

Robust residual-guided iterative reconstruction for sparse-view CT in small animal imaging

Zhang, Jianru; Wang, Zhe; Cao, Tuoyu; Cao, Guohua; Ren, Wuwei; Jiang, Jiahua

DOI:

[10.1088/1361-6560/ad360a](https://doi.org/10.1088/1361-6560/ad360a)

License:

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Zhang, J, Wang, Z, Cao, T, Cao, G, Ren, W & Jiang, J 2024, 'Robust residual-guided iterative reconstruction for sparse-view CT in small animal imaging', *Physics in Medicine & Biology*, vol. 69, no. 10, 105010. <https://doi.org/10.1088/1361-6560/ad360a>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.



PAPER

Robust residual-guided iterative reconstruction for sparse-view CT in small animal imaging

OPEN ACCESS

RECEIVED

15 March 2023

REVISED

27 February 2024

ACCEPTED FOR PUBLICATION

20 March 2024

PUBLISHED

29 April 2024

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Jianru Zhang^{1,5,6} , Zhe Wang^{2,6}, Tuoyu Cao³, Guohua Cao^{2,*}, Wuwei Ren^{1,*} and Jiahua Jiang^{4,5,*} ¹ School of Information Science and Technology, ShanghaiTech University, Shanghai, 201210, People's Republic of China² School of Biomedical Engineering & State Key Laboratory of Advanced Medical Materials and Devices, ShanghaiTech University, Shanghai, 201210, People's Republic of China³ United Imaging Healthcare Co., Ltd, Shanghai, 201807, People's Republic of China⁴ Institute of Mathematical Science, ShanghaiTech University, Shanghai, 201210, People's Republic of China⁵ School of Mathematics, University of Birmingham, Edgbaston, B15 2TT, United Kingdom⁶ These authors contributed equally to this work.

* Authors to whom any correspondence should be addressed.

E-mail: caogh@shanghaitech.edu.cn, renww@shanghaitech.edu.cn and j.jiang.3@bham.ac.uk

Keywords: sparse-view CT, image reconstruction, inverse problem, Golub–Kahan process, regularization

Abstract

Objective. We introduce a robust image reconstruction algorithm named residual-guided Golub–Kahan iterative reconstruction technique (RGIRT) designed for sparse-view computed tomography (CT), which aims at high-fidelity image reconstruction from a limited number of projection views. **Approach.** RGIRT utilizes an inner-outer dual iteration framework, with a flexible least square QR (FLSQR) algorithm implemented in the inner iteration and a restarted iterative scheme applied in the outer iteration. The inner FLSQR employs a flexible Golub–Kahan bidiagonalization method to reduce the size of the inverse problem, and a weighted generalized cross-validation method to adaptively estimate the regularization hyper-parameter. The inner iteration efficiently yields the intermediate reconstruction result, while the outer iteration minimizes the residual and refines the solution by using the result obtained from the inner iteration. **Main results.** The reconstruction performance of RGIRT is evaluated and compared to other reference methods (FBPConvNet, SART-TV, and FLSQR) using projection data from both numerical phantoms and real experimental Micro-CT data. The experimental findings, from testing various numbers of projection views and different noise levels, underscore the robustness of RGIRT. Meanwhile, theoretical analysis confirms the convergence of residual for our approach. **Significance.** We propose a robust iterative reconstruction algorithm for x-ray CT scans with sparse views, thereby shortening scanning time and mitigating excessive ionizing radiation exposure to small animals.

