored (Zackrisson *et al* 2014, Taruttis and Ntziachristos 2015), it is already widely used in preclinical studies to examine small animal anatomy, physiology and pathology (Xia and Wang 2014, Yao and Wang 2014, Jathoul *et al* 2015). For further applications and references we refer to the reviews (Wang 2009, Beard 2011, Nie and Chen 2014).

To obtain high quality three-dimensional (3D) photoacoustic (PA) images with a spatial resolution around one hundred  $\mu\text{m}$ , acoustic waves with a frequency content typically in the range of tens of MHz need to be sampled over cm scale apertures. Hence, satisfying the spatial Nyquist criterion necessitates sampling intervals on a scale of tens of  $\mu\text{m}$ , which requires scanning several thousand detection points. Using sequential scanning schemes, such as the *Fabry–Pérot based PA scanner (FB scanner)* or mechanically scanned piezoelectric receivers, this inevitably results in long acquisition times. In principle, this can be overcome by using an array of detectors. However, a fully sampled array comprising several thousand elements, each with their own signal conditioning electronics and radio frequency analog-to-digital (RF A-D) electronics, would be prohibitively expensive, and efficient multiplexing is challenging due to the low pulse repetition frequency of current excitation lasers. The slow acquisition speed currently limits the use of PAT for applications where movement of the target will cause image artefacts and prohibits the examination of dynamic anatomical and physiological events in high resolution in real time, which is the goal in 4D PAT.

A different approach to accelerate sequential PAT scanners relies on the key observation that, in many situations, the spatial complexity of many of the absorbing tissue structures is rather low, and therefore, data recorded in a conventional, regularly sampled, fashion is highly redundant. It may be possible, therefore, to speed up the data acquisition without a significant loss of image quality by exploiting this redundancy and measuring a subset of the data chosen in such a way as to maximize its non-redundancy. This concept, established as the field of *compressed sensing (CS)* (Candes *et al* 2006, Donoho 2006, Foucart and Rauhut 2013), has been applied to several imaging modalities with success, most notably to *magnetic resonance*

*tomography (MRI)* (Lustig *et al* 2007, Tremoulheac *et al* 2014, von Harbou *et al* 2015) and *computed tomography (CT)* (Ritschl *et al* 2011, Hämläinen *et al* 2013, Jørgensen and Sidky 2015). As the inverse problem of PAT is conceptually similar to these, there has been an increased interest in applying CS to PAT in different ways and for different types of scanners: in Guo *et al* (2010), Meng *et al* (2012a, 2012b), Provost and Lesage (2009) and Zhang *et al* (2012), 2D reconstructions from regularly sub-sampled circular and linear transducer arrays were computed using matrix-based algorithms and first promising results were obtained. An asymmetrical 2D circular sensor arrangement designed based on a low-resolution pre-image was examined in Cao *et al* (2015). In Sandbichler *et al* (2015), the CS-based recovery of suitably transformed integrating line detector data followed by a universal backprojection in 2D was proposed. Another approach, using patterned excitation light, was presented in Liang *et al* (2009) and Sun *et al* (2011), which is a possible approach for photoacoustic microscopy, but is not applicable in PAT where light undergoes strong scattering. In our work, we further explore the potential for acceleration of high resolution 3D PAT using randomized, incoherent encoding of the PA signals and a sparsity-constrained image reconstruction via variational regularization enhanced by Bregman iterations. Extensive studies with simulated data from realistic numerical phantoms, measured data from experimental phantoms and *in vivo* recordings are carried out to demonstrate which conditions have to be fulfilled to obtain high sub-sampling rates. To cope with the immense computational challenges, GPU computing is combined with matrix-free, state-of-the-art optimization approaches.

The remainder of the paper is organized as follows: section 2 provides background on theoretical and practical aspects of PAT. Section 3 describes two general approaches to accelerate sequential scanning schemes by spatial sub-sampling/compressive sensing and exemplifies their implementation with an FP scanner. In section 4, we then describe how images can be reconstructed from the sub-sampled/compressed data. Our methods are exemplified by simulation studies in section 5 and by reconstruction from experimental data in section 6. In section 7, we summarize and discuss the results of our work and point to future directions of research. Table 1 lists all commonly occurring abbreviations for later look-up.

## 2. Background

### 2.1. Basics of photoacoustic tomography

The photoacoustic signal is generated by the coupling of optical and acoustic wave propagation processes through the *photoacoustic effect*: firstly, the tissue is illuminated by a laser pulse with a duration of a few nanoseconds. Inside the tissue, the photons will be scattered and absorbed, the latter predominantly in regions with a high concentration of chromophores, such as haemoglobin. The photoacoustic effect occurs when a sufficient part of the absorbed optical energy is converted to heat (*thermalised*) sufficiently fast and not re-emitted: the induced, local pressure increase  $p_0$  initiates acoustic waves travelling in the tissue on the microsecond timescale. These waves can be measured by ultrasonic transducers at the boundary of the domain.

With several assumptions on the tissue's properties (see Wang and Anastasio (2011) for a detailed discussion), the acoustic part of the signal generation can be modelled by the following initial value problem for the wave equation:

$$(\partial_t - c_0^2 \Delta)p(r, t) = 0, \quad p(r, t = 0) = p_0, \quad \partial_t p(r, t = 0) = 0. \quad (1)$$

The measurement data  $f$  consists of samples of  $p(r, t)$  on the boundary of the domain. See Beard (2011) and Lutzweiler and Razansky (2013) for recent reviews on measurement

**Table 1.** List of commonly occurring abbreviations.

Abbreviation	Meaning	Reference
BP	<i>Back projection</i>	Section 4.4, (5)
BPppTV+ cnv.	BP followed by TV+denoising as post-processing Conventional (data sampling)	See TRppTV+ See section 3.1.1
DMD	<i>Digital micromirror device</i>	Section 4.3
DP	<i>Discrepancy principle</i>	Section 5.2, (15)
FP	<i>Fabry–Pérot interferometer</i>	Section 3.2
L2+	Positivity-constrained $\ell_2$ regularization	Section 4.5, (8)
mxIP	<i>Maximum intensity projection</i>	Section 5.1
PSNR	<i>Peak signal-to-noise ratio</i>	(14)
(Q)PAT	<i>(Quantitative) photoacoustic tomography</i>	Section 1
rSP	<i>Random single point</i> sub-sampling	Section 3.1.1
sHd	<i>Scrambled Hadamard</i> sub-sampling	Section 3.1.2, 4.3
TR	<i>Time reversal</i>	Section 4.4, (6)
TRppTV+	TR followed by TV+denoising as post-processing	(18)
TV+	Positivity-constrained total variation regularization	Section 4.5, (9)
TV+ BR	Bregman iterations applied to TV+	Section 4.6, (11) and (12)

systems. The computation of a high resolution reconstruction of  $p_0$ , usually referred to as the *photoacoustic image* (PA image), from  $f$  is the subject of this paper. Obtaining a high quality PA image is of crucial importance for any subsequent analysis, e.g. for *quantitative photoacoustic tomography* (QPAT) (Cox *et al* 2012), wherein the optical part of the signal generation is inverted based on the PA image.

## 2.2. Nyquist sampling in space and time

Before we discuss how to sub-sample the incident photoacoustic field  $p(r, t)$  in section 3, it is important to understand that a complete sampling requires a certain relation between sampling in  $r$  and  $t$ : imagine we measure the PA signal on the boundary of a domain with homogenous sound speed  $c$  which is band-limited to  $\omega_t^*$ . Firstly, the Nyquist criterion requires us to sample the PA signal with a temporal spacing of  $\delta_t < 1/(2\omega_t^*)$ . Secondly, the PA signal is caused by incident acoustic waves coming from various spatial directions. As an incident wave with a wave vector  $\mathbf{k}$  leads to a PA signal with frequency

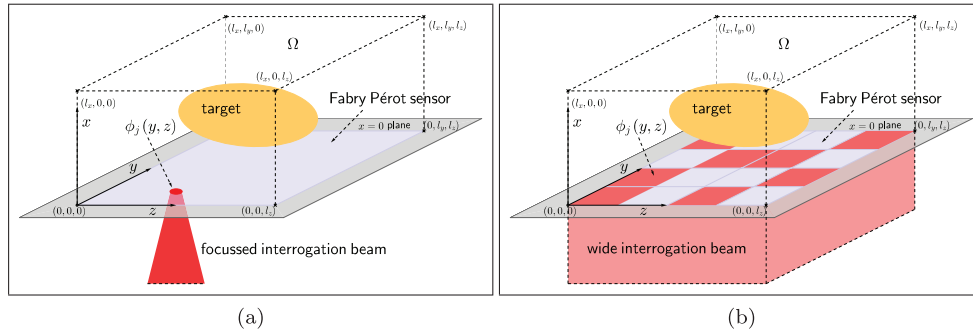
$$\omega_t = c\sqrt{k_x^2 + k_y^2 + k_z^2} = c\|\mathbf{k}\|_2, \quad (2)$$

the band-limit  $\omega_t^*$  of the PA signal implies that the spatial waves are limited by  $\|\mathbf{k}^*\|_2 = \omega_t^*/c$ . To resolve all the spatial information, the Nyquist criterion would require to sample the domain boundary with a spatial spacing of  $\delta_r < 1/(2\|\mathbf{k}^*\|_2) = c/(2\omega_t^*)$ .

## 3. Compressed photoacoustic sensing

### 3.1. Sub-sampling of the photoacoustic field

As discussed in section 1, a drawback of all current 3D PAT systems is that they offer either exquisite image quality or high frame rates but not both, partly due to the difficulty of realizing



**Figure 1.** Conceptual sketch of a Fabry–Pérot interferometer-based PAT setting with different spatial window functions  $\phi_j(y, z)$ : (a) The interrogation beam is focused on a single location, leading to a very localized  $\phi_j(y, z)$ . (b) A wide interrogation beam is used that has been patterned to produce a distributed, binary  $\phi_j(y, z)$ .

a scheme that would complete a scan with a sufficiently small  $(\delta_r, \delta_t)$  in an acceptable acquisition time. In this section, we describe two novel sensing paradigms that aim to accelerate the data acquisition by spatially sub-sampling the incident photoacoustic field  $p(r, t)$ . In this study, the practical realization of these two approaches is achieved via the *Fabry–Pérot* (FP) photoacoustic scanner as described in section 3.2, but they are equally applicable to other sequential scanning systems, such as mechanically scanned piezoelectric detectors.

For simplicity but without loss of generality, we assume that a planar detection surface located at  $x = 0$  is used, that a rectangular area  $[0, l_y] \times [0, l_z]$  on it can be interrogated, and that the target is located in the region  $\Omega = [0, l_x] \times [0, l_y] \times [0, l_z]$ . The extension to other detection geometries is straight forward, but complicates the notation. The concrete measurement process can be modeled in the following way: the incident photoacoustic field on the detection plane,  $p(x = 0, y, z, t)$ , caused by the  $j^{\text{th}}$  pulse of the excitation laser, is first multiplied by a spatial window function  $\phi_j(y, z)$  and then integrated over the whole detection surface:

$$f_j(t) = \int p(x = 0, y, z, t) \phi_j(y, z) dy dz. \quad (3)$$

The spatial sampling is followed by a temporal sampling, e.g. by measuring  $f_j(t)$  at  $t_i = i_t \delta_t$ ,  $i_t = 1, \dots, M_t$ . The setting is sketched in figure 1.

**3.1.1. Single-point scanning.** A standard approach is to try to focus  $\phi_j(y, z)$  on a single point  $(y_j, z_j)$  (see figure 1(a)), which belongs to a set of  $M = M_y \times M_z$  points forming a regular grid with spacings  $\delta_y = l_y/M_y$  and  $\delta_z = l_z/M_z$ . We will refer to the data obtained by this scanning pattern as the *conventional* (cnv.) data. Ideally, these spacings are chosen fine enough to assure that the Nyquist criterion is satisfied in space and time (see section 2.2).

Due to the spatio-temporal characteristics of the wave propagation, the pressure time series recorded at two neighbouring locations on the regular grid provide very similar information. To reduce the coherence of the measured time series, we can instead also sample the photoacoustic field at a smaller number of randomly chosen points  $(y_j, z_j)$ ,  $j = 1, \dots, M_c$ ,  $M_c < M$ , which would yield an acceleration factor of  $M_{\text{sub}} = M/M_c$ . We will denote this sub-sampling

strategy by rSP- $M_{\text{sub}}$ . Choosing random points is firstly motivated by several results in compressed sensing theory (Foucart and Rauhut 2013), that point out the importance of randomness for designing sub-sampling pattern. Secondly, we want to avoid choosing sampling pattern that might lead to systematic biases. We will return to this point in the computational studies in sections 5 and 6.

**3.1.2. Patterned interrogation scanning.** The second idea to accelerate the acquisition is to choose a series of orthogonal pattern  $\phi_j(y, z)$  with each  $\phi_j(y, z)$  being supported all over  $[0, l_y] \times [0, l_z]$  with no particular focus on a single location (see figure 1(b)). Again, choosing  $M_c < M$  pattern would yield an acceleration factor of  $M_{\text{sub}} = M/M_c$  over conventional single point scanning. As we will use a specific type of binary pattern based on scrambled Hadamard matrices described later (section 4.3), we will denote the sub-sampling strategy by sHd- $M_{\text{sub}}$ .

### 3.2. Implementation of sub-sampling schemes using the FB scanner

The planar, interferometry-based Fabry–Pérot photoacoustic scanner provides a convenient implementation of the sub-sampling strategies described above. In the standard set-up, the FP scanner performs the conventional single-point scanning described in section 3.1.1: for each pulse sent in by the excitation laser, the pressure time series at a different location on a grid is measured by interrogating the FP sensor head with an interrogation laser (Zhang *et al* 2008). Due to the planar geometry and the option to introduce the excitation laser through the transparent detection plane, PA signals from a large range of anatomical targets can be scanned in the frequency range from DC to several tens of MHz on a scale of tens of  $\mu\text{m}$  ( $\phi_j(y, z)$  in our model corresponds to the beam profile of the interrogation laser, see figure 1(a)). Superficial features located a few mm below the skin surface can be imaged with high spatial resolution from a conventional FP scan (see, e.g. Jathoul *et al* (2015)). However, as described in section 1, obtaining such an image requires scanning several tens of thousands of locations which, due to the repetition rates of currently available excitation lasers, takes several minutes.

