

PAPER • OPEN ACCESS

## Inverse problems with inexact forward operator: iterative regularization and application in dynamic imaging

To cite this article: Stephanie E Blanke *et al* 2020 *Inverse Problems* **36** 124001

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

You may also like

- [Conditional stability versus ill-posedness for operator equations with monotone operators in Hilbert space](#)  
Radu Ioan Bo and Bernd Hofmann
- [On the interplay of basis smoothness and specific range conditions occurring in sparsity regularization](#)  
Stephan W Anzengruber, Bernd Hofmann and Ronny Ramlau
- [Solving ill-posed inverse problems using iterative deep neural networks](#)  
Jonas Adler and Ozan Öktem



**IOP | ebooks™**

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

# Inverse problems with inexact forward operator: iterative regularization and application in dynamic imaging

Stephanie E Blanke<sup>1</sup>, Bernadette N Hahn<sup>2,\*</sup>  and Anne Wald<sup>3</sup> 

<sup>1</sup> Department of Mathematics, Universität Hamburg, Bundesstr. 55, D-20146 Hamburg, Germany

<sup>2</sup> Department of Mathematics, University of Stuttgart, D-70569 Stuttgart, Germany

<sup>3</sup> Department of Mathematics, Saarland University, PO Box 15 11 50, D-66041 Saarbrücken, Germany

E-mail: [stephanie.blanke@uni-hamburg.de](mailto:stephanie.blanke@uni-hamburg.de), [bernadette.hahn@mathematik.uni-stuttgart.de](mailto:bernadette.hahn@mathematik.uni-stuttgart.de) and [anne.wald@num.uni-sb.de](mailto:anne.wald@num.uni-sb.de)

Received 16 April 2020, revised 29 August 2020

Accepted for publication 7 September 2020

Published 28 October 2020



CrossMark

## Abstract

The classic regularization theory for solving inverse problems is built on the assumption that the forward operator perfectly represents the underlying physical model of the data acquisition. However, in many applications, for instance in microscopy or magnetic particle imaging, this is not the case. Another important example represent dynamic inverse problems, where changes of the searched-for quantity during data collection can be interpreted as model uncertainties. In this article, we propose a regularization strategy for linear inverse problems with inexact forward operator based on sequential subspace optimization methods (SESOP). In order to account for local modelling errors, we suggest to combine SESOP with the Kaczmarz' method. We study convergence and regularization properties of the proposed method and discuss several practical realizations. Relevance and performance of our approach are evaluated at simulated data from dynamic computerized tomography with various dynamic scenarios.

Keywords: model inexactness, dynamic computerized tomography, sequential subspace optimization

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

\* Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

## 1. Introduction

An inverse problem is concerned with extracting a searched-for quantity  $f \in X$  from measured data  $g \in Y$  where  $g = \mathcal{A}f$  with forward operator  $\mathcal{A} : X \rightarrow Y$  modelling the data acquisition process. A characteristic property of an inverse problem is its ill-posedness: in general, a small noise in the measurement can cause a large error in the solution, and therefore, suitable regularization schemes are required.

Due to simplifications in the underlying physical model of the data acquisition process or due to calibration errors, the forward operator is usually not exactly known either. A prominent example arises for instance in the context of microscopic imaging, where the forward operator depends on the point spread function of the microscope which has to be determined experimentally. In magnetic particle imaging, the model-based determination of the system function is an unsolved problem which necessitates an experimental calibration step. Our main motivation to study inverse problems with inexact forward operator is to solve *dynamic inverse problems*, where changes of the searched-for quantity during data collection can be interpreted as model uncertainties.

### 1.1. Motivating example: dynamic inverse problems

In many practical applications, collecting the measurement  $g$  takes a considerable amount of time. In computerized tomography (CT), for instance, it takes time to rotate the radiation source around the investigated object. Within the standard formulation  $\mathcal{A}f = g$ , this can be expressed by specifying the data variable, i.e., the equation more precisely reads

$$\mathcal{A}f(t, y) = g(t, y) \quad \text{for all } (t, y) \in [0, T] \times \Omega_Y \quad (1)$$

with time interval  $[0, T] \subset \mathbb{R}$  covering the time period required for collecting the data.  $\Omega_Y$  could be considered, e.g., as bounded subset of  $\mathbb{R}^m$ .

To simplify the modelling step,  $\mathcal{A}$  is typically derived under the assumption that the searched-for quantity  $f$  is stationary, i.e., independent of time. However, in many applications, the investigated quantity changes during the collection of the data, for instance in medical imaging due to patient or organ motion. In non-destructive testing, the dynamic scenario arises for example when visualizing driven liquid fronts [1] or performing elasticity experiments during the scan to determine material parameters [17]. The respective dynamic data then are inconsistent with the standard model  $\mathcal{A}$ . Thus, simply applying a standard inversion procedure for  $\mathcal{A}$  to such a dynamic data set causes motion artifacts in the reconstruction result which degrade the reconstruction quality. As a consequence, specific regularization methods taking into account the dynamic behaviour of the specimen have to be developed for such *dynamic inverse problems*.

Typically, the individual states of the specimen show a strong temporal correlation. They can be correlated to one reference configuration  $f$  via a deformation model  $\Gamma : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ . In this case, the exact forward model for the dynamic case is given by the operator

$$\mathcal{A}_\Gamma f(t, y) := \mathcal{A}^{\text{stat}}(f \circ \Gamma)(t, y),$$

where  $\mathcal{A}^{\text{stat}}$  refers to the forward model of the static scenario. Thus, in the dynamic case, we have to solve the equation

$$\mathcal{A}_\Gamma f(t, y) = g(t, y) \quad \text{for all } (t, y) \in [0, T] \times \Omega_Y,$$

see, e.g., [12, 18]. In practice, the exact deformation  $\Gamma$ , and therefore the exact dynamic forward operator  $\mathcal{A}_\Gamma$ , will in general be unknown. This either requires a delicate calibration step (so-called motion estimation) prior to the reconstruction [11, 19, 22] or the simultaneous recovery of  $\Gamma$  and  $f$  which is typically computationally demanding [2, 3]. For specific applications and special deformations that preserve the underlying data acquisition geometry, exact analytic reconstruction methods have been derived, especially in CT, where this type of motion includes affine deformations, [6, 7, 10]. In this case, techniques for rebinning the measured data to make them feasible for standard reconstruction methods are proposed as well, [4, 34].

To avoid the challenging task of an explicit motion estimation, several approaches incorporating less strong prior information on the solution have been studied as well. They typically invoke some spatio-temporal regularization term, such as, e.g., temporal smoothness [28, 29]. In [5], a computational method in a Bayesian framework has been proposed, where spatio-temporal covariance matrices are considered. All-at-once and reduced iterative methods to solve time-dependent parameter identification problems are described and analyzed in [16, 25]. Approaches for solving non-linear dynamic tomography problems have been studied in the context of electrical impedance tomography, for instance using an oppositional biogeography-based optimization technique to reconstruct organ boundaries [26].

In this article, we propose a different strategy which requires neither explicitly known deformation fields nor explicit spatio-temporal priors on the solution: instead of the exact but unknown operator  $\mathcal{A}_\Gamma$ , utilize a simplified model, such as the inexact operator  $\mathcal{A}^{\text{stat}}$ , and apply a regularization strategy that accounts for the respective model inexactness.

### 1.2. Inverse problems with inexact forward operator

The aforementioned examples illustrate that uncertainties in the forward model are highly relevant in practice. This inexactness needs to be taken into account within the reconstruction step in order to stably determine  $f$  from noisy data  $g^\delta$  with

$$\|g^\delta - \mathcal{A}f\| \leq \delta,$$

while the reconstruction procedure only uses an available approximate model  $\mathcal{A}^\eta$  with

$$\|\mathcal{A} - \mathcal{A}^\eta\| \leq \eta \quad \in [0, \infty). \quad (2)$$

We refer to the model error (2) in the following as the *global model inexactness*.

In the literature, a few methods have been proposed to deal with an imperfection of the forward model. A well-known technique to compensate modelling errors in the reconstruction process is total least squares regularization [8]. An optimization approach based on partially-ordered spaces was proposed and evaluated for the deblurring problem in [21]. In [20], a reconstruction method for magnetic particle imaging was presented which computes simultaneously the searched-for quantity and deviations of the simplified discrete forward model. Uncertainties in the forward operator are often considered and dealt with in the context of statistical inverse problems, see, e.g., [33].

In this article, we propose a regularization strategy for linear inverse problems with inexact forward operator based on sequential subspace optimization methods. Sequential subspace optimization (or short SESOP) belongs to the class of iterative reconstruction techniques [9, 15, 23, 31, 32, 35, 36]. The underlying idea is to sequentially project from some initial value onto suitable subspaces that contain the searched-for solution. This method has been successfully applied, for example to solve an integral equation of the first kind [30] or in parameter identification [32, 35], particularly in terahertz tomography [37]. The goal in our case is

to choose these subspaces in accordance to the model inexactness such that the sequence of iterates generated by the imperfect model  $\mathcal{A}^\eta$  converges to a solution of  $\mathcal{A}f = g$ .

The above framework can then be used to solve dynamic inverse problems, where the dynamic operator  $\mathcal{A}_\Gamma$  reprises the role of the unknown exact model  $\mathcal{A}$ , i.e., we set  $\mathcal{A}_\Gamma \rightsquigarrow \mathcal{A}$ . As simplified operator, we could choose for instance the static forward operator, i.e.,  $\mathcal{A}^{\text{stat}} \rightsquigarrow \mathcal{A}^\eta$ , provided we have at least an estimate  $\eta$  on the overall model error.

### 1.3. Local inexactness and semi-discrete framework

In most cases, a global error such as (2) will quantify the model deviation too pessimistically. As an example, consider the case of periodic motion, such as respiratory or cardiac motion. Due to the periodicity, it might hold  $\mathcal{A}f(t, \cdot) = \mathcal{A}^\eta f(t, \cdot)$  for certain time instances, i.e., locally, the model error is small. In addition, local deformations, e.g. cardiac motion, affect only a small region of the body and therefore, it can hold  $\mathcal{A}f(t, y) = \mathcal{A}^\eta f(t, y)$  for some  $y \in \Omega_Y$  as well. Thus, we want to consider a framework that takes local model errors depending on the data variable into account.

For this purpose, we focus on a semi-discrete framework

$$A_{k,l}f = g_{k,l}, \quad k \in \{0, \dots, K-1\}, l \in \{0, \dots, L-1\}$$

with a finite sequence of linear operators  $A_{k,l}$  in this article. On the one hand, this framework accounts for the discrete nature of measured data in practice, see also, e.g., [27]. On the other hand, we can then consider a sequence of imperfect forward models  $A_{k,l}^\eta$  with

$$\|A_{k,l} - A_{k,l}^\eta\| \leq \eta_{k,l}, \quad k \in \{0, \dots, K-1\}, l \in \{0, \dots, L-1\}. \quad (3)$$

This allows the level of inexactness, characterized by  $\eta_{k,l}$ , to vary for each data point. Thus, such local error estimates will typically be less pessimistic than a global estimate (2) and further allow to incorporate prior knowledge on the motion (e.g. on periodicity). The indices  $k$  and  $l$  may refer, for example, to time instances and detector points in dynamic imaging.

### 1.4. Outline of the article

The article is organized as follows. Section 2 presents the basic principle of SESOP methods in the continuous setting with exact model operator and noise-free data. We then provide the mathematical framework of our semi-discrete setting. In particular, we suggest to combine SESOP with Kaczmarz' method and discuss the resulting advantages regarding practical applications. Section 2 concludes with the formulation of our regularized version of the SESOP-Kaczmarz method for inexact forward operators. Its convergence and regularization properties are then studied in section 3. In Section 4, we discuss possible variations of our proposed algorithms. Section 5 is devoted to the application in dynamic CT. In particular, our proposed algorithms are evaluated at different simulated CT data for various dynamic scenarios.

## 2. A SESOP-Kaczmarz method for linear inverse problems with inexact forward operator

This section is dedicated to introduce the basic tools that are needed to formulate the SESOP algorithms and their combinations with Kaczmarz' method. Detailed descriptions of SESOP techniques with an exact forward operator can be found in the literature, e.g., [30–32, 35, 36].

### 2.1. Basic principles of sequential subspace optimization

In the following, we provide a brief introduction to SESOP methods in the continuous setting with exact model operator and noise-free data.

Let  $(X, \|\cdot\|_X)$  and  $(Y, \|\cdot\|_Y)$  denote Hilbert spaces equipped with their respective norms. By  $\langle \cdot, \cdot \rangle_X$  we refer to the inner product on  $X$ . If there is no ambiguity, we will omit the indices on the norms and inner products to simplify the notation. Further, let  $\mathcal{A} : X \rightarrow Y$  be a bounded linear operator with

$$\|\mathcal{A}\| \leq \Lambda_{\mathcal{A}} \quad (4)$$

for some  $\Lambda_{\mathcal{A}} > 0$ . By  $\mathcal{A}^*$  we denote the adjoint of  $\mathcal{A}$ . The space

$$\mathcal{R}(\mathcal{A}) := \{\mathcal{A}x : x \in X\} \subset Y$$

is called the *range* of  $\mathcal{A}$  and the space

$$\mathcal{N}(\mathcal{A}) := \{x \in X : \mathcal{A}x = 0\}$$

is called the *null space* of  $\mathcal{A}$ . For given  $g \in Y$ , the set

$$\mathcal{M}_{\mathcal{A},g} := \{f \in X : \mathcal{A}f = g\},$$

denotes the *solution set* of the operator equation  $\mathcal{A}f = g$ .

The underlying idea of SESOP is to create a sequence of iterates  $f_n \in X$  that converges to a solution  $f \in \mathcal{M}_{\mathcal{A},g}$  by sequentially projecting a starting value  $f_0 \in X$  onto suitable convex subsets of  $X$ . The following definition introduces specific types of subsets of  $X$  which play an important role in this context.

**Definition 2.1.** Let  $u \in X \setminus \{0\}$  and  $\alpha, \xi \in \mathbb{R}$  with  $\xi \geq 0$ . We call

$$H(u, \alpha) := \{f \in X : \langle u, f \rangle = \alpha\}$$

a *hyperplane* in  $X$ . The set

$$H_{\leq}(u, \alpha) := \{f \in X : \langle u, f \rangle \leq \alpha\}$$

below  $H(u, \alpha)$  is called a *half-space*. Analogously we define the half spaces  $H_{\geq}(u, \alpha)$ ,  $H_{<}(u, \alpha)$ ,  $H_{>}(u, \alpha)$ . Finally, we define the *stripe*

$$H(u, \alpha, \xi) := \{f \in X : |\langle u, f \rangle - \alpha| \leq \xi\}$$

with *upper bounding hyperplane*  $H(u, \alpha + \xi)$  and *lower bounding hyperplane*  $H(u, \alpha - \xi)$ .

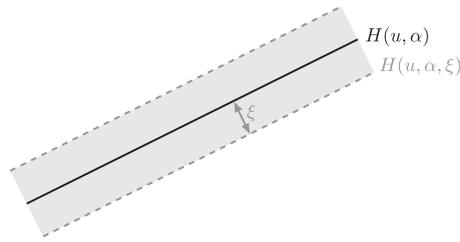
An illustration of the concept of stripes is given in figure 1.

For arbitrary  $w \in Y$ , the solution set  $\mathcal{M}_{\mathcal{A},g}$  is contained in the hyperplane

$$H(\mathcal{A}^*w, \langle w, g \rangle) = \{f \in X : \langle \mathcal{A}^*w, f \rangle = \langle w, g \rangle\}.$$

Indeed, we see that for each  $z \in \mathcal{M}_{\mathcal{A},g}$  we have

$$\langle \mathcal{A}^*w, z \rangle = \langle w, \mathcal{A}z \rangle = \langle w, g \rangle.$$



**Figure 1.** Illustration of the concept of stripes for the case  $\|u\| = 1$ .

Next, we introduce the concept of metric projections, which play a crucial role for SESOP.