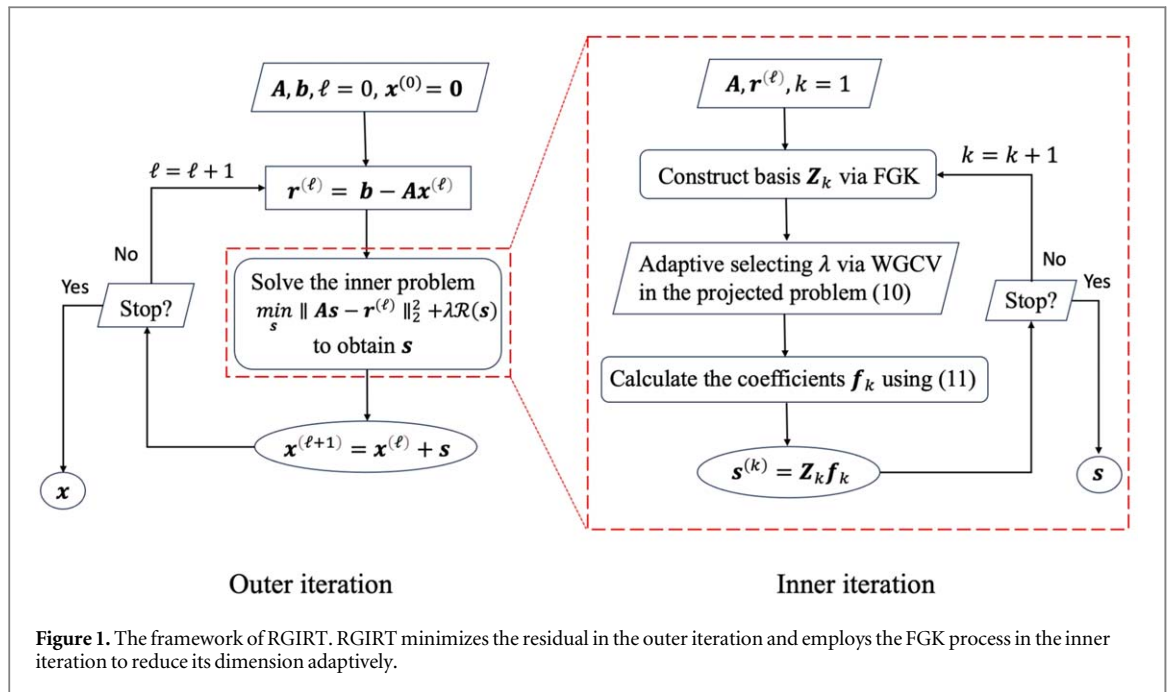
1. Introduction

X-ray computed tomography (CT) has become an indispensable tool in various fields, including industrial inspection, security checks, and medical diagnosis of a wide range of diseases (Wang *et al* 2008). However, traditional CT scanning requires a dataset from a large number of projection angles, which can expose patients to excessive ionizing radiation, increasing the risk of cancer (Brenner and Hall 2007). To lower the CT radiation dose in clinical practice, sparse-view CT scanning with a small number of projection views was proposed. Sparse-view CT has the significant advantage of reducing the radiation dose proportionally to the number of projection views. It also improves scanning speed and hence temporal resolution by shortening data acquisition time, which is particularly useful for some dynamic imaging tasks, such as cardiac micro-CT imaging for small animals (Kudo *et al* 2013).