While the single point sub-sampling described in section 3.1.1 is straight forward to implement with a standard FP scanner, implementing the patterned interrogation scheme requires several modifications: instead of focusing the interrogation beam on a single location, the whole detection plane is illuminated and the reflected beam is patterned before being focused into the photo diode. The spatial modulation is inspired by the working-principle of the celebrated ‘single-pixel Rice camera’ (Duarte *et al* 2008): a *digital micro-mirror device (DMD)* is used to block rectangular sections of the reflected interrogation beam which creates binary pattern. Note that this hardware realization slightly differs from the conceptual sketch in figure 1(b), where the direct interrogation beam is patterned. Further details of the *patterned-illumination* scanner can be found in Huynh *et al* (2014, 2015), Huynh *et al* (2016).

## 4. Image reconstruction from sub-sampled data

In this section, we describe a model of the accelerated data acquisition and how to invert it.

#### 4.1. Continuous forward model

The PAT forward operator,  $A$ , maps a given initial pressure  $p_0$  in the volume  $[0, l_x] \times [0, l_y] \times [0, l_z]$  to the time dependent pressure on the detection plane,  $\bar{p} := p(x=0, y, z, t)$ , as determined by (1):  $\bar{p} = Ap_0$ . A more detailed discussion of the operator  $A$  and its adjoint can be found in Arridge *et al* (2016). In this work, we assume that the target's dynamics are slow enough to be well approximated by a constant within the acquisition time. A sensing operator  $C$  implements (3) to produce the measured data  $f^c \in \mathbb{R}^{M_c M_t}$  from the detection plane pressure  $\bar{p}$ :

$$f^c = C\bar{p} + \varepsilon = CAp_0 + \varepsilon, \quad (4)$$

where we added the term  $\varepsilon$  to account for additive measurement noise. In the following,  $C$  denotes the sampling operator: the conventional point-by-point scan described in section 3.1.1 (where  $M_c = M$ ) or one of the two sub-sampling strategies, the random single-point sub-sampling described in section 3.1.1 or the patterned interrogation described in section 3.1.2.

#### 4.2. Image reconstruction strategies

Given the compressed data  $f^c$ , there are in principle two strategies of how to reconstruct  $p_0$ : in *two-step procedures*, we first reconstruct the detection plane pressure  $\bar{p}$  from  $f^c$  based on  $f^c = C\bar{p}$  (*data reconstruction*) and then use a standard PAT reconstruction for complete planar data (see, e.g. Kuchment and Kunyansky (2011)). In *one-step procedures*,  $p_0$  is reconstructed directly from  $f^c$  using a model-based approach (4). While both approaches have advantages and disadvantages and novel two-step procedures have been introduced in Betcke *et al* (2016), Huynh *et al* (2015) and Sandbichler *et al* (2015), the focus of this work is not to carry out a fair, detailed comparison between one-step and two-step approaches: we rather want to emphasize on the differences between simple, *linear* reconstruction techniques and *variational*, model-based reconstruction techniques, independent of the former being two-step and the latter being one-step procedures. To ease the following presentation, we first introduce the discrete PAT model we will use for numerical computations before discussing the details of the reconstruction techniques.

#### 4.3. Discrete forward model

As all methods we examine directly rely on the wave equation (1), we need a fast numerical method for 3D wave propagation with high spatial and temporal resolution. Our choice is the  $k$ -space pseudospectral time domain method (Mast *et al* 2001, Cox *et al* 2007, Treeby *et al* 2010) implemented in the  $k$ -Wave Matlab Toolbox (Treeby and Cox 2010). For the following, it is only important that  $k$ -Wave discretizes  $\Omega = [0, l_x] \times [0, l_y] \times [0, l_z]$  into  $N = N_x \times N_y \times N_z$  voxels  $\Delta_{x/y/z} = l_{x/y/z}/N_{x/y/z}$  and uses an explicit time stepping. The time step used will always be the same as used in the temporal sampling of the pressure field, i.e.  $\Delta_t = \delta_t$  (see section 3.1). From now on, all variables used are discrete although we will continue to use the same notation for them. For instance, the discretization of the PAT forward operator is now a matrix  $A \in \mathbb{R}^{N_y N_z M_t \times N}$ , mapping the discrete initial pressure  $p_0 \in \mathbb{R}^N$  to the pressure at the first layer of voxels in  $x$  direction at the  $M_t$  discrete time steps, as these voxels represent the detection plane. Note that we cannot construct  $A$  explicitly, but rely on computing matrix-vector products with  $A$  and  $A^T$  using  $k$ -Wave. A detailed discussion of the implementation can be found in Arridge *et al* (2016). The discrete sub-sampling operators  $C$  map from the pressure-time series of the detection plane voxels to  $f^c \in \mathbb{R}^{M_c M_t}$ . The single point sampling operators (see sections 3.1.1 and 3.1.1) simply extract the pressure-time series at the sensor voxels

(which are, in general, only a subset of the detection plane voxels). For the interrogation with binary sensing pattern constructed by the DMD (see section 3.2), (3) can be implemented by multiplying a  $M_c \times N_y N_z$  binary matrix  $H$  with the vector of pressure values at the scanning voxels, separately for each time point  $t$ . Compressed sensing theory suggests that random Bernoulli matrices are optimal for many applications (Foucart and Rauhut 2013). As those are difficult to implement experimentally and computationally, we use scrambled Hadamard matrices which are known to have very similar properties compared to a random Bernoulli matrices and can be implemented very efficiently by a fast-fourier-transform-like operation (Do *et al* 2012, Foucart and Rauhut 2013). Note that as the entries of Bernoulli and Hadamard matrices take the values  $\{-1, 1\}$  and not  $\{0, 1\}$  as implemented by the DMD, we need to demean experimental data in a pre-processing step. Further details can be found in Betcke *et al* (2016), Huynh *et al* (2015, 2016).

#### 4.4. Linear back-projection-type reconstructions

As described above, we will use simple, linear two-step procedures to compare the more sophisticated variational methods to: first, we reconstruct the complete data by the pseudo-inverse of  $C$ :  $C^\dagger f^c$  (see (4)), where for the sub-sampling operators we consider here,  $C^\dagger = C^T$ . Then, we either multiply with  $A^T \in \mathbb{R}^{N \times N_y N_z M_t}$  (called *back-projection (BP)* here), or with the discrete *time reversal (TR)* (Finch and Patch 2004, Xu and Wang 2004, Jathoul *et al* 2015) operator  $A^\dagger \in \mathbb{R}^{N \times N_y N_z M_t}$ :

$$p_{\text{BP}} := A^T C^T f^c \quad (5)$$

$$p_{\text{TR}} := A^\dagger C^T f^c. \quad (6)$$

The difference between BP and TR (including enhanced variants of TR (Stefanov and Uhlmann 2009)), is discussed in more detail in Arridge *et al* (2016). In summary, in the continuous setting, they differ in the way they introduce the time-reversed pressure time series at the detection plane: TR approaches use them as a time dependent Dirichlet boundary condition for the wave equation (1) while the adjoint approach introduces them as a time dependent source term without altering the boundary conditions.

#### 4.5. Variational image reconstruction

*Variational regularization* (Scherzer *et al* 2009) is a popular and well understood approach for approximating the solutions of ill-posed operator equations like (4) in a reasonable and controlled manner: the regularized solution is defined as the minimizer of a suitable energy functional  $\mathcal{E}$ . Assuming that the additive noise,  $\varepsilon$ , is i.i.d. normally distributed, a reasonable approach is to solve

$$p_\lambda := \underset{p}{\operatorname{argmin}} \mathcal{E}(p) = \underset{p}{\operatorname{argmin}} \left\{ \frac{1}{2} \|CAp - f^c\|_2^2 + \lambda \mathcal{J}(p) \right\}, \quad (7)$$

to obtain a regularized solution  $p_\lambda$ . While the first term in the composite functional measures the misfit between measured and predicted data (*data fidelity term*),  $\mathcal{J}(p)$  has to render the minimization problem (7) well-posed by ensuring existence, uniqueness and stability of  $p_\lambda$  (*regularization functional*). Furthermore, its choice can be used to penalize or constrain unwanted features of  $p_\lambda$ , thereby encoding *a-priori* knowledge about the solution. The *regularization parameter*  $\lambda > 0$  controls the balance between both terms.

The first variational method we examine corresponds to classical, zeroth-order Tikhonov regularization, augmented by the physical constraint  $p_0 \geq 0$ :

$$p_{L2+} := \operatorname{argmin}_{p \geq 0} \left\{ \frac{1}{2} \|CAp - f^c\|_2^2 + \lambda \|p\|_2^2 \right\}. \quad (8)$$

As a second functional  $\mathcal{J}(p)$ , we examine the popular *total variation (TV)* energy, which is a discrete version of the *total-variation* seminorm (Rudin et al 1992, Burger and Osher 2013):

$$p_{TV+} := \operatorname{argmin}_{p \geq 0} \left\{ \frac{1}{2} \|CAp - f^c\|_2^2 + \lambda \operatorname{TV}(p) \right\}. \quad (9)$$

The energy  $\operatorname{TV}(p)$  measures the  $\ell_1$  norm of the amplitude of the gradient field of  $p$  (the details of its implementation are given in appendix A) and is a prominent example of non-smooth, *edge-preserving* image reconstruction techniques, and, more generally, of *spatial sparsity constraints*. While TV regularization has been used for PAT before (see, e.g. Huang et al (2013)), our main interest in it arises from its results when applied to sub-sampled data for other imaging modalities (Lustig et al 2007, Brune et al 2011, Ritschl et al 2011, Hämäläinen et al 2013, Mueller 2013, Benning et al 2014, Jørgensen and Sidky 2015, von Harbou et al 2015): TV regularization is often able to recover the object's main feature edges even for high sub-sampling factors. Therefore, we focus on this rather general regularization energy in this first PAT sub-sampling study and examine more specific functionals in future work.

As all of the involved functionals and constraints are convex, a variety of methods exist to solve (7) computationally. In this work, we use an *accelerated proximal gradient*-type method described in appendix B.

#### 4.6. Bregman iterations

A potential drawback of variational techniques like (7) is that they inevitably lead to a systematic bias of the regularized solutions: the solution  $p_\lambda$  moves from an un-biased data-fit towards a minimizer of  $\mathcal{J}(p)$ . Formally, let  $\tilde{f}^c = CAp_0$  be the true, noise-free data and  $\tilde{p}_\lambda$  the solution of (7) for  $f^c = \tilde{f}^c$ . Then, by the minimizing properties of  $\tilde{p}_\lambda$ , we have

$$\frac{1}{2} \|CA\tilde{p}_\lambda - \tilde{f}^c\|_2^2 + \lambda \mathcal{J}(\tilde{p}_\lambda) \leq \frac{1}{2} \|CAp_0 - \tilde{f}^c\|_2^2 + \lambda \mathcal{J}(p_0) = \lambda \mathcal{J}(p_0), \quad (10)$$

and thereby,  $\mathcal{J}(\tilde{p}_\lambda) \leq \mathcal{J}(p_0)$ . For the TV energy, this bias manifests in the well-known, non-linear contrast loss of TV regularized solutions (Burger and Osher 2013). For PAT, this systematic error poses a crucial limitation on the use of TV regularized PA images for quantitative analysis like QPAT studies. To overcome this drawback, an iterative enhancement of variational solutions by the *Bregman iteration* (Bregman 1967) was proposed in Osher et al (2006): for (7), they take the form

$$p_\lambda^{k+1} = \operatorname{argmin}_p \left\{ \frac{1}{2} \|CAp - (f^c + b^k)\|_2^2 + \lambda \mathcal{J}(p) \right\}, \quad (11)$$

$$b^{k+1} = b^k + (f^c - CAp_\lambda^{k+1}), \quad (12)$$

with  $b^0 = 0$ . This iteration has several attractive features (Burger and Osher 2013): it solves the un-regularized problem

$$\min_p \mathcal{J}(p) \quad \text{subject to} \quad p \in \operatorname{argmin}_q \|CAq - f^c\|_2^2 \quad (13)$$

by solving a sequence of well-regularized problems (11) while the residual of iterates,  $\|CA p_\lambda^{k+1} - f^c\|$ , is monotonically decreasing. The potential of Bregman iterations, in particular when used on sub-sampled data, has been demonstrated in Benning *et al* (2014), Brune *et al* (2010), Mueller (2013) and von Harbou *et al* (2015). Note the difference between the use of the Bregman iteration in the *Split Bregman method* (Goldstein and Osher 2009), a method to solve problems like (7) which is also known as the *augmented Lagrangian method*, and the usage here which does not have an equivalent Lagrangian interpretation.

## 5. Simulation studies

We now examine the different inverse methods described in the previous section when applied to sub-sampled data from numerical phantoms.

### 5.1. Realistic numerical phantom

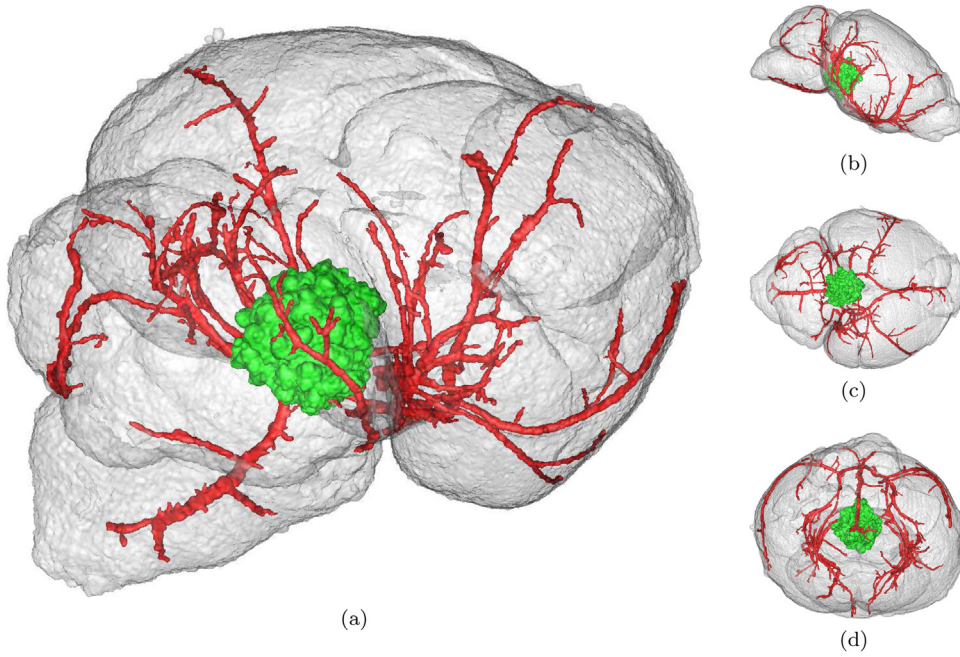
While studies using numerical phantoms composed of simple geometrical objects can provide valuable insights to the basic properties of the inverse problem and reconstruction methods, it is often unclear how their results translate to experimental data from complex targets. In this section, we briefly describe the construction of a realistic numerical phantom that will be used in the main simulation studies.

