Physics in Medicine &
Biology



# Efficient motion-corrected image reconstruction for 3D cardiac MRI through stochastic optimisation

To cite this article before publication: Letizia Protopapa *et al* 2025 *Phys. Med. Biol.* in press https://doi.org/10.1088/1361-6560/adf609

# Efficient motion-corrected image reconstruction for 3D cardiac MRI through stochastic optimisation

**Letizia Protopapa[1,*], Margaret Duff[1], Johannes Mayer[2], Jeanette Schulz-Menger[3,4,5], Kris Thielemans[6,7], Christoph Kolbitsch[2], Edoardo Pasca[1]**

[1] Scientific Computing Department, Rutherford-Appleton Laboratory, UK Research and Innovation, Harwell Campus, Didcot, UK
[2] Physikalisch-Technische Bundesanstalt (PTB), Braunschweig and Berlin, Germany
[3] Charité Medical Faculty University Medicine, Berlin, Germany
[4] Working Group on Cardiovascular Magnetic Resonance, Experimental and Clinical Research Center (ECRC), Charité Humboldt University Berlin, DZHK partner site Berlin, Berlin, Germany
[5] Department of Cardiology and Nephrology, HELIOS Klinikum Berlin Buch, Berlin, Germany
[6] Institute of Nuclear Medicine, University College London, London, UK
[7] UCL Hawkes Institute, University College London, London, UK
[*] Corresponding author: letizia.protopapa@stfc.ac.uk

14 April 2024

**Abstract.**

**Objective**

Motion-corrected image reconstruction (MCIR) allows for fast and efficient cardiac magnetic resonance imaging (MRI) acquisition with predictable scan times. Since data obtained in all phases of respiratory and cardiac motion can be exploited, the duration of the scan is not affected by changes in heart rate or irregular breathing patterns.

Achieving high-quality reconstructions from MCIR data typically requires iterative optimisation algorithms with regularisation. Reconstruction time increases with the number of motion states. This is particularly relevant in cardiac MRI, where both cardiac and respiratory motion corrections are necessary to minimise motion artefacts.

**Approach**

In this work, we present a stochastic optimisation approach for cardio-respiratory MCIR using the Stochastic Primal Dual Hybrid Gradient (SPDHG) algorithm. We compare the convergence rates with deterministic optimisation methods.

**Main Results**

In phantom experiments with simulated motion, we demonstrate the improved convergence rates of SPDHG with respect to deterministic algorithms, while maintaining image quality. Convergence is improved both in terms of reconstruction times and computational effort. We validate the method's effectiveness on an in vivo 3D whole-heart cardiac MR scan. The in vivo method demonstrates that the motion

compensation method we use allows for non-rigid deformation patterns and irregular breathing patterns.

**Significance**

This study demonstrates that stochastic algorithms can converge significantly faster than deterministic algorithms for MCIR, especially for a large number of motion states. With the proposed approach, increasing the number of motion states reduces the number of epochs required to reconstruct the image and therefore it is no longer necessary to balance the competing requirements of accurate motion correction and computational effort.

*Keywords*: motion-corrected image reconstruction, cardiac MRI, stochastic optimisation, MCIR, SPDHG

Submitted to: *Phys. Med. Biol.*

## 1. Introduction

Cardiac Magnetic Resonance Imaging (MRI) is a very versatile medical imaging technique that provides a wide range of different diagnostic information. It can be used to assess cardiac anatomy, myocardial injury such as scar and inflammation, angiography, and fat infiltration [1, 2, 3, 4, 5]. Typically, images are obtained as multiple 2D slices. Although this yields high in-plane image quality, image resolution along the slice direction is often poor. Full coverage of the entire heart is also often not possible. One major challenge of 3D cardiac MRI is its long acquisition times, making it difficult to apply it in clinical routine. The acquisition time for a 3D whole-heart acquisition with high spatial resolution is on the order of several minutes. During this time, the heart constantly moves due to the heartbeat and breathing. This physiological movement can cause severe motion artefacts that make the obtained image data unsuitable for diagnostic purposes [6]. A common approach to minimise motion artefacts is to restrict data acquisition to a predefined respiratory (e.g., end-expiration) and cardiac (e.g., mid-diastole) motion phase [7, 8]. This ensures that all data are acquired in the same motion state of the heart but also strongly increases the overall examination time. In addition, the examination time is strongly dependent on the motion pattern of individual patients, which is challenging for clinical routine [9].

Motion correction approaches have been proposed to minimise motion artefacts while ensuring shorter acquisition times. Data is acquired continuously rather than restricting the data acquisition to certain motion states. In this paper, we use the Motion Corrected Image Reconstruction (MCIR) method, first developed for MRI by Batchelor *et al.* [10], but also developed for other medical imaging modalities, such as Positron Emission Tomography (PET) [11]. In the first step, the movement of each

voxel due to respiratory and / or cardiac motion is determined using image-registration algorithms. In the second step, the obtained motion vector fields are used during image reconstruction to transform all acquired data into the same reference motion state and, therefore, obtain a single motion-corrected diagnostic magnetic resonance (MR) image [10]. Several studies have shown that this approach leads to excellent image quality while ensuring short and predictable scan times [12, 13, 14, 15, 16, 17].

The main limitation of the MCIR technique is the increased computation time for image reconstruction. Especially for cardiac MRI, where both cardiac and respiratory motion need to be corrected, a high number of motion states can be required for an accurate motion correction. In Mayer *et al.* [18], 6 respiratory motion phases and 12 cardiac motion phases were used, which resulted in 72 different motion states in the final reconstruction. Typically, this type of reconstruction process is formulated as an inverse problem and deterministic optimisation schemes such as conjugate gradient, the Fast Iterative Shrinkage-Threshold Algorithm (FISTA) [19], or the Primal Dual Hybrid-Gradient (PDHG) approach [20] are used to solve it. For example, Brown *et al.* [21] perform a joint PET-MR reconstruction, by relying on FISTA for MRI and PDHG for PET, while Zhu *et al.* [22] use PDHG for motion-compensated pulmonary MRI. However, using such iterative optimisation algorithms generally requires regularisation, either for denoising or resolution enhancement [23]. This can lead to reconstruction times of the order of hours, as in the study by Knoll *et al.* [24], where PDHG is used for a joint PET-MR reconstruction on real data, regularized through Total Generalized Variation (TGV), taking about 8 hours in total.

A common method to speed-up inverse problems is only to use a "subset" (or "minibatch") of the data for each update, where the problem geometry normally determines each subset [25]. Dikaios *et al.* [26] extended the subset idea to motion states for MCIR in PET, while Delplancke *et al.* [27] demonstrated in a proof-of-concept setting a theoretical and practical speed-up using stochastic algorithms for MCIR in computed tomography (CT).

In this study, we propose using a stochastic optimisation approach for MRI MCIR, specifically the stochastic primal-dual hybrid-gradient (SPDHG) algorithm [28], which requires very few epochs and hence shorter reconstruction times compared to commonly used optimisation schemes. There are a range of examples in the literature of applying SPDHG to large-scale inverse problems, for example for PET imaging [29], parallel MRI [30] or CT [31]. This paper's novelty consists in applying stochastic methods to MCIR for MRI. We also compare the results of stochastic optimisation with other deterministic methods. We first evaluate the faster convergence in phantom experiments with simulated motion. Importantly, the simulated phantom data allows us to study the behaviour of stochastic and deterministic methods as the number of motion states changes. We then demonstrate stochastic optimisation methods on a cardiac 3D whole-heart scan obtained in vivo, demonstrating the application of the approach to non-rigid and irregular motion.

The paper is organised as follows. In Section 2, we introduce the image

reconstruction problem and the SPDHG algorithm. We also describe the motion correction steps and the software used. In Section 3, we give details about both the in vivo and phantom acquisitions and the motion simulation for the phantom data. Then, we explain how the image reconstructions will be evaluated. In Section 4, we provide the links to data and software necessary to replicate the phantom experiments of this study. We present our results in Section 5 and discuss them in Section 6, together with future work, while in Section 7 we present our conclusions.

## 2. Method

### *2.1. Optimisation Algorithms*

Consider the inverse problem to reconstruct an image, $x \in \mathbb{X}$, from observed data, $b \in \mathbb{Y}$, related by a linear forward operator $K : \mathbb{X} \to \mathbb{Y}$, where $\mathbb{X}$ and $\mathbb{Y}$ are finite-dimensional Hilbert spaces with an inner product. To aid the reader, Table 1 summarises key information on the FISTA, PDHG, and SPDHG algorithms, used in this report.

| Algorithm | Objective it minimises | Function Requirements | Type |
|---|---|---|---|
| FISTA [19] | $x^* = \arg\min_{x\in\mathbb{X}}\{F(x) := f(x) + g(x)\}$ | $f$ is a smooth convex function which has a Lipschitz continuous gradient. $g$ is a continuous convex function which is possibly non-smooth. | Deterministic |
| PDHG [20] | $x^* = \arg\min_{x\in\mathbb{X}}\{F(x) := f(Kx) + g(x)\}$ | Both $f$ and $g$ are both proper closed and convex and their proximal operators can be evaluated, or at least approximated, efficiently. | Deterministic |
| SPDHG [28] | $x^* = \arg\min_{x\in\mathbb{X}} \sum_{i=0}^{M-1} f_i(K_i x) + g(x)$ | All $f_i$ and $g$ are both proper closed and convex and their proximal operators can be evaluated, or at least approximated, efficiently. | Stochastic |

Table 1: Key properties of the three algorithms used in this report.