Despite the potential benefits of sparse-view CT, one significant challenge is image reconstruction due to insufficient projection data. Reconstructing images from under-sampled projection data is an ill-posed inverse problem and imposes a more strict requirement on the robustness of the algorithm in contrast to a full-projection problem (Kim *et al* 2014). In general, CT reconstruction can be primarily classified into three categories: (1) direct or analytical reconstruction, (2) iterative reconstruction, and (3) deep-learning (DL)-based reconstruction (Wang *et al* 2020). The first type of direct reconstruction relies on the mathematical inverse of a forward model. Its embodiment in CT reconstruction is filtered back-projection (FBP) (Holschneider 1991). Despite its high efficiency, FBP may not constitute a proper method in a sparse-view problem, as it is vulnerable to appeared noise and produces highly degraded images with severe streaking artifacts (Singh *et al* 2010). Alternatively, iterative reconstruction combines a numerical forward model with a feedback loop, by progressively reducing the error between predicted sensor data and real measurement (Beister *et al* 2012). Examples include the Kaczmarz family of algorithms (e.g. algebraic reconstruction technique (ART) (Gordon *et al* 1970), simultaneous ART (SART) (Andersen and Kak 1984)), ordered-subsets expectation maximization (OSEM) (Shepp and Vardi 1982, Hudson and Larkin 1994), separable quadratic surrogates (SQS) (Kim *et al* 2013, 2014), penalized weighted least-squares (PWLS) (Fessler 1994, Niu *et al* 2014), and nonlocal means (NLM) algorithm (Kim *et al* 2016). Compared to direct reconstruction, iterative methods demonstrate significant improvement in reconstruction quality in reducing noise and artifacts. Nevertheless, these algorithms exhibit limited robustness regarding reconstruction with sparse-view projection data, attributed to the challenges of balancing data fidelity and regularization terms (e.g. Tikhonov regularization (ℓ_2) (Golub *et al* 1999), total variation (TV) regularization (Lu *et al* 2012, Liu *et al* 2013, Niu *et al* 2014, Kim *et al* 2016), or ℓ_1 regularization (Kim *et al* 2014, Beck and Teboulle 2009)), and, a lack of theoretical convergence analysis (Scherzer 1995, Elfving *et al* 2014, 2017, Magreñán and Argyros 2018). In particular, for some regularizations such as TV and ℓ_1 , it is mathematically challenging to show the theoretical convergence due to the non-smoothness and non-linearity of the regularization terms. More recently, DL-based reconstruction was conceived, showing promising results in sparse-view CT reconstruction (Lee *et al* 2018, Zhang *et al* 2018, 2020, Li *et al* 2022). However, DL models typically lack explainability and generalizability, which implies that these methods tend to be more susceptible to irregularities, biases, and noise inherent in the data, thereby compromising their robustness. Specifically, the effectiveness of supervised learning approaches is significantly influenced by the caliber of their training data, whereas the success of unsupervised learning techniques primarily relies on the convergence ability of the optimization solvers when minimizing their loss functions. Therefore, investigating robust methods capable of consistently producing accurate reconstruction results is of great significance in sparse-view CT reconstruction.

In this paper, we focus on improving the robustness of iterative methods and develop a new approach to tackle two primary challenges: meticulously fine-tuning the regularization parameter and the deficiency of the theoretical convergence analysis. To address the former issue, we consider integrating flexible least squares QR (FLSQR) (Chung and Gazzola 2019), an advanced inverse problem solver to adaptively choose appropriate regularization parameter in an iterative manner. The main ingredient of FLSQR is flexible Golub–Kahan (FGK) process (Chung and Gazzola 2019), an iterative dimension reduction technique which can progressively augments the solution space, thereby enabling the implementation of computationally demanding parameter selection methods (e.g. WGCV (Chung *et al* 2008), UPRE (Vogel 2002), DP (Colton *et al* 1997)). Although FLSQR shows promising simulation results in sparse-view CT, it still requires theoretical convergence guarantee and practical testing. To address the latter problem, we consider adopting an inner-outer scheme (Saad 1993, Eiermann *et al* 2000, Baker *et al* 2005, Ergül *et al* 2010) that uses the inner loop results to refine the solution at each outer loop iteration. This strategy can not only expedite convergence but also facilitate the convergence analysis.