The phantom is based on a segmentation of a micro-CT scan of a mouse brain ( $533 \times 400 \times 346$  voxel) into gray matter, vasculature and dura mater. The vasculature is morphologically closed (dilation followed by erosion) using the 18-neighborhood as a convolution kernel. Thereafter, the vasculature is one connected component with respect to the 26-neighborhood. The whole segmentation is clipped to the bounding box of the vasculature leading to a size of  $306 \times 423 \times 345$ . Next, a gray matter voxel in the central part of the segmentation is chosen uniformly at random. It is used as the seed point for the construction of an artificial cancer tissue inside the gray matter by a stochastic growth process that consists of an iterative application of morphological operations on the surface voxels. Figure 2 shows the final result of this construction. Note that vasculature and tumor tissue are non-intersecting.

For the studies in this work, the clipped volume is embedded into a cubic volume of  $512^3$  voxels, centered in  $y$  and  $z$  direction and in two different heights in  $x$  direction: in the first phantom, called *Tumor1*, the distance between the detection plane  $x = 0$  and the phantom is half of the distance between the plane  $x = l_x$  and the phantom. In the second phantom, called *Tumor2*, it is centered in  $x$  direction, leading to a larger distance between sensor and target (see figure 3). To construct  $p_0$  from the segmentation, all vasculature voxels are given the value of 1, whereas the tumor tissue is given the value of 0.7 for *Tumor1* and 0.3 for *Tumor2*. Next,  $p_0$  is down-sampled to the desired resolution  $[N_x, N_y, N_z]$  by successive sub-averaging over  $2 \times 2 \times 2$  blocks. For *Tumor1*, we modify the resulting  $p_0$  to obtain sharper boundaries:  $p_0$  is normalized such that  $\max(p_0) = 1$  and the intensity  $p_{0,i}$  of all non-zero voxels  $i$  is set to  $(2p_{0,i} + 1)/3$ , thereby ensuring a contrast minimal value of  $1/3$  between background and target. Figure 3 shows *maximum intensity projections (mxIP)* of the resulting  $p_0$ .

### 5.2. Simulation studies with *Tumor1*

We first examine the inverse reconstructions using *Tumor1* with  $N = 128^3$  for ‘inverse crime’ data (Kaipio and Somersalo 2007), which means that we assume that we have exact knowledge of all physical parameters, which are summarized in table 2, and use the same model

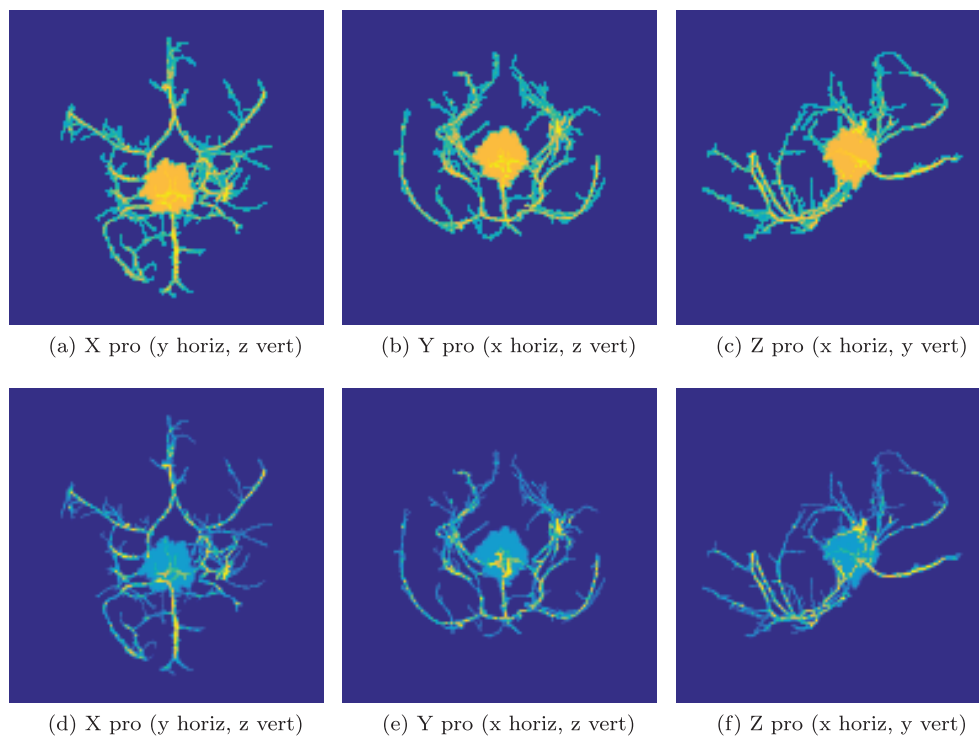


**Figure 2.** 3D visualization (a)–(d): different views) of the segmentation used for the construction of the realistic numerical phantom: vasculature (red), artificial tumor tissue (green) and gray matter (gray). Volume rendering was carried out by SCI Institute (2015).

for both data simulation and reconstruction. In addition, we sampled the data with the spatial spacing given by the Nyquist criterion: the spatial sampling intervals  $\delta_{y/z}$  coincide with the spatial spacing of the computational grid  $\Delta_{y/z}$  and are fine enough to capture all relevant spatial frequencies of the incident photoacoustic field (see section 2.2):  $p_0$  was pre-smoothed to ensure that its discrete approximation by the truncated Fourier basis used in k-Wave was non-oscillatory. This is reflected in a sharp drop of the power spectrum of the pressure time series around 4.8 MHz, which (see section 2.2) corresponds to a spatial Nyquist rate equal to the spatial spacing  $\delta_{y/z} = \Delta_{y/z} = 156.25 \mu\text{m}$ . White noise with a standard deviation of  $\sigma = 0.001$  is added to the clean data  $CAp_0$ , leading to a *signal-to-noise ratio* (SNR) of 18.63 dB.

For all computed solutions  $p$ , voxels with negative pressure and all sensor voxels have been set to 0 in a post-processing step. We will mainly rely on a visual comparison of the results via maximum intensity projections along the  $y$  direction (see figures 3(b) and (e)). Unless stated otherwise, the color scale of each figure is determined independently. It ranges from 0 to the value separating the  $100/256\% \approx 0.39\%$  largest values of  $p$  from the smaller values. This clipping is necessary to avoid that a few large outliers determine the contrast of the image and complicate the comparison between different methods. In addition to the visual comparison, we report the *mean-squared-error* (MSE) between the reconstructed solution  $p$  and  $p_0$ , i.e.  $\|p - p_0\|_2^2/N$ , in the conventional logarithmic scaling termed *peak-signal-to-noise ratio* (PSNR):

$$\text{PSNR}(p, p_0) := -10 \log_{10} \left( \frac{1}{N} \|p - p_0\|_2^2 \right) \quad (14)$$



**Figure 3.** Maximum intensity projections of the realistic numerical phantoms (color map ‘parula’, see figure 7(b)): (a)–(c) Tumor1 (d)–(f) Tumor2. In both cases, the detection plane corresponds to the top edge of the  $Y$  and  $Z$  projections.

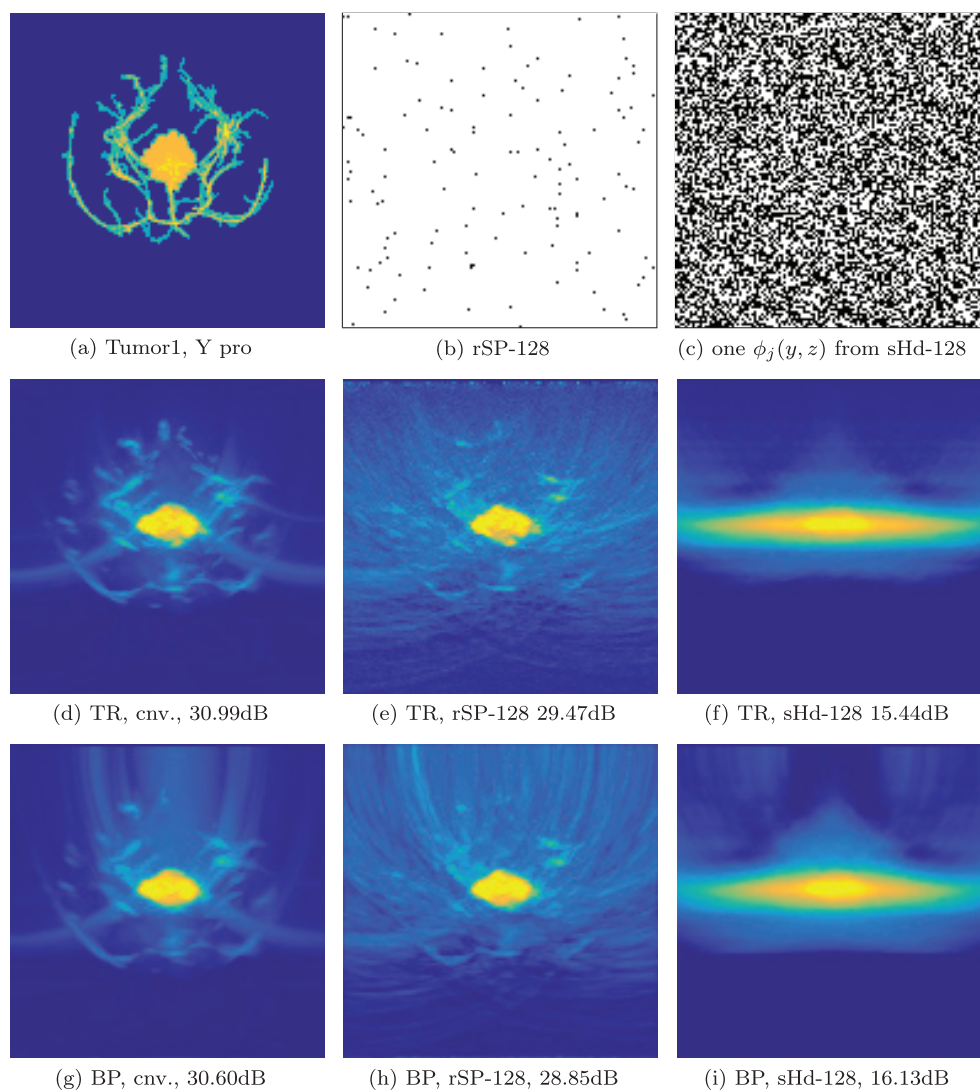
**Table 2.** Parameter of the different inversion models used.

parameter	Tumor1/2	Knot	Hair	Vessels
$(N_x, N_y, N_z)$	128	(44,264,264)	(28,128,128)	(42,282,282)
$(\Delta_x, \Delta_y, \Delta_z)$ ( $\mu\text{m}$ )	156.25	75	(62.12,62.12,68.00)	50
$M_t$	740	391	56	621
$\Delta_t/\delta_t$ (ns)	31.25	12	20	10
$M$	16384	17424	16384	19881
$(\delta_x, \delta_y)$	156.25/312.5	150	/	100
$c_0$ ( $\text{m s}^{-1}$ )	1500	1540	1500	1420

$p$  might have a different scaling compared to  $p_0$  and contain small scale noise. As this typically does not influence the evaluation of a human observer, we also do not want to account for it when computing the PSNR. Therefore, we first rescale and threshold  $p$  and  $p_0$ ,

$$\tilde{p} = \text{thres}\left(\frac{p}{\|p\|_\infty}, 0.1\right), \quad \tilde{p}_0 \text{ accordingly,} \quad \text{where} \quad \text{thres}(v, \alpha) = \begin{cases} v & \text{if } v \geq \alpha, \\ 0 & \text{else} \end{cases},$$

and then compute  $\text{PSNR}(\tilde{p}, \tilde{p}_0)$  by (14).



**Figure 4.** Tumor1 results (mxIP) of linear methods: (a) phantom (see figure 3), detection plane corresponds to the top edge. (b) Visualization of the rSP-128 sub-sampling pattern: while each of the  $M = 128 \times 128$  pixel corresponds to one possible scanning location, all black pixels correspond to one of the  $M_c = 128$  random locations actually scanned. (c) A  $128 \times 128$  Hadamard matrix to symbolize the sHd-128 pattern (the actual matrix is of size  $128 \times 16384$ ) (d)–(i): TR and BP results for conventional data (left column), rSP-128 (middle column) and sHd-128 (right column) and their corresponding PSNR in dB.

Figure 4 shows the results of TR and BP for using rSP-128 or sHd-128 sub-sampling, i.e. accelerated by a factor of  $M_{\text{sub}} = 128$  compared to the conventional scan. This high sub-sampling factor leads to unsatisfactory reconstructions. The different sub-sampling strategies lead to a different appearance of back-propagated noise and sub-sampling artifacts in the

reconstructed images: in the rSP-128 case, both features and noise are back-propagated along the surfaces of spheres centered on the scanning locations while in the sHd-128 case, the non-local nature of the patterned interrogation leads to back-propagation along planes parallel to the detection plane. Figure 5 shows the corresponding results of the classical Tikhonov regularization (8), denoted by ‘L2+’, TV regularization (9), denoted by ‘TV+’ and of the Bregman iteration (11) and (12) applied to TV regularization, denoted by ‘TV+Br’. Despite the high sub-sampling factor, TV+ and TV+Br are still able to reconstruct the main structures of the phantom without excessive image noise.

