

## Research article

## Scale-equivariant deep model-based optoacoustic image reconstruction

Christoph Dehner<sup>a,b</sup>, Ledia Lilaj<sup>a</sup>, Vasilis Ntziachristos<sup>b,c,d</sup>, Guillaume Zahnd<sup>a,b,1</sup>,  
Dominik Jüstel<sup>b,c,e,\*</sup>

<sup>a</sup> iThera Medical GmbH, Munich, Germany

<sup>b</sup> Institute of Biological and Medical Imaging, Helmholtz Zentrum München, Neuherberg, Germany

<sup>c</sup> Chair of Biological Imaging at the Central Institute for Translational Cancer Research (TranslaTUM), School of Medicine and Health, Technical University of Munich, Munich, Germany

<sup>d</sup> Munich Institute of Robotics and Machine Intelligence (MIRMI), Technical University of Munich, Munich, Germany

<sup>e</sup> Institute of Computational Biology, Helmholtz Zentrum München, Neuherberg, Germany

## ARTICLE INFO

## Keywords:

Optoacoustic imaging  
Model-based reconstruction  
Regularization  
Scale-equivariance

## ABSTRACT

Model-based reconstruction provides state-of-the-art image quality for multispectral optoacoustic tomography. However, optimal regularization of in vivo data necessitates scan-specific adjustments of the regularization strength to compensate for fluctuations of the signal magnitudes between different sinograms. Magnitude fluctuations within in vivo data also pose a challenge for supervised deep learning of a model-based reconstruction operator, as training data must cover the complete range of expected signal magnitudes. In this work, we derive a scale-equivariant model-based reconstruction operator that *i*) automatically adjusts the regularization strength based on the  $L^2$  norm of the input sinogram, and *ii*) facilitates supervised deep learning of the operator using input sinograms with a fixed norm. Scale-equivariant model-based reconstruction applies appropriate regularization to sinograms of arbitrary magnitude, achieves slightly better accuracy in quantifying blood oxygen saturation, and enables more accurate supervised deep learning of the operator.

## 1. Introduction

Model-based reconstruction affords state-of-the-art image quality for multispectral optoacoustic tomography (MSOT): scanner impulse response modeling improves image contrast and resolution, non-negativity constraints on the reconstructed image ensure a meaningful interpretation as initial pressure, and a regularization functional mitigates the ill-posedness of the inversion [1–5]. Model-based reconstruction is computationally demanding but real-time imaging is achievable through supervised learning of the reconstruction operator with a deep neural network [6,7].

A remaining practical obstacle to using model-based reconstruction is the choice of the optimal regularization strength to prevent over- or under-regularization. This choice depends on the magnitudes of the input signals, and therefore repeated adjustments are required for in vivo data with inherent signal fluctuations due to different acquisition wavelengths, heterogeneities in the scanned anatomies, or altered imaging hardware. Because such repeated adjustments are impractical when reconstructing large datasets, clinical studies typically reconstruct all input sinograms with a constant predetermined regularization

strength at the expense of a possible quality degradation for some images [8–12]. In addition, the dependence of model-based reconstruction on the magnitudes of the input signals also complicates supervised learning of the reconstruction operator because samples with all possibly expected signal magnitudes must be represented in the training dataset. Compensating for signal magnitude fluctuations during model-based optoacoustic image reconstruction could make MSOT imaging more accurate, reliable, and robust. Fig. 1 summarizes the principal shortcomings of standard model-based reconstruction and introduces the added value of scale-equivariant model-based reconstruction.

In this work we investigate the behavior of optoacoustic image reconstruction under scaling. We define the notion of scale-equivariant image reconstruction and, based on this property, define a scale-equivariant model-based reconstruction operator that automatically adapts the regularization strength to the norm of the input sinogram. Furthermore, we leverage scale-equivariance to simplify supervised learning of the reconstruction operator: Input data is passed to the trainable network layers with a fixed norm to bypass the

\* Corresponding author at: Chair of Biological Imaging at the Central Institute for Translational Cancer Research (TranslaTUM), School of Medicine and Health, Technical University of Munich, Munich, Germany.

E-mail addresses: [christoph.dehner@ithera-medical.com](mailto:christoph.dehner@ithera-medical.com) (C. Dehner), [ledia.lilaj@ithera-medical.com](mailto:ledia.lilaj@ithera-medical.com) (L. Lilaj), [bioimaging.translatum@tum.de](mailto:bioimaging.translatum@tum.de) (V. Ntziachristos), [guillaume.zahnd@ithera-medical.com](mailto:guillaume.zahnd@ithera-medical.com) (G. Zahnd), [dominik.juestel@helmholtz-munich.de](mailto:dominik.juestel@helmholtz-munich.de) (D. Jüstel).

<sup>1</sup> Shared last authorship.

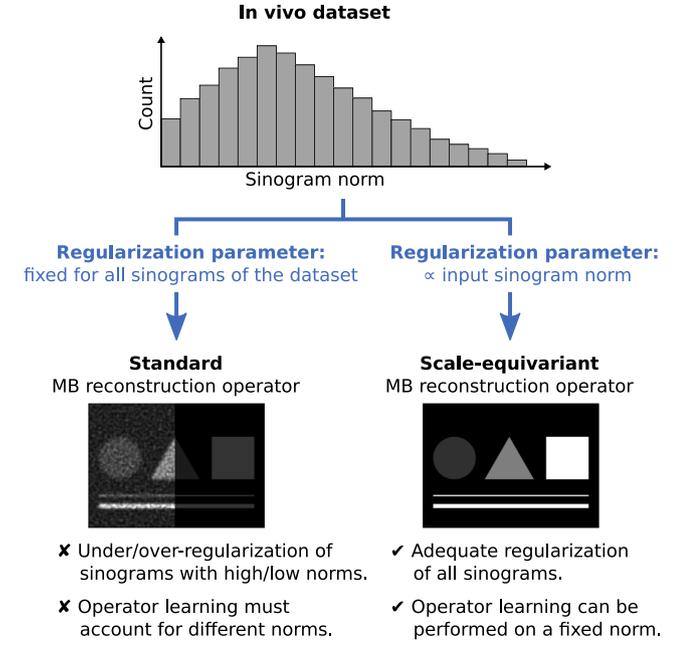


Fig. 1. Schematic representation of the shortcomings associated with standard model-based (MB) reconstruction and the benefits and working principle of scale-equivariant MB reconstruction.

need for manually tuned hyperparameters to select the range of signal magnitudes covered by the training samples. We experimentally demonstrate the benefits of scale-equivariant model-based reconstruction using optoacoustic scans of a phantom as well as in vivo data. Scale-equivariant model-based reconstruction applies appropriate regularization to signals with arbitrary magnitudes, slightly improves blood oxygenation quantification accuracy with multispectral optoacoustic imaging, and enables more accurate supervised learning of the reconstruction operator.

## 2. Theory

### 2.1. Equivariance of optoacoustic reconstruction

We begin with the formal definition of scale-equivariance for optoacoustic image reconstruction. The acoustic part of optoacoustic tomography can be assumed to be linear due to the low magnitudes of the acoustic waves generated by the optoacoustic effect [13]. Let  $\mathcal{A} : L^2(\mathbb{R}^3) \rightarrow L^2(D \times [0, \infty))$ ,  $p_0 \mapsto s$ , be the linear operator that maps an initial pressure distribution  $p_0$  to a sinogram  $s$ .  $D$  is the set of detectors, and the second dimension of  $s$  is time.