Exploiting aspects of both FGK process and inner-outer framework, we propose residual-guided Golub–Kahan iterative reconstruction technique (RGIRT) for sparse-view CT, a more robust method that inherits the benefits of parameter selection in FLSQR, and ensures theoretical convergence of the residual. RGIRT utilizes a unique inner-outer iteration framework, in which FLSQR is implemented in the inner iteration, while a restarted iterative scheme is applied in the outer iteration. The outer iteration is capable of effectively minimizing the residual and refining the solution by integrating the intermediate results acquired from the inner iteration. By incorporating the residual, the theoretical convergence of the whole algorithm does not rely on the convergence of its inner solver, thus circumventing the major bottleneck of FLSQR which lacks a theoretical convergence proof. Based on the theoretical and experimental results, RGIRT offers two notable advantages compared with FLSQR and SART-TV (Lu *et al* 2012) which is a widely used TV-regularized CT reconstruction algorithm: first, RGIRT enhances the precision of the reconstruction under sparse-view conditions. It is important to note that setting the number of outer iterations to one in RGIRT results in a solution equivalent to FLSQR, representing the worst-case scenario for RGIRT, as no correction is implemented; second, RGIRT ensures the theoretical convergence of the residual and experimental convergence to the reference solution. It has a higher convergence rate and costs less memories. On the other hand, compared with FBPCnnNet (Jin *et al*



2017), a representative DL-based method, RGIRT exhibits more robust performance in comparative experiments, especially in increasingly sparse-view condition and escalating noise levels. From an algorithmic perspective of RGIRT, under the correction mechanism applied in the outer iteration, high accuracy is not required during the inner iteration. This increases the tolerance of the algorithm, thereby in turn enhancing its robustness.

2. Method and experiments

2.1. Residual-guided Golub–Kahan iterative reconstruction technique

2.1.1. Inner-outer iterative scheme

The ill-posed nature of sparse-view CT image reconstruction calls for the use of regularization to stabilize the inverse process. The mathematical model can be formulated as

$$\min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \lambda \mathcal{R}(\mathbf{x}), \quad (1)$$

where $\mathbf{x} \in R^N$ is the discrete image array; $\mathbf{b} \in R^M$ is the corresponding measurement data (sinogram) that carries noise; $\mathbf{A} \in R^{M \times N}$ is the system matrix that models the forward process; M is the number of source-detector pairs, N is the number of CT image grids. $\mathcal{R}(\mathbf{x})$ denotes the regularization term, and $\lambda (\lambda > 0)$ is the regularization parameter that balances the data fidelity term $\|\mathbf{Ax} - \mathbf{b}\|_2^2$ and the regularization term $\mathcal{R}(\mathbf{x})$. The primary framework of our method is based on inner-outer iterations (as shown in figure 1), where FLSQR is performed in the inner iterations to solve a subproblem until convergence, and the outer iteration refines the intermediate solution to the main problem (1) heuristically using the results from the inner iteration. Given an initial guess $\mathbf{x}^{(0)}$, we define the inner problem at the $(\ell + 1)$ th outer iteration as

$$\min_{\mathbf{s}} \|\mathbf{As} - \mathbf{r}^{(\ell)}\|_2^2 + \lambda \mathcal{R}(\mathbf{s}), \quad (2)$$

where $\mathbf{r}^{(\ell)} = \mathbf{b} - \mathbf{Ax}^{(\ell)}$ represents the residual at the ℓ th outer iteration and $\mathbf{x}^{(\ell)}$ is the approximated solution at the ℓ th outer iterative; \mathbf{s} represents the correction that will be used for updating $\mathbf{x}^{(\ell)}$ in the outer iteration. Subsequently, we obtain the $(\ell + 1)$ th outer iterative result $\mathbf{x}^{(\ell+1)} = \mathbf{x}^{(\ell)} + \mathbf{s}$. Although there are multiple choices for $\mathcal{R}(\mathbf{s})$, in this work we use ℓ_1 regularization, which is defined as $\mathcal{R}(\mathbf{s}) = \|\mathbf{s}\|_1$. Furthermore, we proved the convergence of residual for our proposed algorithm. We show that incorporating the correction in the outer iteration facilitates a gradual reduction of the residual, thereby mitigating the concern regarding semi-convergence. This can be expressed as

$$\|\mathbf{r}^{(\ell+1)}\|_2 \leq \|\mathbf{r}^{(\ell)}\|_2, \quad (3)$$

where the proof is provided in [appendix](#).