**Definition 2.2.** Let  $C \neq \emptyset$  be a convex subset of  $X$ . Then the *metric projection*  $P_C(x)$  of  $x \in X$  onto  $C$  is defined by

$$P_C(x) := \operatorname{argmin}_{\tilde{x} \in C} \|x - \tilde{x}\|.$$

In Hilbert spaces, the metric projection  $P_{H(u,\alpha)}(x)$  of  $x \in X$  onto a hyperplane  $H(u,\alpha)$  coincides with the orthogonal projection and can be expressed as

$$P_{H(u,\alpha)}(x) = x - \frac{\langle u, x \rangle - \alpha}{\|u\|^2} u. \quad (5)$$

The metric projection onto a halfspace or a stripe corresponds to the metric projection onto the respective bounding hyperplane.

We now can formulate the  $n$ th iteration step ( $n > 0$ ) of the SESOP method for solving  $\mathcal{A}f = g$ , see, e.g., [32]: given a chosen finite index set  $I_n$  and *search directions*  $\mathcal{A}^*w_{n,i}$  with chosen  $w_{n,i} \in Y$ ,  $i \in I_n$ , compute the iterate  $f_{n+1}$  as the metric projection

$$f_{n+1} := P_{H_n}(f_n) \quad (6)$$

of the previous iterate  $f_n$  onto the intersection of hyperplanes

$$H_n := \bigcap_{i \in I_n} H(\mathcal{A}^*w_{n,i}, \langle w_{n,i}, g \rangle).$$

In [31] it has been shown that the metric projection  $P_H(f)$  of  $f \in X$  onto the (non-empty) intersection

$$H := \bigcap_{i \in I} H(u_i, \alpha_i)$$

of a finite set of hyperplanes  $H(u_i, \alpha_i)$ ,  $i \in I$ ,  $|I| < \infty$ , is given by

$$P_H(f) = f - \sum_{i \in I} \tilde{t}_i u_i, \quad (7)$$

where  $\tilde{t} := (\tilde{t}_i)_{i \in I} \in \mathbb{R}^{|I|}$  minimizes the function

$$h(\tilde{t}) := \frac{1}{2} \left\| f - \sum_{i \in I} \tilde{t}_i u_i \right\|^2 + \sum_{i \in I} \tilde{t}_i \alpha_i. \quad (8)$$

If  $\{u_i : i \in I\}$  is linearly independent, the function  $h$  is strictly convex and thus,  $\tilde{f}$  is unique, see also [31].

Thus, the iterates (6) of the SESOP method are given by

$$f_{n+1} = f_n - \sum_{i \in I_n} \tilde{t}_{n,i} \mathcal{A}^* w_{n,i}.$$

Under suitable conditions on the index set  $I_n$ , the search directions  $\mathcal{A}^* w_{n,i}$  and the parameters  $\tilde{t}_{n,i}$ , the sequence  $\{f_n\}_{n \in \mathbb{N}}$  converges to  $f \in \mathcal{M}_{\mathcal{A},g}$ , see [32, 36].

We conclude this subsection by outlining how SESOP can be adapted to a regularization method in case only noisy data  $g^\delta$  with noise level  $\delta$  (i.e.  $\|g - g^\delta\| \leq \delta$ ) and an inexact version  $\mathcal{A}^\eta$  of the forward operator  $\mathcal{A}$  with inexactness  $\eta$  (i.e.  $\|\mathcal{A} - \mathcal{A}^\eta\| \leq \eta$ ) are given. The basic idea consists in replacing the hyperplanes within the projection step by stripes whose width is chosen in accordance to the noise and uncertainty level.

This idea is based on the property that

$$\mathcal{M}_{\mathcal{A},g}^\rho := \mathcal{M}_{\mathcal{A},g} \cap B_\rho(0),$$

where  $B_\rho(0)$  is the open ball with radius  $\rho > 0$  around 0, is contained in the stripe

$$\begin{aligned} H^{\delta,\eta,\rho} &:= H((\mathcal{A}^\eta)^* w, \langle w, g^\delta \rangle, (\delta + \eta\rho)\|w\|) \\ &= \{f \in X : \|\langle (\mathcal{A}^\eta)^* w, f \rangle - \langle w, g^\delta \rangle\| \leq (\delta + \eta\rho)\|w\|\}, \end{aligned} \quad (9)$$

since for all  $z \in \mathcal{M}_{\mathcal{A},g}$  with  $\|z\| \leq \rho$  we have

$$\begin{aligned} \|\langle (\mathcal{A}^\eta)^* w, z \rangle - \langle w, g^\delta \rangle\| &= \|\langle w, (\mathcal{A}^\eta - \mathcal{A})z \rangle + \langle w, \mathcal{A}z \rangle - \langle w, g^\delta \rangle\| \\ &= \|\langle w, (\mathcal{A}^\eta - \mathcal{A})z + g - g^\delta \rangle\| \\ &\leq \|w\| \cdot (\|\mathcal{A}^\eta - \mathcal{A}\|\|z\| + \|g - g^\delta\|) \\ &\leq (\delta + \eta\rho)\|w\|. \end{aligned} \quad (10)$$

This approach has been studied so far for the case of noisy data and exact forward operator (i.e.  $\eta = 0$ ). In particular, combining the respective projection algorithm with the discrepancy principle yields a finite stopping index  $n^*$  and the respective iterate  $f_{n^*}$  constitutes a regularized solution of  $\mathcal{A}f = g$ , see again [32, 36].

One goal of this article is to extend the respective convergence and regularization results to the setting with inexact forward operator. In order to allow incorporating local estimates for the modelling errors, we focus in the following on the more general semi-discrete framework motivated in section 1.

## 2.2. A SESOP-Kaczmarz method for semi-discrete inverse problems

Let  $(Y_{k,l}, \|\cdot\|_{Y_{k,l}})$ , where  $k \in \mathcal{K} := \{0, \dots, K-1\}$ ,  $l \in \mathcal{L} := \{0, \dots, L-1\}$ , with integers  $K, L \geq 1$ , be Hilbert spaces equipped with their respective norms. Further, let  $A_{k,l} : X \rightarrow Y_{k,l}$ ,  $k \in \mathcal{K}$ ,  $l \in \mathcal{L}$ , be bounded linear operators with

$$\|A_{k,l}\| \leq \Lambda_A \quad (11)$$

for some  $\Lambda_A > 0$ . Their Hilbert space adjoints  $A_{k,l}^*$  as well as their range and null space are defined in analogy to the ones in the continuous setting. The solution set of the semi-discrete

inverse problem

$$A_{k,l}f = g_{k,l}, \quad k \in \mathcal{K}, \quad l \in \mathcal{L} \quad (12)$$

is given by

$$\begin{aligned} \mathcal{M}_{A,g}^{\text{sd}} &:= \{f \in X : A_{k,l}f = g_{k,l} \text{ for all } l \in \mathcal{L}, k \in \mathcal{K}\} \\ &= f^+ + \prod_{k \in \mathcal{K}} \prod_{l \in \mathcal{L}} \mathcal{N}(A_{k,l}) \end{aligned} \quad (13)$$

for some solution  $f^+$  of (12).

**Remark 2.3.** Depending on the application, it could be of interest to consider a more general setting with index sets  $\mathcal{K} := \{0, \dots, K-1\}$  and  $\mathcal{L}_k := \{0, \dots, L_k-1\}$ ,  $K, L_k \geq 1$ . In the context of dynamic inverse problems, this might be relevant if the discretization with respect to the variable  $y$  depends on time, i.e., on the index  $k$ . The results presented in this article can be directly transferred to this more general setting. However, to keep the notation as simple as possible, we consider in the following the setting where  $\mathcal{L}$  is independent of  $k$ .

In order to solve the semi-discrete inverse problem (12), we propose to combine the SESOP method with Kaczmarz' method [24].

**Remark 2.4.** A combination of Kaczmarz' method with SESOP allows to incorporate local properties of the inverse problem: on the one hand, SESOP is designed in such a way that the local properties of the inverse problem, i.e., particularly local modelling errors, can be easily included in the regularization. On the other hand, Kaczmarz' method allows to combine the respective subproblems.

We also remark that, in view of a comparison to the Landweber method, the additional computational effort due to solving the required optimization problem in the SESOP iteration is relatively low compared to the reduction in computation time due to this regulation of the step width and using multiple search directions, see also, e.g., [30, 35, 37] for a comparison.

To illustrate the basic idea, we first consider a system of linear operator equations

$$A_k f = g_k, \quad k \in \mathcal{K},$$

with only one index set. The standard Kaczmarz' method reads in this case

$$f_{n+1} = P_{\mathcal{M}_{A_{[n]}, g_{[n]}}} (f_n),$$

where  $[n] := n \bmod K$  and  $f_0 \in X$  is an initial value. Here we use the notation from [13, 14], where the Landweber–Kaczmarz method for nonlinear inverse problems is investigated.

The projection step can now be replaced by a metric projection onto the intersection  $H_n := \bigcap_{i \in I_n} H_{n,i}$  of the hyperplanes  $H_{n,i} := H(A_{[i]}^* w_{n,i}, \langle w_{n,i}, g_{[i]} \rangle)$ . The corresponding SESOP–Kaczmarz method then reads

$$f_{n+1} = f_n - \sum_{i \in I_n} t_{n,i} A_{[i]}^* w_{n,i}. \quad (14)$$

Here,  $t_n := (t_{n,i})_{i \in I_n}$  is calculated as a minimizer of the respective function  $h_n$  from (8) with  $u_{n,i} := A_{[i]}^* w_{n,i}$  and  $\alpha_{n,i} := \langle w_{n,i}, g_{[i]} \rangle$ . Each parameter  $t_{n,i}$  can be interpreted as a regulation of the step width in the direction of the search direction  $u_{n,i}$ .

Regarding our semi-discrete setting, the two discretization indices  $k$  and  $l$  allow for multiple versions of the basic SESOP-Kaczmarz iteration (14). For instance, the Kaczmarz iterations can be executed with respect to both  $k$  and  $l$ . Depending on the nature of the underlying inverse problem, it might be convenient to use search directions that are *averaged* with respect to one of the indices, for instance

$$u_{n,i} := \sum_{l \in \mathcal{L}} A_{[i],l}^* w_{n,i}^l \quad (15)$$

resulting from summing up all search directions with respect to the second discretization index. Respectively, the roles of  $k$  and  $l$  could be interchanged. A detailed overview of these different versions can be found in section 4. Please note that the continuous case, i.e., for the operator equation  $\mathcal{A}f = g$ , is included in the above framework since it corresponds to a setting with  $\mathcal{K} = \mathcal{L} = \{0\}$ .

Throughout sections 2 and 3, we focus on the SESOP-Kaczmarz algorithm with search-directions (15) for the semi-discrete case, since all the above mentioned versions can be treated in this framework, see the discussion in section 4.

In order to be able to provide a solution of  $A_{k,l}f = g_{k,l}$ , the hyperplanes associated to the search directions (15) have to include the solution set  $\mathcal{M}_{A,g}^{sd}$  (13). According to the following lemma, this is indeed the case.

**Lemma 2.5.** *For each iteration index  $n$ , the hyperplanes*

$$H_{n,i}^l := \{f \in X : \langle A_{[i],l}^* w_{n,i}^l, f \rangle = \langle w_{n,i}^l, g_{[i],l} \rangle\} \quad (16)$$

with  $l \in \mathcal{L}$  as well as the hyperplanes

$$H_{n,i} := \left\{ f \in X : \sum_{l \in \mathcal{L}} \langle A_{[i],l}^* w_{n,i}^l, f \rangle = \sum_{l \in \mathcal{L}} \langle w_{n,i}^l, g_{[i],l} \rangle \right\} \quad (17)$$

with  $w_{n,i}^l \in Y_{[i],l}$  for all  $i \in I_n$ ,  $n \in \mathbb{N}$ , and  $l \in \mathcal{L}$  contain the solution set  $\mathcal{M}_{A,g}^{sd}$ .

**Proof.** The proof is analogous to (10). □

### 2.3. Regularizing SESOP-Kaczmarz iteration for inexact forward operators

In this section, we propose a regularizing SESOP-Kaczmarz iteration that is suited for linear semi-discrete inverse problems (12) in Hilbert spaces with inexact forward operator and noisy data. Let  $A_{k,l}^\eta \in \mathcal{L}(X, Y_{k,l})$  denote an inexact version of the forward operator  $A_{k,l}$  with inexactness

$$\|A_{k,l}^\eta - A_{k,l}\| \leq \eta_{k,l} \in [0, \infty). \quad (18)$$

Further, we consider noisy data  $g_{k,l}^\delta \in Y_{k,l}$  with noise level

$$\|g_{k,l}^\delta - g_{k,l}\| \leq \delta_{k,l} \in [0, \infty). \quad (19)$$

**Remark 2.6.** Please note that we label quantities that are influenced by the inexactness  $\eta = (\eta_{k,l})_{k \in \mathcal{K}, l \in \mathcal{L}}$  or the noise level  $\delta = (\delta_{k,l})_{k \in \mathcal{K}, l \in \mathcal{L}}$  by adding the index  $\eta$  or  $\delta$ . For example, we assume that the forward operator  $A_{k,l}$  is inexact with inexactness  $\eta_{k,l}$ , yet we will only write  $A_{k,l}^\eta$  instead of  $A_{k,l}^{\eta_{k,l}}$ .

In analogy to (13), we denote the set of solutions whose norm is bounded by  $\rho > 0$  by

$$\mathcal{M}_{A,g}^{\text{sd},\rho} := \mathcal{M}_{A,g}^{\text{sd}} \cap B_\rho(0).$$

For each element  $z \in \mathcal{M}_{A,g}^{\text{sd},\rho}$ , we obtain the estimate

$$\|A_{k,l}^\eta z - g_{k,l}^\delta\| = \|(A_{k,l}^\eta - A_{k,l})z + A_{k,l}z - g_{k,l}^\delta\| \leq \eta_{k,l}\rho + \delta_{k,l}.$$

We propose to regularize the SESOP-Kaczmarz algorithm with search directions (15), see section 2.2, by replacing the hyperplanes  $H_{n,i}$  by stripes with width in accordance to  $\delta$ , respectively  $\eta$ , and by combining the algorithm with a discrete version of the discrepancy principle. The details are formulated in the following algorithm which is referred to as *RESESOP-Kaczmarz*.

**Algorithm 2.7.** Choose an initial value  $f_0 := f_0^{\eta,\delta} \in B_\rho(0) \subset X$  and constants  $\tau_{k,l} > 1$  for  $k \in \mathcal{K}$ ,  $l \in \mathcal{L}$ . For each iteration index  $n \in \mathbb{N} \cup \{0\}$  set

$$D_n^{\eta,\delta} := \{l \in \mathcal{L} : \|A_{[n],l}^\eta f_n^{\eta,\delta} - g_{[n],l}^\delta\| > \tau_{[n],l} (\eta_{[n],l}\rho + \delta_{[n],l})\}.$$

If  $D_n^{\eta,\delta} = \emptyset$ , set  $f_{n+1}^{\eta,\delta} := f_n^{\eta,\delta}$ . Otherwise, define a finite index set  $I_n^{\eta,\delta} \subset \{0, 1, \dots, n\}$  as well as search directions

$$u_{n,i}^{\eta,\delta} := \sum_{l \in \mathcal{L}} (A_{[i],l}^\eta)^* w_{n,i}^{\eta,\delta,l},$$

where  $w_{n,i}^{\eta,\delta,l} \in Y_{[i],l}$  and  $w_{n,n}^{\eta,\delta,l} = 0$  for all  $l \notin D_n^{\eta,\delta}$ .

For each index  $n = 0, 1, \dots$ , compute the new iterate

$$f_{n+1}^{\eta,\delta} := P_{H_n^{\eta,\delta}}(f_n^{\eta,\delta}) \quad (20)$$

by projecting the current iterate  $f_n^{\eta,\delta}$  onto the intersection

$$H_n^{\eta,\delta} := \bigcap_{i \in I_n^{\eta,\delta}} H(u_{n,i}^{\eta,\delta}, \alpha_{n,i}^{\eta,\delta}, \xi_{n,i}^{\eta,\delta}) \quad (21)$$

of the stripes with the parameters

$$\alpha_{n,i}^{\eta,\delta} := \sum_{l \in \mathcal{L}} \langle w_{n,i}^{\eta,\delta,l}, g_{[i],l}^\delta \rangle,$$

$$\xi_{n,i}^{\eta,\delta} := \sum_{l \in \mathcal{L}} (\eta_{[i],l}\rho + \delta_{[i],l}) \|w_{n,i}^{\eta,\delta,l}\|.$$

Stop iterating, as soon as  $f_n^{\eta,\delta} = f_{n-K}^{\eta,\delta}$  and  $n = 0 \bmod K$ .