Due to the nature of the optoacoustic effect, the initial pressure distribution can be assumed non-negative, such that  $\mathcal{A}$  is in particular positively homogeneous as an operator on the convex cone of non-negative  $L^2$ -functions  $L^2_{\geq 0}(\mathbb{R}^3) := \{f \in L^2(\mathbb{R}^3) \mid f \geq 0\}$ , i.e.,

$$\mathcal{A}(ap_0) = a\mathcal{A}(p_0) \quad \text{for all } p_0 \in L^2_{\geq 0}(\mathbb{R}^3), a > 0. \quad (1)$$

Consequently, a reconstruction operator  $\mathcal{R} : \mathcal{A}(L^2_{\geq 0}(\mathbb{R}^3)) \rightarrow L^2_{\geq 0}(\mathbb{R}^3)$ ,  $s \mapsto p_{\text{rec}}$ , that reconstructs a non-negative initial pressure distribution  $p_{\text{rec}}$  from a given sinogram  $s \in \mathcal{A}(L^2_{\geq 0}(\mathbb{R}^3)) \subset L^2(D \times [0, \infty))$  needs to satisfy

$$\mathcal{R}(as) = a\mathcal{R}(s) \quad \text{for all } s \in \mathcal{A}(L^2_{\geq 0}(\mathbb{R}^3)), a > 0. \quad (2)$$

More generally, given an operator  $\mathcal{T} : V \rightarrow W$  and a group  $(G, \circ)$  that acts on the two spaces  $V$  and  $W$  with actions  $\rho$  and  $\sigma$ , respectively,

the operator is called  $G$ -equivariant (more precisely  $(\rho, \sigma)$ -equivariant), if

$$\mathcal{T}(\rho_g(v)) = \sigma_g(\mathcal{T}(v)) \quad \text{for all } v \in V, g \in G. \quad (3)$$

If we consider the scaling group  $S := ([0, \infty), \cdot)$  that acts on  $\mathcal{A}(L^2_{\geq 0}(\mathbb{R}^3))$  by scaling

$$\rho_a(s) := as \quad \text{for all } s \in \mathcal{A}(L^2_{\geq 0}(\mathbb{R}^3)), a \in S, \quad (4)$$

and on  $L^2_{\geq 0}(\mathbb{R}^3)$  by scaling

$$\sigma_a(p_0) := ap_0 \quad \text{for all } p_0 \in L^2_{\geq 0}(\mathbb{R}^3), a \in S, \quad (5)$$

then we can say that a reconstruction operator is supposed to be  $S$ -equivariant, or scale-equivariant.

### 2.2. Equivariance of model-based reconstruction

Having defined the notion of scale-equivariance, we next derive a model-based reconstruction operator that preserves this property. Let  $\Omega : L^2(\mathbb{R}^3) \rightarrow [0, \infty)$  be a regularization functional, then  $p_{\text{rec}}$  is obtained by solving the optimization problem

$$p_{\text{rec}} = \mathcal{R}(s) := \underset{\substack{p_0 \in L^2(\mathbb{R}^3), \\ p_0 \geq 0}}{\text{argmin}} \|\mathcal{A}(p_0) - s\|_2^2 + \lambda\Omega(p_0), \quad (6)$$

where  $\lambda > 0$  is the regularization parameter.

$G$ -equivariance of  $\mathcal{R}$  implies a condition on the regularizing term:

$$\mathcal{R}(as) = \underset{\substack{p_0 \in L^2(\mathbb{R}^3), \\ p_0 \geq 0}}{\text{argmin}} \|\mathcal{A}(p_0) - as\|_2^2 + \lambda\Omega(p_0) \quad (7)$$

$$= \underset{\substack{p_0 \in L^2(\mathbb{R}^3), \\ p_0 \geq 0}}{\text{argmin}} a^2 \|\mathcal{A}(p_0/a) - s\|_2^2 + \lambda\Omega(p_0) \quad (8)$$

$$= \underset{\substack{p_0 \in L^2(\mathbb{R}^3), \\ p_0 \geq 0}}{\text{argmin}} \|\mathcal{A}(p_0/a) - s\|_2^2 + a^{-2}\lambda\Omega(p_0) \quad (9)$$

$$= \underset{\substack{ap_0 \in L^2(\mathbb{R}^3), \\ ap_0 \geq 0}}{\text{argmin}} \|\mathcal{A}(p_0) - s\|_2^2 + a^{-2}\lambda\Omega(ap_0) \quad (10)$$

$$= a \underset{\substack{p_0 \in L^2(\mathbb{R}^3), \\ p_0 \geq 0}}{\text{argmin}} \|\mathcal{A}(p_0) - s\|_2^2 + a^{-2}\lambda\Omega(ap_0). \quad (11)$$

As a consequence, the model-based reconstruction operator  $\mathcal{R}$  with fixed regularization parameter  $\lambda > 0$  is scale-equivariant if and only if  $\Omega(ap_0) = a^2\Omega(p_0)$  for all  $p_0 \in L^2_{\geq 0}(\mathbb{R}^3)$  and all  $a \in S$ .

If the regularization functional transforms differently under scaling, the regularization parameter needs to be adjusted accordingly. If  $\Omega$  satisfies  $\Omega(ap_0) = a^q\Omega(p_0)$  for  $q \in \mathbb{N}$ , the regularization parameter  $\lambda$  needs to be rescaled to  $a^{2-q}\lambda$ . This is for example the case for regularizers of the form  $\Omega(p_0) = \|\mathcal{L}(p_0)\|_q^q$ , where  $\mathcal{L}$  is a linear operator, e.g., shearlet  $L^1$ -regularization with  $q = 1$  and  $\mathcal{L}$  the shearlet transform.

To implement a scale-equivariant model-based reconstruction operator in this situation, we need to fix a reference scale for the sinogram. Since we work in  $L^2$ -spaces, it is natural to define  $s_0 := \|s\|_2^{-1} s$ . To simplify the notation, we write  $a_s := \|s\|_2$ , such that  $s = a_s s_0$ . With the above considerations, we obtain a scale-equivariant reconstruction operator  $\mathcal{R}_S$  by

$$p_{\text{rec}} = \mathcal{R}_S(s) := \underset{\substack{p_0 \in L^2(\mathbb{R}^3), \\ p_0 \geq 0}}{\text{argmin}} \|\mathcal{A}(p_0) - a_s s_0\|_2^2 + a_s^{2-q}\lambda\Omega(p_0). \quad (12)$$

Instead of varying the regularization parameter,  $\mathcal{R}_S$  is easier implemented by scaling as a preprocessing step and rescaling as a post-processing step. First  $L^2$ -normalize  $s$  to obtain  $s_0$ , then reconstruct an initial pressure with input sinogram  $s_0$  and fixed regularization parameter  $\lambda$ . Then rescale the output initial pressure with  $a_s$  to obtain the

scale-equivariant reconstruction. In other words, the scale-equivariant reconstruction operator can be written in the form:

$$\mathcal{R}(s) = a_s \operatorname{argmin}_{\substack{p_0 \in L^2(\mathbb{R}^3), \\ p_0 \geq 0}} \|\mathcal{A}(p_0) - s_0\|_2^2 + \lambda \Omega(p_0). \quad (13)$$

### 2.3. Data normalization for supervised learning of a scale-equivariant operator

Equivariance of an operator simplifies supervised learning of the operator. Instead of the need to reconstruct inputs on different scales, the model can be trained with normalized data. The correct scale of the output is set by rescaling.