In the results shown, the regularization parameter  $\lambda$  for L2+ and TV+ was chosen by the *discrepancy principle* (DP): for the general variational problem (7), the discrepancy principle selects  $\lambda$  such that

$$\frac{\|CA p_\lambda - f^c\|}{\sqrt{M_c} \sigma} = \kappa, \quad (15)$$

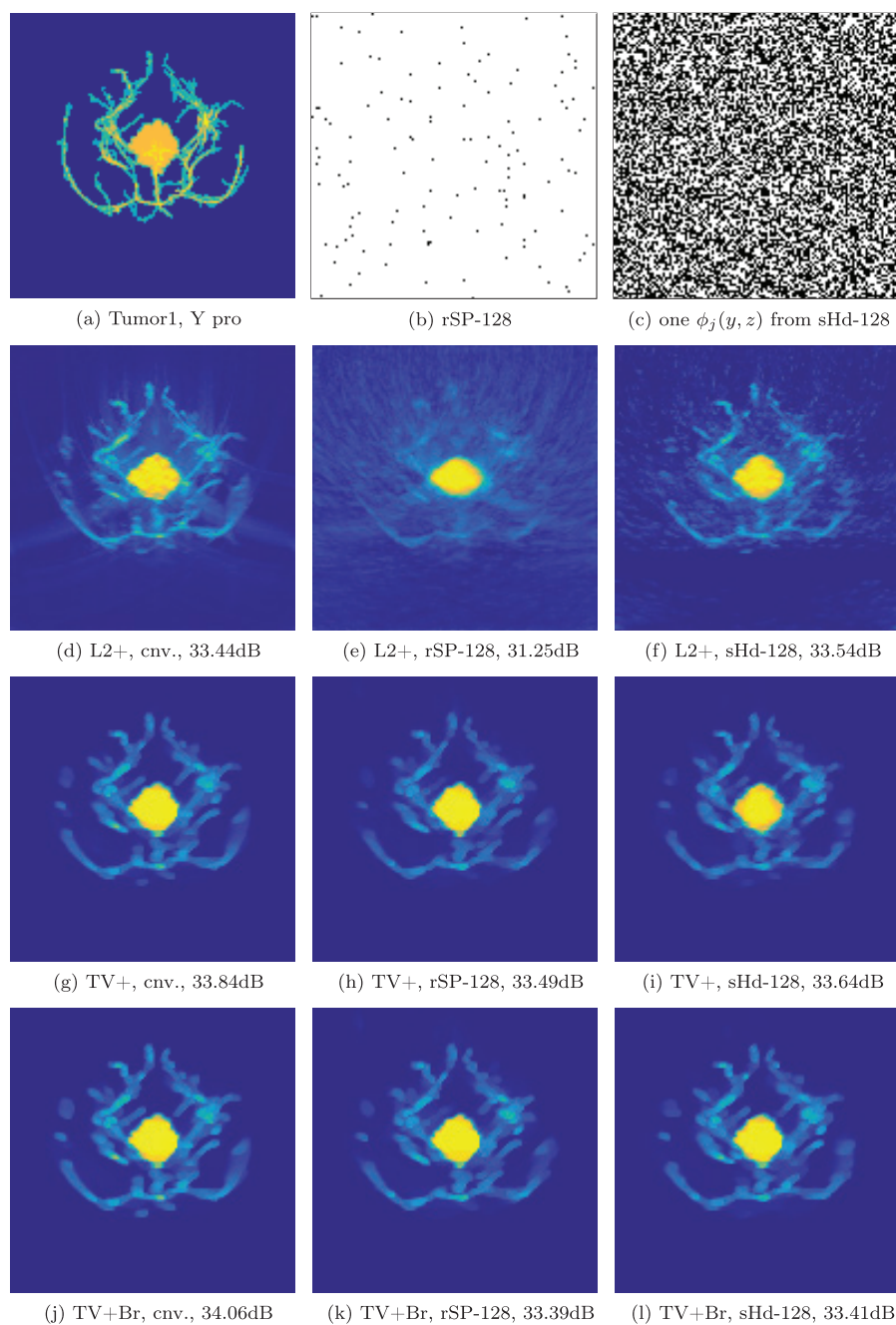
for  $\kappa \geq 1$ . The DP is based on the heuristic argument, that the regularized solution  $p_\lambda$  should explain the data  $f^c$  no more than up to the noise level, which is assumed to be known. As the residual is monotonically increasing in  $\lambda$ , the DP is robust and easy-to-implement. We chose a simple interval-based method that linearly interpolates the left hand side of (15) (the *discrepancy* of the data) in the current search interval. It was terminated when  $\kappa = 1.25$  was reached within a tolerance of 0.01. The Bregman iterations were started with  $\lambda_{\text{TV+Br}} = 10\lambda_{\text{TV+}}$ , where  $\lambda_{\text{TV+}}$  is the regularization parameter found for TV+, and stopped as soon as the discrepancy of  $p_\lambda^{k+1}$  falls below  $\kappa$  (recall that the residual, and thereby the discrepancy monotonically decreases with  $k$ , see section 4.6). This typically happens after about 10 Bregman iterations.

**5.2.1. Enhancement through Bregman iterations.** The Bregman iteration was introduced to compensate for the systematic contrast loss of TV regularized solutions. Figure 6 compares TV+ and TV+Br solutions using the same color scaling and additionally shows mxIPs of the positive and negative parts of the difference between TV+Br and TV+. The difference plots demonstrate that using Bregman iterations especially improves the contrast of the small scale vessel structures and that the benefit is more pronounced for sub-sampled data compared to conventional data.

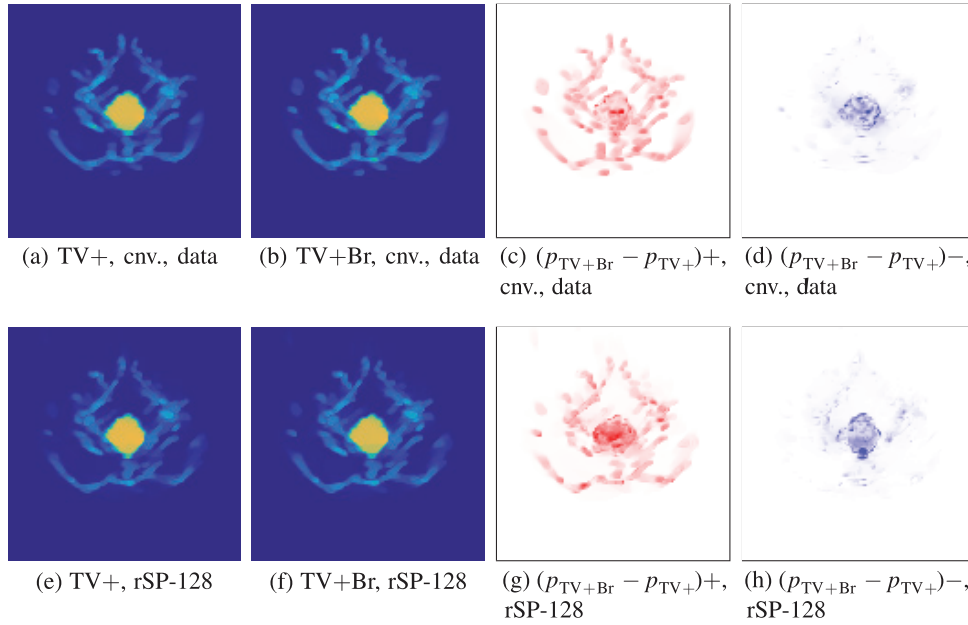
### 5.3. Simulation studies with Tumor2

The good quality of, e.g. the TV+ Br results for the very high sub-sampling factor of 128 have to be interpreted with care: the phantom Tumor1 used was intentionally easy to reconstruct as it was close to the sensors and had high contrast (see figure 3). In addition, the data was created by the same forward model used in the reconstruction, which is known as committing an ‘inverse crime’ (Kaipio and Somersalo 2007), and the conventional data was sampled finely enough to fulfill the Nyquist criterion. The phantom Tumor2 was designed to carry out simulation studies that more accurately reproduce the challenges of experimental data scenarios.

**5.3.1. Inverse crimes.** While inverse crimes are, to a certain extent, unavoidable when carrying out simulation studies, they make it more difficult to extrapolate the results obtained to experimental data. In this section, we discuss how to bridge this gap by modifying the model used for the data generation.



**Figure 5.** Tumor1 results (mxIP) of variational methods: (a)–(c) see figure 4 (d)–(l): L2+, TV+ and TV+Br results for conventional data (left column), rSP-128 (middle column) and sHd-128 (right column) and their corresponding PSNR in dB.



**Figure 6.** Contrast comparison between TV+ and TV+Br solutions (see figure 5): the images in the left two columns share the same color scale. The third and fourth column show mxIPs of the positive (red scale) and negative (blue scale) part of the difference  $p_{\text{TV+Br}} - p_{\text{TV+}}$ .

- The sensitivity of the FP sensor varies spatially (Zhang *et al* 2008). While this effect can be mitigated by calibration and data pre-processing procedures, a residual uncertainty remains that we will model by modifying the discrete (static) data generation model to

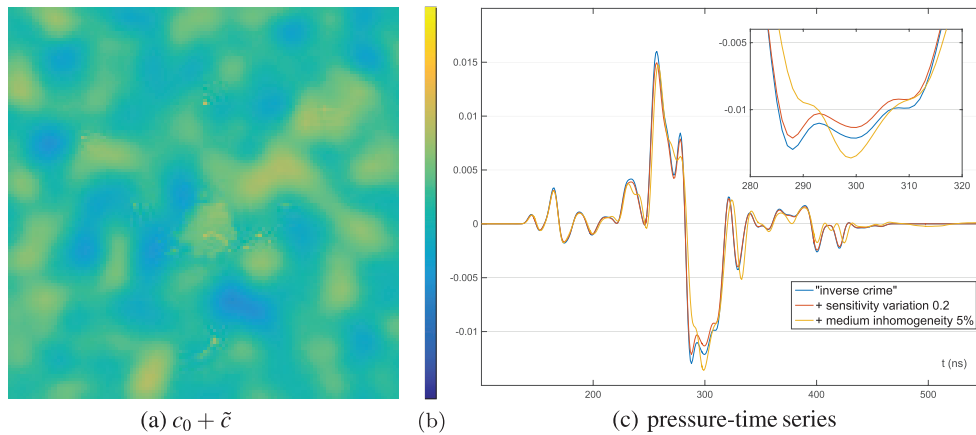
$$f^c = CW_s Ap_0 + \varepsilon \quad (16)$$

where  $W_s$  is a diagonal matrix that multiplies the pressure-time course of each voxel  $i$  in the  $x = 1$  plane by a random variable  $w_i$  following a centered log-normal distribution:

$$w_i = \exp(\sigma_s X_i), \quad X_i \sim \mathcal{N}(0, 1) \quad (17)$$

We choose  $\sigma_s = 0.2$  (90% of  $w_i$  are in  $[0.72, 1.39]$ ).

- In a similar spirit, we assume that we only have a rough estimate of the statistics of  $\varepsilon$ , e.g. from baseline measurements, and can, therefore, only approximately decorrelate the data before the inversion. The residual uncertainty about the noise variance per sensor voxel is modeled by replacing the additional noise term  $\varepsilon$  by  $W_n \varepsilon$ , where  $W_n$  is constructed like  $W_s$  with  $\sigma_n = 0.1$  (90% of  $w_i$  are in  $[0.85, 1.18]$ ). Keeping  $\sigma = 0.001$ , as the standard deviation of  $\varepsilon$ , we end up with an average SNR of 9.51 (the value is considerable lower than for Tumor1 due to the larger distance to the sensors and the lower contrast of Tumor2).
- Although we assume that the medium we image is sufficiently homogeneous to assume a constant sound speed  $c_0$  in the inverse reconstruction, the real sound speed will slightly vary, especially between different tissue types. We use  $c_0 + \tilde{c}$  for the data generation, where  $\tilde{c}$  is constructed by adding a smooth, normalized Gaussian random field and the normalized initial pressure  $p_0$  (as it represents a tissue different from the background).



**Figure 7.** (a)  $x$ -slice of the sound speed used to generate ‘no-inverse-crime’ data for the Tumor2 phantom (b) color scale, range: 1350–1650 m s<sup>-1</sup>, (c) noise-free pressure-time series to demonstrate the effect of sensitivity and sound speed variation (inset zooms into particular section of the plot).

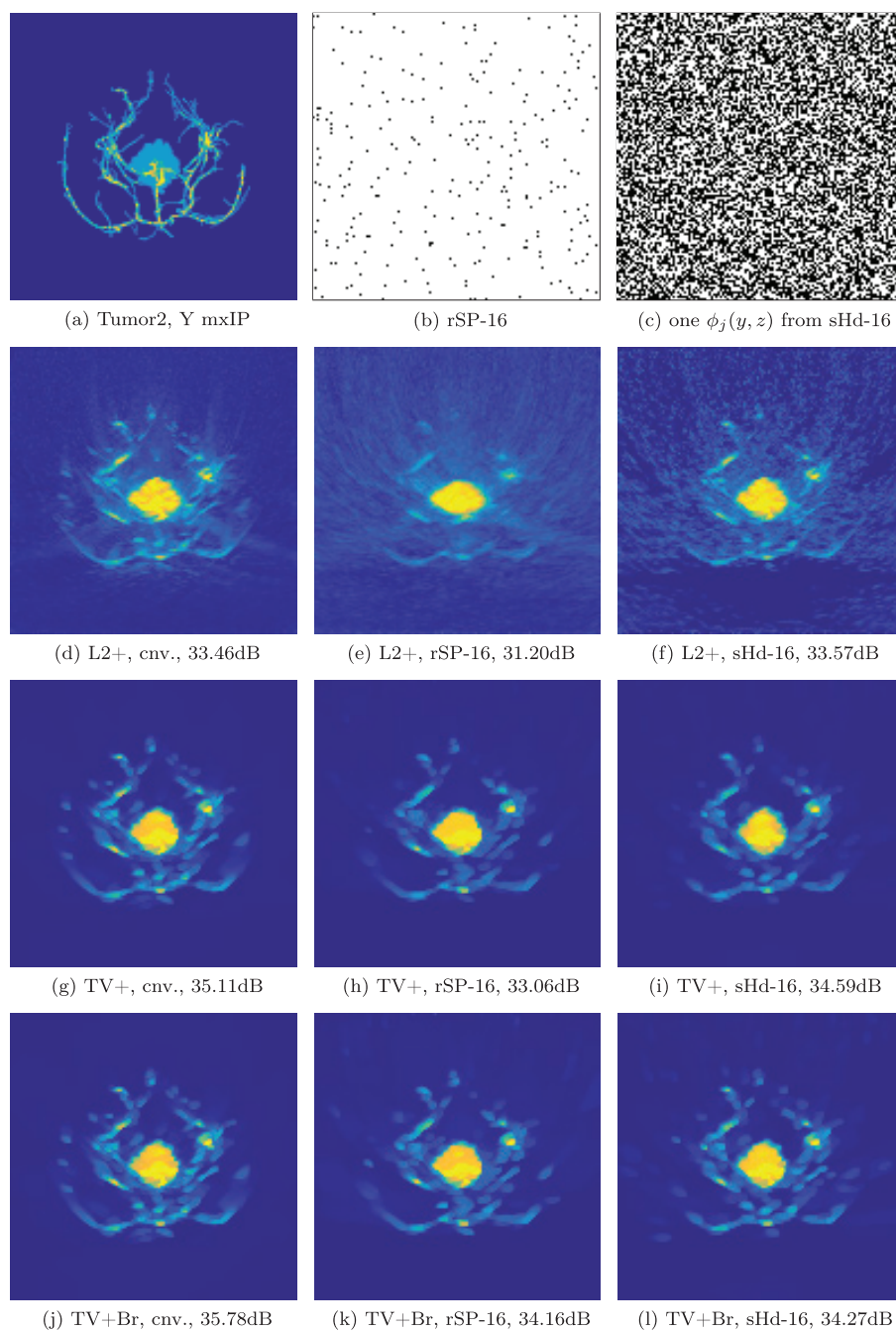
Then,  $\tilde{c}$  is centered and scaled such that its mean is 0 and its maximal absolute value is  $0.05c_0$  (a sound speed variation of 5% is not unusual for soft tissue (Jin and Wang 2006)). The resulting sound speed is shown in figure 7(a).