**Remark 2.8.** Algorithm 2.7 includes an adaption of the discrepancy principle for our semi-discrete setting, which considers the discrepancies of all subproblems separately to increase the accuracy of the reconstruction.

**Definition 2.9.** The iterates  $f_n^{\eta,\delta}$ ,  $n \in \mathbb{N}$ , in algorithm 2.7 are called *auxiliary iterates*. Since algorithm 2.7 is stopped when  $n = r \cdot K$  for some integer  $r \in \mathbb{N}$  (i.e., after a full sweep through all subproblems), we define  $\bar{f}_r^{\eta,\delta} := f_{rK}^{\eta,\delta}$  and call  $\{\bar{f}_r^{\eta,\delta}\}_{r \in \mathbb{N}}$  the sequence of *full iterates*.

**Remark 2.10.** With  $\delta = 0$  and  $\eta = 0$ , our RESESOP-Kaczmarz algorithm coincides with SESOP-Kaczmarz proposed in section 2.2 for the case of exact data and exact forward operator.

### 3. Convergence and regularization

In this section, we study convergence and regularization properties of our RESESOP-Kaczmarz algorithm 2.7. This requires in particular imposing suitable conditions on the index sets  $I_n^{\eta,\delta}$  and the search directions.

Regarding the index set  $I_n^{\eta,\delta}$ , we postulate that

$$n \in I_n^{\eta,\delta} \quad \text{and} \quad I_n^{\eta,\delta} \subset \{n - N + 1, \dots, n\} \cap \mathbb{N} \quad \text{for all } n \in \mathbb{N}, \quad (22)$$

where  $N \in \mathbb{N} \setminus \{0\}$  is fixed.

Regarding the search directions, for each  $n \in \mathbb{N}$ , we define

$$u_n^{\eta,\delta} := u_{n,n}^{\eta,\delta} := \sum_{l \in \mathcal{L}} (A_{[n],l}^\eta)^* w_n^{\eta,\delta,l} \quad (23)$$

with

$$w_n^{\eta,\delta,l} := w_{n,n}^{\eta,\delta,l} := \begin{cases} A_{[n],l}^\eta f_n^{\eta,\delta} - g_{[n],l}^\delta, & \text{if } l \in D_n^{\eta,\delta}, \\ 0, & \text{if } l \notin D_n^{\eta,\delta}. \end{cases}$$

In other words, we define the search direction  $u_{n,n}^{\eta,\delta}$  as the gradient of the functional

$$\Psi_n(f) := \frac{1}{2} \sum_{l \in D_n^{\eta,\delta}} \|A_{[n],l}^\eta f - g_{[n],l}^\delta\|^2$$

evaluated in the current iterate  $f_n^{\eta,\delta}$ .

The section is now organized as follows. We first prove in section 3.1, that the sequence of iterates for the exact forward operator and exact data converges to a solution  $f^+ \in \mathcal{M}_{A,g}^{\text{sd}}$ . We then turn to the case of noisy data and inexact forward operator and show that the respective iteration scheme provides a regularized solution to the semi-discrete inverse problem  $A_{k,l}f = g_{k,l}$ ,  $(k, l) \in \mathcal{K} \times \mathcal{L}$ .

#### 3.1. Convergence

In order to prove the convergence of algorithm 2.7 for an unperturbed forward operator and exact data, we establish in a first step the validity of a descent property.

**Lemma 3.1.** *Let  $\{f_n\}_{n \in \mathbb{N}}$  be the sequence of iterates generated by algorithm 2.7 for  $\eta = \delta = 0$ . If the conditions (22) and (23) are fulfilled, the descent property*

$$\|z - f_{n+1}\|^2 \leq \|z - f_n\|^2 - C \|\bar{w}_n\|^2 \quad (24)$$

with

$$\bar{w}_n := \operatorname{argmax}_{l \in \mathcal{L}} \|w_n^l\| \quad (25)$$

holds for all  $z \in \mathcal{M}_{A,g}^{\text{sd}}$ .

**Proof.** For  $n \in \mathbb{N}$  we set  $\tilde{H}_n := H(u_n, \alpha_n)$  and

$$\tilde{f}_{n+1} := P_{\tilde{H}_n}(f_n) = f_n - \frac{\langle u_n, f_n \rangle - \alpha_n}{\|u_n\|^2} u_n = f_n - \sum_{l \in \mathcal{L}} \frac{\|w_n^l\|^2}{\|u_n\|^2} u_n.$$

Using  $f_{n+1} = P_{H_n}(f_n)$  with  $H_n := \bigcap_{i \in I_n} H(u_{n,i}, \alpha_{n,i}) \subset \tilde{H}_n$ , we obtain

$$\|f_{n+1} - f_n\| \geq \|\tilde{f}_{n+1} - f_n\|.$$

Thanks to (11), we estimate

$$\|u_n\| = \left\| \sum_{l \in \mathcal{L}} A_{[n],l}^* w_n^l \right\| \leq \sum_{l \in \mathcal{L}} \|A_{[n],l}^*\| \cdot \|w_n^l\| \leq \Lambda_A L \|\bar{w}_n\|,$$

which for all  $z \in \mathcal{M}_{A,g}^{sd} \subset H_n$  yields

$$\begin{aligned} \|z - f_{n+1}\|^2 &\leq \|z - f_n\|^2 - \|f_{n+1} - f_n\|^2 \leq \|z - f_n\|^2 - \|\tilde{f}_{n+1} - f_n\|^2 \\ &= \|z - f_n\|^2 - \left\| \sum_{l \in \mathcal{L}} \frac{\|w_n^l\|^2}{\|u_n\|^2} u_n \right\|^2 = \|z - f_n\|^2 \\ &\quad - \frac{1}{\|u_n\|^2} \left( \sum_{l \in \mathcal{L}} \|w_n^l\|^2 \right)^2 \\ &\leq \|z - f_n\|^2 - \frac{\|\bar{w}_n\|^4}{\|u_n\|^2} \leq \|z - f_n\|^2 - \frac{1}{(\Lambda_A L)^2} \|\bar{w}_n\|^2. \end{aligned}$$

The assertion follows with  $C := \frac{1}{(\Lambda_A L)^2}$ . □

As a consequence of lemma 3.1, we obtain the descent property

$$\|z - \bar{f}_{r+1}\|^2 \leq \|z - \bar{f}_r\|^2 - C \sum_{k \in \mathcal{K}} \|\bar{w}_{rK+k}\|^2$$

for the full iterates  $\bar{f}_r := f_{rK}$ ,  $r \in \mathbb{N}$ , i.e., the sequences  $\{\|z - f_n\|\}_{n \in \mathbb{N}}$  and  $\{\|z - \bar{f}_r\|\}_{r \in \mathbb{N}}$  are monotonically decreasing. In addition, since  $z \in \mathcal{M}_{A,g}^{sd}$  is fixed, we deduce from (24) that

$$f_n \in B_\rho(z) \tag{26}$$

for a constant  $\rho > \|z - f_0\| > 0$ . This yields the boundedness of the sequences  $\{f_n\}_{n \in \mathbb{N}}$  and  $\{\bar{f}_r\}_{r \in \mathbb{N}}$ .

The validity of these descent properties are the starting point for the following theorem.

**Theorem 3.2.** *The sequences  $\{f_n\}_{n \in \mathbb{N}}$  and  $\{\bar{f}_r\}_{r \in \mathbb{N}}$ , generated by algorithm 2.7 with (22) and (23), each possess a subsequence that weakly converges to an element of  $\mathcal{M}_{A,g}^{sd}$ .*

**Proof.** We consider the sequence  $\{f_n\}_{n \in \mathbb{N}}$  of auxiliary iterates. From (24) we obtain

$$\begin{aligned} C \sum_{n=0}^N \|\bar{w}_n\|^2 &\leq \sum_{n=0}^N (\|z - f_n\|^2 - \|z - f_{n+1}\|^2) \\ &= \|z - f_0\|^2 - \|z - f_{N+1}\|^2 \leq \|z - f_0\|^2 \end{aligned}$$

for an arbitrary but fixed integer  $N > 0$ . We now let  $N \rightarrow \infty$ . Since the right-hand side of the above estimate is independent of  $N$ , we deduce that the series  $\sum_{n=0}^\infty \|\bar{w}_n\|^2$  is absolutely convergent, and thus  $\{\bar{w}_n\}_{n \in \mathbb{N}}$  is a null sequence. Since  $\{f_n\}_{n \in \mathbb{N}}$  is bounded according to (26), there exists a weakly converging subsequence of  $\{f_n\}_{n \in \mathbb{N}}$  with weak limit  $\tilde{f}$ . The same holds

for each subsequence, such that we have  $f_n \rightharpoonup \tilde{f}$ . In particular, we have  $\bar{f}_r \rightharpoonup \tilde{f}$ . According to

$$\|A_{k,l}\tilde{f} - g_{k,l}\| \leq \liminf_{n \rightarrow \infty} \|A_{k,l}f_n - g_{k,l}\| \leq \liminf_{n \rightarrow \infty} \|\bar{w}_n\| = 0,$$

where we have used the continuity of  $A_{k,l}$  and the weak lower semicontinuity of the norm, we have  $\tilde{f} \in \mathcal{M}_{A,g}^{\text{sd}}$ .  $\square$

**Theorem 3.3 (Convergence of the SESOP-Kaczmarz iteration).**

Let  $\{f_n\}_{n \in \mathbb{N}}$  and  $\{\bar{f}_r\}_{r \in \mathbb{N}}$  be generated by algorithm 2.7 with (22) and

$$w_i^l := w_{n,i}^l := A_{[i],l}f_i - g_{[i],l}, \quad u_i := u_{n,i} := \sum_{l \in \mathcal{L}} A_{[i],l}^* w_i^l \quad \text{for all } i \in I_n \quad (27)$$

from an initial value  $f_0 \in X$ . If the optimization parameters  $|t_{n,i}| \leq t$  with  $i \in I_n$ ,  $n \in \mathbb{N}$  are bounded by some  $t > 0$ , then  $\{f_n\}_{n \in \mathbb{N}}$  and  $\{\bar{f}_r\}_{r \in \mathbb{N}}$  converge strongly to a solution  $f^+$ , and

$$f^+ = P_{\mathcal{M}_{A,g}^{\text{sd}}}(f_0). \quad (28)$$

The proof is given in appendix A.

### 3.2. Regularization

In the previous section we have shown that the SESOP-Kaczmarz iterates converge to a solution of the semi-discrete inverse problem (12) with exact data. The solution is the metric projection of the initial value onto the solution set  $\mathcal{M}_{A,g}^{\text{sd}}$ .

Based on these findings, we now want to proceed to noisy data, i.e., we show that the RESESOP-Kaczmarz iteration, i.e., algorithm 2.7, yields a regularized solution of a semi-discrete system (12) if only noisy data  $g_{k,l}^\delta$  and inexact forward operators  $A_{k,l}^\eta$ ,  $k \in \mathcal{K}$ ,  $l \in \mathcal{L}$ , are given. We assume that the noise level can be estimated as in (19) and the inexactness by (18).

We first remark that if the set  $D_j^{\eta,\delta}$  as defined in algorithm 2.7 is empty for all  $j = n, \dots, n + K - 1$ , then  $D_j^{\eta,\delta} = \emptyset$  and, by consequence,  $f_j^{\eta,\delta} = f_n^{\eta,\delta}$  for all  $j \geq n$ . Thus there exists an  $n = rK$  for some  $r \in \mathbb{N}$  and

$$\|A_{k,l}^\eta \bar{f}_r^{\eta,\delta} - g_{k,l}^{\eta,\delta}\| \leq \tau_{k,l} (\delta_{k,l} + \eta_{k,l} \rho)$$

for all  $k \in \mathcal{K}$ ,  $l \in \mathcal{L}$ . This motivates the following definition.

**Definition 3.4.** We set

$$D_n^{\eta,\delta,\mathcal{K}} := \left\{ j \in \{n, \dots, n + K - 1\} : D_j^{\eta,\delta} \neq \emptyset \right\}$$

and define the *auxiliary stopping index*

$$n_* := n_*(\eta, \delta) := \min \{ n \in \mathbb{N} : D_n^{\eta,\delta,\mathcal{K}} = \emptyset \}.$$

Correspondingly we define the *stopping index* by

$$r_* := r_*(\eta, \delta) := \min \left\{ r \in \mathbb{N} : D_{rK}^{\eta,\delta,\mathcal{K}} = \emptyset \right\} = \frac{n_*(\eta, \delta) - [n_*(\eta, \delta)]}{K} + 1.$$

If  $\{n \in \mathbb{N} : D_n^{\eta,\delta,\mathcal{K}} = \emptyset\} = \emptyset$ , we set  $n_* := \infty$ ,  $r_* := \infty$ .

We proceed as in the case of the SESOP-Kaczmarz algorithm with  $(\eta, \delta) = (0, 0)$  and show that the RESEOP-Kaczmarz algorithm yields a regularized solution  $f_{n_*}^{\eta, \delta}$ , i.e., we consider the iterates  $f_n^{\eta, \delta}$ ,  $n \in \mathbb{N}$ . Due to definition 3.4, the properties of  $\{f_r^{\eta, \delta}\}_{r \in \mathbb{N}}$  will be inherited from  $\{f_n^{\eta, \delta}\}_{n \in \mathbb{N}}$ .

Again we postulate that the conditions (22) and (23) are fulfilled. We can express the search direction  $u_n^{\eta, \delta} := u_{n,n}^{\eta, \delta}$  by

$$u_n^{\eta, \delta} = \sum_{l \in D_n^{\eta, \delta}} (A_{[n],l}^{\eta})^* (A_{[n],l}^{\eta} f_n^{\eta, \delta} - g_{[n],l}^{\delta}).$$

The parameters of the stripes can be reformulated as

$$\alpha_n^{\eta, \delta} := \alpha_{n,n}^{\eta, \delta} = \langle u_n^{\eta, \delta}, f_n^{\eta, \delta} \rangle - \sum_{l \in D_n^{\eta, \delta}} \|w_n^{\eta, \delta, l}\|^2$$

and

$$\xi_n^{\eta, \delta} := \xi_{n,n}^{\eta, \delta} = \sum_{l \in D_n^{\eta, \delta}} (\delta_{[n],l} + \eta_{[n],l} \rho) \cdot \|w_n^{\eta, \delta, l}\|.$$

These conditions assure that the descent property holds:

**Proposition 3.5.** *If  $D_n^{\eta, \delta} \neq \emptyset$ , we have*

$$f_n^{\eta, \delta} \in H_{>} (u_n^{\eta, \delta}, \alpha_n^{\eta, \delta} + \xi_n^{\eta, \delta}) \quad (29)$$

and

$$\|z - f_{n+1}^{\eta, \delta}\|^2 \leq \|z - f_n^{\eta, \delta}\|^2 - \left( \sum_{l \in D_n^{\eta, \delta}} \frac{\|w_n^{\eta, \delta, l}\| (\|w_n^{\eta, \delta, l}\| - (\delta_{[n],l} + \eta_{[n],l} \rho))}{\|u_n^{\eta, \delta}\|} \right)^2 \quad (30)$$

for  $z \in \mathcal{M}_{A,g}^{\text{sd}, \rho}$ .

If  $D_n^{\eta, \delta} = \emptyset$ , we simply have  $\|z - f_{n+1}^{\eta, \delta}\|^2 = \|z - f_n^{\eta, \delta}\|^2$ .