2.1.2. Optimization via iterative reweighted norms

Note that each inner problem of RGIRT can be regarded as a new minimization problem solved by FLSQR method. The implementation of FLSQR can be delineated into two steps: first, employing iterative reweighted norms (IRN) to break the original non-convex optimization problem into a sequence of weighted norm least-square problems; second, utilizing FGK process to perform iterative dimensionality reduction. IRN method has proven effective in solving inverse problems with ℓ_1 regularization constraint, by approximating ℓ_1 regularization with iteratively updated weight matrix (Gorodnitsky and Rao 1992, Rodriguez and Wohlberg 2008). The first step of IRN approach is to reduce ℓ_1 regularized problem to a sequence of least-squares problems involving a weighted ℓ_2 norm.

$$\min_{\mathbf{s}} \|\mathbf{A}\mathbf{s} - \mathbf{r}^{(\ell)}\|_2^2 + \lambda \|\mathbf{D}(\mathbf{s})\mathbf{s}\|_2^2, \quad (4)$$

where

$$\mathbf{D}(\mathbf{s}) = \text{diag} \left\{ [([s]_1)^2 + \tau]^{-\frac{1}{4}}, \dots, [([s]_N)^2 + \tau]^{-\frac{1}{4}} \right\} \in R^{N \times N}, \quad (5)$$

with $[s]_i$ denoting the i th entry of \mathbf{s} and a small threshold $\tau > 0$. To avoid nonlinearities of solving (4), we follow the common practice of approximating $\mathbf{D}(\mathbf{s})$ by the weighted matrix at the k th inner iteration $\mathbf{D}_k = \mathbf{D}(\mathbf{s}^{(k-1)})$, where $\mathbf{s}^{(k-1)}$ is an approximation of the solution at the $(k-1)$ th inner iteration. Notice that seeking $\mathbf{s}^{(k)}$ in (4) requires solving a large optimization problem with N unknowns at each iteration. To ameliorate the computational burden, we adopt the FLSQR method that can reduce problem size and automatically estimate regularization parameter.

2.1.3. Solution for the inner problem via FGK

The essence of dimension reduction lies in an iterative projection scheme, encompassing two main stages in each iteration. First, we generate a basis (a set of vectors) for the solution by exploiting the FGK process (Chung and Gazzola 2019). Second, we compute the coefficients of the basis by solving an optimization problem in the projected subspace, where the regularization parameter can be estimated automatically via WGCV (Chung et al 2008).

During the implementation of the first stage, at the k th iteration we obtain

$$\mathbf{A}\mathbf{Z}_k = \mathbf{U}_{k+1}\mathbf{G}_k \quad (6)$$

$$\mathbf{A}^\top \mathbf{U}_{k+1} = \mathbf{V}_{k+1}\mathbf{T}_{k+1}, \quad (7)$$

where $\mathbf{G}_k \in R^{(k+1) \times k}$ is upper Hessenberg and $\mathbf{T}_{k+1} \in R^{(k+1) \times (k+1)}$ is upper triangular; $\mathbf{Z}_k = [\mathbf{D}_k^{-1}\mathbf{v}_1, \dots, \mathbf{D}_k^{-1}\mathbf{v}_k] \in R^{N \times k}$ whose column vectors are the basis for the solution of $\mathbf{s}^{(k)}$; $\mathbf{V}_{k+1} = [\mathbf{v}_1 \dots \mathbf{v}_{k+1}] \in R^{N \times (k+1)}$ and $\mathbf{U}_{k+1} = [\mathbf{u}_1 \dots \mathbf{u}_{k+1}] \in R^{M \times (k+1)}$ satisfy the orthogonality condition (in exact arithmetic)

$$\mathbf{V}_{k+1}^\top \mathbf{V}_{k+1} = \mathbf{I}_{k+1} \quad \mathbf{U}_{k+1}^\top \mathbf{U}_{k+1} = \mathbf{I}_{k+1} \quad (8)$$