More precisely, if an operator  $\mathcal{T} : V \rightarrow W$  is  $G$ -equivariant, consider the orbit spaces  $V/\rho := \{\rho_G(v) \mid v \in V\}$  and  $V/\sigma := \{\sigma_G(w) \mid w \in W\}$ , where  $\rho_G(v) := \{\rho_g(v) \mid g \in G\}$  is the orbit of  $v \in V$ . We also consider fundamental domains of the actions  $F_\rho$  and  $F_\sigma$ , i.e., collections of elements of the vector spaces, such that one element of each orbit is contained. In the case of the action of  $S$  on  $L^2(D \times [0, \infty))$ , a fundamental domain  $F_S$  is the set of all sinograms  $s$  with  $\|s\|_2 = 1$ , i.e., exactly the set of references that we chose.

Now, the  $G$ -equivariant operator  $\mathcal{T}$  is uniquely determined by its behavior on a fundamental domain of  $V$ . The behavior on any other element is obtained by letting  $G$  act with the according element  $g$  on the output via the action  $\sigma$ . The set of outputs  $\mathcal{T}(F_\rho)$  define a fundamental domain of the action  $\sigma$  on the image space, i.e., only input–output pairs in  $F_\rho \times \mathcal{T}(F_\rho)$  are needed for training. Details on all these considerations are formulated as a theorem in a more general setting in [Appendix A](#).

## 3. Methods

### 3.1. Data acquisition

Validation experiments were conducted using data from a modern handheld MSOT scanner (MSOT Acuity Echo, iThera Medical GmbH), equipped with a multiwavelength laser and a concave array of 256 sensors.

The in vivo dataset comprises scans of six healthy volunteers at various anatomical locations (biceps, thyroid, carotid, calf, elbow, neck, colon, and breast) with wavelengths from the range 700–980 nm in steps of 10 nm, resulting in a total of 4814 sinograms. For each scan, the speed of sound value was manually tuned to obtain the most well-focused reconstructed image. All relevant ethical regulations were observed following the guidelines provided by Helmholtz Center Munich. All participants gave written informed consent upon recruitment. More detailed descriptions of the MSOT scanner used and the data acquisition and preprocessing procedure for the in vivo data are available in a previous study [\[7\]](#).

The oxygenation phantom dataset comprises scans of 33 different  $sO_2$  levels from the range 50%–100% at the wavelengths 700 nm, 730 nm, 760 nm, 800 nm, and 850 nm. The measurements were performed using two phantoms composed of a copolymer in oil material with two embedded PTFE tubes, each located at a depth of 8 mm and having an internal diameter of 1.07 mm. This type of material has previously been investigated as a soft tissue-mimicking material [\[14\]](#). The material acoustic characterization was carried out following the IEC TS 63081 standard [\[15\]](#) on three sets of samples with thicknesses of 1, 2, and 4 cm. The measurement results indicated a mean phase velocity of  $1473 \pm 1$  m/s and an absorption coefficient of  $0.58 \pm 0.11$  dB  $\text{cm}^{-1}$   $\text{MHz}^{-\gamma}$ , where the frequency exponent  $\gamma$  was  $1.55 \pm 0.07$ . The optical characterization was performed on 12 samples using a Dual Integrating Spheres system [\[16\]](#) and the Inverse Adding Doubling algorithm [\[17\]](#). This resulted in an average absorption coefficient of  $0.0144 \pm 0.0009$   $\text{mm}^{-1}$  and an average reduced scattering coefficient of  $0.90 \pm 0.02$   $\text{mm}^{-1}$ . Both the optical and acoustic values are consistent with the expected values for human soft tissues [\[18,19\]](#). The blood used

in the experiment was fresh heparinized porcine blood, from which the red blood cells were separated by centrifugation and diluted in PBS to reach the physiological hemoglobin concentration found in humans. A blood flow circuit connected to the phantom tubes maintained the hemoglobin solution at 37 °C and lowered or increased the oxygen saturation by exposing the solution to varying concentrations of nitrogen or oxygen, respectively. After each optoacoustic acquisition, a sample of blood was analyzed using a blood gas analyzer (Combi Line, Eschweiler GmbH & Co. KG), which measured the hemoglobin concentration and oxygen saturation of the sample in the phantom. Fifty iterations through all wavelengths were acquired for each  $sO_2$  level, resulting in a total of 8250 sinograms.

### 3.2. Image reconstruction

All sinograms were reconstructed using scale-equivariant and standard model-based reconstruction (see Eqs. (13) and (6), respectively) using shearlet  $L^1$ -regularization to tackle the ill-posedness of the inverse problem (i.e.,  $\Omega(p_0) = \|\mathcal{L}(p_0)\|_1$ , where  $\mathcal{L}$  is the shearlet transform). The minimization problem was solved via bound-constrained sparse reconstruction by separable approximation [\[20–22\]](#). All images were reconstructed with a size of  $416 \times 416$  pixels and a field of view of  $4.16 \times 4.16$   $\text{cm}^2$ . Regularization values  $\lambda$  were tuned via L-curves, an established method to select the optimal regularization parameter using the trade-off between the data residual and the regularization term, and set to  $10^{-5}$  and  $10^{-2}$  for scale-equivariant and standard model-based reconstruction, respectively. The selection process is illustrated in Supplementary Figure 2 for the scale-equivariant case. Model-based image reconstruction was implemented in MATLAB and run on a desktop computer (Intel i9-14900K processor, NVIDIA RTX 4900 GPU, 128 GB RAM). The source code is publicly available on GitHub [\[23\]](#).

### 3.3. Supervised learning

The DeepMB framework [\[7\]](#) was used to analyze the benefits of scale-equivariance on the supervised learning of a model-based reconstruction operator. Briefly, DeepMB is a previously validated framework for expressing a model-based reconstruction operator with a deep neural network through training on optoacoustic data synthesized from real-world images. Accurate generalization from synthesized training data to experimental test data is achieved because model-based reconstruction is a well-posed inverse problem, and the network can thus learn a data transformation that is independent of specific features of the ground truth images.

Two DeepMB models were trained: the first model was trained to learn the scale-equivariant model-based operator and the second model was trained to learn the standard model-based operator. Importantly, in the scale-equivariant case, the DeepMB model was trained on normalized data by scaling input sinograms to a fixed norm and rescaling output image to the original range of values (see Section 2.3). For both the scale-equivariant and standard cases, training was otherwise performed according to the originally proposed methodology [\[7\]](#), with the following exception: The smooth  $L^1$  loss was used during network training instead of the mean squared error loss. Training was performed on the same desktop computer as for model-based image reconstruction (see Section 3.2).