**Proof.** Let  $l \in D_n^{\eta, \delta} \neq \emptyset$ . Then

$$\|w_n^{\eta, \delta, l}\| > \tau_{[n],l} (\delta_{[n],l} + \eta_{[n],l} \rho)$$

and it follows

$$\begin{aligned} \langle u_n^{\eta, \delta}, f_n^{\eta, \delta} \rangle &= \sum_{l \in D_n^{\eta, \delta}} \|w_n^{\eta, \delta, l}\|^2 + \alpha_n^{\eta, \delta} \\ &> \sum_{l \in D_n^{\eta, \delta}} (\delta_{[n],l} + \eta_{[n],l} \rho) \|w_n^{\eta, \delta, l}\| + \alpha_n^{\eta, \delta} \\ &= \alpha_n^{\eta, \delta} + \xi_n^{\eta, \delta}, \end{aligned} \quad (31)$$

which proves assertion (29). The descent property (30) follows in the same way as in the unperturbed case, i.e., as in the proof of lemma 3.1, by using the representations of  $u_n^{\eta, \delta}$ ,  $\alpha_n^{\eta, \delta}$  and  $\xi_n^{\eta, \delta}$ .  $\square$

**Remark 3.6.** By iteratively employing (30), we obtain

$$\|z - \bar{f}_{r+1}^{\eta,\delta}\|^2 \leq \|z - \bar{f}_r^{\eta,\delta}\|^2 - \sum_{j \in D_{rK}^{\eta,\delta,\mathcal{K}}} \left( \sum_{l \in D_j^{\eta,\delta}} \frac{\|w_j^{\eta,\delta,l}\| \left( \|w_j^{\eta,\delta,l}\| - (\delta_{[j],l} + \eta_{[j],l}\rho) \right)}{\|u_j^{\eta,\delta}\|} \right)^2$$

for the iterates  $\bar{f}_r^{\eta,\delta}$ ,  $r \in \mathbb{N}$ , and  $z \in \mathcal{M}_{A,g}^{\text{sd},\rho}$ .

**Lemma 3.7.** *The discrepancy principle yields a finite stopping index  $r_* = r_*(\eta, \delta)$ .*

**Proof.** Let us assume that there is no finite stopping index  $r_*$ , i.e.,  $D_{rK}^{\eta,\delta,\mathcal{K}} \neq \emptyset$  for all  $r \in \mathbb{N}$ . Then, for each  $j \in D_{rK}^{\eta,\delta,\mathcal{K}}$ , there is at least one  $l \in \mathcal{L}$  such that  $l \in D_j^{\eta,\delta}$  and

$$0 < \delta_{[j],l} + \eta_{[j],l}\rho < \frac{1}{\tau_{[j],l}} \|w_j^{\eta,\delta,l}\|. \quad (32)$$

We set

$$C_\tau := \min \left\{ \left( 1 - \frac{1}{\tau_{k,l}} \right)^2 : k \in \mathcal{K}, l \in \mathcal{L} \right\}.$$

We plug relation (32) into the descent property from remark 3.6 and obtain

$$\begin{aligned} \|z - \bar{f}_{r+1}^{\eta,\delta}\|^2 &\leq \|z - \bar{f}_r^{\eta,\delta}\|^2 - \sum_{j \in D_{rK}^{\eta,\delta,\mathcal{K}}} \left( \sum_{l \in D_j^{\eta,\delta}} \frac{\|w_j^{\eta,\delta,l}\|^2 \left( 1 - \frac{1}{\tau_{[j],l}} \right)}{\|u_j^{\eta,\delta}\|} \right)^2 \\ &\leq \|z - \bar{f}_r^{\eta,\delta}\|^2 - C_\tau \sum_{j \in D_{rK}^{\eta,\delta,\mathcal{K}}} \frac{\|\bar{w}_j^{\eta,\delta}\|^4}{\|u_j^{\eta,\delta}\|^2}. \end{aligned}$$

In particular we have used our definition of  $\bar{w}_j^{\eta,\delta,l}$  for the last estimate. Together with

$$\begin{aligned} \|u_j^{\eta,\delta}\| &= \left\| \sum_{l \in D_j^{\eta,\delta}} (A_{[j],l}^\eta)^* w_j^{\eta,\delta,l} \right\| \leq \sum_{l \in D_j^{\eta,\delta}} \|(A_{[j],l}^\eta)^*\| \cdot \|w_j^{\eta,\delta,l}\| \\ &\leq \Lambda_{A^\eta} \sum_{l \in D_j^{\eta,\delta}} \|w_j^{\eta,\delta,l}\| \leq \Lambda_{A^\eta} L \|\bar{w}_j^{\eta,\delta}\| \end{aligned}$$

we deduce

$$\|z - \bar{f}_{r+1}^{\eta,\delta}\|^2 \leq \|z - \bar{f}_r^{\eta,\delta}\|^2 - \frac{C_\tau}{\Lambda_{A^\eta}^2 L^2} \sum_{j \in D_{rK}^{\eta,\delta,\mathcal{K}}} \|\bar{w}_j^{\eta,\delta}\|^2.$$

In a similar way as in the proof of theorem 3.2 we find that

$$\sum_{j \in D_{rK}^{\eta,\delta,\mathcal{K}}} \|\bar{w}_j^{\eta,\delta}\|^2 \rightarrow 0 \quad \text{as } r \rightarrow \infty.$$

This is a contradiction to (32).  $\square$

In the following we want to show that the RESESOP-Kaczmarz iteration is a regularization method in case the conditions (22) and

$$w_i^{\eta,\delta,l} := w_{n,i}^{\eta,\delta,l} := \begin{cases} A_{[i],l}^\eta f_i^{\eta,\delta} - g_{[i],l}^\delta, & \text{for } l \in D_i^{\eta,\delta}, \\ 0, & \text{for } l \notin D_i^{\eta,\delta}, \end{cases} \quad (33)$$

$$u_i^{\eta,\delta} := u_{n,i}^{\eta,\delta} := \sum_{l \in \mathcal{L}} (A^\eta)_{[i],l}^* w_i^{\eta,\delta,l}$$

for all  $i \in I_n^{\eta,\delta}$  are fulfilled. We begin by showing that the iterates  $f_n^{\eta,\delta}$ ,  $n \in \mathbb{N}$ , depend continuously on the noise level  $\delta$  and the inexactness  $\eta$ .

**Lemma 3.8.** *Let  $f_0^{\eta,\delta} = f_0 \in X$  be the initial value and  $I_n^{\eta,\delta} = I_n$  for all  $n \in \mathbb{N}$ . Let  $\{f_n^{\eta,\delta}\}_{n \in \mathbb{N}}$  be the sequence of iterates generated by the RESESOP-Kaczmarz algorithm 2.7 with (22), (33). By  $\{f_n\}_{n \in \mathbb{N}}$  we denote the sequence generated by the SESOP-Kaczmarz algorithm 2.7 for  $(\eta, \delta) = (0, 0)$  with (22), (27). If the set of search directions*

$$\{u_{n,i} : i \in I_n\}$$

is linearly independent in each iteration  $n \in \mathbb{N}$ , we have

$$\lim_{(\eta,\delta) \rightarrow (0,0)} f_n^{\eta,\delta} = f_n.$$

**Proof.** See appendix B.  $\square$

We now have assembled all necessary statements to prove the main result for the RESESOP-Kaczmarz algorithm 2.7: if suitable conditions are fulfilled, the RESESOP-Kaczmarz algorithm yields a regularization method for semi-discrete linear inverse problems (12).

**Theorem 3.9.** *Let the conditions from lemma 3.8 be satisfied. We additionally postulate that the optimization parameters  $t_{n,i}^{\eta,\delta}$  fulfill  $|t_{n,i}^{\eta,\delta}| < t$  for all  $i \in I_n^{\eta,\delta}$ ,  $n \in \mathbb{N}$ . Then the RESESOP-Kaczmarz algorithm 2.7 together with the discrepancy principle from definition 3.4 as a stopping criterion yields a regularized solution  $f_{r_*(\eta,\delta)}^{\eta,\delta} \in X$  of the system (12) and*

$$f_{r_*(\eta,\delta)}^{\eta,\delta} \rightarrow f^+ := P_{\mathcal{M}_{A,g}^{\text{sd},\rho}}(f_0). \quad (34)$$

The proof is given in appendix C.

#### 4. Practical realizations of the RESESOP-Kaczmarz algorithm

The semi-discrete setting

$$A_{k,l}f = g_{k,l}, \quad k \in \mathcal{K}, \quad l \in \mathcal{L},$$

allows a range of modifications of our RESESOP-Kaczmarz algorithm 2.7. We recall that algorithm 2.7 first executes an *averaging* of the search directions  $\{u_{[n],i}^l : l \in \mathcal{L}\}$  before executing a Kaczmarz loop with respect to the index  $k$ .

Evidently, the roles of  $k$  and  $l$  could be interchanged, resulting in an algorithmic version that performs the Kaczmarz projection with respect to the index  $l$  and averages regarding the index  $k$ . Alternatively, the Kaczmarz projection could be executed with respect to both  $k$  and  $l$ ,

**Table 1.** Overview of different practical realizations.

	$\mathcal{K}^{\text{alg}}$	$\mathcal{L}^{\text{alg}}$
Algorithm V1 <i>Kaczmarz projection with respect to <math>k</math> and <math>l</math></i>	$\mathcal{K} \times \mathcal{L}$	$\{0\}$
Algorithm V2 <i>Averaging with respect to <math>k</math> and <math>l</math></i>	$\{0\}$	$\mathcal{K} \times \mathcal{L}$
Algorithm V3 <i>Kaczmarz iteration regarding <math>k</math> and averaging regarding <math>l</math></i>	$\mathcal{K}$	$\mathcal{L}$
Algorithm V4 <i>Kaczmarz iteration regarding <math>l</math> and averaging regarding <math>k</math></i>	$\mathcal{L}$	$\mathcal{K}$

omitting any averaging step. In the same spirit, it is possible to average with respect to both  $k$  and  $l$ .

The respective algorithms can be inferred from our general formulation in algorithm 2.7. In order to specify this, we introduce a separate notation for the index sets used in algorithm 2.7 for the Kaczmarz projection,  $\mathcal{K}^{\text{Alg}}$ , and for the averaging,  $\mathcal{L}^{\text{Alg}}$ , respectively. In case of algorithm 2.7, these sets coincide with the respective index sets  $\mathcal{K}$  and  $\mathcal{L}$  given by the semi-discrete inverse problem. With this additional notation, we can infer the modifications described in table 1.

The performance of these four versions of the RESESOP-Kaczmarz regularization is evaluated in section 5.1 at numerical examples from dynamic CT.

Furthermore, our framework includes the setting of an inverse problem which is discretized with respect to one data variable only. This is of particular interest in the context of dynamic inverse problems if we want to take into account model errors  $\eta_k$  only depending on time. In this case, simply set  $\mathcal{L} := \{0\}$  in our setting for semi-discrete inverse problems. In particular, in our RESESOP-Kaczmarz framework, algorithms V1 and V3 coincide (we refer to them as *RESESOP-Kaczmarz iteration for a single discretization parameter*), as well as algorithms V2 and V4 (referred to as *averaged RESESOP iteration for a single discretization parameter* in this case).

**Remark 4.1.** Algorithm V2 with  $\mathcal{K} = \{0\}$  and  $\eta = 0$  has been used in [37] to solve a parameter identification problem in terahertz tomography, where the set  $\mathcal{L}$  corresponds to the different detectors.

As remarked earlier, also the continuous setting with global model inexactness  $\|\mathcal{A} - \mathcal{A}^\eta\| \leq \eta \in [0, \infty)$  is covered by our framework with  $\mathcal{K} = \mathcal{L} = \{0\}$ .

In analogy to [30, 32, 35, 36] we can derive a fast algorithm with two search directions per iteration, in which the new iterate is calculated by at most two metric projections onto (intersections of) bounding hyperplanes. In a Hilbert space setting, these can be formulated explicitly without an evaluation of the underlying minimization problem. This is outlined in more detail in appendix D. For our numerical evaluation in section 5.1, we implemented algorithm 2.7 and the modifications V1–V4 with two search directions as given above.

## 5. Application in dynamic computerized tomography (CT)

In the following, we evaluate our regularization scheme for inverse problems with inexact forward operator at the example of dynamic CT.

In CT, a radiation source rotates around the studied object while emitting x-rays through the specimen to a detector panel. Depending on the density in the interior of the object, the photons

of the x-ray beam get absorbed. According to the Beer–Lambert law, the observed intensity loss in 2D is modelled by the Radon transform which integrates the attenuation coefficient  $f$  of the studied tissue along straight lines (namely the travelling paths of the x-rays). In a semi-discrete setting, this model is given by

$$R_{k,l}f := \int f(s_l\theta_k + \beta\theta_k^\perp) d\beta,$$

$$k \in \{0, \dots, K-1\} =: \mathcal{K}, \quad l \in \{0, \dots, L-1\} =: \mathcal{L},$$

where  $\theta_k \in S^1, k \in \mathcal{K}$ , characterizes the different source positions and  $s_l \in [-1, 1], l \in \mathcal{L}$ , denotes the affected detector points. The goal is to recover  $f$  from the measured CT data

$$g_{k,l} = R_{k,l}f, \quad k \in \mathcal{K}, \quad l \in \mathcal{L},$$

see [24] for a detailed overview.

In modern CT scanners, the time consuming process of data acquisition is the rotation of the x-ray source around the specimen. Hence, each source position—in our model each unit vector  $\theta_k$ —can be interpreted as a time instance. The standard model above is based on the assumption that the investigated object is stationary during the data collection. Thus, if the specimen shows a dynamic behavior during the collection of the data, it is no longer valid.

Correlating the individual states of the specimen to one reference configuration  $f$  via deformation fields  $\Gamma_k : \mathbb{R}^2 \rightarrow \mathbb{R}^2, k \in \mathcal{K}$ , the exact dynamic forward operator is given by

$$R_{\Gamma_{k,l}}f = \int f(\Gamma_k(s_l\theta_k + \beta\theta_k^\perp)) d\beta.$$

While  $R_{k,l}$  integrates along the straight lines

$$C_{k,l} := \{x \in \mathbb{R}^2 : s_l = x^T\theta_k\},$$

the dynamic model  $R_{\Gamma_{k,l}}$  integrates along curved lines

$$C_{\Gamma_{k,l}} := \{x \in \mathbb{R}^2 : s_l = (\Gamma_k^{-1}x)^T\theta_k\}.$$

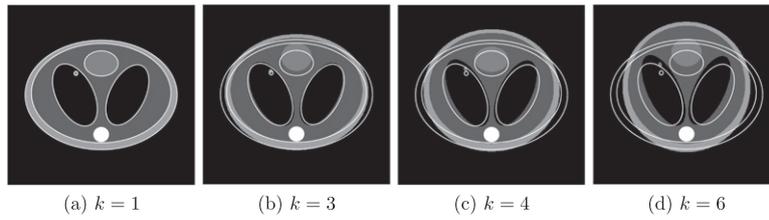
If the deformation fields  $\Gamma_k, k \in \mathcal{K}$  are not explicitly known, neither is  $R_{\Gamma_{k,l}}$ . Therefore, we propose to solve the dynamic inverse problem

$$R_{\Gamma_{k,l}}f = g_{k,l}$$

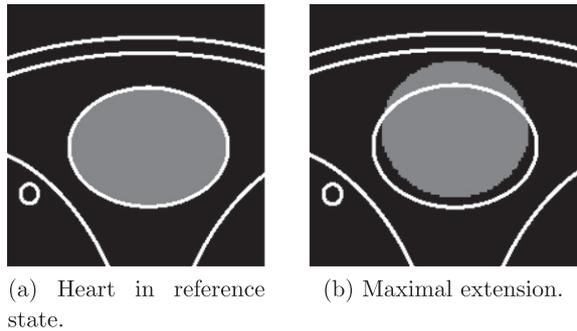
by applying our developed regularization strategy which uses an inexact version of  $R_{\Gamma_{k,l}}$ , for instance the known static operator  $R_{k,l}$ , along with estimates of the local model errors  $\eta_{k,l}$ , i.e.,  $\|R_{k,l} - R_{\Gamma_{k,l}}\| \leq \eta_{k,l}$ .

### 5.1. Numerical results

We evaluate our algorithms for two different numerical phantoms, one undergoing a periodic and affine deformation, the other performing a non-periodic and non-affine motion. The corresponding dynamic Radon data are computed for  $K = 300$  source positions with  $L = 451$  detector points using the composite trapezoidal rule and bilinear interpolation on a  $400 \times 400$  grid. For the iterative reconstruction scheme, we choose the initial value  $f_0 = 0$ , and regarding the discrepancy principle, we use constantly  $\tau = 1,00,001$  for all arising subproblems.