- Among the many other ways to modify the data generation model are to include acoustic absorption, inhomogeneous illumination, acoustic reflections, baseline shifts and drifts in the pressure-time series, correlated noise and corrupted channels. We leave these extensions to future studies.

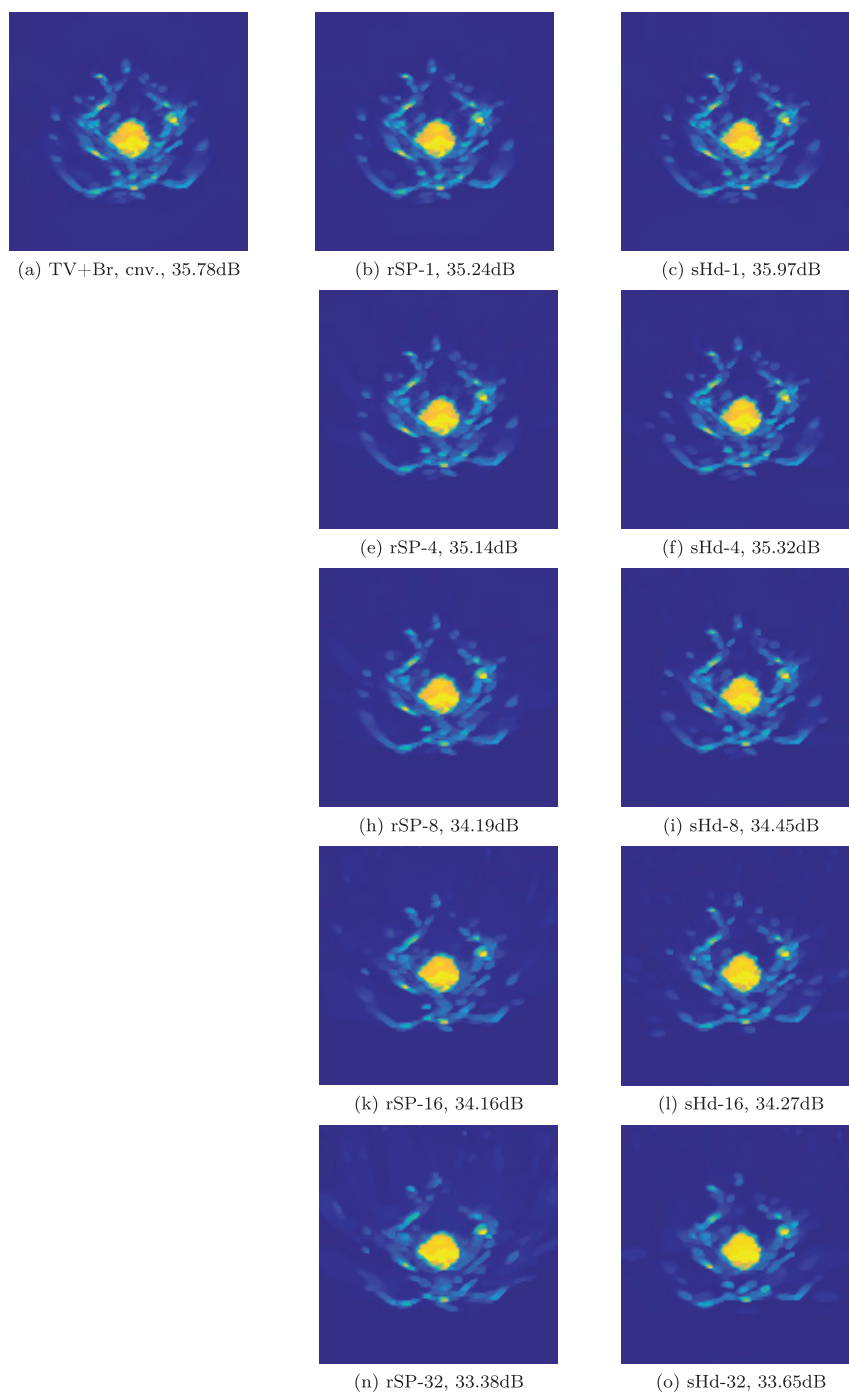
Figure 7(c) compares the noise-free pressure-time series of a single voxel after adding sensitivity and sound speed variations (adding noise and noise variation complicates a visual comparison).

**5.3.2. Nyquist criterion.** In contrast to the previous studies, we now compare to conventional data acquired on a regular grid having a grid spacing corresponding to twice the length determined by the Nyquist criterion (see section 5.2):  $\delta_{y/z} = 2\Delta_{y/z} = 312.5 \mu\text{m}$ , i.e. we model the conventional data by extracting the pressure-time series at a sub-set of the  $128^2$  detection plane voxels forming a regular grid with spacing 2. This is closer to the measured datasets we will examine in section 6. All acceleration factors are defined with respect to the total number of locations of the regular grid which is given by  $M = 128^2/2^2 = 4096$ . Note that it now also makes sense to consider rSP-1 and sHd-1 as ‘sub-sampling’ pattern: although  $M_c = M$ , the data they measure cannot be converted to the conventionally sampled data, as was the case with Tumor1.

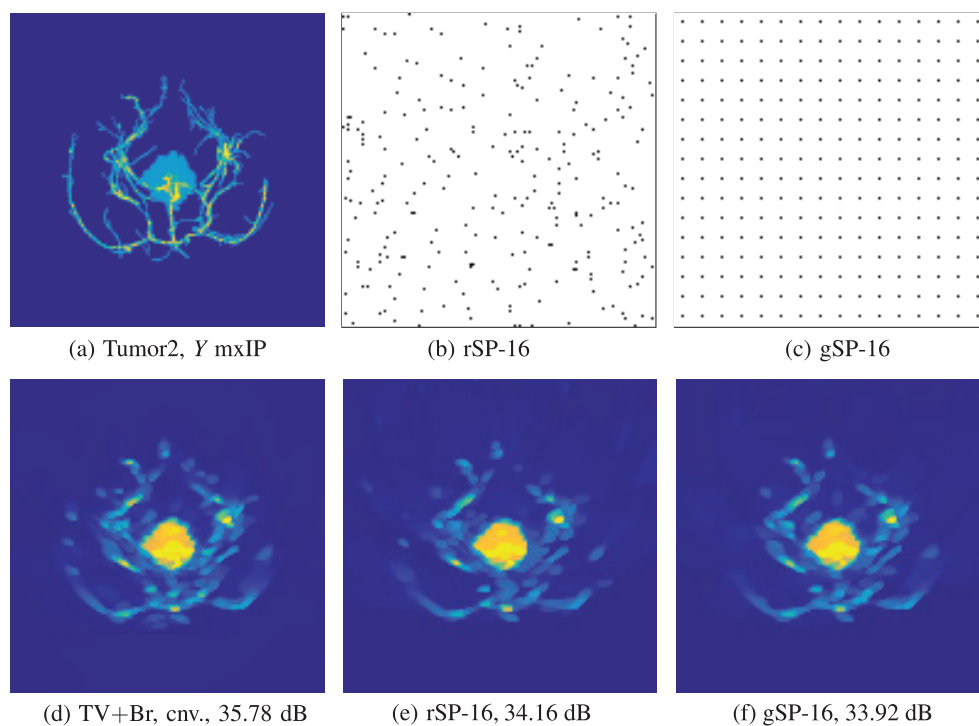
**5.3.3. Reconstruction results.** As the results of TR and BP are of similar quality as for Tumor1 (see figure 4) we omit them here, and concentrate on the results of the variational methods. We again used the DP ( $\kappa = 1.25$ ) to select the regularization parameter. Figure 8 compares them for rSP-16 and sHd-16 sub-sampling: TV+Br, again, leads to the best results by visual impression and PSNR. In figure 9, we therefore examine the influence of  $M_{\text{sub}}$  on the reconstructed images only for TV+Br: up to  $M_{\text{sub}} = 16$ , the rSP-based TV+Br reconstructions only slightly deteriorate. From  $M_{\text{sub}} = 16$  to  $M_{\text{sub}} = 32$ , however, a clear degradation is visible. For the sHd sub-sampling, the image quality remains acceptable up to  $M_{\text{sub}} = 32$ .



**Figure 8.** Tumor2 results (mxIP) of variational methods: (a)–(c) see figure 4 (d)–(l): L2+ TV+ and TV+Br results for conventional data (left column), rSP-16 (middle column) and sHd-16 (right column) and their corresponding PSNR in dB.



**Figure 9.** TV+Br results for Tumor2 phantom, comparison between different sub-sampling factors  $M_{\text{sub}}$ .



**Figure 10.** Influence of the spatial sub-sampling pattern in single point sub-sampling: (a) phantom (b)–(c) visualizations of rSP-16 and gSP-16 (d)–(f) TV+Br reconstructions.

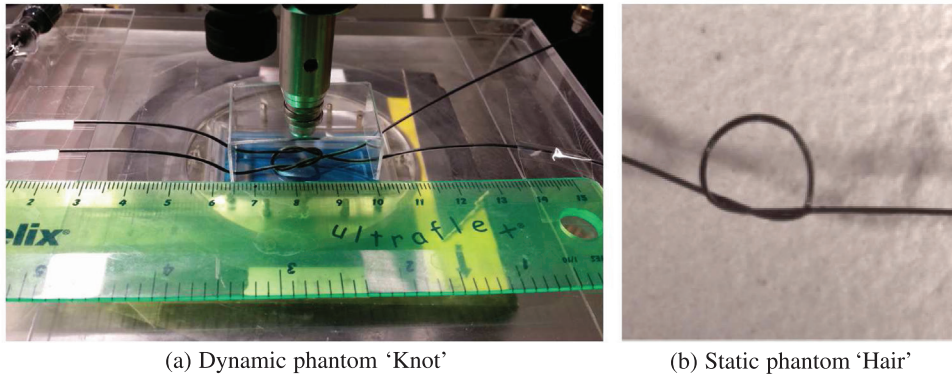
**5.3.4. Influence of the spatial sub-sampling pattern.** As described in section 3.1.1, a random partition of the scanning locations was chosen as the main single point sub-sampling pattern to avoid unintended systematic artifacts like aliasing and was furthermore inspired by several results in compressed sensing theory (Foucart and Rauhut 2013). In figure 10, we compare this random pattern to using a regular sub-sampling pattern based on a coarse grid, which we denote as gSP- $M_{\text{sub}}$ . The results show that the concrete choice of the single point sub-sampling pattern seems to be a minor influence, compared to, e.g. the choice of the inverse method used. We leave a more detailed examination and the design of optimal (dynamic) sampling pattern for further research.

## 6. Experimental data

In this section, we examine the sub-sampling strategies for three example experimental data sets. To have a ground truth, sub-sampling was only carried out artificially: for each experiment, a conventional data set was acquired first, and sub-sampled data sets were produced from this data thereafter.

### 6.1. Single point sub-sampling—dynamic phantom

We start with data acquired by a conventional single point FP scanner, measuring a pseudo-dynamic experimental phantom, which we call ‘Knot’.



**Figure 11.** Experimental phantoms. (a) Dynamic phantom 'Knot'. (b) Static phantom 'Hair'.

**6.1.1. Setup.** Figure 11(a) shows the experimental setup: two polythene tubes were filled with 10% and 100% ink and interleaved to form a knot. One of the loose ends was tied to a motor shaft (top right corner of figure 11(a)) while the other three ends were fixated. The pseudo-dynamic data was acquired in a stop-motion style: a conventional scan (duration  $\sim 15$  min) was performed while the whole arrangement was fixed. Then, the motor shaft was rotated by a stationary angle, causing the knot to tighten and to move into the direction of the motor, and a new scan was performed. This way, a conventional data set comprising 45 frames was acquired. The tubes were immersed in a 1% Intralipid solution with de-ionised water. The excitation laser pulses used had a wavelength of 1064 nm, an energy of around 20 mJ and were delivered a rate of 20 Hz. A conventional scan consisted of  $134 \times 133$  locations ( $\delta_x = \delta_y = 150 \mu\text{m}$ ), measured over  $M_t = 625$  time points with a resolution of  $\delta_t = 12$  ns.

**6.1.2. Preprocessing.** First, the data was clipped to  $132 \times 132$  locations. The baseline of the pressure-time course at each location was estimated by the median of the pre-excitation-pulse time points (1–4) and subtracted. Then, a zero-phase band-pass filter around 0.5–20 MHz was applied (see `applyFilter.m` in the k-Wave toolbox). The 1% locations with the highest variance (computed by the time points 7–13) were excluded from the further analysis. Finally, the remaining data was clipped to the time points 10–400.

**6.1.3. Results.** Table 2 shows the parameters of the acoustic model used for the inversion. Note that the spacing of the spatial grid,  $\Delta_{x/y/z}$ , is 2 times finer than the distance between scanning locations: assuming a sound speed of  $1540 \text{ m s}^{-1}$  the Nyquist criterion would require that we had sampled with  $\delta_{y/z} = 38.5 \mu\text{m}$  in space and  $\delta_t = 25$  ns in time to be able to reconstruct initial pressure distributions leading to signals with a frequency content of 20 MHz. This means that the conventional data is over-sampled in time, but already under-sampled in space, similar to the simulation studies with Tumor2 (see section 2.2). When choosing a finer spatial grid spacing to reconstruct the data as compared to the one in which it was recorded we attempt to recover some spatial resolution from the higher temporal resolution of the pressure time series.