The accuracy of both the scale-equivariant and standard DeepMB models was evaluated by reconstructing all 4814 sinograms from the in vivo dataset and comparing them against the corresponding reference images obtained with scale-equivariant and standard model-based reconstruction, respectively. Reconstruction accuracy was quantified by calculating the (relative) mean absolute errors (MAE and  $\text{MAE}_{\text{rel}}$ ), (relative) mean squared errors (MSE and  $\text{MSE}_{\text{rel}}$ ), and structural similarity indices (SSIM) between the output images and the respective reference images from the model-based reconstruction, defined as

$$\text{MAE} := \|i_{\text{DeepMB}} - i_{\text{MB}}\|_1, \quad (14)$$

$$\text{MAE}_{\text{rel}} := \frac{\|i_{\text{DeepMB}} - i_{\text{MB}}\|_1}{\|i_{\text{MB}}\|_1}, \quad (15)$$

$$\text{MSE} := \|i_{\text{DeepMB}} - i_{\text{MB}}\|_2^2, \quad (16)$$

$$\text{MSE}_{\text{rel}} := \frac{\|i_{\text{DeepMB}} - i_{\text{MB}}\|_2^2}{\|i_{\text{MB}}\|_2^2}, \quad (17)$$

$$\text{SSIM} := \frac{(2\mu_{\text{DeepMB}}\mu_{\text{MB}} + c_1)(2\sigma_{\text{DeepMB,MB}} + c_2)}{(\mu_{\text{DeepMB}}^2 + \mu_{\text{MB}}^2 + c_1)(\sigma_{\text{DeepMB}}^2 + \sigma_{\text{MB}}^2 + c_2)}, \quad (18)$$

where  $i_{\text{DeepMB}}$  and  $i_{\text{MB}}$  are vectorizations of reconstructed DeepMB and model-based images, respectively. The SSIM was calculated as the average over sliding windows of size  $21 \times 21$  pixels, where  $\mu_{\text{DeepMB}}$  and  $\mu_{\text{MB}}$  are the averages of  $i_{\text{DeepMB}}$  and  $i_{\text{MB}}$ ;  $\sigma_{\text{DeepMB}}$  and  $\sigma_{\text{MB}}$  are the variances of  $i_{\text{DeepMB}}$  and  $i_{\text{MB}}$ ;  $\sigma_{\text{DeepMB,MB}}$  is the covariance of  $i_{\text{DeepMB}}$  and  $i_{\text{MB}}$ ; and  $c_1 = (0.01 \cdot \max(i_{\text{MB}}))^2$  and  $c_2 = (0.03 \cdot \max(i_{\text{MB}}))^2$  are two empirical variables to stabilize the division with weak denominators.

The reconstruction fidelity of both DeepMB models was assessed using the data residual norm (R), defined as

$$\text{R} := \frac{\|\mathcal{A}(p_{\text{rec}}) - s\|_2^2}{\|s\|_2^2}, \quad (19)$$

where  $p_{\text{rec}}$  is the reconstructed image,  $\mathcal{A}$  is the acoustic forward imaging operator, and  $s$  is the input sinogram.

### 3.4. Blood oxygen quantification

The blood phantom scans were used to compare the ability to quantify  $\text{sO}_2$  with scale-equivariant and standard model-based reconstruction as well as the corresponding DeepMB models. Reconstructed images were grouped into multispectral stacks. The regions corresponding to the two inclusions were manually segmented, and the average absorption spectra per blood oxygenation level across all scans were extracted. The concentration coefficients for oxygenated and deoxygenated hemoglobin were calculated using linear unmixing. The predicted blood oxygen concentration was computed as

$$m_{\text{sO}_2} := \frac{c_{\text{HbO}_2}}{c_{\text{Hb}} + c_{\text{HbO}_2}}, \quad (20)$$

where  $c_{\text{HbO}_2}$  and  $c_{\text{Hb}}$  are the concentration coefficients for oxygenated and deoxygenated hemoglobin from linear unmixing, respectively. To compare the different reconstruction methods considered, the root mean squared error for predicting blood oxygen saturation ( $\text{RMSE}_{\text{sO}_2}$ ) was computed as

$$\text{RMSE}_{\text{sO}_2} := \sqrt{\frac{1}{N} \sum_{n=1}^N (m_{\text{sO}_2}^{(n)} - \text{gt}_{\text{sO}_2}^{(n)})^2}, \quad (21)$$

where  $N$  is the number of different blood oxygen concentrations scanned,  $\text{gt}_{\text{sO}_2}^{(n)}$  are the ground truth blood oxygen concentrations from the blood gas analyzer, and  $m_{\text{sO}_2}^{(n)}$  are the predicted blood oxygen concentrations from either scale-equivariant or standard model-based reconstruction or any of the corresponding DeepMB models.

## 4. Results

### 4.1. Improved regularization

Scale-equivariant model-based reconstruction enables appropriate regularization of sinograms with arbitrary signal strength without the need to retune the regularization strength. Fig. 2 compares images reconstructed from representative sinograms from the in vivo dataset with high, low, and median  $L^2$  norms. Images obtained via scale-equivariant model-based reconstruction are adequately regularized in all three cases (see Fig. 2a,c,e). Conversely, images obtained via standard model-based reconstruction are too weakly regularized for the

**Table 1**

Comparison of the accuracy for determining blood oxygen saturation ( $\text{sO}_2$ ) with scale-equivariant versus standard model-based (MB) reconstruction. The reported values are the average root mean square errors (RMSE) between the MSOT predictions and the ground truth values from a blood gas analyzer.

	RMSE for predicting $\text{sO}_2$	
	Direct application	Supervised learning
Scale-eq. MB	4.97	5.20
Standard MB	5.00	5.61

sinogram with a high  $L^2$  norm sinogram (see ripple and limited-view artifacts in Fig. 2b), and slightly too strongly regularized for the sinogram with a low  $L^2$  norm (see suppressed image contrast in Fig. 2d). Both the scale-equivariant and standard model-based reconstructions obtain similar images for the sinogram with a median  $L^2$  norm (see Fig. 2e,f). Further examples of scale-equivariant and standard model-based reconstructions for in vivo sinograms with high, low, and median  $L^2$  norms are provided in Supplementary Figure 1.

Subsequently, the ability of scale-equivariant reconstruction to adapt the degree of regularization applied during reconstruction was validated using the full in vivo dataset. The change in regularization between scale-equivariant and standard model-based images was quantified by the differences between their respective regularization terms (a comparatively larger regularization term indicates that images are less regularized, and vice versa). Fig. 3a visualizes the changes in regularization relative to the sinogram norm. The results confirm that the regularization behavior is meaningful across the full dataset: Reconstructions from sinograms with low norms (i.e.,  $L^2$  norm smaller than 1100) are more weakly regularized in the scale-equivariant case than in the standard model-based case, hence the differences in regularization terms are positive. Conversely, images from sinograms with high norms (i.e.,  $L^2$  norm larger than 1100) are more strongly regularized in the scale-equivariant case than in the standard model-based case, hence the differences in regularization terms are negative. Reconstructions from sinograms with median norms (i.e.,  $L^2$  norm around 1100; also see the color-encoding in Fig. 3a, and the histogram in Fig. 3b) are similarly regularized in both cases, hence the difference in regularization terms is close to zero.

At last, the effects of regularization improvements through scale-equivariant model-based reconstruction on MSOT imaging were investigated using the example of blood oxygen saturation ( $\text{sO}_2$ ) measurement. Scale-equivariant model-based reconstruction achieves a slightly lower RMSE for predicting  $\text{sO}_2$  than standard model-based reconstruction (4.97 vs. 5.00, see first column ‘‘Direct application’’ of Table 1).

### 4.2. Improved supervised learning

Scale-equivariant model-based reconstruction also enables more accurate supervised learning of the operator.

Table 2 compares the reconstruction accuracy of DeepMB when learning a scale-equivariant model-based reconstruction operator versus learning a standard model-based reconstruction operator using the in vivo dataset. The average MAE and MSE between DeepMB and model-based images improve by 21% and 44%, respectively. Similar improvements were also observed when assessing ability to quantify  $\text{sO}_2$  with the two DeepMB models. The DeepMB model trained on scale-equivariant model-based images achieves a 7% lower RMSE in predicting  $\text{sO}_2$  than the DeepMB model trained on standard model-based images (5.20 vs. 5.61, see second column ‘‘Supervised learning’’ of Table 1).