More concretely, in this article, we choose $f$ to be the least squares function, so $f_{FISTA}(x) = \|Ax - b\|_2^2$ and $f_{(S)PDHG}(y) = \|y - b\|_2^2$. The regulariser $g$ is the total variation (TV) function (e.g. [32]). For PDHG and SPDHG, this formulation of TV is sometimes called "implicit" because the TV regulariser is included in the function $g$ and not as an additional term combined in $f$. "Implicit" refers to the fact that the TV proximal is not available in closed form, but it can be calculated with a few iterations of a denoising algorithm, for instance, using FISTA or PDHG. In this paper, we chose the "implicit" variation, as we found that it converged more quickly with the default step sizes. For more information on step size choices, see Section 3.5.

### 2.2. SPDHG Background

In this subsection, we expand on the information in Table 1, particularly for our key stochastic method, SPDHG. This can safely be skipped, and, in the next subsection, we will see how stochastic algorithms, such as SPDHG, are naturally suited to motion correction problems.

To solve the inverse problem $Kx = b$, we consider optimisation of the form:

$$x^* = \arg\min_{x\in\mathbb{X}} \{F(x) := f(Kx) + g(x)\} \tag{1}$$

where $f : \mathbb{Y} \to \mathbb{R}_{\geq 0}$ is a data discrepancy term, measuring the distance from the data $b$, and $g : \mathbb{X} \to \mathbb{R}_{\geq 0}$ is a regulariser, encapsulating prior knowledge of the solution. We assume that both $f$ and $g$ are both proper, closed and convex and that their proximal operators can be evaluated, or at least approximated, efficiently.

The primal-dual hybrid-gradient algorithm (PDHG) [33] takes a primal-dual formulation of equation (1) solving instead for the saddle point

$$\max_{y\in\mathbb{Y}} \min_{x\in\mathbb{X}} \{\langle Kx, y\rangle_{\mathbb{Y}} + g(x) - f^*(y)\} \tag{2}$$

where $\langle\cdot,\cdot\rangle_{\mathbb{Y}}$ is the inner product in the space $\mathbb{Y}$ and $f^*$ is the convex conjugate of $f$, $f^*(y) = \sup_{x\in\mathbb{X}}\{\langle x, y\rangle - f(x)\}$. PDHG solves the saddle point problem by alternating in a proximal gradient ascent-like way in the dual variable $y$ and then in a proximal gradient descent-like way in the primal variable $x$.

Now consider the case that $f$ is separable and can be written as $f(y) = \sum_{i=0}^{M-1} f_i(y_i)$. This requires that the Hilbert space $\mathbb{Y}$ can be expressed as product spaces $\mathbb{Y} := \prod_{i=0}^{M-1} \mathbb{Y}_i$ with elements $y = (y_0, ..., y_{M-1})$ and that we can define linear operators $K_i : \mathbb{X} \to \mathbb{Y}_i$, where $K_i x = (Kx)_i$. The optimisation problem in equation (1) becomes

$$x^* = \arg\min_{x\in\mathbb{X}} \sum_{i=0}^{M-1} f_i(K_i x) + g(x). \tag{3}$$

Equation (3) can be expressed as a saddle point problem and solved using PDHG with iterations:

$$x_{k+1} = \text{prox}_{\tau g}\left(x_k - \tau K^* \bar{y}_k\right) \tag{4}$$

$$y_{i,k+1} = \text{prox}_{\sigma_i f_i^*}\left(y_{i,k} + \sigma_i K_i x_k\right), \quad i \in \{0, 1, ..., M-1\} \tag{5}$$

$$\bar{y}_{k+1} = y_{k+1} + \theta\left(y_{k+1} - y_k\right) \tag{6}$$

where $\tau, \sigma_i > 0$ are step sizes, $\theta \in [0, 1]$ is a relaxation parameter, and the algorithm is initialised with $x_0 \in \mathbb{X}$ and $y_0 \in \mathbb{Y}$ with $\bar{y}_0 = y_0$ . Convergence is guaranteed if $\sigma_i \tau \|K\|^2 < 1$ for all $i \in \{0, ..., M-1\}$, where the choice of $\sigma_i$ and $\tau$ greatly affects the convergence speed.

A stochastic version of this algorithm, Stochastic Primal-Dual Hybrid-Gradient (SPDHG), reduces the computational cost and has been shown to converge faster with a greater number of subsets [28]. Instead of updating each dual variable, $y_i$, in each

iteration, only one dual variable is updated for each iteration. The iterations now look like

$$x_{k+1} = \text{prox}_{\tau g}\left(x_k - \tau K^* \bar{y}_k\right) \tag{7}$$

$$\text{Select } s \in \{0, ..., M-1\} \tag{8}$$

$$y_{i,k+1} = \begin{cases} \text{prox}_{\sigma_i f_i^*}\left(y_{i,k} + \sigma_i K_i x_k\right), & i = s \\ y_{i,k+1}, & i \neq s \end{cases} \tag{9}$$

$$\bar{y}_{k+1} = y_{k+1} + \theta Q\left(y_{k+1} - y_k\right) \tag{10}$$

where $Q := \text{diag}(p_1^{-1} I, ..., p_n^{-1} I)$ and $p_i$ is the probability index $i$ which is selected at each iteration. Convergence is guaranteed if $p_i > 0$ for all $i$ (see also equation 12 in [29]) and

$$\|\sigma_i^{1/2} \tau^{1/2} K_i\|^2 < p_i \qquad \forall i \in \{0, ..., M-1\}. \tag{11}$$

### *2.3. Motion Corrected Image Reconstruction*

In the case of MRI, the linear operator $K$ describes the MR acquisition model and $b$ corresponds to the acquired raw data. The raw data in MRI is obtained in the so-called k-space, which is the spatial Fourier domain of the image data. Without motion correction, $K$ can be described as

$$K = SFC \tag{12}$$

where $C$ describes the signal reception with multiple receiver coils, $F$ is the Fourier transform and $S$ is a sampling operator defining which spatial frequencies have been acquired (i.e., the k-space trajectory). In this paper, the raw data $b$ is obtained along a radial-like pattern [34], and thus $F$ is implemented as a non-uniform Fast Fourier transform (NUFFT) [35].

In the case of motion occurring during data acquisition, we can partition raw k-space data into a set $b_i$ for each motion state $i$. There is no synchronisation between the motion (e.g. breathing or heartbeat) and the data acquisition and therefore the data in each motion state $i$ are obtained at different k-space locations. This transforms $K$ to

$$K = \sum_{i=0}^{M-1} K_i = \sum_{i=0}^{M-1} S_i F T_i C \tag{13}$$

for $M$ different motion states. $T_i$ describes the non-rigid transformation of the image due to the underlying motion. If we compare this with equation (3), we can see that motion leads to a problem well-suited for SPDHG as the motion states naturally subset the problem.

A pre-processing step is required to obtain $T_i$ in equation (13). First, images are reconstructed for each motion state $i$ separately. Then, an image registration algorithm is used to estimate the non-rigid transformation $T_i$ between the different motion states. More details can be found below in the Experiments section. $T_i$ is then used in a final MCIR to obtain a high-quality diagnostic image.

For SPDHG to be applicable, and convergence guarantees to hold, we need that the motion correction can be written as a linear transformation $T_i$. This allows for a range of rigid and non-rigid motion.

In this paper, we compare different optimisation algorithms for minimising equation (3), with the operators $K_i$ for the motion-state decomposition defined as in (13). In each iteration of PDHG or FISTA, calculating the sum in equation (3) requires the computation of the forward model $K_i x$ for all $i \in \{0, ..., M-1\}$ performing computations on all the data in each iteration, while SPDHG performs calculations on a subset of the data, just one sampled $K_i$ (and corresponding data $b_i$). Because of this difference, to ensure a fair comparison between the two approaches, convergence is generally plotted and discussed in terms of "epochs" rather than iterations, where one epoch corresponds to one iteration in PDHG and $M$ iterations of SPDHG. Note that we are minimising the same convex objective with each algorithm, so we expect the same solution. The comparison is in how quickly, in terms of epochs, the algorithms reach that solution.

### *2.4. Software*

For the work presented in this article, we rely on two open source software packages: the Synergistic Image Reconstruction Framework (SIRF) developed for synergistic PET and MR image reconstruction, providing a platform to develop and test algorithms on PET and MR patient data [36], and the Core Imaging Library (CIL) [37, 38], a software package for tomographic image reconstruction, which also offers extensive optimisation and regularisation capabilities. Thanks to the integration between the two, we can use regularised PDHG and SPDHG optimisation methods (from CIL) to reconstruct MRI data using acquisition models and non-rigid transformations in SIRF.