In 4D PAT, we can vary the sub-sampling operator  $C$  used in each frame  $i$ . For single-point sub-sampling, one would try to avoid measuring the pressure time series at the same location in subsequent frames as it may contain very similar information<sup>3</sup>. Therefore, we randomly

<sup>3</sup>For patterned interrogation, one would not use the same pattern  $\phi_j$  in subsequent frames.

partitioned the set of all scanned locations into  $M_{\text{sub}}$  subsets, each containing  $M_c = M/M_{\text{sub}}$  different random locations. This yields a sequence of  $M_{\text{sub}}$  sub-sampling operators  $C_i$ ,  $i = 1, \dots, M_{\text{sub}}$  that we periodically apply to the set of all 45 frames, i.e. after  $M_{\text{sub}}$  frames, each locations has been scanned once and the  $M_{\text{sub}} + 1$ -th frame is scanned with  $C_1$ , again. Figure 12 shows the inverse reconstructions of the middle frame 23 for conventional data, rSP-4, rSP-8 and rSP-16 (a movie of the complete reconstruction can be found in the supplementary material ([stacks.iop.org/PMB/61/8908/mmedia](http://stacks.iop.org/PMB/61/8908/mmedia))). Next to TR, TV+ and TV+Br, we also show the result of post-processing the TR solution  $p_{\text{TR}}$  with positivity-constrained TV denoising, which we will denote by ‘TRppTV+’:

$$p_{\text{TRppTV+}} := \underset{p \geq 0}{\operatorname{argmin}} \left\{ \frac{1}{2} \|p - p_{\text{TR}}\|_2^2 + \lambda_{\text{ppTV}} \text{TV}(p) \right\}, \quad (18)$$

where we chose  $\lambda_{\text{pp}}$  large enough to suppress most visible reconstruction artifacts. Solving (18) is discussed in appendix B. The regularization parameter chosen for TV+ was  $\lambda_{\text{TV+}} = 0.02$  for the conventional data while this value was multiplied by  $4/M_{\text{sub}}$  for the sub-sampled data. For TV+Br, we carried out 10 Bregman iterations with  $\lambda_{\text{TV+Br}} = 12.5\lambda_{\text{TV+}}$ .

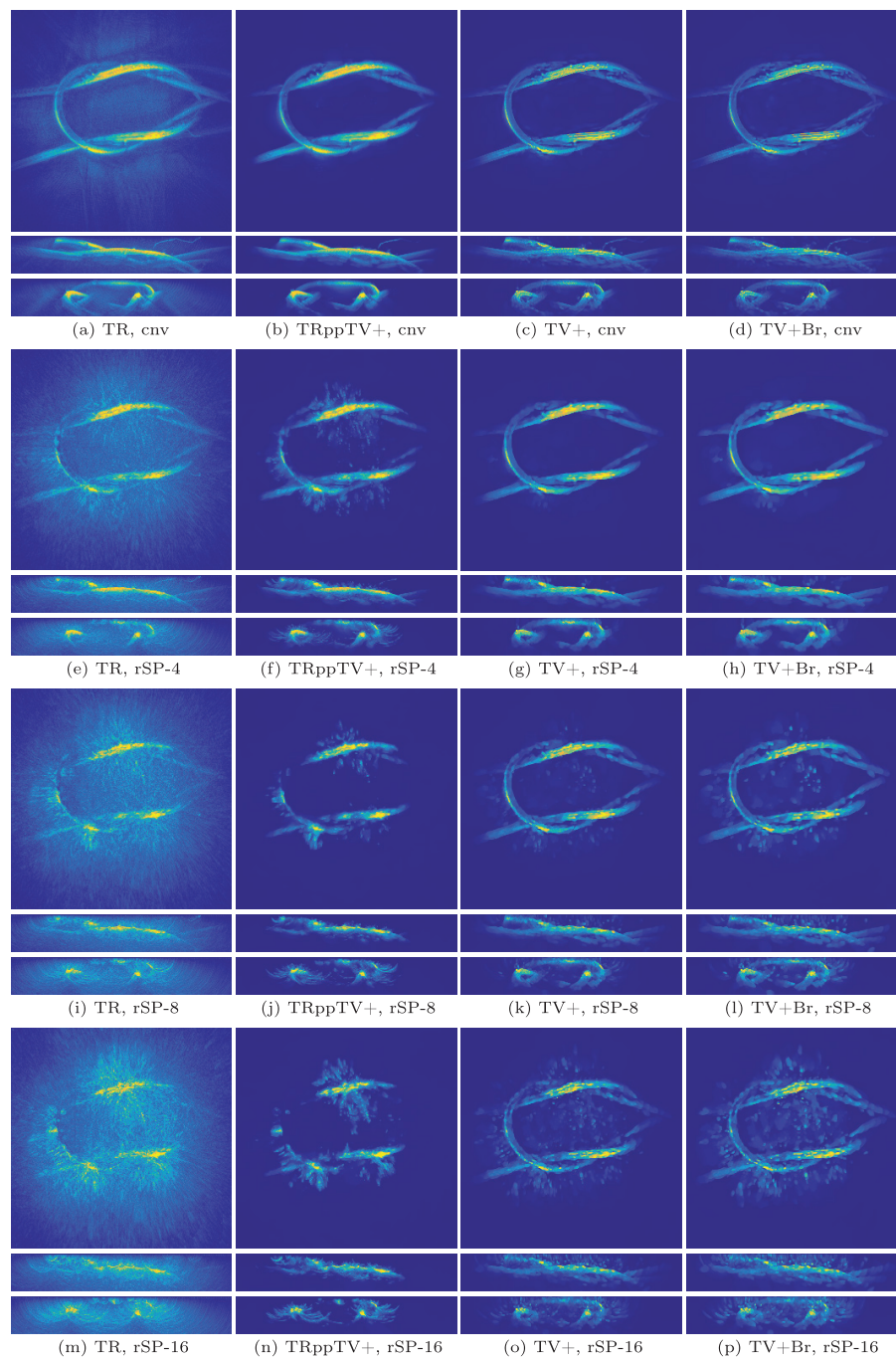
## 6.2. Patterned interrogation—static phantom

Next, we investigate data acquired by the patterned interrogation FP scanner (see figure 1(b)), measuring a static experimental phantom, which we will call ‘Hair’.

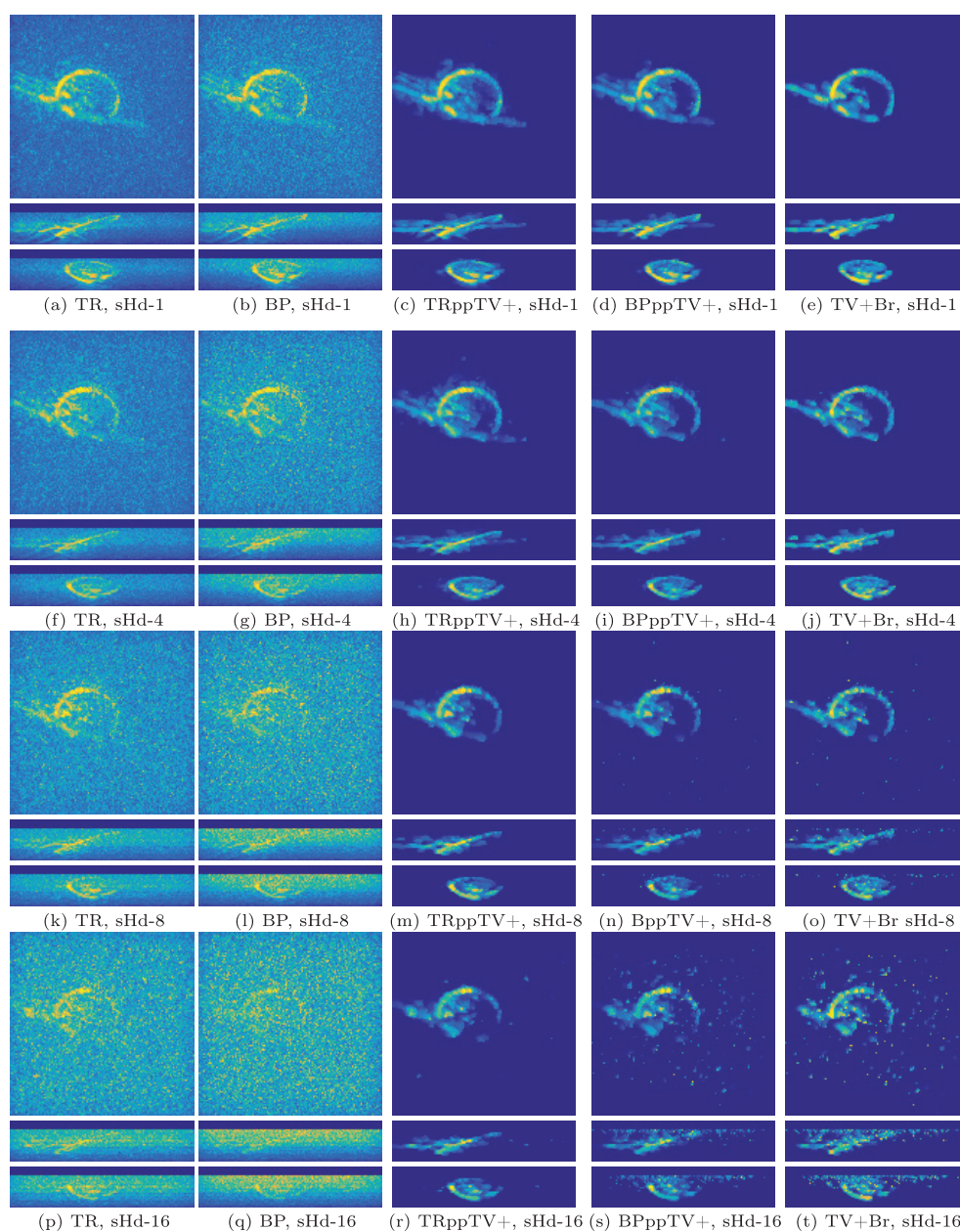
**6.2.1. Setup.** The full technical details of the scan can be found in Huynh *et al* (2014). The target to be scanned was a knotted artificial hair (see figure 11(b), diameter  $\sim 150 \mu\text{m}$ ), immersed in 1% Intralipid solution and positioned approximately 2 mm above the detection plane and 3 mm under the Intralipid surface. On the DMD, an active area of  $640 \times 640$  micromirrors was subdivided by grouping  $5 \times 5$  micromirrors to form one of  $128 \times 128$  ‘pixels’. Due to an angle in the optical path, each pixel corresponded to an area of  $62.12 \times 68 \mu\text{m}$  on the detection plane. Then, the rows of a  $16384 \times 16384$  scrambled Hadamard matrix were used to implement  $128 \times 128$  binary pattern on the DMD.

**6.2.2. Preprocessing.** First, the data needed to be calibrated to fit to our model (see section 4.3). Subsequently, a zero-phase band-pass filter around 3–22.5 MHz was applied and the data was clipped to the time points 85–140.

**6.2.3. Results.** Table 2 shows the parameters of the acoustic model used for the inversion. Note that the conventional data is, again, slightly over-sampled in time but under-sampled in space. For all computed solutions  $p$ , all voxels in the first 6 x-layers were set to 0 in a post-processing step. The latter was done to ease the visualization of the results through maximum intensity projections: the first 12 time points of the signal only seem to contain noise, which means that up to a distance of  $12c_0\Delta_t/\Delta_x = 5.8$  in voxel length,  $p_0$  will only account for noise. For TR and BP solutions, voxels with negative values were also set to 0. Figure 13 shows different reconstructions using the conventional data (sHd-1) and  $M_{\text{sub}} = 4, 8, 16$ : the regularization parameter chosen for TV+Br was  $\lambda_{\text{TV+Br}} = 1.5 \cdot 10^{-4}$  for the conventional data while this value was first divided by  $M_{\text{sub}}$  and then multiplied by  $1/1.2/1.4$  for  $M_{\text{sub}} = 4/8/16$  for the sub-sampled data. A total of 10 Bregman iterations were carried out.



**Figure 12.** Results for frame 23 of the data set Knot: in each sub-figure, maximum intensity projections in X (top), Y (middle) and Z (bottom) direction are shown.



**Figure 13.** Results for data set Hair: in each sub-figure, maximum intensity projections in X (top), Y (middle) and Z (bottom) direction are shown.

### 6.3. *In vivo* measurements—single point sub-sampling

While validating inverse methods on data from experimental phantoms is an important step forward from pure simulation studies, experimental phantoms cannot reproduce all the features of real *in vivo* data sets. As a last example, we therefore investigate a static *in vivo* data set of skin vasculature and subcutaneous anatomy near the right flank of a nude mouse, which we will call ‘Vessels’.

**6.3.1. Setup.** The data was acquired with an excitation wavelength of 590 nm, further technical details and illustrations can be found in Jathoul *et al* (2015). A conventional scan consisted of  $142 \times 141$  locations over a region of size of  $14 \text{ mm} \times 14 \text{ mm}$  ( $\delta_x = \delta_y = 100 \mu\text{m}$ ), measured at  $M_t = 630$  time points with a resolution of  $\delta_t = 10 \text{ ns}$ .

**6.3.2. Preprocessing.** The data was clipped to  $141 \times 141$  locations and to the time points 10–630.

**6.3.3. Results.** Table 2 shows the parameters of the acoustic model used for the inversion. Note that by a similar reasoning about the spatial and temporal sampling intervals, the spatial spacing  $\Delta_{x/y/z}$  is, again, chosen 2 times finer than the distance between scanning locations. Figure 14 shows maximum intensity projections and a slice through  $z = 74$  for TR, TRppTV+ and TV+Br solutions when using the conventional, rSP-4 and rSP-8 data. The denoising parameter  $\lambda_{pp}$ , was, again, chosen large enough to suppress most visible reconstruction artifacts. The regularization parameter chosen for TV+Br was  $\lambda_{TV+Br} = 6.25 \cdot 10^{-2}$  for the conventional data while this value was multiplied by  $4/M_{\text{sub}}$  for the sub-sampled data. A total of 10 Bregman iterations were carried out.

