

## Article

# Learned Block Iterative Shrinkage Thresholding Algorithm for Photothermal Super Resolution Imaging

Jan Christian Hauffen <sup>1,\*</sup> , Linh Kästner <sup>2</sup> , Samim Ahmadi <sup>3</sup>, Peter Jung <sup>1,\*</sup> , Giuseppe Caire <sup>1</sup> and Mathias Ziegler <sup>3</sup>

<sup>1</sup> Communication and Information Theory, Berlin Institute of Technology, 10623 Berlin, Germany; caire@tu-berlin.de

<sup>2</sup> Industry Grade Networks and Clouds, Faculty of Electrical Engineering, and Computer Science, Berlin Institute of Technology, 10623 Berlin, Germany; d.kaestner@tu-berlin.de

<sup>3</sup> Department of Non-Destructive Testing, Bundesanstalt für Materialforschung und -Prüfung, 12489 Berlin, Germany; samim\_193@hotmail.de (S.A.); mathias.ziegler@bam.de (M.Z.)

\* Correspondence: j.hauffen@tu-berlin.de (J.C.H.); peter.jung@tu-berlin.de (P.J.)

**Abstract:** Block-sparse regularization is already well known in active thermal imaging and is used for multiple-measurement-based inverse problems. The main bottleneck of this method is the choice of regularization parameters which differs for each experiment. We show the benefits of using a learned block iterative shrinkage thresholding algorithm (LBISTA) that is able to learn the choice of regularization parameters, without the need to manually select them. In addition, LBISTA enables the determination of a suitable weight matrix to solve the underlying inverse problem. Therefore, in this paper we present LBISTA and compare it with state-of-the-art block iterative shrinkage thresholding using synthetically generated and experimental test data from active thermography for defect reconstruction. Our results show that the use of the learned block-sparse optimization approach provides smaller normalized mean square errors for a small fixed number of iterations. Thus, this allows us to improve the convergence speed and only needs a few iterations to generate accurate defect reconstruction in photothermal super-resolution imaging.

**Keywords:** active thermal imaging; block-sparsity; deep unfolding; defect reconstruction; iterative shrinkage thresholding algorithm; laser thermography; learned optimization; neural network; regularization



**Citation:** Hauffen, J.C.; Kästner, L.; Ahmadi, S.; Jung, P.; Caire, G.; Ziegler, M. Learned Block Iterative Shrinkage Thresholding Algorithm for Photothermal Super Resolution Imaging. *Sensors* **2022**, *22*, 5533. <https://doi.org/10.3390/s22155533>

Academic Editor: Carosena Meola

Received: 28 June 2022

Accepted: 21 July 2022

Published: 25 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Active thermal imaging is a nondestructive testing technique applied in many areas such as in production industries (3D printing, car/aerospace manufacturing [1–3]), in medicine (breast cancer diagnosis [4–6]), or in dentistry [7]. Most of the research in active thermal imaging deals with the location of anomalies or defects in materials. These defects are detected by actively heating up a specimen, e.g., by laser or flash illumination, which yields photothermal imaging [8]. The resulting heat diffusion after heat excitation is observed with an infrared (IR) camera. The observed heat flow by the IR camera is then evaluated using the generated thermal film sequence. Defect indications are clearly visible in these thermal film sequences if the heat accumulates at a void or if changes in the heat flow gradient are visible due to a crack. Instead of photothermal imaging, other thermographic imaging techniques can be applied, where e.g., induction coils are used as a heating source [9]. Apart from active thermography, there are other nondestructive testing techniques such as ultrasound testing (UT), x-ray computed tomography (CT), and eddy current testing (ET). Active thermal imaging outperforms these techniques as it is a contactless technique (UT needs contact), is relatively inexpensive and flexible in terms of specimen sizes (compared to CT), and also enables us to investigate deeper-lying defects (compared to ET). However, active thermography suffers from low spatial resolution due

to the blurring nature of heat diffusion. Therefore, it is of huge interest to overcome the spatial resolution limitations, which can be circumvented by utilizing suitable image processing techniques.

Knowing the resulting heat distribution by active thermal imaging for a defect-free body, the exact defect distribution of the defective body can be found by solving an underlying inverse problem [10]. This severely ill-posed thermal inverse problem can be solved by common techniques such as singular value decomposition or Bayesian approaches [11,12]. However, these techniques are computationally intensive in contrast to sparsity-exploiting techniques such as the least absolute shrinkage and selection operator (LASSO) methods [13].

Sparse signal reconstruction is becoming increasingly popular, especially in the industrial sector. These algorithms are highly attractive as, in the field of compressed sensing, knowledge of the device or target can be exploited for a high-quality reconstruction [14,15]. Consequently, in the last few years, optimization algorithms like ISTA, FISTA [16], learned algorithms such as LISTA [17], and adaptations like ALISTA [18] or LISTA-AT [19] gained much attention, especially in the field of computational imaging. Unfolding methods are already established in research and are used to solve inverse problems [20]. Especially for multiple-measurement-vector (MMV) problems, group LASSO as a block-regularization method can be applied to find an accurate solution [21,22]. Block-sparse recovery has been extensively studied with performance guarantees, also in fields of applied mathematics [23–25] and has been recently applied to active thermal imaging for defect detection in nondestructive testing [26–32]. Using multiple and different blind illumination patterns for heating in active thermal imaging results in an MMV problem. Blind illumination patterns assume that the exact position of illumination is unknown in industrial applications [29]. This occurs, for example, when the heat source or the specimen is held by a robot that has a certain amount of positional noise. Unfortunately, so far, one needs to carefully choose the regularization parameters for the block optimization problem due to the sensitivity of the thresholding [31], which made this approach tedious to implement. In addition, it is very likely that an optimal manual choice of these parameters differs for different investigated specimens. Additionally, it is unclear whether the optimal parameters have been really found. Finally, the algorithm needs hundreds of iterations to reach convergence even if the optimal regularization parameters have been found.

These algorithms are analytically well understood, but difficult to apply in practice. The main disadvantage of these classical approaches to solving the inverse thermal problem is the high computational effort, e.g., due to many iterations or expensive computations. Additionally, these algorithms depend on parameters, such as the block sparse regularizer, which directly affects the reconstruction quality or stability of the respective method. They are therefore impractical, especially in the manufacturing industry where one wants to reconstruct defects quickly and accurately to see if a large number of samples are defective. In this work, a solution is provided for the aforementioned problem by utilizing a learned iterative joint sparsity approach called the learned block iterative shrinkage thresholding algorithm (LBISTA). This approach applies neural network learning, e.g., a suitable choice of the regularization parameters or the training weights. In our approach, the training data are generated synthetically using uniformly random distributed defect distribution and corresponding thermal film sequences. Utilizing LBISTA for photothermal super resolution (SR) imaging avoids the manual choice of regularization parameters and allows for a higher convergence speed than the block iterative shrinkage thresholding algorithm (Block-ISTA or BISTA). It should be noted that LBISTA needs training times of a few hours, depending on the used hardware, to establish a trained network. Once the network is trained, LBISTA can be applied to photothermal SR imaging and generates super-resolved images much faster than BISTA. Thus, this could be highly attractive in production industries such as in 3D printing, where thermal imaging is used to monitor the processing [33–36], avoiding further post-inspections after the 3D printing has finished [26] and thus enables smaller processing time and costs. In general, such a learned block-sparse recovery approach is

not new as learned group LASSO approaches already exist [37–39]. In [40] we were able to apply this idea to various BISTA-type algorithms but we did not focus on the theoretical aspects of this approach. Thus, the novel and main contributions of this manuscript are:

- implementation of LBISTA as a method to solve the inverse problem in photothermal SR imaging for nondestructive testing without a manual choice of regularization parameters;
- results of applying LBISTA to synthetic and experimental test data based on experiments performed with a specimen made of steel S235JR to examine defects that are not resolvable with conventional flash thermography;
- comparison of LBISTA with state-of-the-art BISTA for photothermal SR imaging;
- parameter studies of (i) tied and (ii) untied LBISTA: (1) studies of the hyperparameters, (2) studies of the parameters set to generate the synthetic training data (forward model) for photothermal SR imaging. In the tied case we train the same particular trainable parameter for each layer and in the untied case we can train it for each layer individually.