## 3. Experiments

### *3.1. Data acquisition*

MR data acquisition was carried out with a Golden Radial Phase Encoding (GRPE) sampling scheme [39, 34] which combines non-Cartesian phase encoding with Cartesian frequency encoding. The phantom scan was performed on a 3T Siemens Verio scanner and the in vivo scan was carried out on a 1.5T Siemens Avanto scanner. Signal receptions were carried out with a 32-channel receive coil. The field-of-view was a cube of size 288 $mm^3$ and an isotropic resolution of 1.5 $mm^3$. A single echo was acquired for the phantom scans. For in vivo scans, three echoes were acquired to perform Dixon-based fat-water separation. For this study, we only used data from the first echo. Information about the coil sensitivity maps was obtained from the data directly without needing an additional coil reference scan.

All in vivo imaging was carried out in accordance with the Declaration of Helsinki. Subjects gave their informed written consent to participate in the study. A $T_1$ contrast

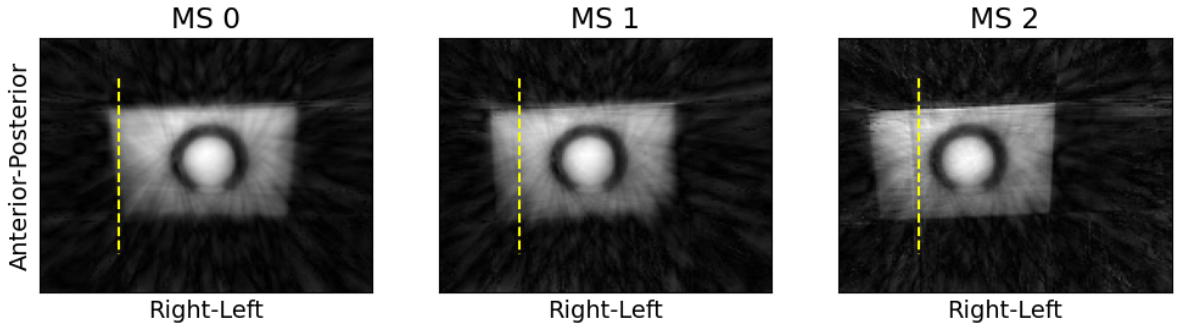


Figure 1: Example of three simulated motion states, obtained by applying a phase shift on the acquired raw k-space data, which results in a shift in the right-left direction. The figures show a slice through the volumes, reconstructed using a simple least-squares objective function without regularisation. More specifically, from the left-end side (MS0) to the right-end side (MS2), the figures show: the untranslated phantom, the phantom translated by half the maximum shift and the phantom translated by the maximum shift. The yellow dashed line is placed at a fixed position along the Right-Left axis and serves as a reference to better appreciate the effect of the translation.

agent was administered before the GRPE scan as part of the standard clinical protocol.

### *3.2. Motion simulation for the phantom*

To study the effect of different numbers of motion states on the convergence behaviour of the proposed SPDHG-MCIR approach, simulations were carried out. Starting from the acquired MR raw data of a static phantom, acquisition with different numbers of translational motion states was simulated by applying a phase factor to the acquired raw k-space data. Based on the Fourier shift theorem, a linear phase shift in the frequency space can describe a translational shift of the object in the image space. The shift amplitudes had a maximum of 20 pixels (i.e. 30 mm) and were linearly spaced along the right-left axis, as shown in Figure 1. We investigated three cases where the number of motion states was 6, 30 and 60 (which we will refer to as 6MS, 30MS, and 60MS). Ground-truth motion vector fields were created based on the known shifts and used in the subsequent MCIR image reconstruction.

### *3.3. Motion estimation in vivo*

For the GRPE trajectory used in this study, the k-space centre was acquired repeatedly. From the k-space centre, a self-navigator was calculated as a respiratory surrogate. External ECG signals recorded during the acquisition of MR data were used as cardiac surrogates. Based on these surrogates, the acquired data was separated into 6 respiratory and 12 cardiac motion states. The way this method separates the data is such that each subset is of equal length in time, independently of any heart rate or breathing irregularities. TV-regularised iterative image reconstruction was used to reconstruct

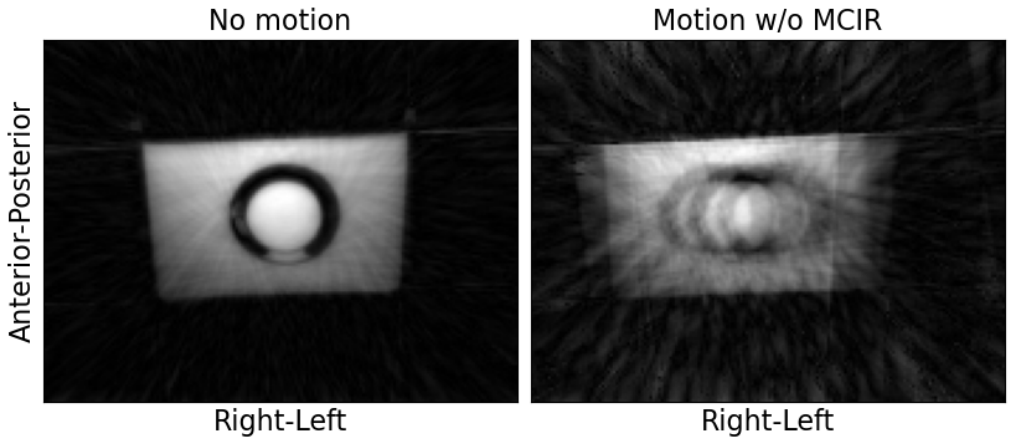


Figure 2: To illustrate the effect of motion, we show a reconstruction obtained without simulating motion (on the left) and a reconstruction obtained using 3 simulated motion states, without motion compensation (on the right).

images of the different motion states. For cardiac-resolved images, respiratory motion correction was applied. TV regularisation was carried out in the spatial image domain and along the cardiac/respiratory motion dimension. A non-rigid spline-based image registration algorithm was used to obtain 3D motion vector fields from these motion-resolved images for both cardiac and respiratory motion. As a final step, the cardiac and respiratory motion fields were combined into 72 cardio-respiratory motion vector fields. More details on the motion estimation can be found in Mayer *et al.* [18]. Animations showing the respiratory and cardiac motion fields are included in the supplementary material, as figure 1 and 2, respectively.

### *3.4. Evaluation*

We compare the convergence of SPDHG against the deterministic algorithms PDHG and FISTA to optimise equation (3).

For the phantom data, we investigated the three cases 6MS, 30MS and 60MS. For in vivo data, image reconstructions were performed with respiratory motion correction (6 motion states) and with cardiorespiratory motion correction (72 motion states). All image reconstructions were run for a large number of epochs to ensure that each algorithm had converged.

To compare the convergence rate between SPDHG and PDHG, we use the normalised root-mean-squared-error (NRMSE), defined as:

$$NRMSE = \frac{\mathrm{RMSE}(U_i, \overline{U})}{\mathrm{RMSE}(\overline{U}, 0)} \tag{14}$$

where $U_i$ is the reconstruction at epoch $i$ and $\overline{U}$ is the reference image. For the phantom data, the reference was obtained by running PDHG for 200 epochs on the

data with the relevant number of motion states. For in vivo data, the reference image was reconstructed with PDHG for 800 epochs.

We define the point of convergence as the epoch in which the NRMSE fell below 5%. This value was chosen based on visual inspection of the reconstructed images. Different thresholds are expected to lead to a different number of epochs required to meet the point of convergence; however, the relative difference in epochs to convergence between PDHG and SPDHG is expected to remain similar.

### *3.5. Algorithm Parameter Choices*

A step-size strategy was chosen for fair comparison, taking into account the functions and operators in the optimisation objective and ensuring mathematical convergence guarantees are met. Additionally, again for comparison purposes, we choose a step-size strategy for PDHG and SPDHG to ensure that, in the limit where $M = 1$, the step size for PDHG and SPDHG would be equal. Further adjustments to these values may enhance convergence rates and will be discussed in the conclusions.

For PDHG, we take the dual and primal step sizes $\sigma$ and $\tau$ to be equal to $\frac{1}{\|K\|}$, and choose the relaxation parameter $\theta = 1$. This is chosen to be as large as possible while ensuring that $\sigma\tau\|K\|^2 < 1$ and makes a default assumption that we want similar-sized steps in the primal and dual space. Similarly, for SPDHG, we take the relaxation parameter $\theta = 1$, the dual step-size $\sigma_i = \frac{1}{\|K_i\|}$, and the primal step-size $\tau = \min_i \left(\frac{p_i}{\|K_i\|}\right)$. This satisfies equation (11) and again encourages similar-sized steps in the primal and dual space. We take uniform probabilities, $p_i = \frac{1}{M}$ and the indices are selected randomly with replacement. Each motion state contains data from an equal length time period and therefore we weight the contribution from each motion state equally. For FISTA, we consider the same objective function as in equation (3). Given $f_{\text{FISTA}}(x) = \|Kx - b\|_2^2$ , we take the step size to be the reciprocal of the Lipschitz constant of $f_{\text{FISTA}}$, chosen as the largest step size that is provably convergent [19].