**Figure 2.** Dynamic behaviour of the phantom during one half of a breathing cycle with contours of the reference state (white solid line).



**Figure 3.** Visualization of the heart beat with contours of the reference state (white solid line).

*Periodic and affine deformation modelling respiratory and cardiac motion.* As first example, we consider a chest phantom performing 30 breathing cycles and 100 heart beats during the data collection. The dynamic behavior of the phantom during one half of a breathing cycle is illustrated in figure 2. The local motion of the heart—without the effect of the breathing—is further highlighted in figure 3.

In terms of explicit formulas, this dynamic behavior is described as follows. We choose the initial state, shown in figure 2(a), as reference configuration. Each ellipse, labelled consecutively by  $i = 1, \dots, 7$  as in table 2, performs an affine deformation described by  $\Gamma_k^i: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,

$$\Gamma_k^i x := A_k^i x + b_k^i.$$

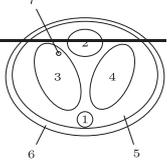
With the auxiliary variables

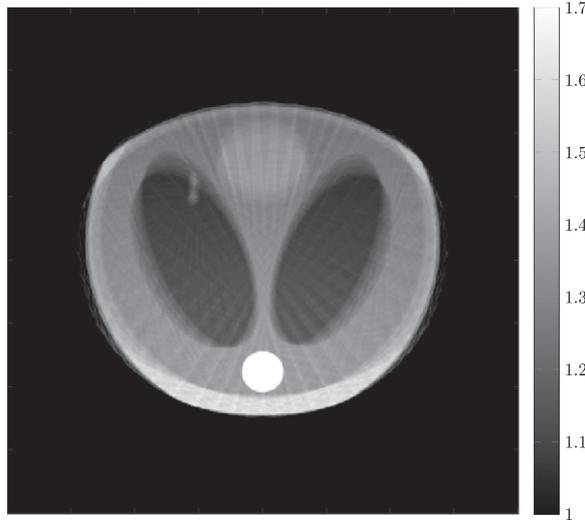
$$a_k := k \bmod 10, \quad \alpha_k := \frac{\pi}{80} \begin{cases} a_k, & a_k \leq 4 \\ 9 - a_k, & \text{otherwise} \end{cases},$$

$$d_k := 0.08 \cos\left(\frac{k}{5}\pi\right) + 0.92, \quad d_k^{\text{heart}} := \begin{cases} 1, & k \bmod 3 \leq 1 \\ 0.9, & \text{otherwise} \end{cases},$$

we define the matrices

**Table 2.** Chest phantom.

$i$		$A_k^i$	$b_k^i$	
1	Spine	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	
2	Heart	$S_k^{\text{heart}}$	$b_k$	
3, 7	Left lung, tumor	$S_k U(\alpha_k)$	$b_k$	
4	Right lung	$S_k U(-\alpha_k)$	$b_k$	
5, 6	Surrounding tissue	$S_k$	$b_k$	

**Figure 4.** SESOP reconstruction result using the static forward operator without taking the model inexactness into account.

$$U(\alpha_k) := \begin{pmatrix} \cos(\alpha_k) & \sin(\alpha_k) \\ -\sin(\alpha_k) & \cos(\alpha_k) \end{pmatrix}, \quad S_k := \begin{pmatrix} \frac{1}{d_k} & 0 \\ 0 & d_k \end{pmatrix},$$

$$S_k^{\text{heart}} := \begin{pmatrix} \frac{1}{d_k^{\text{heart}}} & 0 \\ 0 & d_k^{\text{heart}} \end{pmatrix}.$$

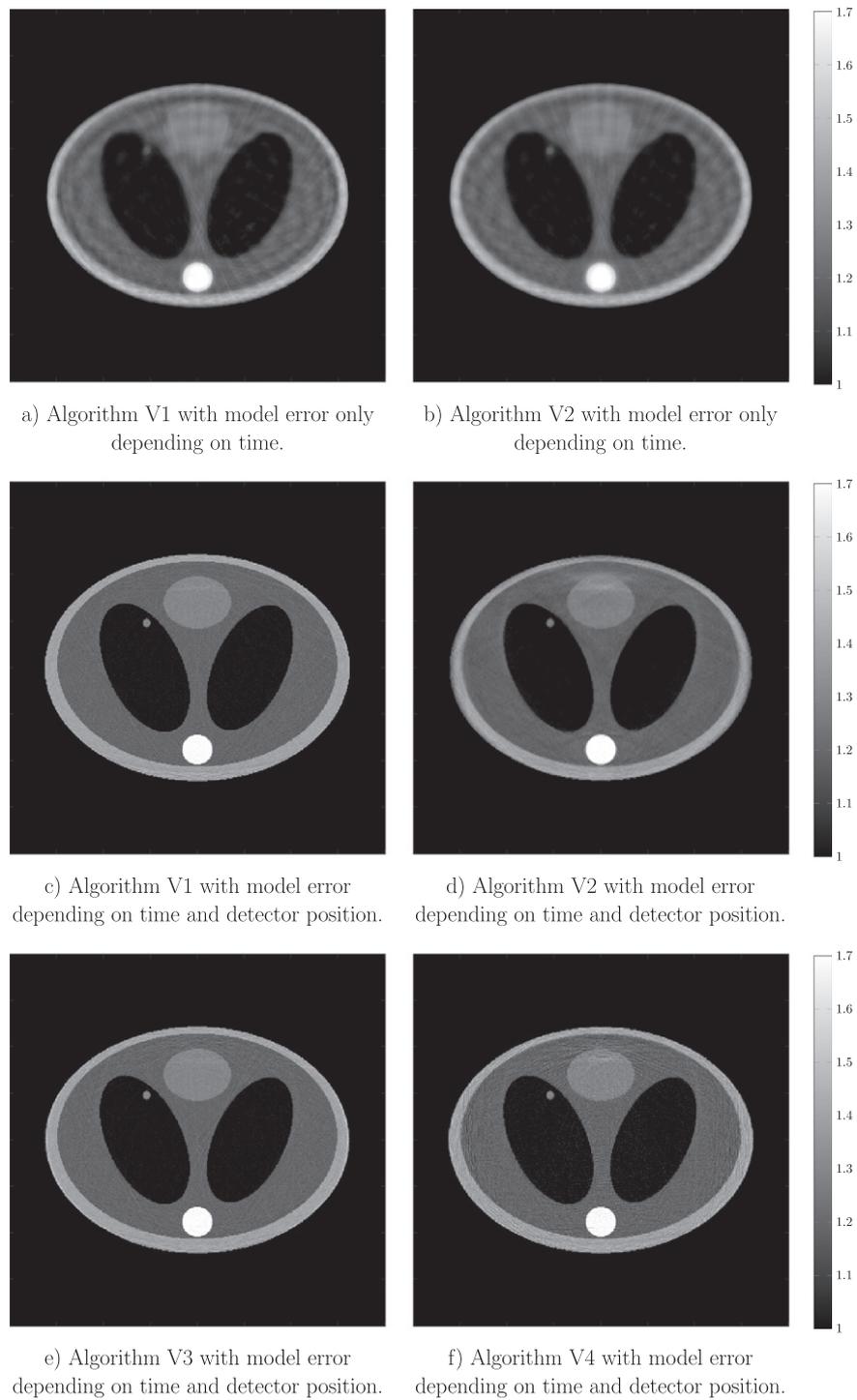
Together with the translation vector

$$b_k := (0, 0.44(d_k - 1))^T$$

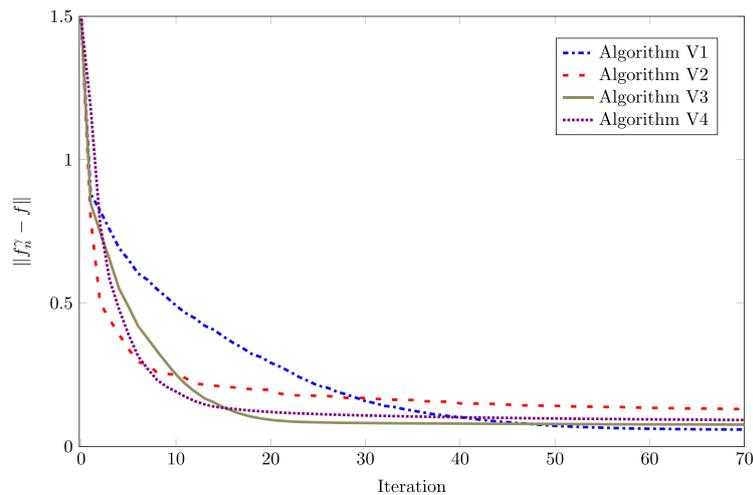
the individual motion functions  $\Gamma_k^i x = A_k^i x + b_k^i$  are chosen according to table 2.

Our goal is to recover an image of the initial state from the dynamic CT data. Since the dynamic forward operators  $R_{\Gamma_k, l}$  are in general unknown, we only want to use the known static operator  $R_{k, l}$  within the reconstruction step.

Figure 4 shows the result of the classic SESOP algorithm that does not take the model inexactness into account. The naive use of the static forward operator yields a blurred and



**Figure 5.** Results of the reconstruction algorithms incorporating the model inexactness.



**Figure 6.** Plot of error reduction.

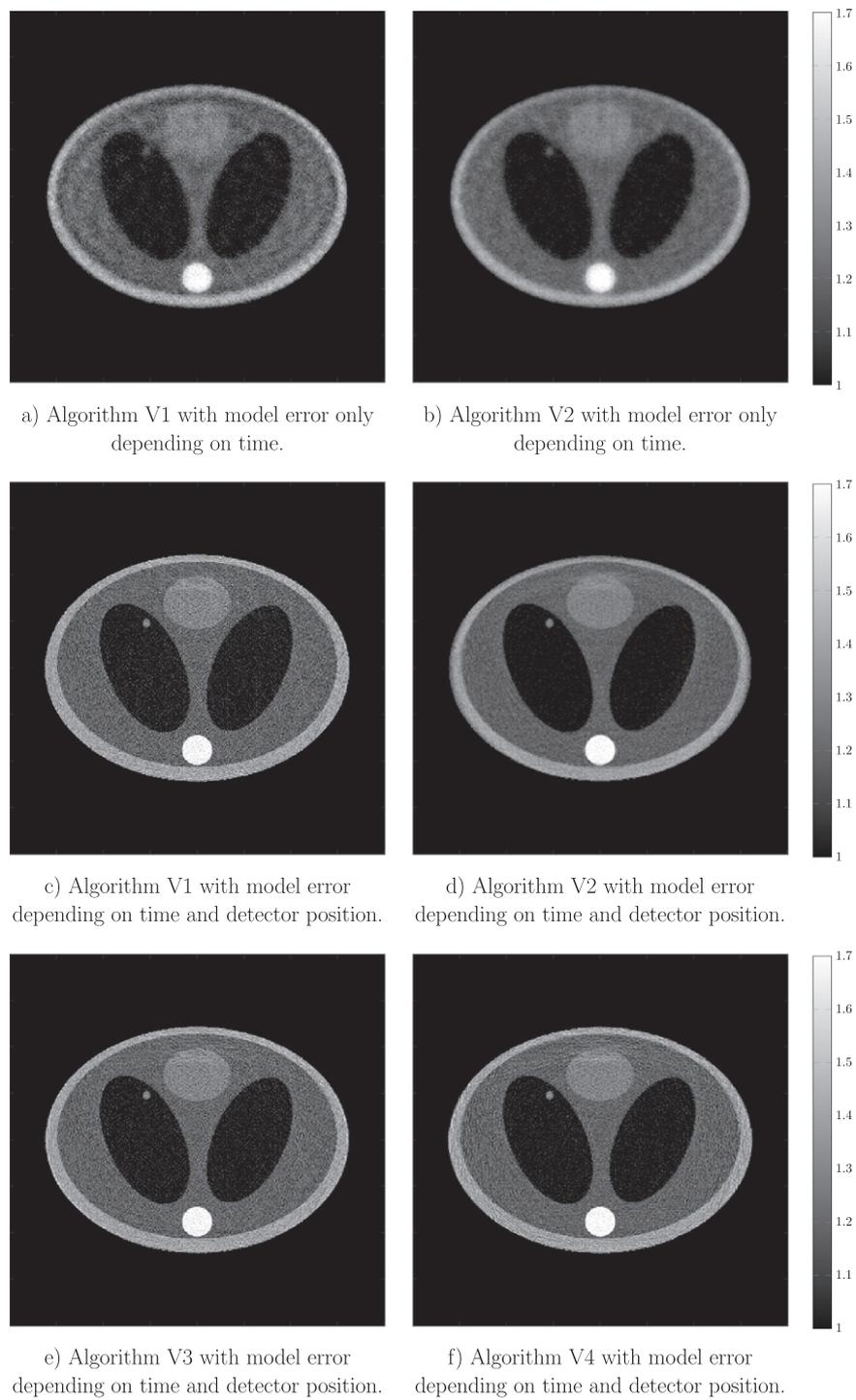
deformed reconstruction, motivating that the model inexactness needs to be incorporated into the reconstruction step.

Therefore, we next apply our derived regularization technique for inverse problems with inexact forward operator. In a first step, we want to use a good approximation for the local model errors  $\eta_{k,l}$  which we determine numerically as the discrepancy of computed static and dynamic data.

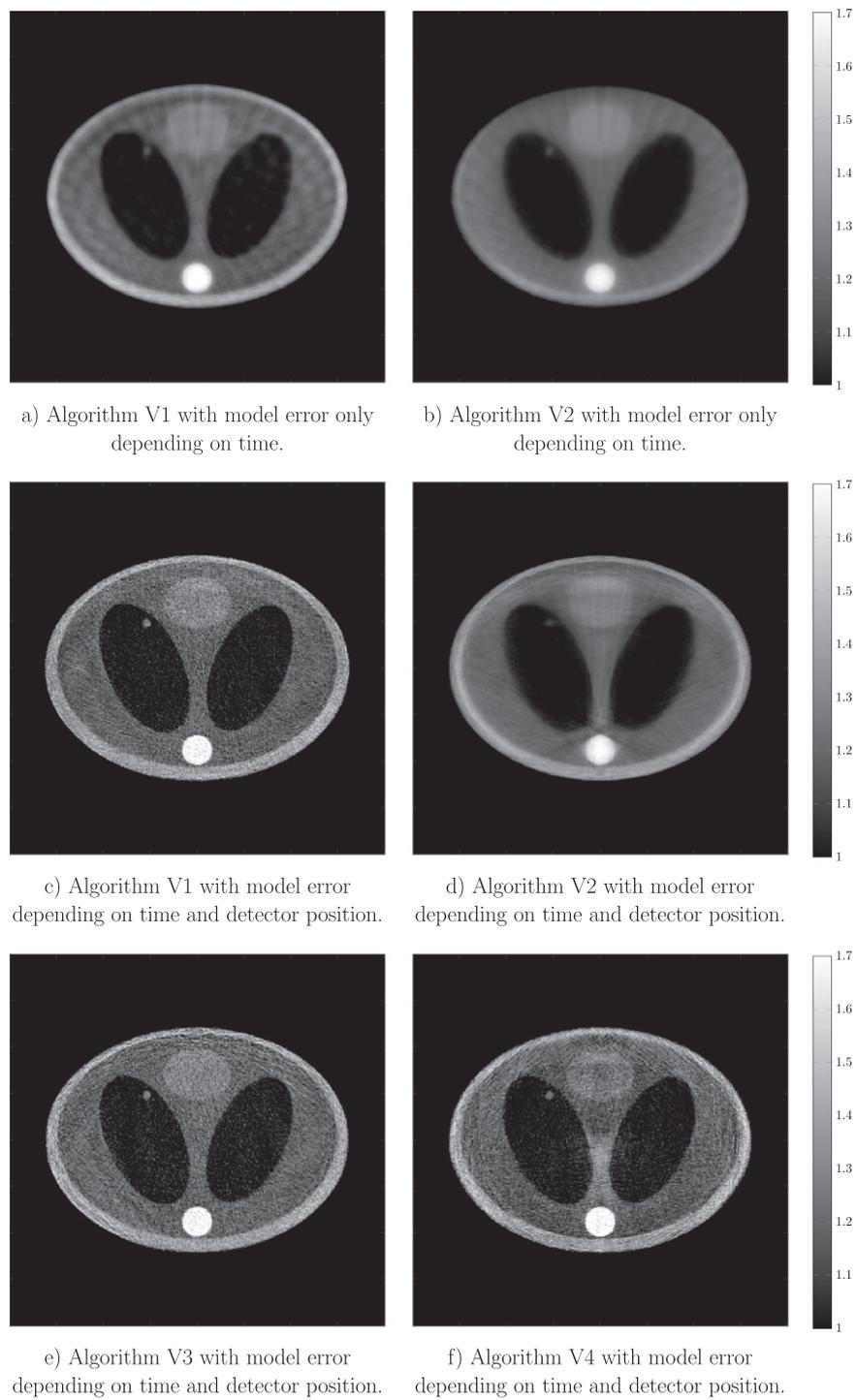
The reconstruction results of the individual algorithms presented in section 4 are shown in figure 5. Each of these algorithms provides an image of the interior structure at the reference time without motion artefacts. Thus, the suggested approach to treat the unknown dynamics as uncertainty in the forward model is indeed capable of compensating for the motion.