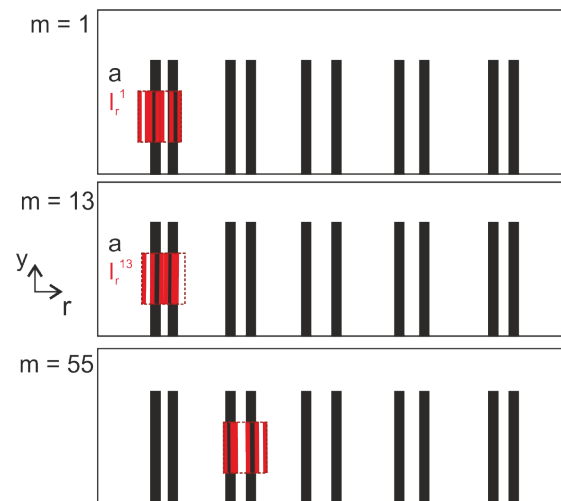
The outline of this manuscript is as follows. Section 2 explains the mathematical model in active thermal imaging and in photothermal SR imaging. In this section, it is shown how photothermal SR imaging can be cast into a block-sparse recovery problem. In Section 3, the implementation of LBISTA is shown, i.e., the description of the training data and the description of the proposed algorithms (the code is included in [https://github.com/BAMresearch/Photothermal\\_SR\\_Net](https://github.com/BAMresearch/Photothermal_SR_Net), accessed on 23 January 2020). Section 4 presents the results after applying LBISTA to synthetic data and to experimental data acquired from thermal imaging during structured laser illumination measurements. The results are compared with state-of-the-art BISTA in conventional photothermal SR imaging (where Block-ISTA or Block-FISTA are often used as optimization techniques within the iterative joint sparsity approach, see e.g., [32]). In addition, parameter studies are shown indicating the influence of the hyperparameters and the training parameters set to generate synthetic training data on the performance of LBISTA. Finally, in Section 5, the main achievements with LBISTA are listed and some future perspectives are given.

## 2. Mathematical Model in Active Thermal Imaging

In these studies, we work with a defect pattern assuming it does not change over the height (lines as defects, see Figure 1) so that we can simplify our model by calculating the mean over the vertically arranged pixels (see dimension  $y$  in Figure 1). The IR camera measures temperature values that can be described in active thermal imaging by a convolution in space and time (denoted in the following as  $*_{r,t}$ ) of the fundamental solution of the heat diffusion equation and the absorbed excitation energy (see Green's function solution approach in [10]). In the following, these measured temperature values are described by a discrete temperature matrix

$$\mathbf{T} = \boldsymbol{\phi}_{\text{PSF}} *_{r,t} \mathbf{X} \quad (1)$$

with  $\mathbf{T} \in \mathbb{R}^{N_r \times N_t}$ .  $\boldsymbol{\phi}_{\text{PSF}} \in \mathbb{R}^{N_r \times N_t}$  represents the discrete equivalent of the fundamental solution of the heat diffusion equation, where  $N_r$  stands for the number of measured pixels in the dimension  $r$  and  $N_t$  for the number of measured images in time domain. PSF indicates that the values in the matrix refer to the well-known thermal point spread function (PSF). The discrete equivalent of the absorbed heat flux density is designated as  $\mathbf{X} \in \mathbb{R}^{N_r \times N_t}$ . It considers the irradiance in discrete space ( $r$ ) and time ( $t$ ) dimension denoted by  $\mathbf{I}_{r,t} \in \mathbb{R}^{N_r \times N_t}$  as well as the absorption coefficient of the material under investigation in discrete space denoted by  $\mathbf{a} \in \mathbb{R}^{N_r}$ , so that  $\mathbf{X} = \mathbf{I}_{r,t} \circ \mathbf{A}$ , where  $\circ$  is the Hadamard-product and  $\mathbf{A} = [\mathbf{a} \dots \mathbf{a}] \in \mathbb{R}^{N_r \times N_t}$



**Figure 1.** Exemplary specimen with defects shown as blackened stripes. A laser line array with twelve laser lines is used. Here  $y$  stands for the spatial vertical dimension and  $r$  stands for the spatial horizontal dimension. The illumination pattern differs for each measurement in the illuminated position and in the number of laser lines (indicated by the red area) which are switched on (randomly chosen). The dashed frame around the pattern indicates the covered illuminated area if all laser lines are switched on.

As described in the introduction, we are dealing with blind structured illumination, which means that we do not know the exact illuminated positions at the specimen under investigation using a laser as a heat source. This results in a small but relevant change of Equation (1) into:

$$\mathbf{T}^m = \boldsymbol{\phi}_{\text{PSF}} *_{r,t} \mathbf{X}^m = \boldsymbol{\phi}_{\text{PSF}} *_{r,t} (\mathbf{I}_{r,t}^m \circ \mathbf{A}), \quad (2)$$

with  $m = 1 \dots N_{\text{meas}}$ , where  $N_{\text{meas}}$  denotes the number of measurements. The spatial and temporal distribution of the absorbed heat flux density  $\mathbf{X}^m = \mathbf{I}_{r,t}^m \circ \mathbf{A}$  varies for each measurement  $m$  as the illuminated spatial position varies for each measurement  $m$ . This represents an MMV problem in the forward problem formulation.

### 2.1. Defect Detection and Reconstruction

In our case, defects on a material are physically described as a change of optical absorption coefficient since the defect region (high value:  $\max\{\mathbf{a}\} \sim 0.95$ ) differs from the defect-free region (low value:  $\min\{\mathbf{a}\} \sim 0.15$ ) in its material properties. Thus, the absorption pattern matrix  $\mathbf{A}$  (see Equation (2)) represents the defect pattern that is of interest. Since we assume blind structured illumination, we cannot separate the illumination  $\mathbf{I}_{r,t}^m$  from the absorption pattern matrix  $\mathbf{A}$ . However, we can reformulate Equation (2) by extracting the known illumination pulse duration  $\mathbf{i}_t \in \mathbb{R}^{N_t}$ , which does not change over the measurements, from  $\mathbf{I}_{r,t}^m = \mathbf{i}_r^m \otimes \mathbf{i}_t$  with  $\mathbf{i}_r^m \in \mathbb{R}^{N_r}$  yielding:

$$\begin{aligned} \mathbf{T}^m &= \boldsymbol{\phi}_{\text{PSF}} *_{r,t} [(\mathbf{i}_r^m \otimes \mathbf{i}_t) \circ \mathbf{A}] \\ &= (\boldsymbol{\phi}_{\text{PSF}} *_{t} \mathbf{i}_t) *_{r} [(\mathbf{i}_r^m \otimes \mathbf{1}) \circ \mathbf{A}] \\ &= \boldsymbol{\phi} *_{r} \mathbf{X}_r^m, \end{aligned} \quad (3)$$

where  $\boldsymbol{\phi} \in \mathbb{R}^{N_r \times N_t}$  represents the thermal PSF, that considers the illumination pulse length. Moreover,  $\mathbf{X}_r^m \in \mathbb{R}^{N_r \times N_t}$ .

### 2.2. Photothermal Super Resolution

The photothermal SR technique refers to structured illumination in space and promotes high resolvability by e.g., illuminating a specimen using a narrow laser line. The realization

of photothermal SR is possible if a lot of measurements/illuminations are performed with the narrow laser line resulting in an MMV problem. In our previous studies, we have shown that one can scan step by step with a single laser line with submillimeter position shifts to provide high resolvability [31]. This technique relies on spatial frequency mixing of the illumination pattern and the target pattern, here the absorption pattern. Spatial frequency mixing allows us to generate higher frequency components, enabling super-resolution. Since  $i_r^m$  has a certain width in space according to the used narrow laser line width, the use of photothermal SR results in equal spatial distributions for  $\sum_m i_r^m$  and  $\mathbf{a}$  (see [31]). Hence, the initial goal is to determine the variable  $\mathbf{X}_r^m$ . We measure  $\mathbf{T}^m$  with the IR camera and we can determine  $\phi$  analytically. This underlying inverse problem has been considered many times in our studies [29–32] and a promising approach to obtain  $\mathbf{X}_r$  was block-regularization.