To calculate the proximal of the Total Variation function, we use the fast gradient projection algorithm implemented in the CCPi Regularisation Toolkit (version 22.0.0) [40].

## 4. Reproducibility

The results from the phantom data can be reproduced using the notebook `https://github.com/paskino/SIRF-Contribs/blob/MR_MCIR_2023/src/notebooks/implicit_TV_simulated_motion.ipynb`. This notebook uses SIRF version 3.6.0 [41], CIL version 23.1.0 [42] and CCPi Regularisation Toolkit version 22.0.0. The notebook is also compatible with more recent SIRF, CIL and the CCPi Regularisation toolkit versions: 3.8.0, 24.2.0 and 24.0.1 respectively, which are available at `https://github.com/SynerBI/SIRF`, `https://github.com/TomographicImaging/CIL` and `https://github.com/TomographicImaging/CCPi-Regularisation-Toolkit`. The dataset of

the phantom is available on Zenodo at `https://zenodo.org/record/7903282`.

## 5. Results

### *5.1. Phantom*

In Figure 3, we plot the objective value against epochs for 6MS, 30MS, and 60MS for FISTA, PDHG and SPDHG on the simulated motion phantom data. We see that SPDHG exhibits the fastest convergence compared to the deterministic algorithms. In addition, this convergence speed increases with a larger number of motion states. Note that a unique minimiser exists as the objective we are minimising is proper and strictly convex. Therefore, we expect all algorithms to converge to the same point, we are just comparing the speed at which they converge.

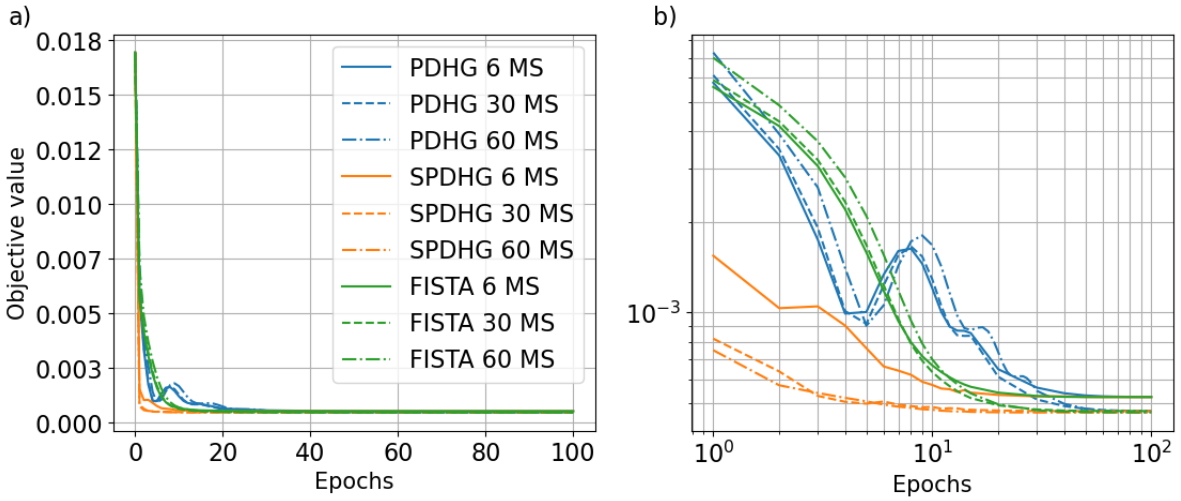


Figure 3: Objective value for 6, 30 and 60 motion states (MS) for the FISTA (green), PDHG (blue) and SPDHG (orange) algorithms presented in linear (a) and log-log (b) scale.

In Figure 4, we plot the NRMSE as defined in equation (14), for the 60MS case. A decreasing NRMSE means that the solution obtained by the algorithm is becoming more and more similar to the reference one. As can be seen in Figure 4, SPDHG is the fastest algorithm to reach the arbitrary convergence threshold of 0.05, which requires 10 epochs, while FISTA and PDHG require 60 and 70 epochs, respectively. In Figure 5, in the first and third row, we plot reconstructions of the phantom data in the 60MS case at 10 and 70 epochs respectively, while in the second and fourth row, we present error maps between each reconstruction and the reference image, which is shown in the top right corner. The error maps are obtained as the absolute value of the difference between the reference image and each reconstruction, normalised by the maximum value in the reference image. For each reconstruction, we also show the NRMSE with respect to the reference image, as defined in equation (14). From the figure, we can see that the

quality of the SPDHG reconstruction at 10 epochs is already similar to that achieved at 70 epochs. On the other hand, FISTA and PDHG at 10 epochs have clearly not converged yet. Only by 70 epochs do the deterministic algorithms achieve good image quality.

From a breakdown of the wall clock time per epoch for the 6MS to 60MS cases, we observed that the time per epoch in PDHG and SPDHG was dominated by the operators $K_i$ in equation (13), whose direct and adjoint evaluations are required for each iteration. The time taken for the computation of the direct and adjoint of each motion state is about 10 seconds. This needs to be multiplied by the total number of motion states to obtain the time taken by the operator computation per epoch. We measured that approximately 90% of the epoch time is taken by these operators in the case of PDHG, while this percentage is reduced to about 75% for SPDHG, confirming the effect of reduced computational effort in the evaluation of the operator $K$ for the latter.

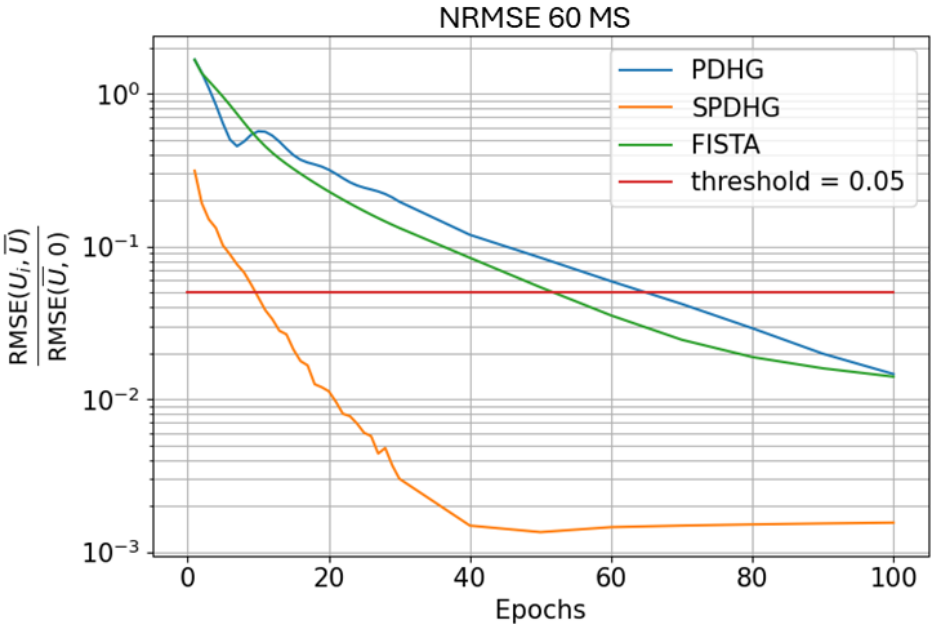


Figure 4: The NRMSE as computed in equation (14) for PDHG, SPDHG and FISTA. Lower values indicate solutions that are closer to the reference solution $\overline{U}$. SPDHG reaches a threshold of 0.05 at epoch 10, while PDHG does so at epoch 70 and FISTA at epoch 60.

The measured wall clock time per epoch of SPDHG is longer than that of PDHG, by approximately 15% for the 6MS example, 25% for 30MS and 35% for 60MS in our implementation. However, in Figure 4 it can be seen that, in the 60MS case, SPDHG takes only about 20% of the epochs required by PDHG and FISTA to satisfy the convergence criterion outlined in Section 3.4 (through equation (14)).