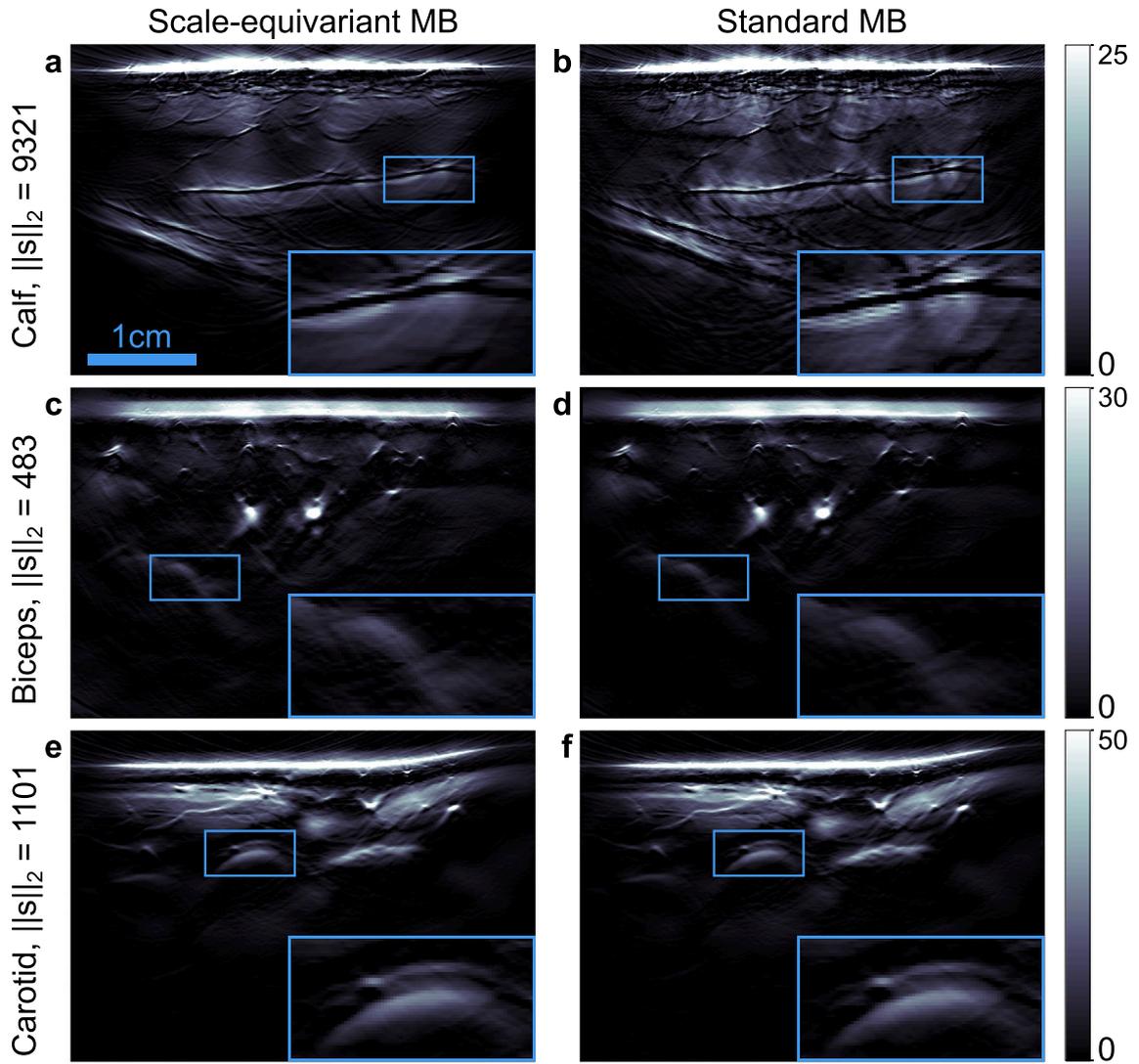


Fig. 2. Optoacoustic images obtained using scale-equivariant and standard model-based (MB) reconstruction for in vivo sinograms with high (a, b), low (c, d), and median (e, f)  $L^2$  norms ( $\|s\|_2$ ).

Table 2

Comparison of the reconstruction accuracy of DeepMB when learning a scale-equivariant versus a standard model-based (MB) reconstruction operator across 4814 in vivo sinograms. The accuracy of the scale-equivariant and standard DeepMB models is quantified against ground-truth images reconstructed using the reference scale-equivariant and standard MB algorithms, respectively. The table shows the mean values and in square brackets the 25th and 75th percentiles. The arrow symbols ( $\uparrow$ ,  $\downarrow$ ) indicate for each metric whether a higher or lower value is better. MAE,  $MAE_{rel}$ : (relative) mean absolute error, MSE,  $MSE_{rel}$ : (relative) mean squared error, SSIM: structural similarity, R: data residual norm. The average data residual norm of both model-based datasets is 0.141.

	DeepMB for scale-equivariant MB	DeepMB for standard MB
MAE ( $\downarrow$ )	0.59 [0.41, 0.60]	0.75 [0.47, 0.77]
$MAE_{rel}$ (%) ( $\downarrow$ )	13.24 [11.69, 14.61]	15.65 [14.07, 17.22]
MSE ( $\downarrow$ )	3.32 [0.44, 1.59]	5.98 [0.58, 2.19]
$MSE_{rel}$ (%) ( $\downarrow$ )	0.85 [0.52, 1.05]	1.14 [0.74, 1.31]
SSIM ( $\uparrow$ )	0.98 [0.98, 0.99]	0.98 [0.97, 0.99]
R ( $\downarrow$ )	0.146 [0.083, 0.181]	0.149 [0.082, 0.185]

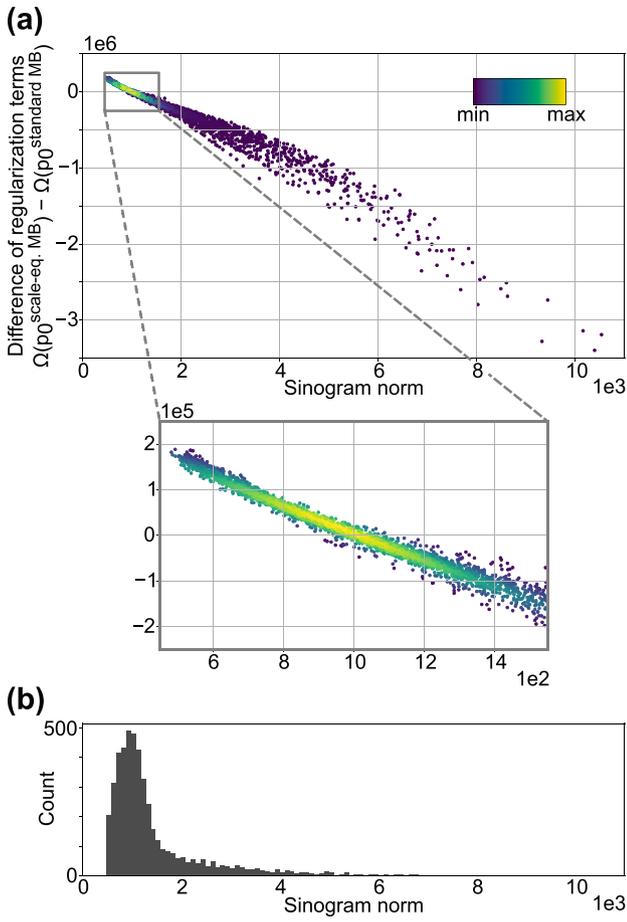
## 5. Discussion

In this work we propose a method for adequate regularization of model-based optoacoustic image reconstruction. Scale-equivariant model-based reconstruction is able to adapt the regularization strength

based on the norm of the input sinogram and can thus account for magnitude variations within in vivo data that are generally caused by the anatomical region, acquisition wavelength, or hardware characteristics.