with identity matrix $\mathbf{I}_{k+1} \in R^{(k+1) \times (k+1)}$. Then, we can convert problem (4) to

$$\min_{\mathbf{f} \in R^k} \|\mathbf{A}\mathbf{Z}_k \mathbf{f} - \mathbf{r}^{(\ell)}\|_2^2 + \lambda \|\mathbf{D}_k \mathbf{Z}_k \mathbf{f}\|_2^2 \quad (9)$$

and approximate the solution of (4) via the relations $\mathbf{s}^{(k)} = \mathbf{Z}_k \mathbf{f}_k$ where the basis coefficient $\mathbf{f}_k \in R^k$ is the solution to (9). To further simplify the problem (9), we perform QR factorization of $\mathbf{D}_k \mathbf{Z}_k = \mathbf{Q}_{Z,k} \mathbf{R}_{Z,k}$ Demmel (1997), where $\mathbf{Q}_{Z,k}$ has orthonormal columns and $\mathbf{R}_{Z,k}$ is an upper triangular matrix. Plugging the factorization results from equations (6) and (8) into problem (9), we arrive at a small-sized optimization problem in the projected subspace at the k th iteration

$$\min_{\mathbf{f} \in R^k} \|\mathbf{G}_k \mathbf{f} - \beta \mathbf{e}_1\|_2^2 + \lambda \|\mathbf{R}_{Z,k} \mathbf{f}\|_2^2, \quad (10)$$

Where $\beta = \|\mathbf{r}^{(\ell)}\|_2$ is a scalar, $\mathbf{e}_1 = [1, 0, \dots, 0]^\top \in R^k$. Note that the size of the projected problem (10) is $(k+1)$ by k . Therefore, RGIRT alleviates the computational complexity by transforming the challenge of solving N unknowns of \mathbf{s} in the full-sized problem (4) into that of solving k unknowns of \mathbf{f}_k in the small-sized projected problem (10). After applying WGCV in (10) to estimate the regularization parameters λ , we can determine the optimal coefficient

$$\mathbf{f}_k = (\mathbf{G}_k^\top \mathbf{G}_k + \lambda \mathbf{R}_{Z,k}^\top \mathbf{R}_{Z,k})^{-1} \mathbf{G}_k^\top \beta \mathbf{e}_1. \quad (11)$$

Finally, when the stop criterion of the inner iteration is satisfied, we obtain the solution of problem (1) at the $(\ell+1)$ th iteration

Algorithm 1. Residual-guided Golub–Kahan iterative reconstruction technique (RGIRT).

Initialization: $A \in R^{M \times N}$, $\mathbf{b} \in R^M$, δ_{tol} , ζ_{in} , ζ_{out} , $k = 1$
while $\|\mathbf{r}^{(\ell)}\|_2 / \|\mathbf{b}\|_2 < \delta_{\text{tol}}$ and $\ell < \zeta_{\text{out}}$ **do** \triangleright outer iteration
 $\mathbf{r}^{(\ell)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(\ell)}$
Initialize $\mathbf{D}_1 \in R^{N \times N}$,
for $k = 1$ **to** ζ_{in} **do** \triangleright inner iteration
 $\mathbf{w} = \mathbf{A}^\top \mathbf{u}_k$, $t_{j,k} = \mathbf{w}^\top \mathbf{v}_j$ for $j = 1, \dots, k - 1$
 $\mathbf{w} = \mathbf{w} - \sum_{j=1}^{k-1} t_{j,k} \mathbf{v}_j$, $t_{k,k} = \|\mathbf{w}\|_2$, $\mathbf{v}_k = \mathbf{w} / t_{k,k}$
 $\mathbf{z}_k = \mathbf{D}_k^{-1} \mathbf{v}_k$
 $\mathbf{w} = \mathbf{A} \mathbf{z}_k$, $g_{j,k} = \mathbf{w}^\top \mathbf{u}_j$ for $j = 1, \dots, k$
 $\mathbf{w} = \mathbf{w} - \sum_{j=1}^k g_{j,k} \mathbf{u}_j$, $g_{k+1,k} = \|\mathbf{w}\|_2$, $\mathbf{u}_{k+1} = \mathbf{w} / g_{k+1,k}$
Do QR factorization $\mathbf{D}_k \mathbf{Z}_k = \mathbf{Q}_{Z,k} \mathbf{R}_{Z,k}$
Obtain λ via WGCV
Calculate \mathbf{f}_k using (11),
Update $\mathbf{s}^{(k)} = \mathbf{Z}_k \mathbf{f}_k$
 $\mathbf{D}_{k+1} = \mathbf{D}(\mathbf{s}^{(k)})$, $k = k + 1$
end for
Obtain $\mathbf{x}^{(\ell)}$ using (12)
 $\ell = \ell + 1$
end while
return final reconstruction $\mathbf{x}^{(\ell)}$