Therefore, SPDHG converges faster to the solution also in terms of wall clock time

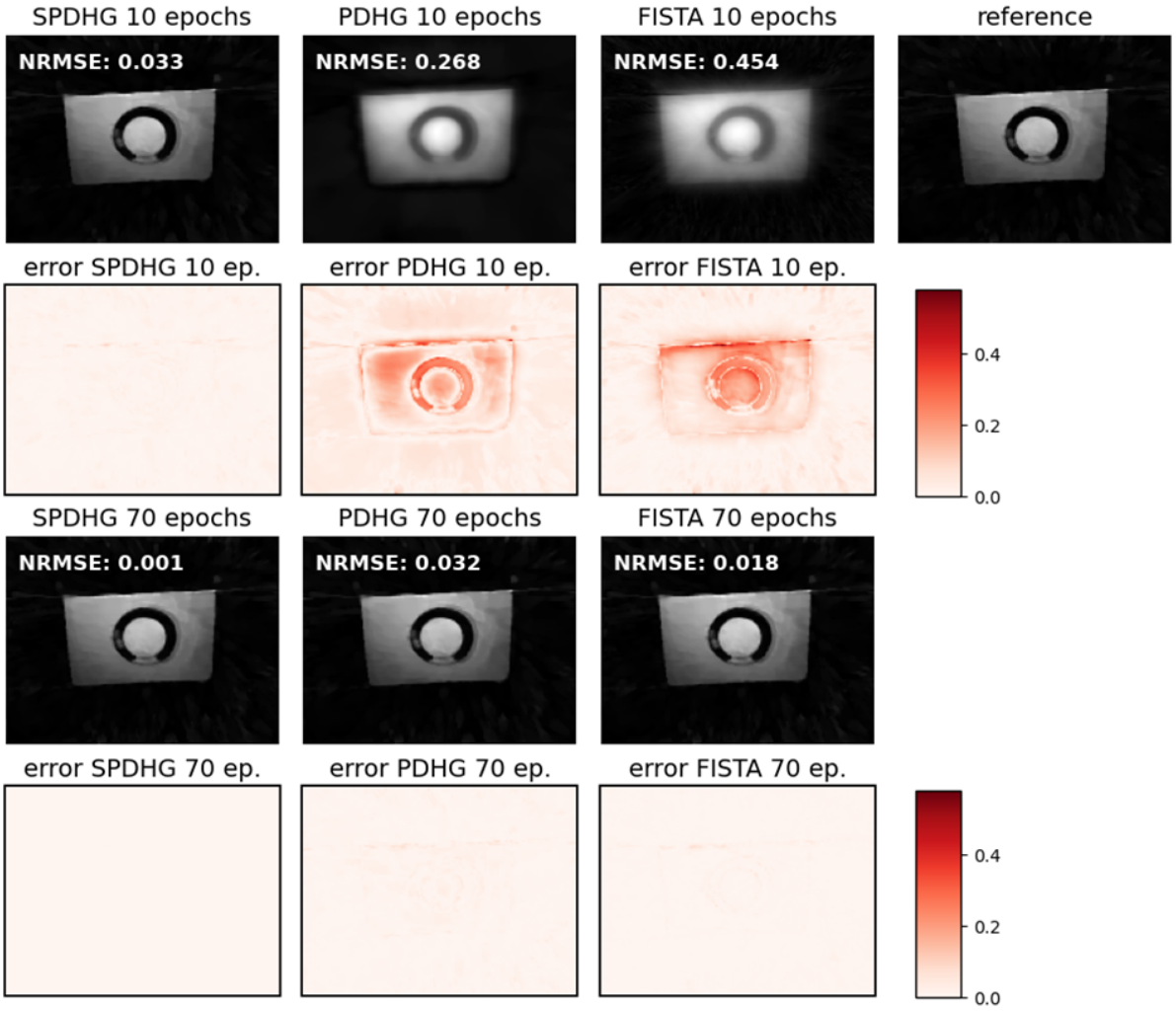


Figure 5: Phantom data reconstructions for FISTA, PDHG and SPDHG for 60MS displayed on the axial view, obtained at 10 epochs (first row) and 70 epochs (third row), as well as error maps (second and fourth row) showing the absolute value of the normalised difference between each reconstruction and the reference image (shown in the top right corner). The plots demonstrate that at 10 epochs SPDHG is already close to convergence while the deterministic algorithms have not yet converged. Only by 70 epochs do FISTA and PDHG achieve a good reconstruction quality.

because it requires much fewer epochs than PDHG. To show this, in Table 2 we report the reconstruction times (times for the algorithms to reach the NRMSE threshold in Figure 4) for PDHG and SPDHG for 6, 30, and 60 motion states. For each case, we also show the SPDHG speedup, which was obtained as a ratio between PDHG time and SPDHG time, and therefore refers to the acceleration allowed by SPDHG with respect to PDHG. These reconstruction times were obtained by running the experiments on a virtual machine from the STFC (Science and Technology Facilities Council) cloud with 28 virtual CPUs, 180 GB RAM, and an NVidia RTX-4000 used only for the evaluation

of the total variation proximal.

| | 6MS | 30MS | 60MS |
|---|---|---|---|
| PDHG time | 1h 10min | 5h 29min | 10h 50min |
| SPDHG time | 42min | 1h 22min | 2h 27min |
| SPDHG speedup | 1.65 | 3.98 | 4.43 |

Table 2: Reconstruction times (times for the algorithms to reach the NRMSE threshold) for PDHG and SPDHG for 6, 30, and 60 motion states, for the phantom data. The SPDHG speedup is obtained as a ratio between PDHG time and SPDHG time and therefore refers to the acceleration allowed by SPDHG with respect to PDHG.

### *5.2. In vivo data*

The points of convergence for SPDHG were 80 epochs and 20 epochs for respiratory and cardio-respiratory motion correction, respectively. For PDHG the point of convergence was reached after 200 epochs for both types of motion correction. As expected, the convergence rate of SPDHG increases with the number of motion states: when 72 motion states are used (cardio-respiratory MCIR), SPDHG converges in fewer epochs, as compared to when 6 motion states are used (respiratory MCIR).

Figure 6 shows the results of PDHG-MCIR and SPDHG-MCIR using only respiratory motion correction. Respiratory motion correction reduces blurring, especially in the lower parts of the heart and abdomen, because the motion amplitudes are highest in these areas. Even for 80 epochs, small details, such as the coronary arteries, are visible in SPDHG-MCIR. Small details are not yet visible in PDHG-MCIR for the same number of epochs because the algorithm has not yet converged to a satisfactory solution. Figure 7 shows the corresponding convergence curves.

The results for cardio-respiratory MCIR with 72 motion states are shown in Figure 8 in addition to the corresponding convergence curves in Figure 9. The convergence behaviour of PDHG-MCIR does not depend on the number of motion states and is comparable to the respiratory MCIR case shown in Figure 6. Again, we see that SPDHG-MCIR shows faster convergence as a result of the higher number of motion states compared to respiratory MCIR. Even for 20 epochs, fine details such as the coronary arteries are clearly visible.

## 6. Discussion and Future Work

This study demonstrated that SPDHG-MCIR can converge significantly faster than PDHG-MCIR, especially for a large number of motion states. SPDHG-MCIR was applied for the correction of cardio-respiratory motion in a 3D whole-heart magnetic resonance scan; however, the proposed approach can be used for any other MCIR

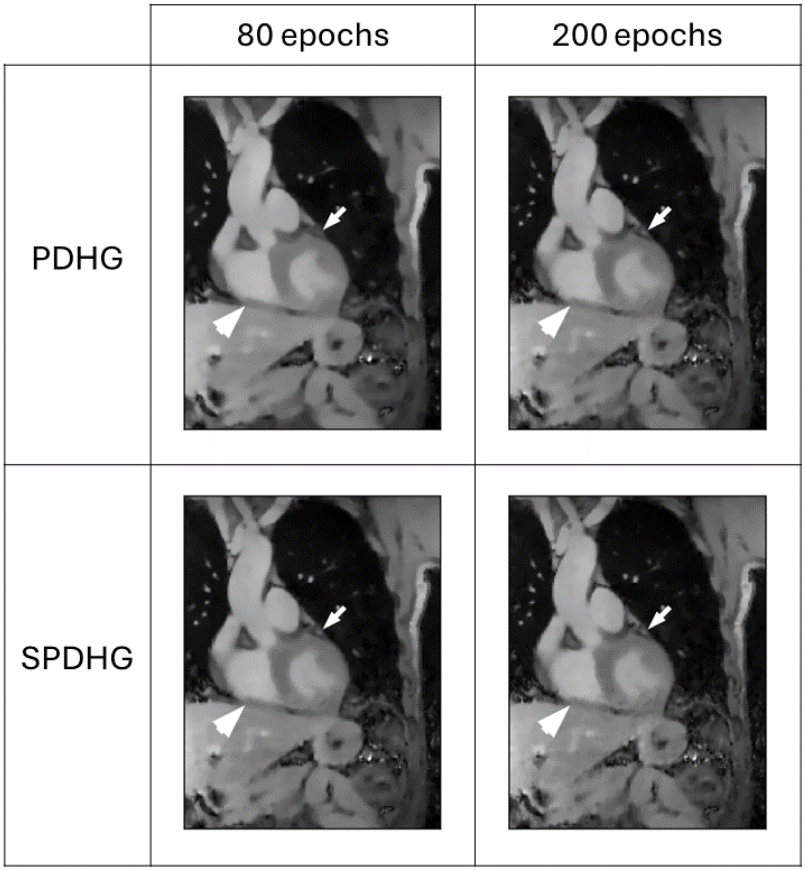


Figure 6: Respiratory MCIR. Reconstructed images using PDHG-MCIR and SPDHG-MCIR for 80 and 200 epochs are displayed on the coronal view. Respiratory MCIR reduces blurring visible, for example, at the liver-heart interface (arrow head). In contrast to PDHG-MCIR, for SPDHG-MCIR fine details such as the left coronary artery (arrow) are already clearly visible at 80 epochs.