Scale-equivariant model-based reconstruction achieves a slightly higher image quality and quantification accuracy than standard model-based reconstruction when evaluated on a large and diverse in vivo dataset as well as on scans of an oximetry phantom. The extent of the improvement observed depends on the distribution of sinogram norms in the dataset under consideration. Scale-equivariant and standard model-based reconstructions lead to similar results if the norm of the sinograms within a dataset follow a narrow distribution. However, the broader the distribution of sinogram norms, the more superior scale-equivariant is compared standard model-based reconstruction. Scale-equivariance also enables more accurate supervised learning of the reconstruction operator using a deep neural network: During training, the data manifold is more densely sampled because input sinograms have a fixed norm, resulting in improved prediction accuracy.

Overall, improvements in image quality through scale-equivariance are incremental: A scale-equivariant reconstruction moderately alters the image appearance and enables a gradual enhancement in quantification accuracy of multispectral optoacoustic tomography. Thereby, the method can contribute to obtaining more comprehensive and precise diagnoses using multispectral optoacoustic tomography in clinical



**Fig. 3.** (a) Differences between the regularization terms of scale-equivariant and standard model-based images from the in vivo dataset,  $\Omega(p_0^{\text{scale-eq. MB}}) - \Omega(p_0^{\text{standard MB}})$ , relative to the sinogram  $L^2$  norm, for all 4814 samples of the in vivo dataset. The colors in the scatter plot encode point density. (b) Histogram of the sinogram  $L^2$  norm.

practice. By improving image reconstruction for sinograms with high or low sinogram values, scale-equivariant reconstruction can also help to account for anatomical heterogeneities within a study cohort, for example, with regard to body type or skin color. Beyond direct effects on clinical imaging, precise reconstructions of initial pressure achieved through scale equivariant reconstruction can also benefit scientific research and serve as a building block for the development of more advanced image processing algorithms for quantitative optoacoustic imaging [24–27].

Another crucial advantage of scale-equivariant reconstruction is improved practicality: Scale-equivariant reconstruction enables the use of the same regularization value for multiple datasets (acquired with the same scanner type) because input sinograms are  $L^2$ -normalized before the reconstruction, and rescaled after. Scale-equivariance therefore eliminates the need to adapt the regularization strength to signal intensity variations caused, for example, by the scanned anatomy or the laser wavelength. In addition, for the supervised learning of the model-based reconstruction operator, scale-equivariance eliminates the need to manually determine a hyperparameter to scale the value range of training data to the range of experimental test data. These practical improvements can significantly advance the clinical implementation of multispectral optoacoustic tomography by reducing maintenance efforts, simplifying data analysis, and facilitating standardization.

While scale equivariance enables more accurate optoacoustic imaging, it still leaves important areas for improvement open. Clinical optoacoustic imaging commonly uses limited-angle acquisition schemes,

resulting in the loss of image features perpendicular to the acquisition probe orientation. Model-based reconstruction can mitigate the resulting ill-posedness of the inversion problem using regularization; however, only rigid regularization functionals are used at the current state (e.g., in this work a L1-based regularization functional was used to promote sparse encoding with a Shearlet basis). The development of more complex tissue priors could allow for a more comprehensive compensation for the information lost through limited-angle acquisition and thus further increase the accuracy of optoacoustic images. Another open challenge in optoacoustic imaging is the optical inversion. Recent work focuses on optical inversion using deep learning models trained on synthesized data [26,27]. Scale equivariance might be a sub-step in such an approach, helping to bridge the domain gap between such synthesized data and in vivo test data, but the development and comprehensive validation of such approaches is still the subject of ongoing research.

In summary, scale-equivariant model-based reconstruction is strictly advantageous over standard model-based reconstruction. Scale-equivariance does not increase computational complexity, requires only minor implementation adjustments, and can be used in any application area of optoacoustic tomography.

### CRedit authorship contribution statement

**Christoph Dehner:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation. **Ledia Lilaj:** Writing – review & editing, Data curation. **Vasilis Ntziachristos:** Writing – review & editing, Resources. **Guillaume Zahnd:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation. **Dominik Jüstel:** Writing – review & editing, Writing – original draft, Conceptualization.

### Funding

This project has received funding from the Bavarian Ministry of Economic Affairs, Energy and Technology (StMWi) (DIE-2106-0005//DIE0161/01 and 02, DeepOPUS).

### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Christoph Dehner reports a relationship with iThera Medical GmbH that includes: employment. Ledia Lilaj reports a relationship with iThera Medical GmbH that includes: employment. Guillaume Zahnd reports a relationship with iThera Medical GmbH that includes: employment. Vasilis Ntziachristos reports a relationship with iThera Medical GmbH that includes: equity or stocks. Christoph Dehner has patent #EP22177153.8 and PCT/EP2023/064714 pending to. Guillaume Zahnd has patent #EP22177153.8 and PCT/EP2023/064714 pending to. Dominik Juestel has patent #EP22177153.8 and PCT/EP2023/064714 pending to. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Appendix A

We provide a theorem on equivariant regularization that is applicable in a very general setting.

**Theorem 1 (Equivariant Regularization).** Let  $T, W$  be sets,  $A : T \rightarrow W$ , and  $S := A(T) \subseteq W$ . Let  $(G, \circ)$  be a group that acts on  $T$  and  $S$  via actions  $\rho$  and  $\sigma$ , respectively. Let  $d : S \times S \rightarrow [0, \infty)$ . Define the function  $\mathcal{R} : S \rightarrow T$ , defined by

$$\mathcal{R}(s) := \operatorname{argmin}_{p_0 \in T} d(A(p_0), s) + \lambda(s)\Omega(p_0), \quad s \in S, \quad (22)$$

where we assume the existence of a regularization functional  $\Omega : T \rightarrow [0, \infty)$  and a regularization parameter function  $\lambda : S \rightarrow [0, \infty)$ , such that  $\mathcal{R}$  is well-defined.

Let  $d$  and  $\Omega$  be  $G$ -equivariant, i.e., there are group homomorphisms  $f_d : G \rightarrow [0, \infty)$  and  $f_\Omega : G \rightarrow [0, \infty)$ , such that

$$d(\sigma_g s_1, \sigma_g s_2) = f_d(g)d(s_1, s_2), \quad g \in G, s_1, s_2 \in S, \quad (23)$$

$$\Omega(\rho_g t) = f_\Omega(g)\Omega(t), \quad g \in G, t \in T. \quad (24)$$

Let  $F_S$  be a fundamental domain of the action  $\sigma$  on  $S$ . For  $s \in S$ , denote by  $s_0 \in F_S$  the element in the orbit of  $s$ , and by  $g(s) \in G$  the group element that satisfies  $s = \sigma_{g(s)} s_0$ .

If  $\mathcal{A}$  is  $(\rho, \sigma)$ -equivariant, then

(i) If there is a  $\lambda_{\text{fix}} > 0$ , such that  $\mathcal{R}$  is well-defined for  $\lambda \equiv \lambda_{\text{fix}}$ , then

$\mathcal{R}$  is  $(\sigma, \rho)$ -equivariant

$$\Leftrightarrow f_\Omega(g) = f_d(g) \quad \forall g \in G. \quad (25)$$

(ii) If  $\mathcal{R}$  is well-defined on  $F_S$  for  $\lambda|_{F_S} \equiv \lambda_0 > 0$ , and

$$\lambda(s) := \frac{f_d(g(s))}{f_\Omega(g(s))} \lambda_0, \quad s \in S, \quad (26)$$

then  $\mathcal{R}$  is  $(\sigma, \rho)$ -equivariant.