### 2.3. Block-Minimization Problem

Before using block-regularization, we eliminate the time dimension to reduce the data size. The straightforward way shown in our previous studies [30–32] is to extract one thermogram per measurement in the time domain, which exhibits the highest SNR—this is the so-called maximum thermogram (MT) method. We can reformulate Equation (3) after applying the MT method separately to each measurement  $m$  yielding the reduced (reduc) data with:

$$\mathbf{T}_{\text{reduc}}^m = \phi_{\text{reduc}} *_{r} \mathbf{x}_{\text{reduc}}^m \tag{4}$$

with  $\mathbf{T}_{\text{reduc}}^m$ ,  $\phi_{\text{reduc}}$ , and  $\mathbf{x}_{\text{reduc}}^m \in \mathbb{R}^{N_r}$ . To cast this problem into a block-sparse problem, we use the knowledge that the absorption pattern remains the same for each measurement, i.e., for each  $m = 1, \dots, N_{\text{meas}}$  we have  $\text{supp}(\mathbf{x}_{\text{reduc}}^m) \subset \text{supp}(\mathbf{a})$ , where

$$\text{supp}(\mathbf{x}_{\text{reduc}}^m) = \{k \mid |x_{\text{reduc}}^m[k]| \neq 0, k = 1, \dots, N_r\}.$$

By casting these measurements into a matrix

$$\mathbf{X}_{\text{reduc}} = (\mathbf{x}_{\text{reduc}}^1, \dots, \mathbf{x}_{\text{reduc}}^m, \dots, \mathbf{x}_{\text{reduc}}^{N_{\text{meas}}}) \in \mathbb{R}^{N_r \times N_{\text{meas}}} \tag{5}$$

we obtain the block-sparse vector  $\tilde{\mathbf{x}} = \text{vec}(\mathbf{X}_{\text{reduc}}^T) \in \mathbb{R}^{N_r \cdot N_{\text{meas}}}$  with  $N_r$  blocks of length  $N_{\text{meas}}$ . Using e.g., Block-ISTA then tries to minimize the following term:

$$\hat{\mathbf{X}}_{\text{reduc}} = \arg \min_{\mathbf{X}_{\text{reduc}}} \sum_{m=1}^{N_{\text{meas}}} \sum_{k=1}^{N_r} |(\phi_{\text{reduc}} *_{r} \mathbf{x}_{\text{reduc}}^m)[k] - \mathbf{T}_{\text{reduc}}^m[k]|^2 + \lambda \|\mathbf{X}_{\text{reduc}}\|_{2,1} \tag{6}$$

with the block- or joint-sparsity inducing  $\ell_{2,1}$ -norm  $\|\mathbf{X}_{\text{reduc}}\|_{2,1} = \sum_{k=1}^{N_r} \sqrt{\sum_{m=1}^{N_{\text{meas}}} |x_{\text{reduc}}^m[k]|^2}$  this is important as we are working with blind structured illumination. Block-ISTA is stated as the following fixed-point iteration (inspired by [27,41]) and were also used in our previous studies [30–32] to solve (6),

$$\hat{\mathbf{x}}_{(i)}^m = \eta_{\lambda} \left( \hat{\mathbf{x}}_{(i-1)}^m - 2\gamma \phi_{\text{reduc}} *_{r} \left( \phi_{\text{reduc}} *_{r} \hat{\mathbf{x}}_{(i-1)}^m - \mathbf{T}_{\text{reduc}}^m \right) \right), \tag{7}$$

$i = 1, \dots, N_{\text{iter}}$ . The number of iterations  $N_{\text{iter}}$  is chosen such that convergence is reached and the soft block-thresholding operator  $\eta_{\lambda}$  can be computed by:

$$\eta_{\lambda}(\hat{\mathbf{x}}^m)[k] = \max \left\{ 0, 1 - \frac{\lambda}{\sqrt{\sum_{m=1}^{N_{\text{meas}}} |\hat{\mathbf{x}}^m[k]|^2}} \right\} \hat{\mathbf{x}}^m[k]. \tag{8}$$

The step size is defined by  $\gamma \in (0, \frac{1}{L}]$ , where  $L$  is the Lipschitz constant of the gradient of the data fitting term. The value for  $L$  can be determined empirically or by  $L = 2\pi \|\hat{\phi}_{\text{reduc}}\|_{\infty}$  where  $\|\cdot\|_{\infty}$  denotes the supremum/infinity norm and  $\hat{\phi}_{\text{reduc}}$  denotes the Fourier transform of  $\phi_{\text{reduc}}$  [27]. The value for  $\lambda$  is chosen empirically for (7)—Block-ISTA, e.g.,  $\lambda = 4 \times 10^{-3}$ . The same value was used in our previous work, where we utilized Block-FISTA to examine the same specimen [30].

### 3. Learned Block Iterative Shrinkage Thresholding Algorithm

The success of LBISTA is strongly dependent on the choice of the training data and on the implementation of the training which will be explained in detail in the following subsections.

#### 3.1. Training Data

To create the training data, we follow the forward problem shown in Equation (4). To define  $\mathbf{x}_{\text{reduc}}^m$ , we have different parameters, which can be varied: the defect width (see the width of a black stripe in Figure 1), the laser line width (see the width of a single laser line shown in Figure 1), the absorption coefficient for a defective and a defect-free region, the number of defects (probability of nonzero in space for  $\mathbf{a}$  if we assume that the defect-free regions have an absorption coefficient of 0, i.e., the element  $a_i$  of  $\mathbf{a}$  is equal to 1 with a given probability indicating a defect - we then widen these defects according to the defect width, i.e., setting the respective elements of  $\mathbf{a}$  also to 1), and the number of laser lines (probability of nonzero in space for  $\mathbf{i}_r^m$  if we assume that  $\mathbf{i}_r^m$  equals zero at the nonilluminated positions, works similar to generating  $\mathbf{a}$ ). Hence, we implemented  $\mathbf{x}_{\text{reduc}}^m$  as  $\mathbf{i}_r^m \circ \mathbf{a}$ , determined  $\phi_{\text{reduc}}$  analytically and calculated  $\mathbf{T}_{\text{reduc}}^m$ .

#### 3.2. Training Implementation

Inspired by the code implementation of [42] we consider layer-wise learning instead of end-to-end learning for LBISTA. Each network depends on a set  $\Theta$  of trainable variables and on a number of layers  $K$ . We define the layers for tied LBISTA as follows

$$\hat{\mathbf{x}}_{(i)}^m = \eta_{\lambda^{(i-1)}} \left( \mathbf{s} * \hat{\mathbf{x}}_{(i-1)}^m + \mathbf{b} * \mathbf{y}^m \right) \quad (9)$$

and for untied LBISTA

$$\hat{\mathbf{x}}_{(i)}^m = \eta_{\lambda^{(i-1)}} \left( \mathbf{s}^{(i-1)} * \hat{\mathbf{x}}_{(i-1)}^m + \mathbf{b}^{(i-1)} * \mathbf{y}^m \right), \quad (10)$$

$m = 1, \dots, N_{\text{meas}}$ . Note that we simplified the fixed point representation with  $* := *_r$  and  $\mathbf{y}^m := \mathbf{T}_{\text{reduc}}^m \mathbf{y}$ . We define the set of trainable variables  $\Theta_{\text{tied}} = \{\mathbf{s}, \mathbf{b}\} \cup \bigcup_{i=1}^K \Theta_{\text{tied}}^{(i)}$  for tied LBISTA (9), where

$$\Theta_{\text{tied}}^{(i)} = \left\{ \lambda^{(i-1)} \right\}, i = 1, \dots, K.$$

We initialize these variables with  $\mathbf{b} = 2\gamma\phi_{\text{reduc}} \in \mathbb{R}^{N_r}$ ,  $\mathbf{s} = \mathbf{e} - \mathbf{b} * \phi_{\text{reduc}} \in \mathbb{R}^{N_r}$ , where  $\mathbf{e} = [1, 0, \dots, 0]^T \in \mathbb{R}^{N_r}$  for some step size  $\gamma \in (0, \frac{1}{L}]$ . In addition, we use the same initialization for the regularization parameter in each layer  $i = 0, \dots, K - 1$  with  $\lambda^{(i)} = 4 \times 10^{-3}$ .