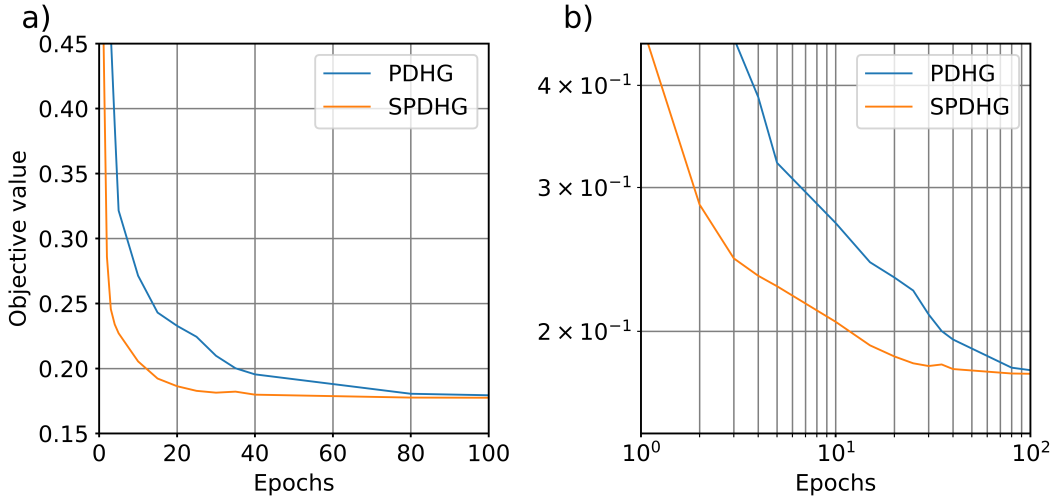


Figure 7: Convergence of respiratory MCIR in linear scale (a) and log-log scale (b).

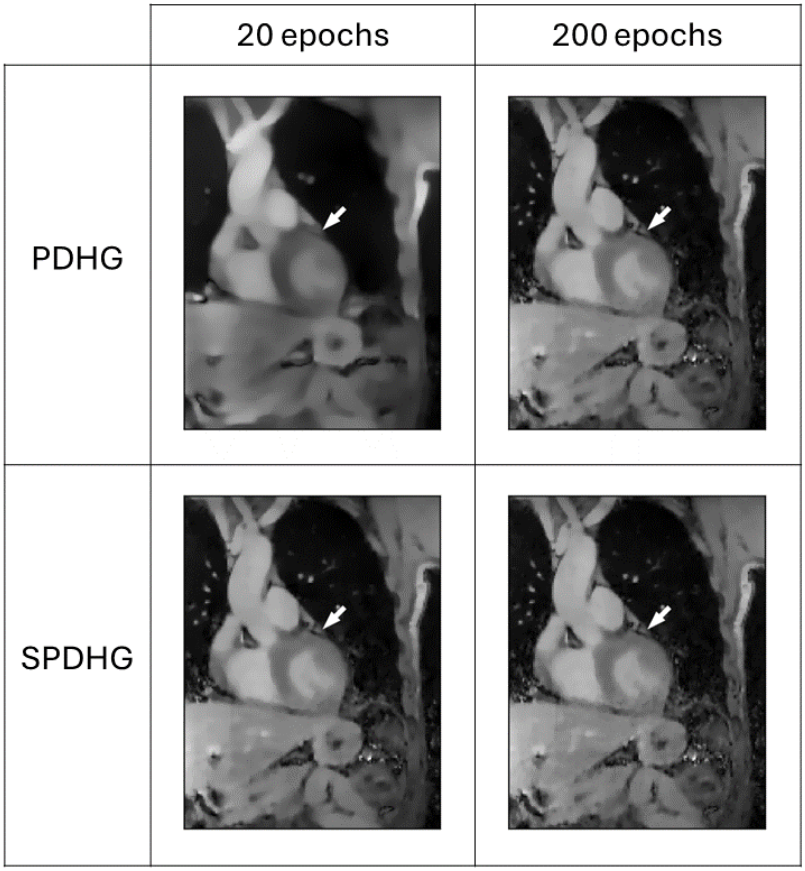


Figure 8: Cardio-respiratory MCIR. Reconstructed images using PDHG-MCIR and SPDHG-MCIR for 20 and 200 epochs are displayed on the coronal view. Cardiorespiratory MCIR improves image quality and reduces motion artefacts. Even for 20 epochs, small details, such as the left coronary artery (arrow), are already clearly visible for SPDHG-MCIR.

application. Typically, the number of motion states is optimised to balance accurate motion correction, ensuring high image quality and reconstruction time, as in [22], where the authors explain how 6 is the optimal number of motion states for respiratory motion correction on their data. With the proposed approach, this is no longer necessary, as increasing the number of motion states reduces the number of epochs required to reconstruct the image.

To exploit the faster convergence, one needs a reliable stopping criterion. Throughout this paper, we have compared reconstructions with a reference image and, in Figure 4, we used the NRMSE with respect to this reference to suggest a convergence threshold. In practice, a reference image will not be available. Further work could investigate stopping criteria or consider whether a fixed number of epochs could provide good-quality images for different acquisitions using the same MR protocol.

One limitation of this work is that we only evaluated the approach on a single in vivo

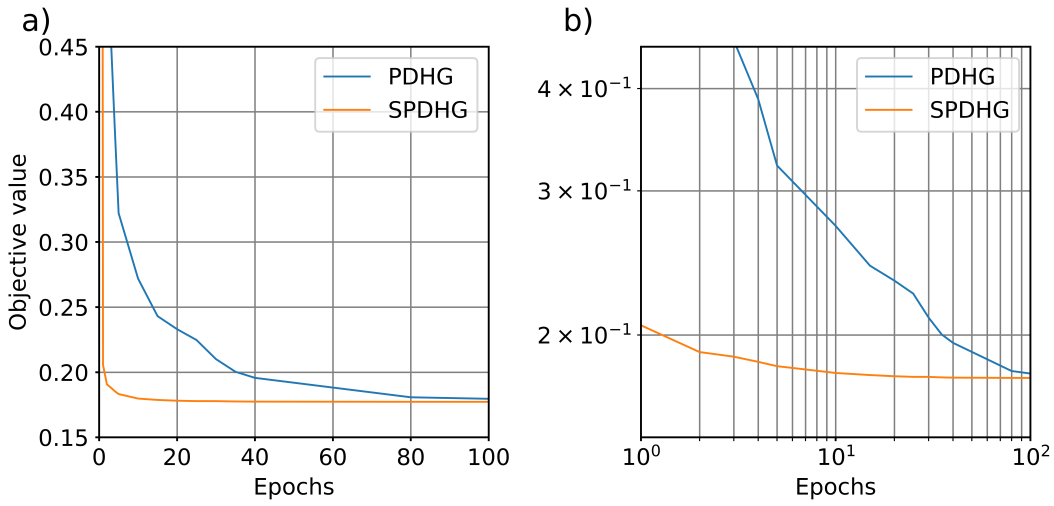


Figure 9: Convergence of cardio-respiratory MCIR in linear scale (a) and log-log scale (b).

dataset. Nevertheless, we used the same settings for SPDHG-MCIR for the phantom and the in vivo datasets and achieved excellent performance in both cases. As the image content, type of motion, matrix size and scan duration was very different between these two datasets, we would expect that SPDHG-MCIR performs well for different patients.

While the SPDHG-MCIR provides a dramatic reduction in reconstruction time, our current implementation needs further reductions for clinical application. An initial investigation indicated that the computational time is dominated by the NUFFT and non-rigid transformations calculations, as explained in section 5.1. This is consistent with the findings by Knoll *et al.* In [43], where they use PDHG for TGV-regularised brain MRI and explain how the NUFFT constitutes most of the computational effort in their experiments, while the time taken by the regularisation is minimal in comparison. By investing in software engineering efforts to speed up forward operator calculations, for instance, using GPUs, we would hope to see clinically useful reconstruction times. In the literature, [44] and [45] see up to 10x acceleration for niftyreg and NUFFT transformations. We expect similar acceleration to be achievable for the motion compensation.

In this study, we used a step-size strategy for the stochastic and deterministic algorithms, which prioritised fair comparisons and mathematical convergence guarantees, taking into account the functions and operators in the optimisation objective. We experimented with two very different datasets, and in both cases, these settings saw stochastic methods outperforming deterministic methods. We expect that with further experimentation, the parameters for both deterministic and stochastic methods could be fine-tuned to increase convergence speed, but this was out of the scope of this study. Future work could consider the use of grid search or bayesian methods [46] for identifying better algorithms' parameters, or the use of adaptive step-size methods,

as in [31].

Due to the way the in vivo data was obtained, each motion state was of equal length in time and so, for the stochastic methods, we chose equal probability for all motion states. Future work could also consider different methods for partitioning into motion states, such as considering the same magnitude of change in each motion state, or potentially different choices of probabilities. Again, we would still expect the stochastic methods to outperform the deterministic methods.