Drawing a comparison among the results of the different algorithms shows that considering local modelling errors  $\eta_{k,l}$  depending on source as well as detector position provides a higher resolution and sharper images, see figures 5(c)–(f) in comparison to figures 5(a) and (b)). This is to be expected since this approach incorporates the richest information on the dynamics. In particular, the local motion of the heart is more efficiently compensated for. We further observe that the algorithms involving an averaging part (figures 5(b) and (d), respectively) provide slightly smoother results than the true Kaczmarz algorithms (see figures 5(a) and (c)). When incorporating local model errors and combining Kaczmarz with an averaging part regarding the source position, i.e., regarding different time instances, can result in small artefacts in the region of local deformations (in our case in the upper part of the heart), see figures 5(d) and (f)).

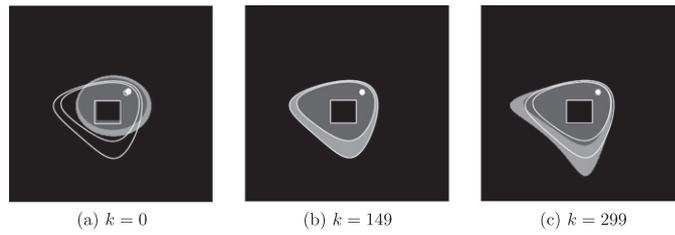
In order to gain further insights into the performance of the four algorithms V1–V4 proposed for incorporating source and detector dependent model uncertainties, figure 6 displays the error of the generated iterates and the ground truth. With an increasing number of iterations, we clearly observe the monotonicity proved in proposition 3.5 and the consecutive remark. However, the slope and the magnitude slightly differ for each algorithm. This is due to the different number of subproblems that are considered for the averaging part. More precisely, the number of subproblems is



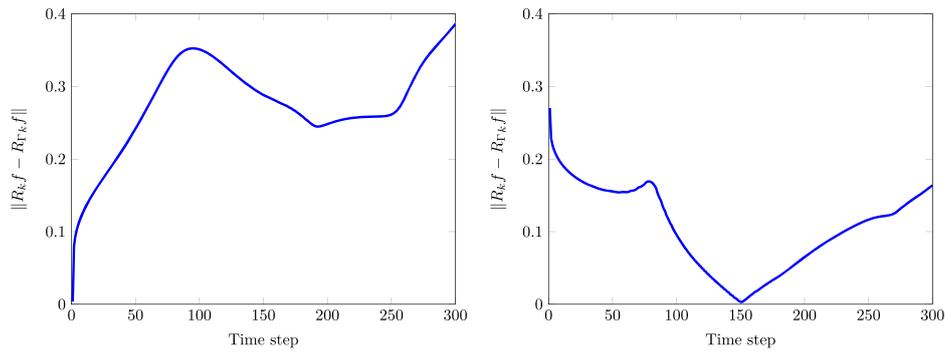
**Figure 7.** Reconstruction results for noisy data.



**Figure 8.** Reconstruction results with overestimated uncertainty level.



**Figure 9.** Images of the phantom at the beginning, the middle, and the end of motion. Contours of the reference state (determined by  $\theta_{149}$ ) are illustrated.



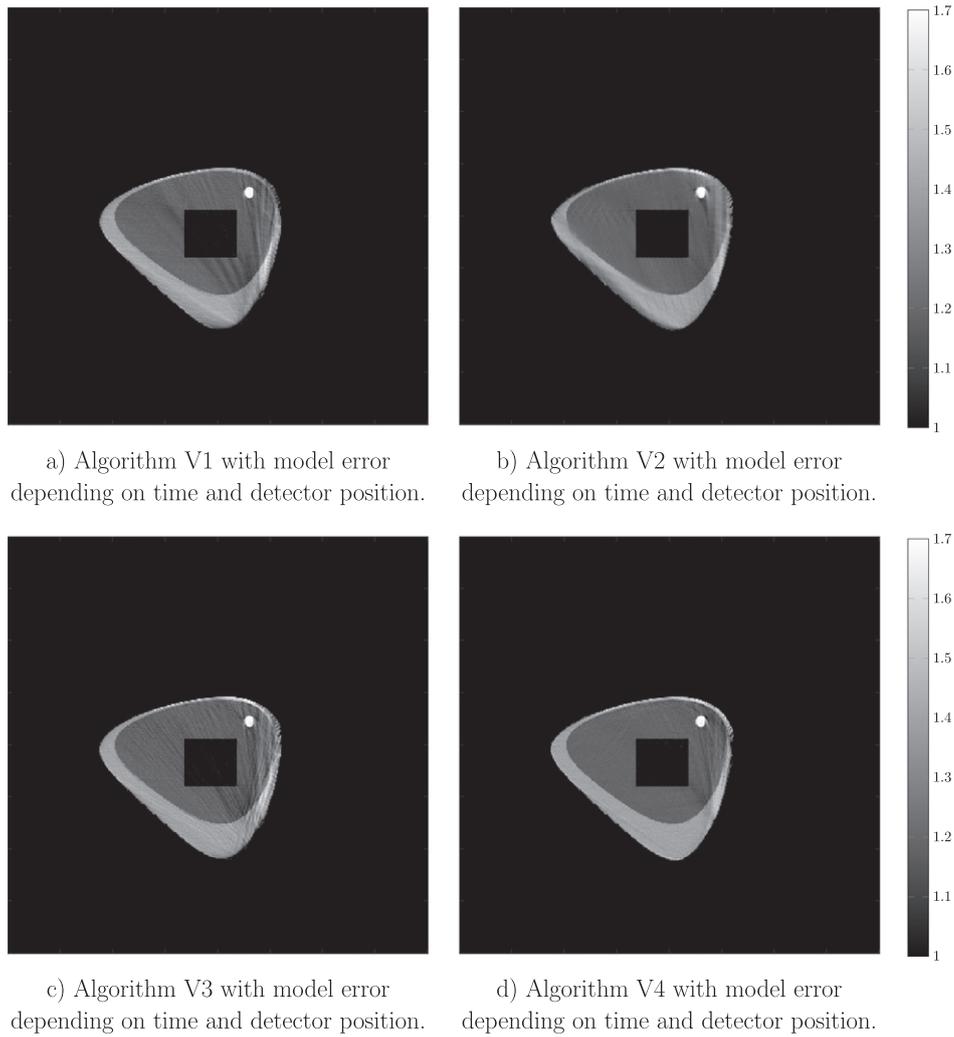
**Figure 10.** Computed model errors with respect to the initial state (left) and to the state at time instance  $k = 149$  (right) chosen as reference configuration.

- 1 in case of algorithm V1 (dashdotted blue),
- $L = 451$  for algorithm V3 (solid green),
- $K = 300$  for algorithm V4 (dotted violet) and
- $K \cdot L$  for algorithm V2 (dashed red).

The more subproblems are considered, the faster the error decreases in the beginning (since more information are included in the first iteration steps). However, the averaging in each subproblem has an additional smoothing effect resulting in larger errors for higher iteration indices.

In order to test the stability of our method with respect to perturbations in the measurement, we add noise samples uniformly distributed in  $[-0.02, 0.02]$  to our dynamic Radon data. Again, we obtain good results with our algorithms as shown in figure 7. This illustrates that our approach can in fact handle both noise in the data and modelling errors. In particular, the additional smoothing property of the algorithms involving an inner summation part yields a better noise suppression than their respective counterparts.

So far, we used uncertainty levels computed numerically as the discrepancy of dynamic and static data. In practice, however, this is not feasible since a static data set will not be available. Determining good estimates for these error levels will be part of future research (see also the discussion in section 6). Nevertheless, we want to test the sensitivity of our method regarding overestimated uncertainty levels  $\eta_{k,l}$ . For this purpose, we add a sample of noise uniformly distributed in  $[0.02, 0.06]$  to the computed values, which corresponds to an overestimation of

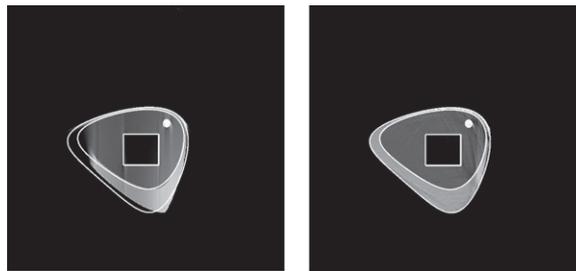


**Figure 11.** Reconstruction results for non-periodic and non-affine motion.

the model errors by up to 5%. The respective reconstruction results shown in figure 8 illustrate that our method still reconstructs a good approximation to the actual reference state. Comparing the results for the different algorithms reveals that the algorithms of type V1 (V3) are more stable with respect to overestimations of the uncertainty.

*Non-periodic and non-affine deformation.* The dynamic behaviour considered so far corresponds to an affine motion of periodic nature. Due to the high periodicity, the phantom returns to its reference state multiple times during the data collection, and, as a consequence,  $\|R_{k,l} - R_{\Gamma_{k,l}}\|$  is small for these time instances  $k$ . Now, we want to study how our method performs in case of a non-periodic motion. Our test phantom and its (non-affine) dynamic behavior are illustrated in figure 9. The deformation is described by the motion functions

$$(\Gamma_k x)_i := x_i \left( 0.2(a_i(k)x_i)^4 + (a_i(k)x_i)^3 + 2(a_i(k)x_i)^2 + 2(a_i(k)x_i) + 1 \right),$$



(a) Static reconstruction without considering the model inexactness. (b) Reconstruction with Algorithm V4.

**Figure 12.** Comparison of reconstruction results (after  $n = 200$  iterations) ignoring (left) and incorporating (right) the model inexactness with contours of the reference state.

for  $i = 1, 2$  with

$$a_i(k) := \sqrt[4]{5m_i(k)}, \quad m_1(k) := \sin\left(0.015\frac{k}{K}\right), \quad m_2(k) := \sin\left(0.021\frac{k}{K}\right),$$

where  $x_i$  refers to the  $i$ th component of the vector  $x \in \mathbb{R}^2$ .

We now aim to recover the state of the phantom at time instance  $k = 149$  from the dynamic data. This is achieved by adapting the model errors  $\eta_{k,l}$  such that  $\eta_{149,l} = 0$  for all  $l$ . Our choice serves two purposes: on the one hand, we want to illustrate that we are not restricted to recovering the initial state. In addition, in case of a monotone motion, considering the initial state will typically be disadvantageous with respect to the magnitude of the model errors, which will increase monotonically up to the maximum value at the end of the scanning. This is illustrated in figure 10 (left) which shows the computed model errors with respect to the initial state as reference configuration. However, if we choose a state during the scanning, then the model errors will first decrease and then increase, see figure 10 (right). In particular, the maximum value is lower than for the initial state which corresponds to overall smaller stripe widths.

In order to recover the selected reference state from the dynamic data, we apply our method with the simplified static operator accounting for the respective modelling errors. Due to the strong non-linearity of the motion which affects the lower left part of the phantom much more than the upper right part, we incorporate modelling errors  $\eta_{k,l}$  depending on both source position and detector point. The respective reconstruction results for our respective algorithms are shown in figure 11. A comparison with the reference state in figure 9(b)) reveals that our approach can also compensate for a non-periodic and highly non-linear dynamic behavior, whereas the naive approach ignoring the model errors results in a distorted visualization of the interior structure, see figure 12.

## 6. Conclusion

In this article, we presented a regularization strategy for linear inverse problems with inexact forward operator based on SESOP. The suggested combination with Kaczmarz' method in particular allows to account for local modelling errors. This plays an essential role when dealing with dynamic inverse problems since characteristics of the motion such as periodicity

or locality can be incorporated in this way. Our numerical examples from dynamic CT illustrate that our iterative reconstruction method compensates for the object motion during the data acquisition without requiring the explicit deformation fields.

Studying suitable prior information on the motion in more detail will be subject to future research together with the application on real data. Instead of performing our approach with the static operator, it is further possible to use a dynamic forward operator with simplified motion model and to account for modelling and estimation errors regarding the deformation fields. In particular, we expect further insights regarding the relevant information on the model inexactness by taking into account specifics of the individual application.

## Acknowledgments

The work of the second author was supported by the Deutsche Forschungsgemeinschaft under Grant HA 8176/1-1. The work of the third author was funded under the Grant 05M16TSA by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF). She would also like to acknowledge support from the Hermann und Dr Charlotte Deutsch Stiftung.

## Appendix A. Proof of theorem 3.3

**Proof.** According to our assumptions (22) and (27), the absolute convergence of the series  $\sum_{n=0}^{\infty} \|\bar{w}_n\|^2$ , which we have shown in the proof of theorem 3.2, implies the absolute convergence of the series

$$\sum_{n \in \mathbb{N}} \sum_{i \in I_n} \|\bar{w}_i\|^2, \quad \sum_{n \in \mathbb{N}} \sum_{k \in \mathcal{K}_n} \sum_{i \in I_k} \|\bar{w}_i\|^2, \quad \sum_{n \in \mathbb{N}} \|\bar{W}_n\|^2,$$

where  $\mathcal{K}_n := \{n - K, \dots, n - 1\} \cap \mathbb{N}$  and

$$\|\bar{W}_n\| := \max \{ \|\bar{w}_i\| : i \in \{n - K - N + 1, \dots, n - 1\} \cap \mathbb{N} \}.$$

Note that

$$\sum_{n \in \mathbb{N}} \|\bar{W}_n\|^2 \leq \sum_{n \in \mathbb{N}} \sum_{k \in \mathcal{K}_n} \sum_{i \in I_k} \|\bar{w}_i\|^2.$$

Our previous results show that we can find a subsequence  $\{f_{n_j}\}_{j \in \mathbb{N}}$  with the following properties:

- (a) The sequence  $\{\|f_{n_j}\|\}_{j \in \mathbb{N}}$  converges,
- (b) The sequence  $\{f_{n_j}\}_{j \in \mathbb{N}}$  is weakly convergent,
- (c)  $\lim_{j \rightarrow \infty} \|\bar{W}_{n_j}\| = 0$  and  $\|\bar{W}_{n_j}\| \leq \|\bar{W}_n\|$  for all  $n \leq n_j$ .