For untied LBISTA (10) we define  $\Theta_{\text{untied}} = \bigcup_{i=1}^K \Theta_{\text{untied}}^{(i)}$  where

$$\Theta_{\text{untied}}^{(i)} = \left\{ \mathbf{s}^{(i-1)}, \mathbf{b}^{(i-1)}, \lambda^{(i-1)} \right\}, i = 1, \dots, K.$$

Untied means that each layer has its own set of trainable variables  $\mathbf{s}, \mathbf{b}$ , we use the same initialization for each layer. Therefore, we initialize  $\mathbf{b}^{(i)} = 2\gamma\phi_{\text{reduc}}$ ,  $\mathbf{s}^{(i)} = \mathbf{e} - \mathbf{b} * \phi_{\text{reduc}}$  in each layer  $i = 0, \dots, K - 1$ . These variables are initialized such that we would obtain the original Block-ISTA if we do not apply the training.



The training follows Algorithm 1 and can take a large amount of time depending on the choice of the termination condition and therefore the number of iterations for each layer and each refinement. Hence, the training time also depends on the number of refinements and on the number of layers. The calculation time for one iteration is around 150 ms. Within a layer, we usually have 10,000 iterations and with each refinement, an additional 1000 iterations. We have used the GPU Quadro RTX 8000 to perform the training.

---

**Algorithm 1:** LBISTA, implementation of training
 

---

**Input** : Training rate  $t_r$ , refinements  $f$ , exact solution  $\mathbf{x}^{*,m}$  and trainable Variables  $V$  with  $case = tied$  or  $case = untied$

```

for  $i = 1, \dots, K$  do
   $loss = \frac{1}{2} \sum_{m=1}^{N_{meas}} \|\hat{\mathbf{x}}_{(i)}^m - \mathbf{x}^{*,m}\|_2^2$ 
  for  $t = 1, \dots, Maxiter$  do
    AdamOptimizer( $t_r$ ).min( $loss$ , var list= $\Theta_{case}^{(i)}$ )
  end
  for  $f_m$  in refinements  $f$  do
    for  $t = 1, \dots, Maxiter$  do
      AdamOptimizer( $t_r \cdot f_m$ ).min( $loss$ , var list= $\Theta_{case}$ )
    end
  end
end

```

---

## 4. Evaluation of LBISTA

### 4.1. Numerical Results

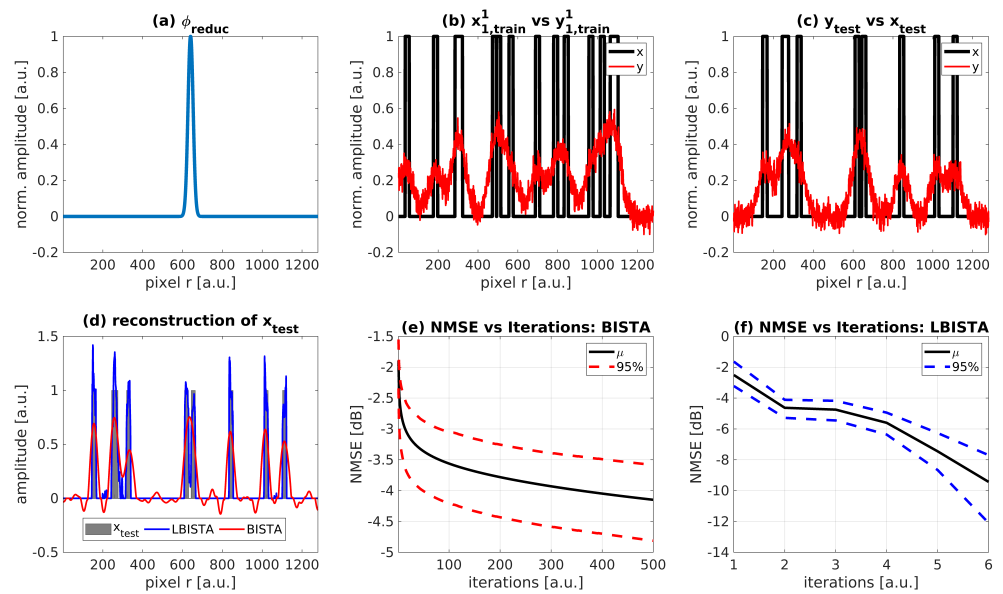
To study the performance of the LBISTA, we first show numerical studies considering the block-sparse convolution problem  $\mathbf{T}_{reduc}^m = \boldsymbol{\phi}_{reduc} * \mathbf{x}_{reduc}^m$ . There is also a theoretical analysis of this approach in the setting of sparse recovering—in [18,43] a lower and upper bound for the reconstruction error is derived. This could also be applied to the block-setting and especially to the setting of photothermal SR imaging, but this is part of ongoing work.

In the following, we use the same notation as shown in Equation (4) with  $\mathbf{y}_{train}^m = \boldsymbol{\phi}_{reduc} * \mathbf{x}_{train}^m$  and  $\mathbf{y}_{train}^m, \mathbf{x}_{train}^m \in \mathbb{R}^{N_r}$ . An analytical formulation of  $\boldsymbol{\phi}_{reduc}$  can be found in [30]. Note that  $\boldsymbol{\phi}_{reduc}$  is generated after applying the MT method. Table 1 shows the used parameters for training with synthetic data based on  $\mathbf{x}_{train}^m = \mathbf{i}_{train}^m \circ \mathbf{a}$ . More precisely, the cells corresponding to defect pattern determine the values in  $\mathbf{a}$  and the cells corresponding to illumination pattern the values in  $\mathbf{i}_{train}^m \in \mathbb{R}^{N_r}$ . Further, Gaussian noise is added to generate  $\mathbf{y}_{train}^m$  based on the shown SNR value in the table. For the determination of Gaussian noise variance  $\sigma_{noise}^2$ , we have used:  $SNR = \mu^2 / \sigma_{noise}^2 = 8$  dB with the squared signal mean  $\mu^2 = PNZ$ . PNZ denotes the probability of nonzeros. LBISTA is trained for the dimensions  $N_r = 1280$  (number of pixels),  $N_{meas} = 150$  (number of measurements based on differently chosen illumination and therefore differently generated  $\mathbf{x}_{reduc}^m$ ). The maximum number of iterations  $Maxiter$  has been decreased for untied LBISTA since updating more variables causes higher computation times for each iteration.

In the following, the training data are determined by  $\mathbf{x}_{i,train}^m, \mathbf{y}_{i,train}^m \in \mathbb{R}^{N_r}$  and  $i \in \{1, \dots, N_B\}$ ,  $m \in \{1, \dots, N_{meas}\}$ , so that  $(\mathbf{X}_{i,train})_{i=1}^{N_B} = (\mathbf{x}_{i,train}^1, \dots, \mathbf{x}_{i,train}^{N_{meas}})_{i=1}^{N_B} \in \mathbb{R}^{N_B \times N_r \times N_{meas}}$  and similar for  $(\mathbf{Y}_{i,train})_{i=1}^{N_B} \in \mathbb{R}^{N_B \times N_r \times N_{meas}}$ , e.g., Figure 2b. LBISTA reaches a smaller normalized mean square error (NMSE) much faster than BISTA, Figure 2e,f. According to the curves in (d), LBISTA outperforms BISTA in terms of reconstruction using  $\boldsymbol{\phi}_{reduc}$  (see (a)) and  $\mathbf{y}_{test}$  (see (c)). Of course, it strongly depends on the choice of  $\lambda$  how well BISTA reconstructs. Therefore, we chose the same initial values of  $\lambda$  for BISTA and LBISTA to have a better comparison.