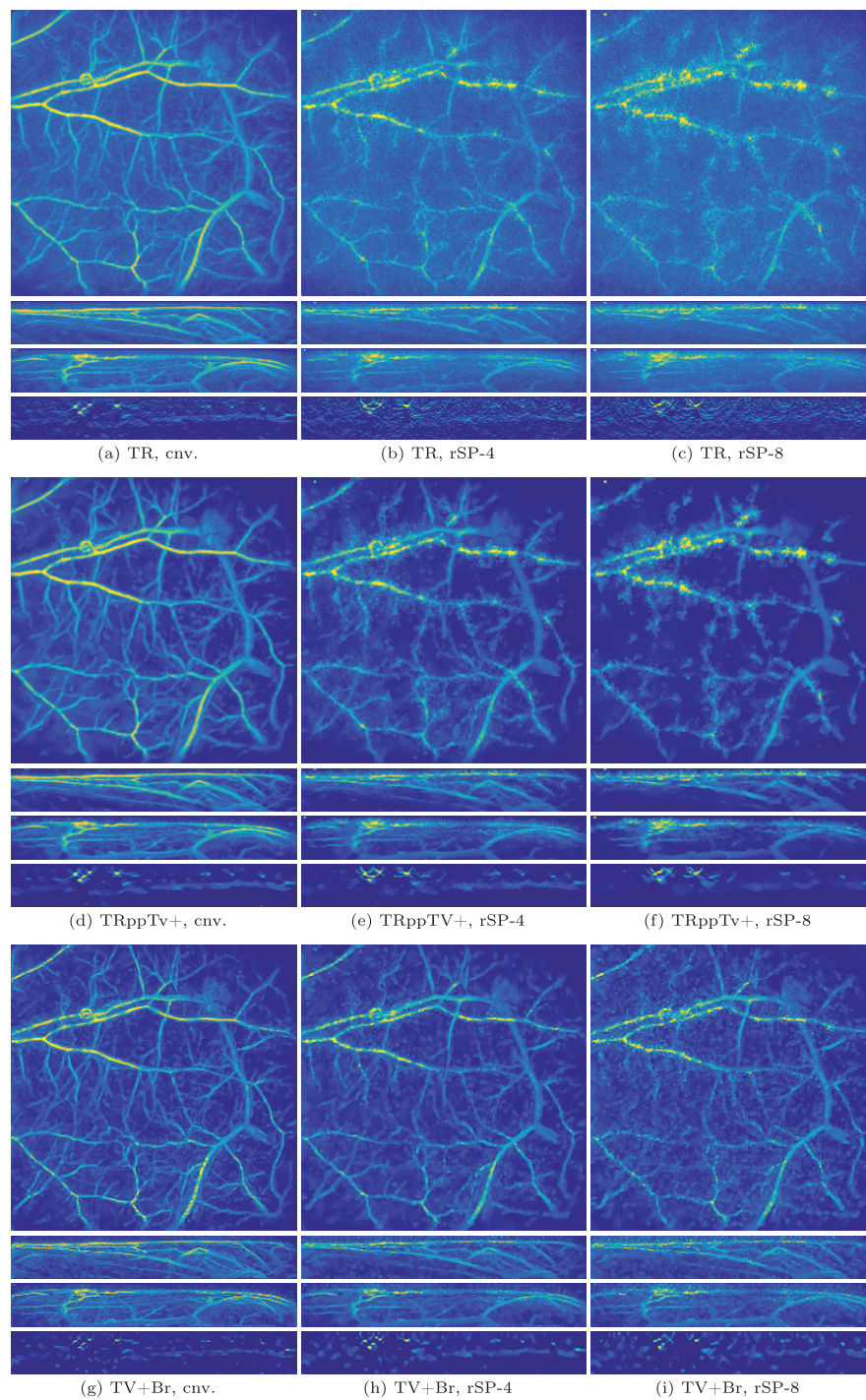
## 7. Discussion, outlook and conclusion

### 7.1. Discussion

The main results of the simulation studies (section 5) can be summarized as follows.

- Using model-based variational reconstruction methods employing spatial sparsity constraints, such as TV+, (9), is essential for obtaining good quality PA images from sub-sampled PAT data. The linear methods, TR (6) and BP (5), and the L2+ method (8) could not produce images of acceptable quality in any setting.
- The results obtained for Tumor1 and Tumor2 demonstrate that the image quality obtained for a certain sub-sampling rate  $M_{\text{sub}}$  varies strongly: the ‘inverse crime’ data of the more superficial, high contrast target Tumor1 can be up-sampled up to  $M_{\text{sub}} = 128$  without a significant loss of image quality. On the other hand, a bad *model-fit*, i.e. a mismatch between of the models used for data generation and inversion, combined with a more challenging target, such as Tumor2, significantly impairs the image quality beyond a certain sub-sampling rate ( $M_{\text{sub}} = 16$ –32 in our particular example).
- Using Bregman iterations improves upon conventional variational approaches for PAT. Most importantly, the systematic contrast loss of small scale structures such as blood vessels is mitigated which is a crucial prerequisite for QPAT studies.
- Sub-sampling by patterned interrogation (sHd) was slightly more efficient than single point sub-sampling (rSP).

The sub-sampling rates we achieved with experimental data were similar to those obtained with simulated data in the more realistic Tumor2 scenario. However, unexpectedly, the qualitative difference between the linear methods TR and BP and the variational methods TV+ and TV+Br was not as dramatic as in the simulation studies (see figures 4, 5, 12 and 13). In many cases, post-processing the TR solution with TV+ denoising, (18), comes remarkably close to the TV+ solution, which is not to be expected from the simulation studies (see, e.g. the TR sHd-128 solution in figure 4). In addition, Bregman iterations could not improve much over conventional variational approaches,



**Figure 14.** Results for *in vivo* data set Vessels: in each sub-figure, from top to bottom: Maximum intensity projections in  $X,Y,Z$  direction and slice through  $z = 74$ .

again in contrast to our findings in section 5.2.1. A partial explanation for both findings is a bad model fit: while some pre-processing routines (e.g. baseline correction, band-pass filtering and a detection and deletion of corrupt channels) were implemented and carried out to align the data with the model used, other known model-mismatches (e.g. the inhomogeneous sensitivity of the FP sensor and the non-whiteness of the measurement noise) were not accounted for.

A closer examination of the ‘Knot’ data reveals several flaws which deteriorated our results: firstly, the optical excitation was inhomogeneous in lateral direction (see figure 12), leading to an inhomogeneous initial pressure distribution in regions consisting of the same materials. The TV energy is not well-suited to recover such targets. Secondly, the baseline shifts in the data are more complex than what we corrected for and a spatio-temporal visualization of the data shows several artifacts on the sensor that our automatic channel deletion procedure cannot fully remove. And lastly, the acoustic properties of the polythene tubes lead to reflections we did not account for in our model. For these reasons, the decrease in image quality for sub-sampled data is faster compared to the simulation studies. While we see a clear advantage of using TV+ as opposed to the simpler TRppTV+, using Bregman iterations does not seem to lead to a better image quality.

The ‘Hair’ data was acquired with the novel patterned FP scanner which still suffers from several major technical difficulties, e.g. the widening of the interrogation beam leads to a significant loss in SNR and systematic shifts in signal power can be observed that need to be examined more carefully. Furthermore, we suspect that the hair knot might have moved during the acquisition (see figure 13(c)). For these reasons, already the reconstructions from the sHd-1 data are not of very good quality. If we accept these as a ground truth nonetheless, we see that we can reach sub-sampling rates of around  $M_{\text{sub}} = 8$  without a significant further loss of image quality. If we compare the different methods, we find that TRppTV+ yields the visually most appealing results while BPppTV+ and TV+Br look very similar. As BPppTV+ is more or less the first iterate of the optimization scheme we use to compute TV+Br (see appendix B), the latter might, again, point to a bad model-fit (see above).

For the ‘Vessel’ data set, the visual impression of the conventional data reconstructions (see figure 14) and the fact that we did not have to perform any pre-processing suggest that the *in vivo* data examined is of good quality and we have a good model-fit. A comparison with the results from experimental phantoms further reveals that the diffusive nature of biological tissue leads to a more even lateral illumination. Now, TV+Br clearly outperforms TR and TRppTV+. For instance, from the slice view we can see that TR seems to overestimate the diameter of the blood vessels. The reason for not achieving high sub-sampling rates despite the good data quality is the apparent mismatch between the target geometry and the spatial sparsity constraints employed by the TV energy: the PA image is exceptionally rich in vasculature, but TV regularization tends to break up such anisotropic, line-like structures (see figure 14(i)).

## 7.2. Outlook

As the simulation studies show that a model misfit can severely decrease the sub-sampling rates achievable, we need to improve the accuracy of the acoustic forward model to obtain better results for experimental data: data pre-processing aligns the data with the forward

model, *model calibration* can determine some uncertain parameters of the forward model (such as the FP sensitivity distribution or the noise statistics), and Bayesian *model selection* (Toussaint 2011) or Bayesian *approximation error modeling* (Arridge *et al* 2006, Kaipio and Somersalo 2007, Tarvainen *et al* 2013) can reduce or account for the uncertainty in other parameters, such as  $c_0$ . To improve the results for *in vivo* data (see figure 14) acoustic absorption models of biological tissue (Treeby *et al* 2010, Treeby 2013) need to be incorporated.

While we exploited spatial sparsity to accelerate the acquisition of a single scan here, the next step to enable 4D PAT imaging with both high spatial and temporal resolution (see section 6.1) is to extend the frame-by-frame inversion methods examined here to *full spatio-temporal* variational models that also exploit the temporal redundancy of data generated by dynamics of low complexity.

We used the TV energy as a generic, well-understood first example of a spatial sparsity constraint. However, as discussed above, it is, e.g. not very suitable to recover thin vasculature. Higher sub-sampling rates could be reached by employing more sophisticated regularization functionals designed to recover such anisotropic structures (Kutyniok and Lim 2012).

Choosing random locations as single-point sub-sampling and scrambled Hadamard pattern as patterned interrogation was based on results obtained for similar applications such as CT and MRI. The optimal choice for PAT applications is yet to be determined. For instance, Schmid *et al* (2016) has shown that a non-uniform distribution of the sampling locations in single-point sub-sampling can be used to focus into a specific area (at the expense of the resolution elsewhere). A theoretical examination, e.g. through *micro-local analysis* (Friel and Quinto 2015), could help to gain new insights on this.

### 7.3. Conclusion

In this study, we investigated different possibilities to sub-sample the incident photoacoustic field in order to accelerate the acquisition of high resolution PAT. In simulation studies, we demonstrated that PAT wave fields generated by targets with a low spatial complexity can indeed be highly compressible and identified under which conditions this feature can be exploited to obtain high quality images from highly sub-sampled data: firstly, variational image reconstruction methods employing sparsity constraints that match the structure of the target have to be used. Secondly, using an accurate forward model well-aligned with the data is crucial. We furthermore applied the methods developed to three experimental data sets from experimental phantoms and *in vivo* recordings. While obtaining promising first results, we also identified several challenges for realizing the full potential of data sub-sampling, most notably obtaining a good model-fit as discussed above. While we focused on sub-sampling the data using the Fabry–Pérot based scanner, the novel reconstruction strategies offer new opportunities to dramatically increase the acquisition speed of other photoacoustic scanners that employ point-by-point sequential scanning as well as reducing the channel count of parallelized schemes that use detector arrays.

### Acknowledgments

We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Tesla K40 GPU used for this research.

The numerical phantom is based upon  $\mu$ CT data that was kindly provided by Simon Walker-Samuel at the Centre for Advanced Biomedical Imaging, University College

London, and a segmentation thereof kindly provided by Roman Hochuli, University College London.

This work was supported by the Engineering and Physical Sciences Research Council, UK (EP/K009745/1), the European Union project FAMOS (FP7 ICT, Contract 317744) and the National Institute of General Medical Sciences of the National Institutes of Health under grant number P41 GM103545-18.

## Appendix A. Discrete total variation energy

Let the voxels of the 3D pressure  $p \in \mathbb{R}^N$ , with  $N = N_x N_y N_z$  be indexed by  $(i, j, k)$ ,  $i = 1, \dots, N_x$ ,  $j = 1, \dots, N_y$ ,  $k = 1, \dots, N_z$ . Using finite forward differences, the most commonly used discretization of the total variation seminorm with Neumann boundary conditions is given by

$$\text{TV}(p) = \sum_{(i,j,k)} \sqrt{(p_{(i+1,j,k)} - p_{(i,j,k)})^2 + (p_{(i,j+1,k)} - p_{(i,j,k)})^2 + (p_{(i,j,k+1)} - p_{(i,j,k)})^2},$$

where  $p_{(N_x+1,j,k)} := p_{(N_x,j,k)}$ ,  $p_{(i,N_y+1,k)} := p_{(i,N_y,k)}$  and  $p_{(i,j,N_z+1)} := p_{(i,j,N_z)}$ . Additional terms have to be added to account for different boundary conditions: in the simulation studies, we use Dirichlet boundary conditions which requires to add the jumps at the domain boundaries. For the experimental data, we impose Dirichlet boundary conditions at the detection plane voxels  $x = 0$  while Neumann boundary conditions are applied on all other faces of the image cube.

## Appendix B. Optimization

The optimization problem given by (7) consists of the minimizing the sum of two functionals  $\mathcal{E}(p) = \mathcal{D}(p) + \lambda \mathcal{J}(p)$ , where we can compute the gradient of the strictly convex, smooth functional  $\mathcal{D}(p)$ ,

$$\mathcal{D}(p) = \frac{1}{2} \|CAp - f^c\|_2^2, \quad \nabla \mathcal{D}(p) = A^T C^T (CAp - f^c), \quad (\text{B.1})$$

but no higher derivatives, and we know how to compute the *proximal operator* of the convex, potentially non-differentiable functional  $\mathcal{J}(p)$ :

$$\text{prox}_{\mathcal{J},\alpha}(p) := \underset{q}{\operatorname{argmin}} \left\{ \mathcal{J}(q) + \frac{1}{2\alpha} \|q - p\|_2^2 \right\}. \quad (\text{B.2})$$

In the case of  $\mathcal{J}(p)$  being the positivity-constrained TV energy, the proximal operators simply solves a positivity-constrained TV denoising problem (18).

A wide range of *semi-smooth, first order optimization* algorithms for image reconstruction have been developed over recent years (Burger *et al* 2014), each of them advantageous for a specific scenario. In our case, the special feature of PAT is that the application of  $A$  and  $A^T$  requires considerably more computation time than solving most proximal operators (B.2), including the 3D positivity-constrained TV denoising problem (18), up to a high numerical precision. Under these circumstances, the rather simple *proximal gradient descent* scheme:

$$p^{k+1} = \text{prox}_{\mathcal{J},\eta\lambda}(p^k - \eta A^T C^T (CAp^k - f^c)), \quad p^0 = 0, \quad k = 1, \dots, K \quad (\text{B.3})$$

turns out to be most efficient if tuned carefully (see Goldstein *et al* (2014) for an extensive overview).