We show that  $\{f_{n_j}\}_{j \in \mathbb{N}}$  is a Cauchy sequence. For all  $j, m \in \mathbb{N}$ ,  $j > m$ , we have

$$\begin{aligned} \|f_{n_m} - f_{n_j}\|^2 &= \|f_{n_m}\|^2 - \|f_{n_j}\|^2 - 2 \langle f_{n_m} - f_{n_j}, f_{n_j} \rangle \\ &\rightarrow -2 \langle f_{n_m} - f_{n_j}, f_{n_j} \rangle \quad \text{for } m \rightarrow \infty \end{aligned}$$

due to (a). Furthermore, using (b), we obtain

$$\begin{aligned} \langle f_{n_m} - f_{n_j}, f_{n_j} \rangle &= \langle f_{n_m} - f_{n_j}, f^+ \rangle + \langle f_{n_m} - f_{n_j}, f_{n_j} - f^+ \rangle \\ &\rightarrow \langle f_{n_m} - f_{n_j}, f_{n_j} - f^+ \rangle \quad \text{for } m \rightarrow \infty. \end{aligned}$$

We apply the iteration from algorithm 2.7, i.e.,

$$f_{n_j} = f_{n_m} - \sum_{n=n_m}^{n_j-1} \sum_{i \in I_n} t_{n,i} u_i,$$

and estimate

$$\begin{aligned} |\langle f_{n_m} - f_{n_j}, f_{n_j} - f^+ \rangle| &= \left| \sum_{n=n_m}^{n_j-1} \sum_{i \in I_n} t_{n,i} \langle u_i, f_{n_j} - f^+ \rangle \right| \\ &\leq \sum_{n=n_m}^{n_j-1} \sum_{i \in I_n} t_{n,i} |\langle u_i, f_{n_j} - f^+ \rangle| \end{aligned} \tag{A.1}$$

as well as

$$|\langle u_i, f_{n_j} - f^+ \rangle| = \left| \sum_{l \in \mathcal{L}} \langle A_{[i],l}^* w_l^i, f_{n_j} - f^+ \rangle \right| \leq \sum_{l \in \mathcal{L}} \|w_l^i\| \cdot \|A_{[i],l}(f_{n_j} - f^+)\|. \tag{A.2}$$

We want to estimate the term  $\|A_{[i],l}(f_{n_j} - f^+)\|$  from (A.2), i.e., the index  $i, j$  and  $l$  are fixed. Now let  $\vartheta_j \in \mathbb{N}$  be defined by

$$\vartheta_j := \begin{cases} n_j - [n_j] + [i] - K, & \text{if } [i] \geq [n_j], \\ n_j - [n_j] + [i], & \text{if } [i] < [n_j], \end{cases}$$

such that  $[\vartheta_j] = [i]$  and  $n_j - K \leq \vartheta_j \leq n_j - 1$ . We then obtain, analogously to our previous step,

$$\|f_{n_j} - f_{\vartheta_j}\| = \left\| \sum_{n=\vartheta_j}^{n_j-1} \sum_{i \in I_n} t_{n,i} \sum_{l \in \mathcal{L}} A_{[i],l}^* w_l^i \right\| \leq t \cdot \Lambda_A \sum_{n=\vartheta_j}^{n_j-1} \sum_{i \in I_n} L \cdot \|\bar{w}_i\|.$$

Together with  $\vartheta_j \in I_{\vartheta_j}$  and  $\|w_{\vartheta_j}^l\| \leq \|\bar{w}_{\vartheta_j}\|$  (by definition, see (25)) the above estimate yields

$$\begin{aligned} \|A_{[i],l}(f_{n_j} - f^+)\| &= \|A_{[i],l}(f_{n_j} - f_{\vartheta_j} + f_{\vartheta_j} - f^+)\| \\ &\leq \|A_{[i],l}(f_{n_j} - f_{\vartheta_j})\| + \|A_{[i],l}(f_{\vartheta_j} - f^+)\| \\ &\leq t \Lambda_A^2 \sum_{n=\vartheta_j}^{n_j-1} \sum_{i \in I_n} L \cdot \|\bar{w}_i\| + \|\bar{w}_{\vartheta_j}\| \\ &\leq t \Lambda_A^2 (L + 1) \sum_{n=\vartheta_j}^{n_j-1} \sum_{i \in I_n} \|\bar{w}_i\| \\ &\leq t \Lambda_A^2 N \cdot K (L + 1) \|\bar{W}_{n_j}\|. \end{aligned}$$

With  $C := t\Lambda_A^2 N \cdot K(L+1)$  we thus have

$$\|A_{[l],l}(f_{n_j} - f^+)\| \leq C \|\bar{W}_{n_j}\|. \quad (\text{A.3})$$

We plug this into (A.1) and arrive at

$$\begin{aligned} |\langle f_{n_m} - f_{n_j}, f_{n_j} - f^+ \rangle| &\leq \sum_{n=n_m}^{n_j-1} \sum_{i \in I_n} t |\langle u_i, f_{n_j} - f^+ \rangle| \\ &\leq Ct \sum_{n=n_m}^{n_j-1} \sum_{i \in I_n} \sum_{l \in \mathcal{L}} \|w_i^l\| \cdot \|\bar{W}_{n_j}\| \\ &\leq CLt \sum_{n=n_m}^{n_j-1} \sum_{i \in I_n} \|\bar{w}_i\| \cdot \|\bar{W}_{n_j}\| \\ &\leq CLNt \sum_{n=n_m}^{n_j-1} \|\bar{W}_{n+1}\| \cdot \|\bar{W}_{n_j}\| \\ &\leq \tilde{C} \sum_{n=n_m}^{n_j-1} \|\bar{W}_{n+1}\|^2, \end{aligned}$$

where  $\tilde{C} := CLNt$ .

Due to the absolute convergence of  $\sum_{n \in \mathbb{N}} \|\bar{W}_n\|^2$ , the right-hand side of the above inequality converges to 0 as  $m \rightarrow \infty$ . Altogether, this shows that  $\{f_{n_j}\}_{j \in \mathbb{N}}$  is indeed a Cauchy sequence, converging to some  $\tilde{f} \in X$ . The continuity of the norm and the forward operators  $A_{k,l}$ ,  $k \in \mathcal{K}$ ,  $l \in \mathcal{L}$ , yield

$$\|A_{k,l}(\tilde{f} - f^+)\| = \lim_{j \rightarrow \infty} \|A_{k,l}(f_{n_j} - f^+)\| \leq \lim_{j \rightarrow \infty} C \|\bar{W}_{n_j}\| = 0, \quad (\text{A.4})$$

i.e.,  $\tilde{f} \in \mathcal{M}_{A,g}^{\text{sd}}$ . As before, we conclude  $f_n \rightarrow \tilde{f}$  for  $n \rightarrow \infty$  and  $\bar{f}_r \rightarrow \tilde{f}$  for  $r \rightarrow \infty$ .

It remains to show that  $\tilde{f} = f^+ := P_{\mathcal{M}_{A,g}^{\text{sd}}}(f_0)$ . From (A.4) we derive

$$(\tilde{f} - f^+) \in \bigcap_{k \in \mathcal{K}} \bigcap_{l \in \mathcal{L}} \mathcal{N}(A_{k,l}).$$

On the other hand we have

$$f_n - f_0 \in \sum_{k \in \mathcal{K}} \sum_{l \in \mathcal{L}} \overline{\mathcal{R}(A_{k,l}^*)}$$

and thus

$$\tilde{f} - f_0 \in \overline{\sum_{k \in \mathcal{K}} \sum_{l \in \mathcal{L}} \mathcal{R}(A_{k,l}^*)} = \overline{\sum_{k \in \mathcal{K}} \sum_{l \in \mathcal{L}} \mathcal{N}(A_{k,l})^\perp} = \left( \bigcap_{k \in \mathcal{K}} \bigcap_{l \in \mathcal{L}} \mathcal{N}(A_{k,l}) \right)^\perp.$$

We finally use proposition 3.7 from [31] for Hilbert spaces to obtain

$$\tilde{f} = P_{f^+} \bigcap_{k \in \mathcal{K}} \bigcap_{l \in \mathcal{L}} \mathcal{N}(A_{k,l})(f_0) = P_{\mathcal{M}_{A,g}^{\text{sd}}}(f_0),$$

i.e., the result  $\tilde{f}$  of the SESOP-Kaczmarz iteration corresponds to the metric projection of the initial value  $f_0$  onto the solution set  $\mathcal{M}_{A,g}^{\text{sd}}$ .  $\square$

## Appendix B. Proof of lemma 3.8

**Proof.** First we note that our assumptions on the inexactness and the continuity of the forward operators  $A_{k,l}$ ,  $k \in \mathcal{K}$ ,  $l \in \mathcal{L}$ , yield  $w_i^{\eta,\delta,l} \rightarrow w_i^l$  for  $(\eta, \delta) \rightarrow (0, 0)$  if  $f_i^{\eta,\delta} \rightarrow f_i$  for  $(\eta, \delta) \rightarrow (0, 0)$ , since

$$\begin{aligned} \|w_i^{\eta,\delta,l} - w_i^l\| &= \|A_{[i],l}^\eta f_i^{\eta,\delta} - g_{[i],l}^\delta - (A_{[i],l} f_i - g_{[i],l})\| \\ &\leq \|A_{[i],l}^\eta f_i^{\eta,\delta} - A_{[i],l} f_i\| + \|g_{[i],l}^\delta - g_{[i],l}\| \\ &\leq \|(A_{[i],l}^\eta - A_{[i],l}) f_i^{\eta,\delta}\| + \|A_{[i],l} (f_i^{\eta,\delta} - f_i)\| + \delta_{[i],l} \\ &\leq \eta_{[i],l} \|f_i^{\eta,\delta}\| + \Lambda_A \|f_i^{\eta,\delta} - f_i\| + \delta_{[i],l} \\ &\rightarrow 0 \quad \text{for } (\eta, \delta) \rightarrow (0, 0). \end{aligned}$$

Similarly, we then also obtain

$$\lim_{(\eta,\delta) \rightarrow (0,0)} u_i^{\eta,\delta} = u_i, \quad \lim_{(\eta,\delta) \rightarrow (0,0)} \alpha_i^{\eta,\delta} = \alpha_i, \quad \lim_{(\eta,\delta) \rightarrow (0,0)} \xi_i^{\eta,\delta} = \xi_i = 0.$$

We now prove by induction that  $\lim_{(\eta,\delta) \rightarrow (0,0)} f_n^{\eta,\delta} = f_n$  for all  $n \in \mathbb{N}$ . For  $n = 0$  we have  $f_0^{\eta,\delta} = f_0$  and  $I_0^{\eta,\delta} = I_0 = \{0\}$ . For

$$f_1^{\eta,\delta} = \begin{cases} f_0, & \text{if } D_0^{\eta,\delta} = \emptyset, \\ f_0 - \frac{\langle u_0^{\eta,\delta}, f_0 \rangle - \alpha_0^{\eta,\delta} - \xi_0^{\eta,\delta}}{\|u_0^{\eta,\delta}\|^2} u_0^{\eta,\delta}, & \text{if } D_0^{\eta,\delta} \neq \emptyset \end{cases}$$

we directly obtain  $f_1^{\eta,\delta} \rightarrow f_1$  as  $(\eta, \delta) \rightarrow (0, 0)$  according to our initial considerations.

Now suppose that  $\lim_{(\eta,\delta) \rightarrow (0,0)} f_j^{\eta,\delta} = f_j$  for all  $j \leq n$ . The iterate

$$f_{n+1}^{\eta,\delta} = \begin{cases} f_n^{\eta,\delta}, & \text{if } D_n^{\eta,\delta} = \emptyset, \\ f_n^{\eta,\delta} - \sum_{k \in I_n^{\eta,\delta}} t_{n,k}^{\eta,\delta} u_k^{\eta,\delta}, & \text{if } D_n^{\eta,\delta} \neq \emptyset \end{cases} \quad (\text{B.1})$$

fulfills  $f_{n+1}^{\eta,\delta} \in H_n^{\eta,\delta}$ , i.e., we have

$$\alpha_i^{\eta,\delta} - \xi_i^{\eta,\delta} \leq \langle u_i^{\eta,\delta}, f_{n+1}^{\eta,\delta} \rangle \leq \alpha_i^{\eta,\delta} + \xi_i^{\eta,\delta}$$

for all  $i \in I_n^{\eta,\delta}$ . We plug (B.1) into the above inequality and arrive at

$$\langle u_i^{\eta,\delta}, f_n^{\eta,\delta} \rangle - \alpha_i^{\eta,\delta} - \xi_i^{\eta,\delta} \leq \sum_{k \in I_n^{\eta,\delta}} t_{n,k}^{\eta,\delta} \langle u_i^{\eta,\delta}, u_k^{\eta,\delta} \rangle \leq \langle u_i^{\eta,\delta}, f_n^{\eta,\delta} \rangle - \alpha_i^{\eta,\delta} + \xi_i^{\eta,\delta}. \quad (\text{B.2})$$

By defining the Gram matrix

$$G_n^{\eta,\delta} := \left( \left\langle u_i^{\eta,\delta}, u_k^{\eta,\delta} \right\rangle \right)_{i,k \in I_n^{\eta,\delta}}$$

as well as the vectors

$$t_n^{\eta,\delta} := \left( t_{n,i}^{\eta,\delta} \right)_{i \in I_n^{\eta,\delta}}, \quad b_n^{\eta,\delta} := \left( \left\langle u_i^{\eta,\delta}, f_n^{\eta,\delta} \right\rangle - \alpha_i^{\eta,\delta} \right)_{i \in I_n^{\eta,\delta}}, \quad d_n^{\eta,\delta} := \left( \xi_i^{\eta,\delta} \right)_{i \in I_n^{\eta,\delta}},$$

we reformulate (B.2) as

$$b_n^{\eta,\delta} - d_n^{\eta,\delta} \leq G_n^{\eta,\delta} t_n^{\eta,\delta} \leq b_n^{\eta,\delta} + d_n^{\eta,\delta} \quad (\text{B.3})$$

(the inequality is to be understood componentwise). The matrix  $G_n^{\eta,\delta}$  is invertible if and only if  $\{u_i^{\eta,\delta} : i \in I_n^{\eta,\delta}\}$  is linearly independent. Thus,  $G_n := (\langle u_i, u_k \rangle)_{i,k \in I_n}$  is invertible. Since the search directions  $u_i^{\eta,\delta}$ ,  $i \in I_n^{\eta,\delta}$ , depend continuously on  $(\eta, \delta)$ , and the inner product is continuous, the matrix  $G_n^{\eta,\delta}$  is invertible if  $(\eta, \delta)$  is close to  $(0, 0)$ . We now let  $(\eta, \delta) \rightarrow (0, 0)$  in (B.3). Due to our previous considerations, we obtain

$$\lim_{(\eta,\delta) \rightarrow (0,0)} t_n^{\eta,\delta} = \lim_{(\eta,\delta) \rightarrow (0,0)} (G_n^{\eta,\delta})^{-1} b_n^{\eta,\delta} = t_n.$$

This yields  $f_{n+1}^{\eta,\delta} \rightarrow f_{n+1}$  for  $(\eta, \delta) \rightarrow (0, 0)$ , which concludes the proof.  $\square$

### Appendix C. Proof of Theorem 3.9

**Proof.** For all  $k \in \mathcal{K}, l \in \mathcal{L}$  let  $\{\delta_{k,l}^i\}_{i \in \mathbb{N}}$  with  $\delta_{k,l}^i \geq 0$  for all  $i \in \mathbb{N}$  and  $\{\eta_{k,l}^j\}_{j \in \mathbb{N}}$  with  $\eta_{k,l}^j \geq 0$  for all  $j \in \mathbb{N}$  be null sequences such that

$$\|g_{k,l} - g_{k,l}^{\delta_{k,l}^i}\| \leq \delta_{k,l}^i, \quad \|A_{k,l} - A_{k,l}^{\eta_{k,l}^j}\| \leq \eta_{k,l}^j.$$

Note that the sequence  $\{(\eta_{k,l}^j, \delta_{k,l}^i)\}_{(i,j) \in \mathbb{N} \times \mathbb{N}}$  is a null sequence if and only if all subsequences  $\{(\eta_{k,l}^{\tilde{j}(p)}, \delta_{k,l}^{\tilde{i}(p)})\}_{p \in \mathbb{N}}$  with strictly increasing functions  $\tilde{i}, \tilde{j} : \mathbb{N} \rightarrow \mathbb{N}$  are null sequences. In the following, we set

$$(\eta, \delta)_k^{i,j} := (\eta_{k,l}^j, \delta_{k,l}^i)_{l \in \mathcal{L}}, \quad (\eta, \delta)^{i,j} := \left( (\eta, \delta)_k^{i,j} \right)_{k \in \mathcal{K}}$$

for all  $i, j \in \mathbb{N}$ . We further write

$$(\eta, \delta)^p := (\eta, \delta)^{\tilde{i}(p), \tilde{j}(p)}$$

for  $p \in \mathbb{N}$ .

Let  $n_*$   $((\eta, \delta)^p)$  denote the auxiliary stopping index as defined in definition 3.4. The sequence  $\{n_*((\eta, \delta)^p)\}_{p \in \mathbb{N}}$  either has a finite accumulation point  $n \in \mathbb{N}$  or  $n_*((\eta, \delta)^p) \rightarrow \infty$  for  $p \rightarrow \infty$ .