**Table 1.** Parameters used in LBISTA for problem definition and training.

<b>Gaussian Noise in <math>y_{\text{train}}^m</math></b>	SNR = 8 dB
<b>Defect pattern</b>	defect width = 1 mm
	defect sparsity (PNZ) = 0.01
	absorption coefficient = $\{0, 1\}$
<b>Illumination pattern</b>	laser line width = 0.8 mm
	illumination sparsity (PNZ) = 0.01
<b>Training parameters</b>	refinements $f$ : $f_m = \{0.5, 0.1, 0.05\}$
	training rate $t_r = 0.001$
	initial lambda $\lambda = 0.004$
	number of layers $K = 6$
	step size $\gamma = \frac{1}{\sqrt{2}}$
	batch number $N_B = 150$
tied LBISTA	max. number of iterations $Maxiter = 10^5$
untied LBISTA	max. number of iterations $Maxiter = 10^4$



**Figure 2.** Exemplary synthetic training and synthetic test data as well as NMSE performance studies. (a): shape of used  $\phi_{\text{reduc}}$  in training, (b,c): comparison of the curve shapes for exemplary training and test datasets with e.g.,  $x_{1,\text{train}}^1 \in \mathbb{R}^{N_r}$ , (d): reconstruction of  $x_{\text{test}}$  using convolution-based tied LBISTA (6 iterations/layers) vs. BISTA (1000 iterations), (e): NMSE over iterations for BISTA with 95% confidence interval, (f): NMSE over iterations for tied LBISTA with 95% confidence interval.

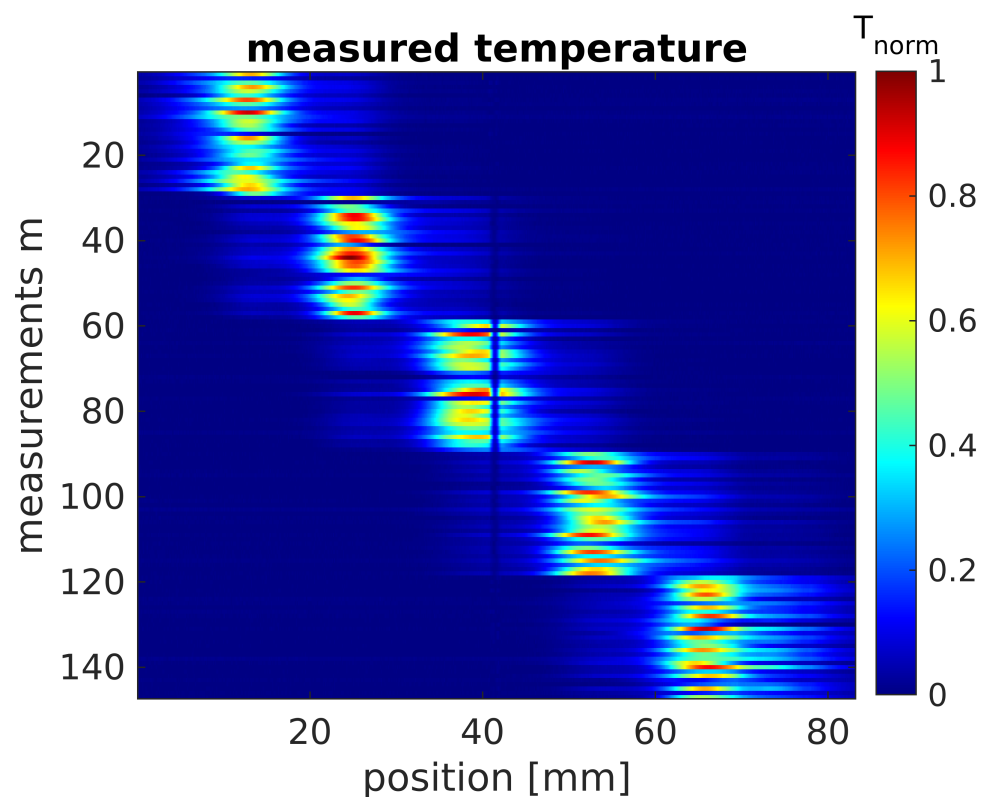
Figure 2 shows the results with synthetic test data  $y_{\text{test}} \in \mathbb{R}^{N_r}$  and ground truth  $x_{\text{test}} \in \mathbb{R}^{N_r}$ . In the following section we use real measurement test data instead of  $y_{\text{test}}$  on LBISTA, but still trained with synthetic training data.

#### 4.2. Evaluation of LBISTA with Experimental Data from Active Thermal Imaging

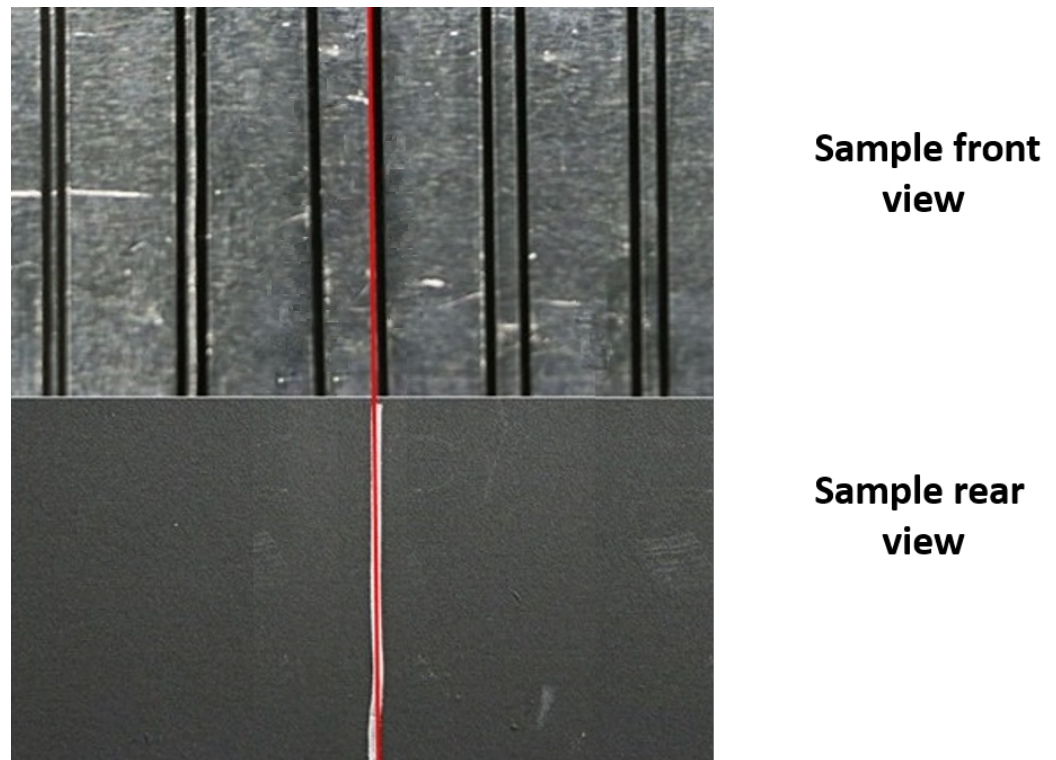
Instead of creating test data synthetically as in the previous section, we apply LBISTA to real experimental data, which were measured with the IR camera (InfraTec ImageIR 9300,  $1280 \times 1024$  pixels full frame, spectral range: 3–5  $\mu\text{m}$ ). We have used the same dataset as in a previous publication (see [30] as a reference). We have used the IR camera in transmission configuration (illumination on the front of the specimen and observation of the back side



with the IR camera). The specimen has a thickness of 3 mm and 5 blackened stripe pairs on the front, see Figure 1. All in all, we performed around 150 measurements, resulting in  $m = 1 \dots 150$ , 30 measurements per pair. Within these 30 measurements, we shifted the position by 0.4 mm twice: once after the first and again after the second ten measurements. After 30 measurements the position was shifted to the next pair. This is repeated until we scanned the whole specimen (Figure 1). Each measurement differs by randomly (uniform distribution) switching on a certain number of laser lines out of twelve (at least one laser line is turned on for each measurement). Figure 3 shows the experimental test data after removing the dimension  $y$  by calculating the mean over the vertically arranged pixels and eliminating the time dimension by applying the MT method ([30–32]). Nevertheless, there is a disturbance in the data at position 41 mm represented by a straight line over all measurements. This comes from the fact that the specimen had a marker line on the rear side to have an orientation about the defects in thermographic transmission configuration, Figure 4.