$$\mathbf{s} = \mathbf{Z}_k \mathbf{f}_k, \mathbf{x}^{(\ell+1)} = \mathbf{x}^{(\ell)} + \mathbf{s} \quad (12)$$

A general overview of RGIRT is illustrated in figure 1. Meanwhile, the implementation details of RGIRT are provided in algorithm 1, where δ_{tol} is the tolerance given by the user, and ζ_{in} , ζ_{out} are the maximum number of the inner and outer iterations, respectively.

2.2. Numerical simulation based on a Shepp–Logan phantom

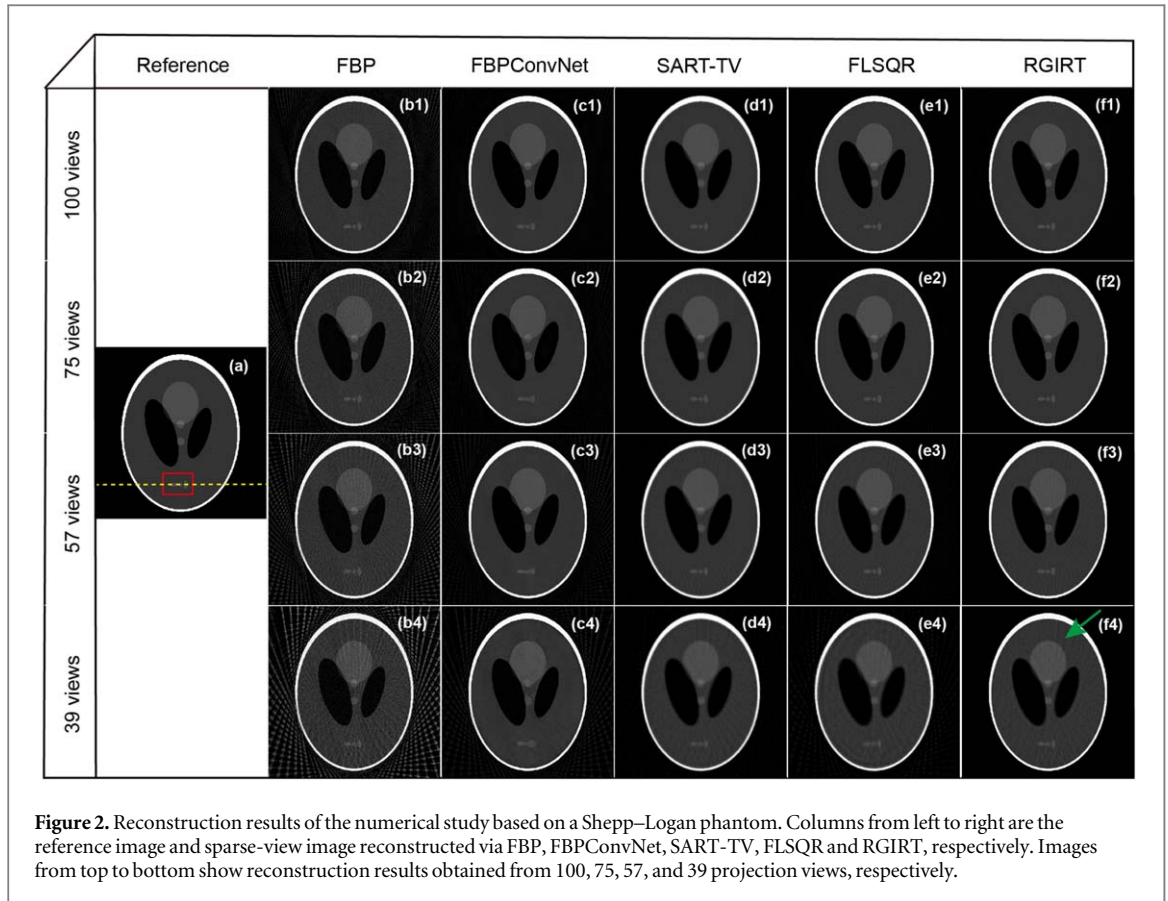
We utilized a Shepp–Logan phantom to simulate sparse-view CT reconstruction, as it incorporates both high-contrast and low-contrast structures. The sinogram data were generated using ASTRA toolbox (van Aarle *et al* 2016). The detector row comprises 1024 elements with a pixel pitch of 0.05 mm. The distance from the focal spot of the x-ray source to the detector was set to 185.03 mm, while the distance to the system iso-centre was set to 141.52 mm. The detector array consists of 1024 elements. The dimension of each element is 0.05 mm. The matrix size of the reconstruction image is 512×512 , and each pixel covers an area of $0.0765 \times 0.0765 \text{ mm}^2$. A total of 400 projections were evenly distributed over 200° in the fan-beam geometry. The photon number was set to 1×10^6 for simulating Poisson noise in the sinogram. Different projection views (100, 75, 57, and 39 views) were sampled evenly among the 400 views and tested with different reconstruction methods for validation purposes.