- The step-size  $\eta$  is set to  $1.8/L$ , where  $L$  is an approximation of the Lipschitz constant of  $A^T C^T C A$ . For a given setting and sub-sampling scheme,  $L$  can be computed quite efficiently by a simple power iteration and then stored in a look-up table.
- We use a gradient extrapolation modification (*fast or accelerated gradient methods*) ensuring a quadratic convergence. The concrete technique we use is the *FISTA* algorithm (Beck and Teboulle 2009), where we restart the acceleration if an increase in  $\mathcal{E}(p^k)$  is detected and switch to a normal gradient for this iterations  $k$ , followed by up to 5 back-tracking steps if necessary ( $\eta$  is, however, not changed for future iterations).
- For the 3D PAT problems considered in this work, the iterates from  $k = 11$  onwards are usually visually not distinguishable. However, we computed a maximum of  $K = 50$  iterations for all except for Knot, where we only computed  $K = 20$  iterations. We terminated the iteration earlier if  $p^k$  did not change or  $\min_k \mathcal{E}(p^k)$  did not decrease for 5 times in a row.

The proximal operator for (8) can be computed component-wise and explicitly:

$$\tilde{p} = \operatorname{argmin}_{q \geq 0} \left\{ \|q\|_2^2 + \frac{1}{2\alpha} \|q - p\|_2^2 \right\} \Leftrightarrow \tilde{p}_i = \max\left(0, \frac{q_i}{1 + \alpha}\right). \quad (\text{B.4})$$

The positivity-constrained TV denoising is implemented by a *primal-dual hybrid gradient* algorithm as described in Chambolle and Pock (2011).

## Appendix C. Implementation

All routines have been implemented in as part of a larger, Matlab toolbox for PAT image reconstruction which will be made available in near future. The toolbox relies on the k-Wave toolbox (see Treeby and Cox (2010), [www.k-wave.org/](http://www.k-wave.org/)) to implement  $A$  and  $A^T$ , which allows to use highly optimized C++ and CUDA code to compute the 3D wave propagation on parallel CPU or GPU architectures. To give an idea about the range of different computations times, computing one application of  $A$  for the *in vivo* Vessels scenario (see table 2) in single precision takes 15 s using the optimized CUDA code on a Tesla K40 GPU (counting only the GPU run-time), 51 s using the Matlab code on the same GPU, 47 s/6 min 36 s using the optimized C++ code on 12/1 cores of an Intel Xeon CPU (2.70 GHz) (counting only the CPU run-time) and 4 min 3 s/26 min 48 s using the Matlab code on 12/1 cores of the same CPU.

## References

- Arridge S R, Betcke M M, Cox B, Lucka F and Treeby B 2016 On the adjoint operator in photoacoustic tomography *Inverse Problems* **32** 115012
- Arridge S R, Kaipio J P, Kolehmainen V, Schweiger M, Somersalo E, Tarvainen T and Vauhkonen M 2006 Approximation errors and model reduction with an application in optical diffusion tomography *Inverse Problems* **22** 175
- Arridge S R and Scherzer O 2012 Imaging from coupled physics *Inverse Problems* **28** 080201
- Beard P 2011 Biomedical photoacoustic imaging *Interface Focus* **1** 602–31
- Beck A and Teboulle M 2009 Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems *IEEE Trans. Image Process.* **18** 2419–34
- Benning M 2011 Singular regularization of inverse problems *PhD Thesis* University of Muenster
- Benning M, Gladden L, Holland D, Schönlieb C and Valkonen T 2014 Phase reconstruction from velocity-encoded MRI measurements—a survey of sparsity-promoting variational approaches *J. Magn. Reson.* **238** 26–43

- Betcke M M, Cox B T, Huynh N, Zhang E Z, Beard P C and Arridge S R 2016 Acoustic wave field reconstruction from compressed measurements with application in photoacoustic tomography arXiv:[1609.02763](#)
- Bregman L M 1967 The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming {USSR} *Computat. Math. Math. Phys.* **7** 200–17
- Brune C, Sawatzky A and Burger M 2011 Primal and dual Bregman methods with application to optical nanoscopy *Int. J. Comput. Vis.* **92** 211–29
- Burger M and Osher S 2013 A guide to the TV zoo *Level Set and PDE Based Reconstruction Methods in Imaging (Lecture Notes in Mathematics)* (Berlin: Springer) pp 1–70
- Burger M, Sawatzky A and Steidl G 2014 First order algorithms in variational image processing (arXiv:[1412.4237](#)) 60
- Candes E, Romberg J and Tao T 2006 Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information *IEEE Trans. Inf. Theory* **52** 489–509
- Cao M, Yuan J, Du S, Xu G, Wang X, Carson P L and Liu X 2015 Full-view photoacoustic tomography using asymmetric distributed sensors optimized with compressed sensing method *Biomed. Signal Process. Control* **21** 19–25
- Chambolle A and Pock T 2011 A first-order primal-dual algorithm for convex problems with applications to imaging *J. Math. Imaging Vis.* **40** 120–45
- Cox B, Laufer J G, Arridge S R and Beard P C 2012 Quantitative spectroscopic photoacoustic imaging: a review *J. Biomed. Opt.* **17** 061202
- Cox B T, Kara S, Arridge S R and Beard P C 2007 k-space propagation models for acoustically heterogeneous media: application to biomedical photoacoustics *J. Acoust. Soc. Am.* **121** 3453–64
- Do T, Gan L, Nguyen N and Tran T 2012 Fast and efficient compressive sensing using structurally random matrices *IEEE Trans. Signal Process.* **60** 139–54
- Donoho D L 2006 Compressed sensing *IEEE Trans. Inf. Theory* **52** 1289–306
- Duarte M, Davenport M, Takhar D, Laska J, Sun T, Kelly K and Baraniuk R 2008 Single-pixel imaging via compressive sampling *IEEE Signal Process. Mag.* **25** 83–91
- Finch D and Patch S K 2004 Determining a function from its mean values over a family of spheres *SIAM J. Math. Anal.* **35** 1213–40
- Foucart S and Rauhut H 2013 *A Mathematical Introduction to Compressive Sensing* (Basel: Birkhäuser)
- Friel J and Quinto E T 2015 Artifacts in incomplete data tomography with applications to photoacoustic tomography and sonar *SIAM J. Appl. Math.* **75** 703–25
- Goldstein T and Osher S 2009 The Split Bregman method for L1-regularized problems *SIAM J. Imaging Sci.* **2** 323–43
- Goldstein T, Studer C and Baraniuk R 2014 A field guide to forward-backward splitting with a FASTA implementation arXiv:[1411.3406](#)
- Guo Z, Li C, Song L and Wang L V 2010 Compressed sensing in photoacoustic tomography *in vivo* *J. Biomed. Opt.* **15** 021311
- Hämäläinen K, Kallonen A, Kolehmainen V, Lassas M, Niinimäki K and Siltanen S 2013 Sparse tomography *SIAM J. Sci. Comput.* **35** B644–65
- Huang C, Wang K, Nie L, Wang L and Anastasio M 2013 Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media *IEEE Trans. Med. Imaging* **32** 1097–110
- Huynh N, Zhang E, Betcke M, Arridge S, Beard P and Cox B 2014 Patterned interrogation scheme for compressed sensing photoacoustic imaging using a Fabry P  rot planar sensor *Proc. SPIE* **8943** 894327
- Huynh N, Zhang E, Betcke M, Arridge S, Beard P and Cox B 2015 A real-time ultrasonic field mapping system using a Fabry-P  rot single pixel camera for 3D photoacoustic imaging *Proc. SPIE* **9323** 932310
- Huynh N, Zhang E, Betcke M, Arridge S, Beard P and Cox B 2016 Single-pixel optical camera for video rate ultrasonic imaging *Optica* **3** 26–9
- Jathoul A P et al 2015 Deep *in vivo* photoacoustic imaging of mammalian tissues using a tyrosinase-based genetic reporter *Nat. Photon.* **9** 239–46
- Jin X and Wang L V 2006 Thermoacoustic tomography with correction for acoustic speed variations *Phys. Med. Biol.* **51** 6437
- J  rgensen J S and Sidky E Y 2015 How little data is enough? Phase-diagram analysis of sparsity-regularized x-ray computed tomography *Phil. Trans. R. Soc. A* **373** 2043

- Kaipio J P and Somersalo E 2007 Statistical inverse problems: discretization, model reduction and inverse crimes *J. Comput. Appl. Math.* **198** 493–504
- Kuchment P and Kunyansky L 2011 Photoacoustic and thermoacoustic tomography: image formation principles *Handbook of Mathematical Methods in Imaging* ed O Scherzer (New York: Springer) pp 817–65
- Kutyniok G and Lim W Q 2012 Chapter: Image separation using wavelets and shearlets *Curves and Surfaces: 7th Int. Conf. (Avignon, France, 24–30 June 2010)* (Berlin: Springer) pp 416–30 (revised selected papers)
- Liang D, Zhang H F and Ying L 2009 Compressed-sensing photoacoustic imaging based on random optical illumination *Int. J. Funct. Inf. Personalised Med.* **2** 394–406
- Lustig M, Donoho D and Pauly J M 2007 Sparse MRI: the application of compressed sensing for rapid MR imaging *Magn. Reson. Med.* **58** 1182–95
- Lutzwiler C and Razansky D 2013 Optoacoustic imaging and tomography: reconstruction approaches and outstanding challenges in image performance and quantification *Sensors* **13** 7345
- Mast T, Souriau L, Liu D L, Tabei M, Nachman A and Waag R 2001 A k-space method for large-scale models of wave propagation in tissue *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **48** 341–54
- Meng J, Wang L V, Liang D and Song L 2012a *In vivo* optical-resolution photoacoustic computed tomography with compressed sensing *Opt. Lett.* **37** 4573–75
- Meng J, Wang L V, Ying L, Liang D and Song L 2012b Compressed-sensing photoacoustic computed tomography *in vivo* with partially known support *Opt. Express* **20** 16510–23
- Mueller J 2013 Advanced image reconstruction and denoising—bregmanized (higher order) total variation and application in PET *PhD Thesis* Institute for Computational and Applied Mathematics, University of Münster
- Nie L and Chen X 2014 Structural and functional photoacoustic molecular tomography aided by emerging contrast agents *Chem. Soc. Rev.* **43** 7132–70
- Osher S, Burger M, Goldfarb D, Xu J and Yin W 2006 An iterative regularization method for total variation-based image restoration *Multiscale Model. Simul.* **4** 460–89
- Provost J and Lesage F 2009 The application of compressed sensing for photo-acoustic tomography *IEEE Trans. Med. Imaging* **28** 585–94
- Ritschl L, Bergner F, Fleischmann C and Kachelrieß M 2011 Improved total variation-based CT image reconstruction applied to clinical data *Phys. Med. Biol.* **56** 1545
- Rudin L I, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms *Phys. D: Nonlinear Phenom.* **60** 259–68
- Sandbichler M, Krahmer F, Berer T, Burgholzer P and Haltmeier M 2015 A novel compressed sensing scheme for photoacoustic tomography *SIAM J. Appl. Math.* **75** 2475–94
- Scherzer O, Grasmair M, Grossauer H, Haltmeier M and Lenzen F 2009 *Variational Methods in Imaging* vol 320 (Berlin: Springer)
- Schmid J, Glatz T, Zabihian B, Liu M, Drexler W and Scherzer O 2016 Nonequispaced grid sampling in photoacoustics with a nonuniform fast Fourier transform *J. Biomed. Opt.* **21** 015005
- SCI Institute 2015 SCIRun: a scientific computing problem solving environment, Scientific Computing and Imaging Institute (SCI), download from: [www.scirun.org](http://www.scirun.org)
- Stefanov P and Uhlmann G 2009 Thermoacoustic tomography with variable sound speed *Inverse Problems* **25** 075011
- Sun M, Feng N, Shen Y, Shen X, Ma L, Li J and Wu Z 2011 Photoacoustic imaging method based on arc-direction compressed sensing and multi-angle observation *Opt. Express* **19** 14801–6
- Taruttis A and Ntziachristos V 2015 Advances in real-time multispectral optoacoustic imaging and its applications *Nat. Photon.* **9** 219–27
- Tarvainen T, Pulkkinen A, Cox B, Kaipio J and Arridge S 2013 Bayesian image reconstruction in quantitative photoacoustic tomography *IEEE Trans. Med. Imaging* **32** 2287–98
- Toussaint U 2011 Bayesian inference in physics *Rev. Mod. Phys.* **83** 943–99
- Treeby B E 2013 Acoustic attenuation compensation in photoacoustic tomography using time-variant filtering *J. Biomed. Opt.* **18** 036008
- Treeby B E and Cox B T 2010 k-wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave fields *J. Biomed. Opt.* **15** 021314
- Treeby B E, Zhang E Z and Cox B T 2010 Photoacoustic tomography in absorbing acoustic media using time reversal *Inverse Problems* **26** 115003
- Tremoulheac B, Dikaios N, Atkinson D and Arridge S 2014 Dynamic MR image reconstruction—separation from undersampled (k,t)-space via low-rank plus sparse prior *IEEE Trans. Med. Imaging* **33** 1689–701

- von Harbou E, Fabich H T, Benning M, Tayler A B, Sederman A J, Gladden L F and Holland D J 2015 Quantitative mapping of chemical compositions with MRI using compressed sensing *J. Magn. Reson.* **261** 27–37
- Wang L V 2009 Multiscale photoacoustic microscopy and computed tomography *Nat. Photon.* **3** 503–9
- Wang K and Anastasio M 2011 Photoacoustic and thermoacoustic tomography: image formation principles *Handbook of Mathematical Methods in Imaging* ed O Scherzer (New York: Springer) pp 781–815
- Xia J and Wang L V 2014 Small-animal whole-body photoacoustic tomography: a review *IEEE Trans. Biomed. Eng.* **61** 1380–9
- Xu Y and Wang L V 2004 Application of time reversal to thermoacoustic tomography *Biomedical Optics 2004 (Int. Society for Optics and Photonics)* (Bellingham, WA: SPIE) pp 257–63
- Yao J and Wang L V 2014 Photoacoustic brain imaging: from microscopic to macroscopic scales *Neurophotonics* **1** 011003
- Zackrisson S, van de Ven S M W Y and Gambhir S S 2014 Light in and sound out: emerging translational strategies for photoacoustic imaging *Cancer Res.* **74** 979–1004
- Zhang E, Laufer J and Beard P 2008 Backward-mode multiwavelength photoacoustic scanner using a planar Fabry–Pérot polymer film ultrasound sensor for high-resolution three-dimensional imaging of biological tissues *Appl. Opt.* **47** 561–77
- Zhang Y, Wang Y and Zhang C 2012 Total variation based gradient descent algorithm for sparse-view photoacoustic image reconstruction *Ultrasonics* **52** 1046–55