**Figure 3.** Test data from photothermal structured illumination measurements. To obtain this image, averaging over the vertically arranged pixels as well as applying the MT method (see [30–32]) to eliminate the time dimension is necessary. The blue marker at position = 41 mm refers to the marker line on the investigated specimen (Figure 4).



**Figure 4.** Sample front and rear view to understand origin of white marker line at position 41 mm, Figure 3. The rear side of the sample has been observed by the IR camera and the front side has been excited by a laser array. The defects on the front side are shown by blackened stripes (5 pairs). The marker line on the rear side is in the same position as the left edge of the right stripe from the third pair. The distances between the pairs are (left to right): 0.5, 1, 3, 2, 1.3 mm. A stripe is 1 mm wide and the distance of each pair is around 10 mm.

In the following, we compare the result after applying BISTA or LBISTA. Calculating the sum over all measurements as explained in Section 2.2 results in an approximation of the defect pattern, as shown in Figure 5. The dashed blue curve indicates the outcome of the BISTA algorithm and shows only more or less good indications for the third pair with the largest distance between the stripes (around position 37 . . . 44 mm). The result of block fast ISTA (Block-FISTA), the purple dashed line, also shows only a more or less good indication for the defects, but still improves the resolution and even shows the fourth defect. Block-FISTA is also a block iterative shrinkage algorithm but with a better convergence rate than BISTA [16,27], but also has the main drawback of finding optimal regularization parameters empirically. In contrast, the application of LBISTA results in very good indications for three of five pairs. Only the stripes with the largest and closest distance to each other (position 11 . . . 14 mm) cannot be clearly recognized. The other three stripes are very well reconstructed, even the stripe width can be recognized. Since the marker line at position 41 mm disturbs the pattern recognition, the stripes with the largest distance to each other are hard to resolve. For now, parameters from Table 1 have been used, but it is still unclear whether these are the best choice. Therefore, we varied some of these to study their impact on the result of tied and untied LBISTA, see Figure 6 and Figure 7, respectively.

Figure 6:

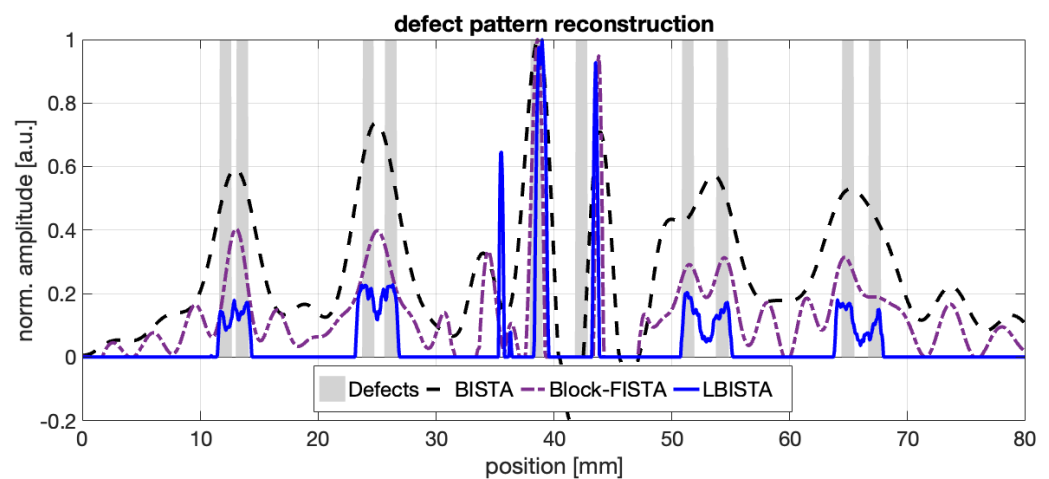
- Training batch size (a,b): Varying the number of batches between  $N_B = 100$  and  $N_B = 200$  does not really change the result.
- Defect pattern (c,d,e,f): In contrast, changing the sparsity of the defects from  $PNZ \in [0.005, 0.03]$  (c.f. (e,f)) or the defect width (see (c,d)) leads to significant changes in the results such that the result in (d) is even able to clearly detect both stripes in the first pair with the smallest distance between the stripes. In (f), obviously a too small

sparsity has been used as the PNZ is quite high in comparison to the chosen PNZ in (e). Thus, it is obviously beneficial to know roughly the sparsity.

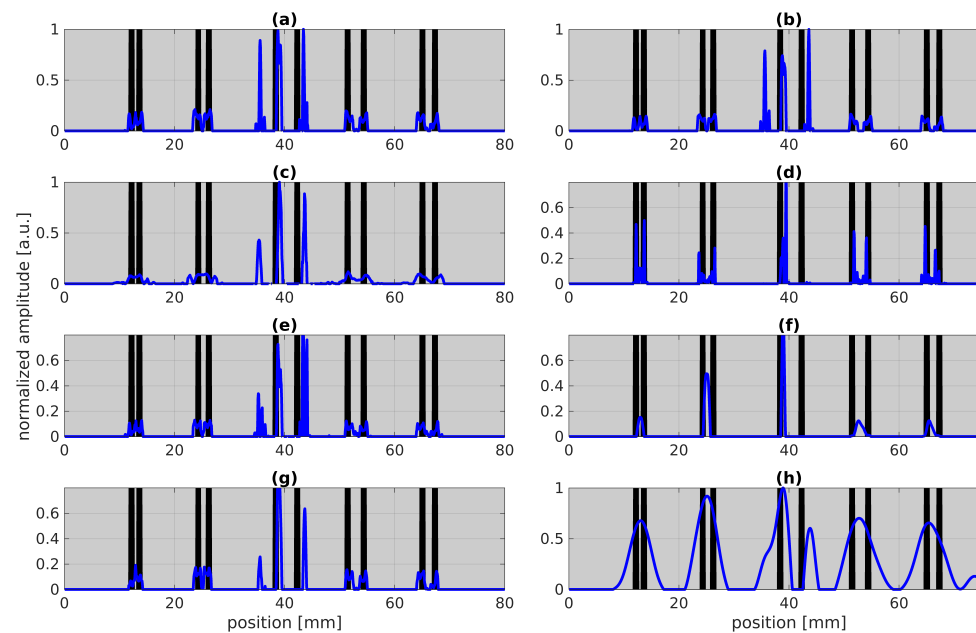
- Training rate and number of layers (g,h): In (g) we get similarly good results just by using one refinement instead of three as used in Figure 5. Only using one refinement could save a lot of time during training. In (h), a rather bad result is shown, where we used one refinement and only three layers. Thus, these parameters should be chosen high enough so that we reach convergence within training.

Figure 7:

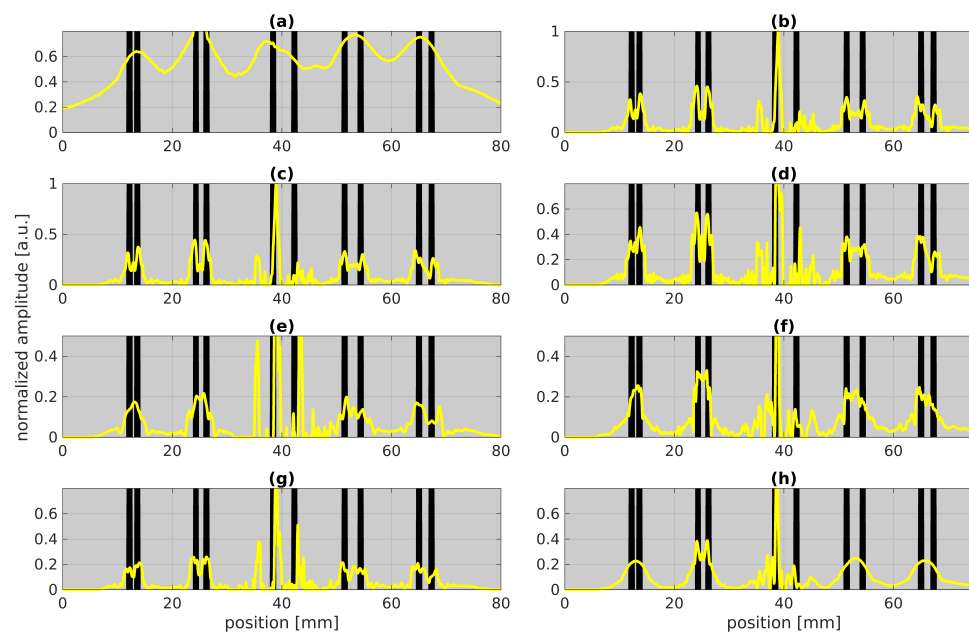
- Absorption coefficient and training iterations (a,b): In contrast to tied LBISTA, untied LBISTA exhibits significant changes by varying the absorption coefficient in training. According to our studies, an absorption coefficient of  $\{0, 1\}$  is a good choice. Further, increasing the maximum number of iterations as shown in (b) can enhance the reconstruction quality so that all stripes could be indicated very well except for the middle stripe pair (most likely due to the marker line).
- Training batch size (c,d): The result in (c) using fewer batches and iterations as in (b) shows that similarly good results can be achieved. Increasing the number of batches to  $N_B = 200$  as shown in (d) can further enhance the reconstruction quality as now even the middle defect pair could be clearly resolved.
- Defect width (e,f): The variation of the defect width rather degrades the reconstruction result as shown in (e,f). This means that the default choice performs best for untied LBISTA. However, with tied LBISTA (Figure 6c,d) we could see improvements by changing the parameter of the defect width.
- Defect sparsity (g,h): The variation of the sparsity in untied LBISTA in (g,h) confirms our investigations in tied LBISTA (see (e,f)), deterioration by using a too small sparsity of  $PNZ = 0.03$ .