A common approach to accelerate primal-dual methods is preconditioning, in particular diagonal preconditioning [47, 48], which can significantly improve convergence without affecting computational complexity. Although future work could investigate the use of preconditioning to speed up computations, it would improve convergence of all algorithms, and therefore we would still expect SPDHG to converge faster than the deterministic methods.

Algorithmic improvements could lead to additional speed-up. For example, when iteratively calculating the TV proximal, the number of required iterations could be reduced by initialising with the result from the previous call to the TV proximal. We could also consider only calling the TV proximal once per epoch, as suggested in [49]. To further reduce the number of required iterations, future work could also focus on different sampling probabilities, $p_i$ [28]. While there are other stochastic algorithms in the literature, for example, see the review paper in [25], here we considered only PDHG and SPDHG, because they can be fairly compared. However, we expect similar speed-up results with other stochastic algorithms.

## 7. Conclusion

In this study, we presented a stochastic primal-dual hybrid-gradient (SPDHG) approach for motion-corrected image reconstruction (MCIR) in cardiac MRI. We demonstrated its superior convergence rate compared to traditional deterministic methods, such as PDHG and FISTA, using both phantom and in vivo data. Our results show that SPDHG significantly reduces reconstruction times while maintaining high image quality, which is essential for clinical applications where speed and accuracy are critical. Future work will explore further optimisations of the algorithm, including adaptive step size strategies and parallelisation techniques to further reduce reconstruction times.

**Acknowledgements**

This work was funded by the UK EPSRC grants the "Computational Collaborative Project in Synergistic Reconstruction for Biomedical Imaging" (CCP SyneRBI) EP/T026693/1; "Collaborative Computational Project in tomographic imaging" (CCPi) EP/M022498/1 and EP/T026677/1. This work used computational support from CoSeC, the Computational Science Centre for Research Communities, through CCP SyneRBI and CCPi. This work was supported by the Ada Lovelace Centre. The authors acknowledge funding from the German Research Society (GRK 2260, BIOQIC). This work used computational resources from STFC Scientific Computing Department's Cloud Operations Group.

We acknowledge helpful discussions on this project with Matthias J. Ehrhardt at the University of Bath and Evangelos Papoutsellis at Finden Ltd.



## References

[1] Quinn Counseller and Yasser Aboelkassem. "Recent technologies in cardiac imaging". In: *Frontiers in Medical Technology* 4 (2023). DOI: 10.3389/fmedt.2022.984492.

[2] Emily Aherne, Kelvin Chow, and James Carr. "Cardiac T1 mapping: Techniques and applications". In: *Journal of Magnetic Resonance Imaging* 51.5 (2020), pp. 1336–1356. DOI: 10.1002/jmri.26866.

[3] Andreas Seraphim, Kristopher D. Knott, Joao Augusto, Anish N. Bhuva, Charlotte Manisty, and James C. Moon. "Quantitative cardiac MRI". In: *Journal of Magnetic Resonance Imaging* 51.3 (2020), pp. 693–711. DOI: 10.1002/jmri.26789.

[4] Karen G Ordovas and Charles B Higgins. "Delayed contrast enhancement on MR images of myocardium: past, present, future." In: *Radiology* 261.2 (2011), pp. 358–374. DOI: 10.1148/radiol.11091882.

[5] Bharath Ambale-Venkatesh and João a. C. Lima. "Cardiac MRI: a central prognostic tool in myocardial fibrosis". In: *Nature Reviews Cardiology* 12.1 (2014), pp. 18–29. DOI: 10.1038/nrcardio.2014.159.

[6] Pedro F. Ferreira, Peter D. Gatehouse, Raad H. Mohiaddin, and David N. Firmin. "Cardiovascular magnetic resonance artefacts". In: *Journal of Cardiovascular Magnetic Resonance* 15.1 (2013), pp. 1–39. DOI: 10.1186/1532-429X-15-41.

[7] Andrew D. Scott, Jennifer Keegan, and David N. Firmin. "Motion in cardiovascular MR imaging." In: *Radiology* 250.2 (2009), pp. 331–351. DOI: 10.1148/radiol.2502071998.

[8] Tevfik F. Ismail, Wendy Strugnell, Chiara Coletti, Masa Bozic-Iven, Sebastian Weingaertner, Kerstin Hammernik, Teresa Correia, and Thomas Kuestner. "Cardiac MR: From Theory to Practice". In: *Frontiers in Cardiovascular Medicine* 9 (2022). DOI: 10.3389/fcvm.2022.826283.

[9] Christoph Kolbitsch, Claudia Prieto, Jouke Smink, and Tobias Schaeffter. "Highly Efficient Whole-heart Imaging using Radial Phase Encoding - Phase Ordering with Automatic Window Selection". In: *Magnetic Resonance in Medicine* 66.4 (2011), pp. 1008–1018. DOI: 10.1002/mrm.22888.

[10] Philip G Batchelor, David Atkinson, P Irarrazaval, D L G Hill, J Hajnal, and D Larkman. "Matrix description of general motion correction applied to multishot images". In: *Magnetic Resonance in Medicine* 54.5 (2005), pp. 1273–1280. DOI: 10.1002/mrm.20656.

[11] Irene Polycarpou, Charalampos Tsoumpas, Andrew P. King, and Paul K. Marsden. "Quantitative Evaluation of PET Respiratory Motion Correction Using MR Derived Simulated Data". In: *IEEE Transactions on Nuclear Science* 62.6 (2015), pp. 3110–3116. DOI: 10.1109/TNS.2015.2494593.

[12] Johannes F. M. Schmidt, Martin Buehrer, Peter Boesiger, and Sebastian Kozerke. "Nonrigid retrospective respiratory motion correction in whole-heart coronary MRA". In: *Magnetic Resonance in Medicine* 66.6 (2011), pp. 1541–1549. DOI: 10.1002/mrm.22939.

[13] Holden H. Wu, Paul T. Gurney, Bob S. Hu, Dwight G. Nishimura, and Michael V. McConnell. "Free-breathing multiphase whole-heart coronary MR angiography using image-based navigators and three-dimensional cones imaging". In: *Magnetic Resonance in Medicine* 69.4 (2013), pp. 1083–1093. DOI: 10.1002/mrm.24346.

[14] Gastao Cruz, David Atkinson, Markus Henningsson, Rene M. Botnar, and Claudia Prieto. "Highly efficient nonrigid motion-corrected 3D whole-heart coronary vessel wall imaging". In: *Magnetic Resonance in Medicine* 77.5 (2017), pp. 1894–1908. DOI: 10.1002/mrm.26274.

[15] Christoph Kolbitsch, Mark A. Ahlman, Cynthia Davies-Venn, Robert Evers, Michael Hansen, Devis Peressutti, Paul Marsden, Peter Kellman, David A. Bluemke, and Tobias Schaeffter. "Cardiac and Respiratory Motion Correction for Simultaneous Cardiac PET/MR". In: *Journal of Nuclear Medicine* 58.5 (2017), pp. 846–852. DOI: 10.2967/jnumed.115.171728.

[16] Aurélien Bustin, Giulia Ginami, Gastão Cruz, Teresa Correia, Tevfik F. Ismail, Imran Rashid, Radhouene Neji, René M. Botnar, and Claudia Prieto. "Five-minute whole-heart coronary MRA with sub-millimeter isotropic resolution, 100% respiratory scan efficiency, and 3D-PROST reconstruction". In: *Magnetic Resonance in Medicine* January (2018), pp. 1–14. DOI: 10.1002/mrm.27354.

[17] Andrew Phair, Gastão Cruz, Haikun Qi, René M. Botnar, and Claudia Prieto. "Free-running 3D whole-heart T1 and T2 mapping and cine MRI using low-rank reconstruction with non-rigid cardiac motion correction". In: *Magnetic Resonance in Medicine* 89.1 (2023), pp. 217–232. DOI: 10.1002/mrm.29449.

[18] Johannes Mayer, Edyta Blaszczyk, Alberto Cipriani, Giulio Ferrazzi, Jeanette Schulz-Menger, Tobias Schaeffter, and Christoph Kolbitsch. "Cardio-respiratory motion-corrected 3D cardiac water-fat MRI using model-based image reconstruction". In: *Magnetic Resonance in Medicine* 88.4 (2022), pp. 1561–1574. DOI: 10.1002/mrm.29284.

[19] Amir Beck and Marc Teboulle. "A fast iterative shrinkage-thresholding algorithm for linear inverse problems". In: *SIAM journal on imaging sciences* 2.1 (2009), pp. 183–202.

[20] Antonin Chambolle and Thomas Pock. "A first-order primal-dual algorithm for convex problems with applications to imaging". In: *Journal of mathematical imaging and vision* 40 (2011), pp. 120–145.

[21] Richard Brown et al. "Motion estimation and correction for simultaneous PET/MR using SIRF and CIL". In: *Phil. Trans. R. Soc. A 20200208* (2021). DOI: 10.1098/rsta.2020.0208.