We begin with the first case and assume, without loss of generality, that  $n_*((\eta, \delta)^p) = n$  for all  $p \in \mathbb{N}$ . We then have

$$\left\| A_{k,l}^{\tilde{\eta}^{(p)}} f_n^{(\eta,\delta)^p} - g_{k,l}^{\tilde{\delta}^{(p)}} \right\| \leq \tau_{k,l} \left( \tilde{\delta}_{k,l}^{(p)} + \tilde{\eta}_{k,l}^{(p)} \cdot \rho \right)$$

for all  $p \in \mathbb{N}$  and  $k \in \mathcal{K}$ ,  $l \in \mathcal{L}$ . By lemma 3.8 we have  $f_n^{(\eta,\delta)^p} \rightarrow f_n$  for  $p \rightarrow \infty$ . In addition, we obtain

$$\begin{aligned} \|A_{k,l} (f_n^{(\eta,\delta)^p} - f^+)\| &\leq \left\| A_{k,l}^{\tilde{\eta}^{(p)}} f_n^{(\eta,\delta)^p} - g_{k,l}^{\tilde{\delta}^{(p)}} \right\| + \left\| A_{k,l} - A_{k,l}^{\tilde{\eta}^{(p)}} \right\| \cdot \|f_n^{(\eta,\delta)^p}\| \\ &\quad + \left\| g_{k,l}^{\tilde{\delta}^{(p)}} - g_{k,l} \right\| \\ &\leq \tau_{k,l} \left( \tilde{\delta}_{k,l}^{(p)} + \tilde{\eta}_{k,l}^{(p)} \cdot \rho \right) + \tilde{\eta}_{k,l}^{(p)} \cdot \|f_n^{(\eta,\delta)^p}\| + \tilde{\delta}_{k,l}^{(p)} \end{aligned}$$

for all  $k \in \mathcal{K}$ ,  $l \in \mathcal{L}$ . The right-hand side converges to 0 and the left-hand side to  $\|A_{k,l} (f_n - f^+)\|$  for  $p \rightarrow \infty$ . We thus have

$$(f_n - f^+) \in \bigcap_{k \in \mathcal{K}} \bigcap_{l \in \mathcal{L}} \mathcal{N}(A_{k,l}).$$

Similarly as in the proof of theorem 3.3, we also obtain

$$f_n - f_0 \in \left( \bigcap_{k \in \mathcal{K}} \bigcap_{l \in \mathcal{L}} \mathcal{N}(A_{k,l}) \right)^\perp$$

and thus  $f_n = f^+$  and

$$\lim_{p \rightarrow \infty} \|f_n^{(\eta,\delta)^p} - f^+\| = 0.$$

Together with lemma 3.8 we obtain  $f_n^{(\eta,\delta)^p} \rightarrow f^+$  for  $p \rightarrow \infty$ .

Now let us assume that  $n_*((\eta, \delta)^p) \rightarrow \infty$  as  $p \rightarrow \infty$ . From lemma 3.1 we know that the sequence  $\{\|f_n - f^+\|^2\}_{n \in \mathbb{N}}$  is a bounded, monotonically decreasing null sequence. Hence, for all  $\epsilon > 0$  there exists  $N_0(\epsilon) \in \mathbb{N}$  such that

$$\|f_{N_0(\epsilon)} - f^+\|^2 < \frac{\epsilon}{2}$$

and  $p_0 \in \mathbb{N}$  with  $n_*((\eta, \delta)^p) \geq N_0(\epsilon)$  for all  $p \geq p_0$ . Due to lemma 3.8 and the continuity of the norm we find a  $p_1 \in \mathbb{N}$  with

$$\left| \|f_{N_0(\epsilon)}^{(\eta,\delta)^p} - f^+\|^2 - \|f_{N_0(\epsilon)} - f^+\|^2 \right| < \frac{\epsilon}{2}$$

for all  $p \geq p_1$ . We finally conclude

$$\begin{aligned} \|f_{n_*((\eta,\delta)^p)}^{(\eta,\delta)^p} - f^+\|^2 &\leq \|f_{N_0(\epsilon)}^{(\eta,\delta)^p} - f^+\|^2 \\ &\leq \|f_{N_0(\epsilon)} - f^+\|^2 + \left| \|f_{N_0(\epsilon)}^{(\eta,\delta)^p} - f^+\|^2 - \|f_{N_0(\epsilon)} - f^+\|^2 \right| \\ &< \epsilon \end{aligned}$$

for all  $p \geq \max\{p_0, p_1\}$ , which finishes our proof.

Since we have  $r_*(\eta, \delta) \cdot K \geq n_*(\eta, \delta)$ , we also have  $\bar{f}_{r^*}^{\eta, \delta} = f_{n_*(\eta, \delta)}^{\eta, \delta}$  and we immediately obtain

$$\lim_{(\eta, \delta) \rightarrow (0, 0)} \bar{f}_{r^*}^{\eta, \delta} = f^+.$$

□

## Appendix D. RESESOP-Kaczmarz algorithm with two search directions

**Algorithm appendix D.1 (RESESOP-Kaczmarz with two search directions).** In algorithm 2.7, set  $I_n^{\eta, \delta} := \{n_-, n\}$  with  $n_- := \max \{i \in \{-1, \dots, n-1\} : D_i^{\eta, \delta} \neq \emptyset\}$ . Define  $H_{-1}^{\eta, \delta} := X$  and  $D_{-1}^{\eta, \delta} := \mathcal{K}$ .

Whenever  $D_n^{\eta, \delta} \neq \emptyset$ , we have

$$f_n^{\eta, \delta} \in H_{>}(u_n^{\eta, \delta}, \alpha_n^{\eta, \delta} + \xi_n^{\eta, \delta}) \cap H_{n_-}^{\eta, \delta}. \quad (\text{D.1})$$

Compute  $f_{n+1}^{\eta, \delta}$  according to the following two steps:

(a) Determine

$$\begin{aligned} \tilde{f}_{n+1}^{\eta, \delta} &:= P_{H_n^{\eta, \delta}}(f_n^{\eta, \delta}) = P_H(u_n^{\eta, \delta}, \alpha_n^{\eta, \delta} + \xi_n^{\eta, \delta})(f_n^{\eta, \delta}) \\ &= f_n^{\eta, \delta} - \sum_{l \in D_n^{\eta, \delta}} \frac{\|w_n^{\eta, \delta, l}\| \cdot (\|w_n^{\eta, \delta, l}\| - (\delta_{[n], l} - \eta_{[n], l} \rho))}{\|u_n^{\eta, \delta}\|^2} u_n^{\eta, \delta}. \end{aligned}$$

If now  $\tilde{f}_{n+1}^{\eta, \delta} \in H_{n_-}^{\eta, \delta}$ , set  $f_{n+1}^{\eta, \delta} := \tilde{f}_{n+1}^{\eta, \delta}$ . Otherwise, proceed with step (b):

(b) Decide whether

1.  $\tilde{f}_{n+1}^{\eta, \delta} \in H_{>}(u_{n_-}^{\eta, \delta}, \alpha_{n_-}^{\eta, \delta} + \xi_{n_-}^{\eta, \delta})$  or
2.  $\tilde{f}_{n+1}^{\eta, \delta} \in H_{<}(u_{n_-}^{\eta, \delta}, \alpha_{n_-}^{\eta, \delta} - \xi_{n_-}^{\eta, \delta})$ .

Calculate accordingly

$$\begin{aligned} f_{n+1}^{\eta, \delta} &:= P_{H(u_n^{\eta, \delta}, \alpha_n^{\eta, \delta} + \xi_n^{\eta, \delta}) \cap H_{\geq}(u_{n_-}^{\eta, \delta}, \alpha_{n_-}^{\eta, \delta} \pm \xi_{n_-}^{\eta, \delta})}(\tilde{f}_{n+1}^{\eta, \delta}) \\ &= \tilde{f}_{n+1}^{\eta, \delta} + \langle u_n^{\eta, \delta}, u_{n_-}^{\eta, \delta} \rangle T_n^{\eta, \delta} - \|u_n^{\eta, \delta}\|^2 T_n^{\eta, \delta}, \end{aligned}$$

where

$$T_n^{\eta, \delta} := \frac{\langle u_{n_-}^{\eta, \delta}, \tilde{f}_{n+1}^{\eta, \delta} \rangle - (\alpha_{n_-}^{\eta, \delta} \pm \xi_{n_-}^{\eta, \delta})}{\|u_n^{\eta, \delta}\|^2 \cdot \|u_{n_-}^{\eta, \delta}\|^2 - \langle u_n^{\eta, \delta}, u_{n_-}^{\eta, \delta} \rangle^2}.$$

Then  $f_{n+1}^{\eta, \delta} = P_{H_n^{\eta, \delta} \cap H_{n_-}^{\eta, \delta}}$ .

**Remark appendix D.2.** Algorithm appendix D.1 is inspired by the fast algorithm that was first presented in [30] and then adapted in [35, 36] for nonlinear inverse problems. The statements given in algorithm appendix D.1 such as (D.1) can be proved in a straightforward way, see also [30, 35, 36].

## ORCID iDs

Bernadette N Hahn  <https://orcid.org/0000-0002-4876-8188>

Anne Wald  <https://orcid.org/0000-0001-6149-8576>

## References

- [1] Boutchko R, Rayz V L, Vandehey N T, O'Neil J P, Budinger T F, Nico P S, Moses W W, Druhan J L, Saloner D A and Gullberg G T 2012 Imaging and modelling of flow in porous media using clinical nuclear emission tomography systems and computational fluid dynamics *J. Appl. Geophys.* **76** 74–81
- [2] Burger M, Dirks H, Frerking L, Hauptmann A, Helin T and Siltanen S 2017 A variational reconstruction method for undersampled dynamic x-ray tomography based on physical motion models *Inverse Problems* **33** 124008
- [3] Chen C, Gris B and Öktem O 2019 A new variational model for joint image reconstruction and motion estimation in spatiotemporal imaging *SIAM J. Imaging Sci.* **12** 1686–719
- [4] Chen D, Li H, Wang Q, Zhang P and Zhu Y 2015 Computed tomography for high-speed rotation object *Opt. Express* **23** 13423–42
- [5] Chung J, Saibaba A K, Brown M and Westman E 2018 Efficient generalized Golub–Kahan based methods for dynamic inverse problems *Inverse Problems* **34** 024005
- [6] Crawford C, King K, Ritchie C and Godwin J 1996 Respiratory compensation in projection imaging using a magnification and displacement model *IEEE Trans. Med. Imaging* **15** 327–32
- [7] Desbat L, Roux S and Grangeat P 2007 Compensation of some time dependent deformations in tomography *IEEE Trans. Med. Imaging* **26** 261–9
- [8] Golub G H, Hansen P C and O'Leary D P 1999 Tikhonov regularization and total least squares *SIAM J. Matrix. Appl.* **21** 185–94
- [9] Gu R, Han B and Chen Y 2019 Fast subspace optimization method for nonlinear inverse problems in Banach spaces with uniformly convex penalty terms *Inverse Problems* **35** 125011
- [10] Hahn B 2014 Reconstruction of dynamic objects with affine deformations in dynamic computerized tomography *J. Inverse Ill-Posed Problems* **22** 323–39
- [11] Hahn B 2017 Motion estimation and compensation strategies in dynamic computerized tomography *Sens. Imaging* **18** 1–20
- [12] Hahn B N 2014 Efficient algorithms for linear dynamic inverse problems with known motion *Inverse Problems* **30** 035008
- [13] Haltmeier M, Kowar R, Leitao A and Scherzer O 2007 Kaczmarz methods for regularizing nonlinear ill-posed equations II: applications *Inverse Problems Imaging* **1** 507–23
- [14] Haltmeier M, Leitao A and Scherzer O 2007 Kaczmarz methods for regularizing nonlinear ill-posed equations I: convergence analysis *Inverse Problems Imaging* **1** 289–98
- [15] Heber F, Schöpfer F and Schuster T 2019 Acceleration of sequential subspace optimization in Banach spaces by orthogonal search directions *J. Comput. Appl. Math.* **345** 1–22
- [16] Kaltenbacher B 2017 All-at-once versus reduced iterative methods for time dependent inverse problems *Inverse Problems* **33** 064002
- [17] Kastner J, Plank B and Heinzl C 2015 Advanced x-ray computed tomography methods: high resolution CT, phase contrast CT, quantitative CT and 4DCT *Proc. Digital Industrial Radiology and Computed Tomography (DIR 2015)* (Ghent, Belgium)
- [18] Katsevich A 2010 An accurate approximate algorithm for motion compensation in two-dimensional tomography *Inverse Problems* **26** 065007
- [19] Katsevich A, Silver M and Zamayatin A 2011 Local tomography and the motion estimation problem *SIAM J. Imaging Sci.* **4** 200–19
- [20] Kluth T and Maass P 2017 Model uncertainty in magnetic particle imaging: nonlinear problem formulation and model-based sparse reconstruction *Int. J. Magn. Part. Imaging* **3** 1707004
- [21] Korolev Y and Lellmann J 2018 Image reconstruction with imperfect forward models and applications in deblurring *SIAM J. Imaging Sci.* **11** 197–218
- [22] Lu W and Mackie T R 2002 Tomographic motion detection and correction directly in sinogram space *Phys. Med. Biol.* **47** 1267–84

- [23] Narkiss G and Zibulevsky M 2005 Sequential subspace optimization method for large-scale unconstrained optimization *Technical Report* Technion-The Israel Institute of Technology, Department of Electrical Engineering
- [24] Natterer F 1986 *The Mathematics of Computerized Tomography* (Braunschweig: Vieweg)
- [25] Nguyen T 2019 Landweber–Kaczmarz for parameter identification in time-dependent inverse problems: all-at-once versus reduced version *Inverse Problems* **35** 035009
- [26] Rashid A, Kim S, Liu D and Kim K 2016 A dynamic oppositional biogeography-based optimization approach for time-varying electrical impedance tomography *Physiol. Meas.* **37** 820–42
- [27] Rieder A and Schuster T 2000 The approximate inverse in action with an application to computerized tomography *SIAM J. Numer. Anal.* **37** 1909–29
- [28] Schmitt U and Louis A 2002 Efficient algorithms for the regularization of dynamic inverse problems: I. Theory *Inverse Problems* **18** 645–58
- [29] Schmitt U, Louis A, Wolters C and Vauhkonen M 2002 Efficient algorithms for the regularization of dynamic inverse problems: II. Applications *Inverse Problems* **18** 659–76
- [30] Schöpfer F and Schuster T 2009 Fast regularizing sequential subspace optimization in Banach spaces *Inverse Problems* **25** 015013
- [31] Schöpfer F, Schuster T and Louis A K 2008 Metric and Bregman projections onto affine subspaces and their computation via sequential subspace optimization methods *J. Inverse Ill-Posed Problems* **16** 479–206
- [32] Schuster T, Kaltenbacher B, Hofmann B and Kazimierski K S 2012 *Regularization Methods in Banach Spaces* (Berlin: de Gruyter & Co)
- [33] Tarvainen T, Kolehmainen V, Pulkkinen A, Vauhkonen M, Schweiger M, Arridge S R and Kaipio J P 2009 An approximation error approach for compensating for modelling errors between the radiative transfer equation and the diffusion approximation in diffuse optical tomography *Inverse Problems* **26** 015005
- [34] Van Nieuwenhove V, De Beenhouwer J, De Schryver T, Van Hoorebeke L and Sijbers J 2017 Data-driven affine deformation estimation and correction in cone beam computed tomography *IEEE Trans. Image Process.* **26** 1441–51
- [35] Wald A 2018 A fast subspace optimization method for nonlinear inverse problems in Banach spaces with an application in parameter identification *Inverse Problems* **34** 085008
- [36] Wald A and Schuster T 2016 Sequential subspace optimization for nonlinear inverse problems *J. Inverse Ill-Posed Problems* **25** 99–117
- [37] Wald A and Schuster T 2018 Tomographic terahertz imaging using sequential subspace optimization *New Trends in Parameter Identification (Springer Series Trends in Mathematics)* ed B Hofmann, A Leitao and J P Zubelli (Basel: Birkhäuser) (<https://doi.org/10.1007/978-3-319-70824-9>)