**Figure 5.** Pattern reconstruction of blackened stripes as shown in Figure 4 using BISTA, Block-FISTA, and tied LBISTA. For LBISTA the parameters in Table 1 were used. Hence, synthetic training data have been used. The BISTA and Block-FISTA result has been created by using  $\gamma = \frac{1}{\sqrt{2}}$ ,  $\lambda = 0.004$ ,  $N_{\text{iter}} = 500$  (selected from [30]). The marker line can be seen in the BISTA curve exhibiting normalized amplitude values  $< 0$  at position 41 mm.



**Figure 6.** Parameter studies for tied LBISTA applied to experimental test data shown in Figure 3. The blackened areas represent the blackened stripes for better orientation. The blue curves stand for the result after applying the tied LBISTA to the experimental test data. All images are based on the parameter choice shown in Table 1. The following parameters have been changed, respectively: (a)  $N_B = 100$ , (b)  $N_B = 200$ , (c) defect width = 0.5 mm, (d) defect width = 2 mm, (e) defect sparsity PNZ = 0.005, (f) defect sparsity PNZ = 0.03, (g) only one refinement  $f_m = \{0.5\}$ , (h) only three layers  $K = 3$  and one refinement  $f_m = \{0.01\}$ .



**Figure 7.** Parameter studies for untied LBISTA applied to experimental test data shown in Figure 3. The blackened areas represent the blackened stripes for better orientation. The yellow curves stand for the result after applying the untied LBISTA to the experimental test data. All images are based on the parameter choice shown in Table 1. The following parameters have been changed, respectively: (a) absorption coefficient =  $\{0.3, 0.7\}$ , (b)  $Maxiter = 10^5$ , (c)  $N_B = 100$ , (d)  $N_B = 200$ , (e) defect width = 0.5 mm, (f) defect width = 2 mm, (g) defect sparsity PNZ = 0.005, (h) defect sparsity PNZ = 0.03.

## 5. Conclusions and Outlook

The application of LBISTA leads to much more promising results than the state-of-the-art BISTA. With LBISTA, we can omit manually selected regularization parameters and obtain the results faster and with higher reconstruction qualities/smaller NMSE for a small fixed number of iterations. This paper showcases the improvements in reconstruction for different cases: (1) convolution measurements with synthetic test data (see Figure 2); (2) convolution measurements with experimental test data from photothermal measurements (see outstanding results for tied LBISTA in Figures 5 and 6d and for untied LBISTA e.g., in Figure 7d). These are based on the same training dataset. In all cases, it could be observed that untied LBISTA provides more reliable results than tied LBISTA in terms of reconstruction quality. Moreover, the parameter studies encouraged us to use as many batches, iterations, and layers as possible to achieve high reconstruction qualities. Further, a very precise model of the experiment is necessary to provide accurate reconstruction results.

Thus, the application of learned regularization algorithms, such as the proposed LBISTA, is highly recommended and attractive for industrial applications where the user does not have to choose regularization parameters manually and benefits from the remarkable speed of convergence of LBISTA. LBISTA therefore enables a less complicated evaluation of the photothermal SR data and offers the possibility of reliable in situ inspections.

As an outlook, we will study other learned block-regularization techniques based on unfolding algorithms such as FISTA or Elastic-Net. In addition, we will further study how to increase the performance of the training to train with larger datasets where we do not have to eliminate the time dimension or calculate the mean over the vertically arranged pixels.

**Author Contributions:** Conceptualization, J.C.H., L.K., S.A., P.J., G.C. and M.Z.; methodology, J.C.H., S.A.; software, J.C.H., L.K. and S.A.; validation, J.C.H., P.J. and G.C.; formal analysis, J.C.H., S.A., P.J. and G.C.; investigation, J.C.H.; resources, S.A. and M.Z.; data curation, J.C.H., S.A. and M.Z.; writing—original draft preparation, J.C.H., S.A. and L.K.; writing—review and editing, J.C.H., L.K. and S.A.; visualization, J.C.H., S.A. and L.K.; supervision, G.C., P.J. and M.Z.; project administration, P.J.; funding acquisition, G.C. and M.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** Gefördert durch die Deutsche Forschungsgemeinschaft (DFG)—400857558, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—400857558.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** We would like to thank Osman Musa from TU Berlin for discussing the studies and assisting in the implementation of LBISTA. Further, we are very grateful to work on experimental results with the specimen provided by Günther Mayr and his team from FH Oberösterreich.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Vollmer, M.; Möllmann, K.-P. *Infrared Thermal Imaging: Fundamentals, Research and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2017.
2. Ortega, J.M.; Golobic, M.; Sain, J.D.; Lenhardt, J.M.; Wu, A.S.; Fisher, S.E.; Perez, L.X.P.; Jaycox, A.W.; Smay, J.E.; Duoss, E.B.; et al. Active mixing of disparate inks for multimaterial 3D printing. *Adv. Mater. Technol.* **2019**, *4*, 1800717. [[CrossRef](#)]
3. Ciampa, F.; Mahmoodi, P.; Pinto, F.; Meo, M. Recent advances in active infrared thermography for non-destructive testing of aerospace components. *Sensors* **2018**, *18*, 609. [[CrossRef](#)] [[PubMed](#)]
4. Strakowska, M.; Strakowski, R.; Strzelecki, M.; De Mey, G.; Więcek, B. Thermal modelling and screening method for skin pathologies using active thermography. *Biocybern. Biomed. Eng.* **2018**, *38*, 602–610. [[CrossRef](#)]
5. Gonzalez-Hernandez, J.L.; Recinella, A.N.; Kandlikar, S.G.; Dabydeen, D.; Medeiros, L.; Phatak, P. Technology, application and potential of dynamic breast thermography for the detection of breast cancer. *Int. J. Heat Mass Transf.* **2019**, *131*, 558–573. [[CrossRef](#)]
6. Wiecek, B. Review on thermal image processing for passive and active thermography. In Proceedings of the IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, 17–18 January 2006.