(iii) The equivariant operator  $\mathcal{R}$  defined in (ii) satisfies

$$\begin{aligned} \mathcal{R}(s) &= \rho_{g(s)} \left( \underset{p_0 \in T}{\operatorname{argmin}} d(\mathcal{A}(p_0), s_0) \right. \\ &\quad \left. + \lambda_0 \Omega(p_0) \right). \end{aligned} \quad (27)$$

**Proof.** Consider the following calculation for  $g \in G$  and  $s \in S$ :

$$\mathcal{R}(\sigma_g s) = \underset{p_0 \in T}{\operatorname{argmin}} d(\mathcal{A}(p_0), \sigma_g s) + \lambda(\sigma_g s) \Omega(p_0) \quad (28)$$

$$\begin{aligned} &= \underset{p_0 \in T}{\operatorname{argmin}} d(\sigma_g \mathcal{A}(\rho_{g^{-1}} p_0), \sigma_g s) \\ &\quad + \lambda(\sigma_g s) \Omega(p_0) \end{aligned} \quad (29)$$

$$\begin{aligned} &= \underset{p_0 \in T}{\operatorname{argmin}} f_d(g) d(\mathcal{A}(\rho_{g^{-1}} p_0), s) \\ &\quad + \lambda(\sigma_g s) \Omega(p_0) \end{aligned} \quad (30)$$

$$\begin{aligned} &= \rho_g \left( \underset{p_0 \in T}{\operatorname{argmin}} f_d(g) d(\mathcal{A}(p_0), s) \right. \\ &\quad \left. + \lambda(\sigma_g s) \Omega(\rho_g p_0) \right) \end{aligned} \quad (31)$$

$$\begin{aligned} &= \rho_g \left( \underset{p_0 \in T}{\operatorname{argmin}} f_d(g) d(\mathcal{A}(p_0), s) \right. \\ &\quad \left. + \lambda(\sigma_g s) f_\Omega(g) \Omega(p_0) \right) \end{aligned} \quad (32)$$

$$\begin{aligned} &= \rho_g \left( \underset{p_0 \in T}{\operatorname{argmin}} d(\mathcal{A}(p_0), s) \right. \\ &\quad \left. + \lambda(\sigma_g s) f_d(g)^{-1} f_\Omega(g) \Omega(p_0) \right), \end{aligned} \quad (33)$$

where we used the definition of  $\mathcal{R}$  in the first equality, equivariance of  $\mathcal{A}$  is the second equality, equivariance of  $d$  in the third equality, transformed the argument via  $\rho_g$  in the fourth equality, used equivariance of  $\Omega$  in the fifth equality, and divided the minimized functional by  $f_d(g)$  in the sixth equality. The last step does not change the result, because we multiply by a non-negative constant.

Now, (i) “ $\Leftarrow$ ” follows from this calculation, because if  $f_\Omega(g) = f_d(g)$ , then

$$\lambda(\sigma_g s) f_d(g)^{-1} f_\Omega(g) = \lambda(\sigma_g s) \equiv \lambda_{\text{fix}}. \quad (34)$$

The argument of  $\rho_g$  thus simplifies to  $\mathcal{R}(s)$ , showing equivariance of  $\mathcal{R}$ .

The other way round ((i) “ $\Rightarrow$ ”), if  $\mathcal{R}$  is equivariant, the argument of  $\rho_g$  needs to be equal to  $\mathcal{R}(s)$ . The factor in front of  $\Omega(p_0)$  needs to be equal to  $\lambda_{\text{fix}}$ , which implies  $f_\Omega(g) = f_d(g)$  for all  $g \in G$ .

(iii) follows, when choosing  $s = s_0$  and  $g = g(s)$ . Then, using the definition of  $\lambda(s)$  in (ii),

$$\lambda(s) f_d(g(s))^{-1} f_\Omega(g(s)) = \lambda_0, \quad (35)$$

showing (iii).

To show (ii), note that for  $h \in G$ ,

$$\sigma_{g(\sigma_h s)} s_0 = \sigma_h s \quad (36)$$

$$\Rightarrow \sigma_{h^{-1} \circ g(\sigma_h s)} s_0 = s \quad (37)$$

$$\Rightarrow h^{-1} \circ g(\sigma_h s) = g(s) \quad (38)$$

$$\Rightarrow g(\sigma_h s) = h \circ g(s). \quad (39)$$

Now, replacing the argument of  $\mathcal{R}$  in (iii) by  $\sigma_h s$ , and using this identity, we get

$$\begin{aligned} \mathcal{R}(\sigma_h s) &= \rho_h \left( \underset{p_0 \in T}{\operatorname{argmin}} d(\mathcal{A}(p_0), s_0) \right. \\ &\quad \left. + \lambda_0 \Omega(p_0) \right). \end{aligned} \quad (40)$$

Thus, it remains to show that the argument of  $\rho_h$  on the right hand side equals  $\mathcal{R}(s)$ . But that is exactly statement (iii), showing (ii).  $\square$

## Appendix B. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.pacs.2025.100727>.

## Data availability

In vivo data from two of the six scanned volunteers are provided online [28]. All other data cannot be shared due to privacy, consent, or confidentiality restrictions.

## References

- [1] K.B. Chowdhury, J. Prakash, A. Karlas, D. Jüstel, V. Ntziachristos, A synthetic total impulse response characterization method for correction of hand-held optoacoustic images, *IEEE Trans. Med. Imaging* 39 (10) (2020) 3218–3230.
- [2] K.B. Chowdhury, M. Bader, C. Dehner, D. Jüstel, V. Ntziachristos, Individual transducer impulse response characterization method to improve image quality of array-based handheld optoacoustic tomography, *Opt. Lett.* 46 (1) (2021) 1–4.
- [3] X.L. Dean-Ben, D. Razansky, A practical guide for model-based reconstruction in optoacoustic imaging, *Front. Phys.* 10 (2022) 1028258.
- [4] A. Buehler, A. Rosenthal, T. Jetzfellner, A. Dima, D. Razansky, V. Ntziachristos, Model-based optoacoustic inversions with incomplete projection data, *Med. Phys.* 38 (3) (2011) 1694–1704.
- [5] K. Wang, R. Su, A.A. Oraevsky, M.A. Anastasio, Investigation of iterative image reconstruction in three-dimensional optoacoustic tomography, *Phys. Med. Biol.* 57 (17) (2012) 5399–5423.
- [6] A. Hauptmann, F. Lucka, M. Betcke, N. Huynh, J. Adler, B. Cox, P. Beard, S. Ourselin, S. Arridge, Model-based learning for accelerated, limited-view 3-D photoacoustic tomography, *IEEE Trans. Med. Imaging* 37 (6) (2018) 1382–1393.
- [7] C. Dehner, G. Zahnd, V. Ntziachristos, D. Jüstel, A deep neural network for real-time optoacoustic image reconstruction with adjustable speed of sound, *Nat. Mach. Intell.* 5 (10) (2023) 1130–1141.
- [8] J. Kukačka, S. Metz, C. Dehner, A. Muckenhuber, K. Paul-Yuan, A. Karlas, E.M. Fallenberg, E. Rummeny, D. Justel, V. Ntziachristos, Image processing improvements afford second-generation handheld optoacoustic imaging of breast cancer patients, *Photoacoustics* 26 (2022) 100343.
- [9] D. Jüstel, H. Irl, F. Hinterwimmer, C. Dehner, W. Simson, N. Navab, G. Schneider, V. Ntziachristos, Spotlight on nerves: portable multispectral optoacoustic imaging of peripheral nerve vascularization and morphology, *Adv. Sci.* 10 (19) (2023) 2301322.
- [10] J. Vonk, J. Kukačka, P.J. Steinkamp, J.G. de Wit, F.J. Voskuil, W.T.R. Hooghiemstra, M. Bader, D. Jüstel, V. Ntziachristos, G.M. van Dam, M.J.H. Witjes, Multispectral optoacoustic tomography for in vivo detection of lymph node metastases in oral cancer patients using an EGFR-targeted contrast agent and intrinsic tissue contrast: A proof-of-concept study, *Photoacoustics* 26 (2022) 100362.