[22] Xucheng Zhu, Marilynn Chan, Michael Lustig, Kevin Johnson, and Peder Larson. "Iterative motion-compensation reconstruction ultra-short TE (iMoCo UTE) for high-resolution free-breathing pulmonary MRI". In: *Magnetic Resonance in Imaging* (2019). DOI: 10.1002/mrm.27998.

[23] Matthias J. Ehrhardt, Ferdia A. Gallagher, Mary A. McLean, and Carola-Bibiane Schönlieb. "Enhancing the spatial resolution of hyperpolarized carbon-13 MRI of human brain metabolism using structureguidance". In: *Magnetic Resonance in Medicine* 87 (3 2021), pp. 1301–1312. DOI: 10.1002/mrm.29045.

[24] Florian Knoll, Martin Holler, Thomas Koesters, Ricardo Otazo, Kristian Bredies, and Daniel Sodickson. "Joint MR-PET reconstruction using a multi-channel image regularizer". In: *IEEE Trans Med Imaging* 36 (2017). DOI: 10.1109/TMI.2016.2564989.

[25] Matthias J. Ehrhardt, Zeljko Kereta, Jingwei Liang, and Junqi Tang. *A Guide to Stochastic Optimisation for Large-Scale Inverse Problems*. 2024. arXiv: 2406.06342 [math.NA].

[26] N Dikaios and T D Fryer. "Acceleration of motion-compensated PET reconstruction: ordered subsets-gates EM algorithms and a priori reference gate information". In: *Physics in Medicine and Biology* 56.6 (2011), p. 1695. DOI: 10.1088/0031-9155/56/6/011.

[27] Claire Delplancke, Matthias J. Ehrhardt, and Kris Thielemans. "Accelerated Convergent Motion Compensated Image Reconstruction". In: *2021 IEEE Medical Imaging Conference (MIC)*. 2021. DOI: https://petpp.github.io/data/slides/2021_MIC_Delplancke.pdf.

[28] Antonin Chambolle, Matthias J. Ehrhardt, Peter Richtarik, and Carola-Bibiane Schoenlieb. "Stochastic primal-dual hybrid gradient algorithm with arbitrary sampling and imaging applications". In: *SIAM Journal on Optimization* 28.4 (2018), pp. 2783–2808. DOI: 10.1137/17M1134834.

[29] Matthias J. Ehrhardt, Pawel Markiewicz, Antonin Chambolle, Peter Richtarik, Jonathan Schott, and Carola-Bibiane Schoenlieb. "Faster PET reconstruction with a stochastic primal-dual hybrid gradient method". In: *SPIE proceedings* (2017). DOI: 10.1117/12.2272946.

[30] Eric B Gutiérrez, Claire Delplancke, and Matthias J Ehrhardt. "Convergence properties of a randomized primal-dual algorithm with applications to parallel mri". In: *International Conference on Scale Space and Variational Methods in Computer Vision*. Springer. 2021, pp. 254–266.

[31] Antonin Chambolle, Claire Delplancke, Matthias J Ehrhardt, Carola-Bibiane Schönlieb, and Junqi Tang. "Stochastic primal–dual hybrid gradient algorithm with adaptive step sizes". In: *Journal of Mathematical Imaging and Vision* 66.3 (2024), pp. 294–313.

[32] Stanley Osher, Martin Burger, Donald Goldfarb, Jinjun Xu, and Wotao Yin. "An iterative regularization method for total variation-based image restoration". In: *Multiscale Modeling & Simulation* 4.2 (2005), pp. 460–489.

[33] Antonin Chambolle and Thomas Pock. "An introduction to continuous optimization for imaging". In: *Acta Numerica* 25 (2016), 161–319. DOI: 10.1017/S096249291600009X.

[34] Claudia Prieto, Sergio Uribe, Reza Razavi, David Atkinson, and Tobias Schaeffter. "3D undersampled golden-radial phase encoding for DCE-MRA using inherently regularized iterative SENSE". In: *Magnetic Resonance in Medicine* 64 (2010), pp. 514–526. DOI: 10.1002/mrm.22446.

[35] J D O'Sullivan and J D O Sullivan. "A fast sinc function gridding algorithm for fourier inversion in computer tomography." In: *IEEE Trans Med Imaging* 4.4 (1985), pp. 200–207. DOI: 10.1109/TMI.1985.4307723.

[36] Evgueni Ovtchinnikov et al. "SIRF: Synergistic Image Reconstruction Framework". In: *Computer Physics Communications* 249 (2020), p. 107087. DOI: https://doi.org/10.1016/j.cpc.2019.107087.

[37] Jakob S Jørgensen, Evelina Ametova, Genoveva Burca, Gemma Fardell, Evangelos Papoutsellis, Edoardo Pasca, Kris Thielemans, Martin Turner, Ryan Warr, William RB Lionheart, et al. "Core Imaging Library-Part I: a versatile Python framework for tomographic imaging". In: *Philosophical Transactions of the Royal Society A* 379.2204 (2021), p. 20200192. DOI: 10.1098/rsta.2020.0192.

[38] Evangelos Papoutsellis, Evelina Ametova, Claire Delplancke, Gemma Fardell, Jakob S Jørgensen, Edoardo Pasca, Martin Turner, Ryan Warr, William RB Lionheart, and Philip J Withers. "Core Imaging Library-Part II: multichannel reconstruction for dynamic and spectral tomography". In: *Philosophical Transactions of the Royal Society A* 379.2204 (2021), p. 20200193. DOI: 10.1098/rsta.2020.0193.

[39] Redha Boubertakh, C Prieto, Philip G Batchelor, S Uribe, David Atkinson, H Eggers, T S Sørensen, M S Hansen, R S Razavi, and T Schaeffter. "Whole-heart imaging using undersampled radial phase encoding (RPE) and iterative sensitivity encoding (SENSE) reconstruction." In: *Magnetic Resonance in Medicine* 62.5 (2009), pp. 1331–1337. DOI: 10.1002/mrm.22102.

[40] Daniil Kazantsev, Edoardo Pasca, Martin J. Turner, and Philip J. Withers. "CCPi-Regularisation toolkit for computed tomographic image reconstruction with proximal splitting algorithms". In: *SoftwareX* 9 (2019), pp. 317–323. DOI: https://doi.org/10.1016/j.softx.2019.04.003.

[41] Evgueni Ovtchinnikov et al. *SIRF Synergistic Image Reconstruction Framework*. Version v3.5.0. 2023. DOI: 10.5281/zenodo.2707911.

[42] Edoardo Pasca, Jakob Sauer Jørgensen, Evangelos Papoutsellis, Evelina Ametova, Gemma Fardell, Kris Thielemans, Laura Murgatroyd, Margaret Duff, Casper da Costa-Luis, and Hannah Robarts. *Core Imaging Library (CIL)*. Version v23.1.0. 2023. DOI: 10.5281/zenodo.8398330.

[43] Florian Knoll, Kristian Bredies, Thomas Pock, and Rudolf Stollberger. "Second Order Total Generalized Variation (TGV) for MRI". In: *Magnetic Resonance in Medicine* 65 (2 2011), pp. 480–491.

[44] Marc Modat, Gerard R. Ridgway, Zeike A. Taylor, Manja Lehmann, Josephine Barnes, David J. Hawkes, Nick C. Fox, and Sébastien Ourselin. "Fast free-form deformation using graphics processing units". In: *Computer Methods and Programs in Biomedicine* 98.3 (2010). HP-MICCAI 2008, pp. 278–284. DOI: https://doi.org/10.1016/j.cmpb.2009.09.002.

[45] Yu-hsuan Shih, Garrett Wright, Joakim Anden, Johannes Blaschke, and Alex H. Barnett. " cuFINUFFT: a load-balanced GPU library for general-purpose nonuniform FFTs ". In: *2021 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*. Los Alamitos, CA, USA: IEEE Computer Society, 2021, pp. 688–697. DOI: 10.1109/IPDPSW52791.2021.00105.

[46] Hussain Alibrahim and Simone A. Ludwig. "Hyperparameter Optimization: Comparing Genetic Algorithm against Grid Search and Bayesian Optimization". In: *2021 IEEE Congress on Evolutionary Computation (CEC)*. 2021. DOI: 10.1109/CEC45853.2021.9504761.

[47] Yu-Jung Tsai, Alexandre Bousse, Matthias J. Ehrhardt, Charles W. Stearns, Sangtae Ahn, Brian F. Hutton, Simon Arridge, and Kris Thielemans. "Fast Quasi-Newton Algorithms for Penalized Reconstruction in Emission Tomography and Further Improvements via Preconditioning". In: *IEEE Transactions on Medical Imaging* 37.4 (2018), pp. 1000–1010. DOI: 10.1109/TMI.2017.2786865.

[48] Thomas Pock and Antonin Chambolle. "Diagonal Preconditioning for First Order Primal-Dual Algorithms in Convex Optimization". In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 2011, pp. 1762–1769. DOI: 10.1109/ICCV.2011.6126441.

[49] Laurent Condat and Peter Richtárik. "RandProx: Primal-dual optimization algorithms with randomized proximal updates". In: *arXiv preprint arXiv:2207.12891* (2022).