7. Samadi, N.; Thapa, D.; Salimi, M.; Parkhimchyk, A.; Tabatabaei, N. Low-cost active thermography using cellphone infrared cameras: From early detection of dental caries to quantification of THC in oral fluid. *Sci. Rep.* **2020**, *10*, 7857. [[CrossRef](#)]
8. Almond, D.P.; Patel, P.M. *Photothermal Science and Techniques*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 1996; Volume 10.
9. Hung, Y.Y.; Chen, Y.S.; Ng, S.P.; Liu, L.; Huang, Y.H.; Luk, B.L.; Ip, R.W.L.; Wu, C.M.L.; Chung, P.S. Review and comparison of shearography and active thermography for nondestructive evaluation. *Mater. Sci. Eng. R Rep.* **2009**, *64*, 73–112. [[CrossRef](#)]
10. Cole, K.; Beck, J.; Haji-Sheikh, A.; Litkouhi, B. *Heat Conduction Using Greens Functions*; Taylor & Francis: Abingdon, UK, 2010.
11. Bryan, K.; Caudill, L.F., Jr. An inverse problem in thermal imaging. *SIAM J. Appl. Math.* **1996**, *56*, 715–735. [[CrossRef](#)]
12. Wang, J.; Zabarvas, N. Hierarchical Bayesian models for inverse problems in heat conduction. *Inverse Probl.* **2004**, *21*, 183. [[CrossRef](#)]
13. Kukreja, S.L.; Löfberg, J.; Brenner, M.J. A least absolute shrinkage and selection operator (LASSO) for nonlinear system identification. *IFAC Proc.* **2006**, *39*, 814–819. [[CrossRef](#)]
14. Donoho, D.L. Compressed Sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306. [[CrossRef](#)]
15. Baraniuk, R.G. Compressive Sensing. *IEEE Signal Process. Mag.* **2007**, *24*, 187–229. [[CrossRef](#)]
16. Beck, A.; Teboulle, M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2009**, *2*, 183–202. [[CrossRef](#)]
17. Gregor, K.; LeCun, Y. Learning Fast Approximations of Sparse Coding. In Proceedings of the 27th International Conference on Machine Learning, Haifa, Israel, 21–24 June 2010; pp. 399–406.
18. Liu, J.; Chen, X.; Wang, Z.; Yin, W. ALISTA: Analytic Weights Are As Good As Learned Weights in LISTA. In Proceedings of the International Conference on Learning Representations, New Orleans, LA, USA, 6–9 May 2019.
19. Kim, D.; Park, D. Element-Wise Adaptive Thresholds for Learned Iterative Shrinkage Thresholding Algorithms. *IEEE Access* **2020**, *8*, 45874–45886. [[CrossRef](#)]
20. Chun, I.Y.; Huang, Z.; Lim, H.; Fessler, J.A. Momentum-Net: Fast and convergent iterative neural network for inverse problems. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *1*. [[CrossRef](#)]
21. Lv, X.; Bi, G.; Wan, C. The Group Lasso for Stable Recovery of Block-Sparse Signal Representations. *IEEE Trans. Signal Process.* **2011**, *59*, 1371–1382. [[CrossRef](#)]
22. Eldar, Y.C.; Bolcskei, H. Block-sparsity: Coherence and efficient recovery. In Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, 19–24 April 2009; pp. 2885–2888.
23. Elhamifar, E.; Vidal, R. Block-sparse recovery via convex optimization. *IEEE Trans. Signal Process.* **2012**, *60*, 4094–4107. [[CrossRef](#)]
24. Lin, J.H.; Li, S. Block sparse recovery via mixed  $l_{2/1}$  minimization. *Acta Math. Sin. Engl. Ser.* **2013**, *29*, 1401–1412. [[CrossRef](#)]
25. Boufounos, P.; Kutyniok, G.; Rauhut, H. Sparse recovery from combined fusion frame measurements. *IEEE Trans. Inf. Theory* **2011**, *57*, 3864–3876. [[CrossRef](#)]
26. Ahmadi, S.; Thummerer, G.; Breitwieser, S.; Mayr, G.; Lecomagnon, J.; Burgholzer, P.; Jung, P.; Caire, G.; Ziegler, M.; Multidimensional Reconstruction of Internal Defects in Additively Manufactured Steel Using Photothermal Super Resolution Combined With Virtual Wave-Based Image Processing. *IEEE Trans. Ind. Inform.* **2021**, *17*, 7368–7378. [[CrossRef](#)]
27. Murray, T.W.; Haltmeier, M.; Berer, T.; Leiss-Holzinger, E.; Burgholzer, P. Super-resolution photoacoustic microscopy using blind structured illumination. *Optica* **2017**, *4*, 17–22. [[CrossRef](#)]
28. Burgholzer, P.; Berer, T.; Gruber, J.; Mayr, G. Super-resolution thermographic imaging using blind structured illumination. *Appl. Phys. Lett.* **2017**, *111*, 031908. [[CrossRef](#)]
29. Burgholzer, P.; Berer, T.; Ziegler, M.; Thiel, E.; Ahmadi, S.; Gruber, J.; Mayr, G.; Hendorfer, G. Blind structured illumination as excitation for super-resolution photothermal radiometry. *Quant. Infrared Thermogr. J.* **2020**, *7*, 268–278. [[CrossRef](#)]
30. Ahmadi, S.; Burgholzer, P.; Mayr, G.; Jung, P.; Caire, G.; Ziegler, M. Photothermal super resolution imaging: A comparison of different thermographic reconstruction techniques. *NDT E Int. J.* **2020**, *111*, 102228. [[CrossRef](#)]
31. Ahmadi, S.; Burgholzer, P.; Jung, P.; Caire, G.; Ziegler, M. Super resolution laser line scanning thermography. *Opt. Lasers Eng.* **2020**, *134*, 106279. [[CrossRef](#)]
32. Ahmadi, S.; Lecomagnon, J.; Hirsch, P.D.; Burgholzer, P.; Jung, P.; Caire, G.; Ziegler, M. Laser excited super resolution thermal imaging for nondestructive inspection of internal defects. *Sci. Rep.* **2020**, *10*, 22357. [[CrossRef](#)]
33. Seppala, J.E.; Migler, K.D. Infrared thermography of welding zones produced by polymer extrusion additive manufacturing. *Addit. Manuf.* **2016**, *12*, 71–76. [[CrossRef](#)]
34. Mohr, G.; Altenburg, S.J.; Ulbricht, A.; Heinrich, P.; Baum, D.; Maierhofer, C.; Hilgenberg, K. In-situ defect detection in laser powder bed fusion by using thermography and optical tomography—Comparison to computed tomography. *Metals* **2020**, *10*, 103. [[CrossRef](#)]
35. Altenburg, S.J.; Straße, A.; Gumenyuk, A.; Maierhofer, C. In-situ monitoring of a laser metal deposition (LMD) process: Comparison of MWIR, SWIR and high-speed NIR thermography. *Quant. Infrared Thermogr. J.* **2020**, *19*, 97–114. [[CrossRef](#)]
36. Metz, C.; Franz, P.; Fischer, C.; Wachtendorfer, V. Active thermography for quality assurance of 3D-printed polymer structures. *Quant. Infrared Thermogr. J.* **2021**, *18*, 50–71. [[CrossRef](#)]
37. Wang, J.; Zhang, H.; Wang, J.; Pu, Y.; Pal, N.R. Feature Selection Using a Neural Network With Group Lasso Regularization and Controlled Redundancy. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 1110–1123. [[CrossRef](#)]
38. Wen, W.; Wu, C.; Wang, Y.; Chen, Y.; Li, H. Learning Structured Sparsity in Deep Neural Networks. In Proceedings of the Advances in Neural Information Processing Systems 29 (NIPS 2016), Barcelona, Spain, 5–10 December 2016.



39. Huang, G.; Liu, S.; Van der Maaten, L.; Weinberger, K.Q. Condensenet: An efficient densenet using learned group convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
40. Ahmadi, S.; Kästner, L.; Hauffen, J.C.; Jung, P.; Ziegler, M. Photothermal-SR-Net: A Customized Deep Unfolding Neural Network for Photothermal Super Resolution Imaging. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 4502509. [[CrossRef](#)]
41. Haltmeier, M. Block-sparse analysis regularization of ill-posed problems via L<sub>2,1</sub>-minimization. In Proceedings of the 18th International Conference on Methods & Models in Automation & Robotics (MMAR), Miedzyzdroje, Poland, 26–29 August 2013.
42. Borgerding, M.; Schniter, P.; Rangan, S. AMP-Inspired Deep Networks for Sparse Linear Inverse Problems. *IEEE Trans. Signal Process.* **2017**, *65*, 4293–4308. [[CrossRef](#)]
43. Chen, X.; Liu, J.; Wang, Z.; Yin, W. Theoretical linear convergence of unfolded ISTA and its practical weights and thresholds. *arXiv* **2018**, arXiv:1808.10038.