- [11] W. Roll, N.A. Markwardt, M. Masthoff, A. Helfen, J. Claussen, M. Eisenblätter, A. Hasenbach, S. Hermann, A. Karlas, M. Wildgruber, V. Ntziachristos, M. Schafers, Multispectral optoacoustic tomography of benign and malignant thyroid disorders: A pilot study, *J. Nucl. Med.* 60 (10) (2019) 1461–1466.
- [12] C. Dehner, I. Olefir, K.B. Chowdhury, D. Justel, V. Ntziachristos, Deep-learning-based electrical noise removal enables high spectral optoacoustic contrast in deep tissue, *IEEE Trans. Med. Imaging* 41 (11) (2022) 3182–3193.
- [13] T. Tarvainen, B. Cox, Quantitative photoacoustic tomography: modeling and inverse problems, *J. Biomed. Opt.* 29 (S1) (2024) S11509–S11509.
- [14] L. Hacker, J. Joseph, A.M. Ivory, M.O. Saed, B. Zeqiri, S. Rajagopal, S.E. Bohndiek, A copolymer-in-oil tissue-mimicking material with tuneable acoustic and optical characteristics for photoacoustic imaging phantoms, *IEEE Trans. Med. Imaging* 40 (12) (2021) 3593–3603.
- [15] IECTS63081:2019, Ultrasonics – methods for the characterization of the ultrasonic properties of materials, 2019, <https://webstore.iec.ch/publication/32282>.
- [16] P. Lemaillet, J.-P. Bouchard, J. Hwang, D.W. Allen, Double-integrating-sphere system at the national institute of standards and technology in support of measurement standards for the determination of optical properties of tissue-mimicking phantoms, *J. Biomed. Opt.* 20 (12) (2015) 121310.
- [17] S. Prahl, Everything I think you should know about inverse adding-doubling, *Or. Med. Laser Cent. St. Vincent Hosp.* 1344 (2011) 1–74.
- [18] W.C. Vogt, C. Jia, K.A. Wear, B.S.M.D. Garra, T.J. Pfefer, Biologically relevant photoacoustic imaging phantoms with tunable optical and acoustic properties, *J. Biomed. Opt.* 21 (10) (2016) 101405.
- [19] IEC60601-2-37:2007, Particular requirements for the basic safety and essential performance of ultrasonic medical diagnostic and monitoring equipment, 2007, <https://webstore.iec.ch/publication/2652>.
- [20] S.J. Wright, R.D. Nowak, M.A.T. Figueiredo, Sparse reconstruction by separable approximation, *IEEE Trans. Signal Process.* 57 (7) (2009) 2479–2493.
- [21] R. Chartrand, B. Wohlberg, Total-variation regularization with bound constraints, in: 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 766–769.
- [22] G. Kutyniok, W.-Q. Lim, R. Reisenhofer, ShearLab 3D: Faithful digital shearlet transforms based on compactly supported shearlets, *ACM Trans. Math. Software* 42 (1) (2016) 5.
- [23] C. Dehner, G. Zahnd, Reconstruction toolbox for MSOT v1.0.0, 2023, <https://doi.org/10.5281/zenodo.8170989>.
- [24] B. Cox, J.G. Laufer, S.R. Arridge, P.C. Beard, Quantitative spectroscopic photoacoustic imaging: a review, *J. Biomed. Opt.* 17 (6) (2012) 061202–061202.
- [25] Y. Gao, T. Feng, H. Qiu, Y. Gu, Q. Chen, C. Zuo, H. Ma, 4D spectral-spatial computational photoacoustic dermoscopy, *Photoacoustics* 34 (2023) 100572.
- [26] J. Gröhl, K. Yeung, K. Gu, T.R. Else, M. Golinska, E.V. Bunce, L. Hacker, S.E. Bohndiek, Distribution-informed and wavelength-flexible data-driven photoacoustic oximetry, *J. Biomed. Opt.* 29 (S3) (2024) S33303–S33303.
- [27] X. Wei, T. Feng, Q. Huang, Q. Chen, C. Zuo, H. Ma, Deep learning-powered biomedical photoacoustic imaging, *Neurocomputing* (2023) 127207.
- [28] C. Dehner, G. Zahnd, DeepMB v1.0.0, 2023, <https://doi.org/10.5281/zenodo.8169175>.



**Christoph Dehner** is Senior Data Scientist at iThera Medical GmbH, where he is co-developing the next generation of optoacoustic imaging systems. He studied Computer Science at the Technical University of Munich and received his doctoral degree at the Chair of Biological Imaging at the Technical University of Munich about accurate and real-time imaging with multispectral optoacoustic tomography by means of deep learning.



**Ledia Lilaj** is Senior R&D Engineer at iThera Medical GmbH, where she designs and develops performance quantification test objects and methods, and calibration tools for optoacoustic devices. She is a member of the International Photoacoustic Standardization Consortium's leadership team, contributing to the development of test objects and methods. Dr. Lilaj holds an M.Sc. in Biomedical Engineering from Politecnico di Milano and a Ph.D. from Charité Universitätsmedizin Berlin, focusing on Magnetic Resonance Elastography.



**Vasilis Ntziachristos** studied Electrical Engineering at Aristotle University in Thessaloniki, Greece and received his M.Sc. and Ph.D. from the Bioengineering Department of the University of Pennsylvania. He served as assistant professor and director of the Laboratory for Bio-Optics and Molecular Imaging at Harvard University and Massachusetts General Hospital. Currently, he is the Director of the Institute for Biological and Medical Imaging at the Helmholtz Zentrum in Munich, Germany, as well as a Professor of Electrical Engineering, Professor of Medicine and Chair for Biological Imaging at the Technical University Munich, Germany. His work focuses on novel innovative optical and optoacoustic imaging modalities for studying biological processes and diseases as well as the translation of these findings into the clinic.



**Guillaume Zahnd** is Director of Imaging at iThera Medical GmbH. He obtained the PhD Degree in Physics from the University of Lyon in France. His current research interests include algorithms for optoacoustic and ultrasound image acquisition, reconstruction, and analysis towards clinical applications.



**Dominik Jüstel** studied mathematics and informatics at the Technical University of Munich (TUM), graduating at the Chair for Mathematical Modeling of Biological Systems in association with the Institute for Biomathematics and Biometry at Helmholtz Munich. He received his doctoral degree (Dr. rer. nat.) in mathematics with highest distinction at the Chair for Analysis at TUM. After a postdoc at the mathematics faculty of TUM, Dr. Jüstel joined the Chair for Biological Imaging at TUM and the Institute for Biological and Medical Imaging at Helmholtz Munich. With his group for computational optoacoustics, he works on quantitative multispectral optoacoustic tomography and advanced data analysis for various optical and optoacoustic imaging modalities.