2.3. Mouse cardiac micro-CT data

The sinogram data used in this study were experimentally collected from a mouse cardiac scan using a customized micro-CT scanner (Cao *et al* 2010). The scanning parameters were 50 kVp anode voltage, 2 mA anode current, 1024 detector elements of 0.05 mm pixel pitch for each detector row. The source-to-detector distance and source-to-object distance are 185.03 mm and 141.52 mm, respectively. A total of 400 projections were acquired with a step angle of 0.5° , resulting in a total scan angle of 200° for a short-scan mode in the fan-beam geometry. The 400 projection views were acquired in the step-and-shoot mode with prospective gating to both the respiratory and cardiac signals of the mouse subject under free-breathing condition, so that all the projection views were acquired at the same phase of the cardiac and respiratory cycles. Similar to the Shepp–Logan phantom experiment, we designate the FBP reconstruction results obtained from 400 projection views as the reference image, and employ the reconstruction outcomes from 100 (or 75, 57, and 39) projection views to evaluate the performance of our algorithm.

2.4. Performance evaluation

We chose four different methods for comparisons including FBP, FBPCNN (Jin *et al* 2017), SART-TV (Lu *et al* 2012), and FLSQR (Chung and Gazzola 2019). FBP relies on the radon transform and its inverse, usually viewed as much faster than iterative methods. FBPCNN is a post-processing DL model based on convolutional neural network (CNN). To ensure a fair comparison, we trained four models for each experiment,



with each model representing a distinct sparsity sampling condition (100, 75, 57, and 39 views, respectively). For each sampling rate, the training dataset consisted of 300 images of mouse with an image size of 512×512 pixels. Meanwhile, we used mean square error (MSE) as the loss function and employed the Adam (Kingma and Ba 2014) optimizer during network training. The MSE loss function can be written as

$$L = \|Y - F(X)\|_2^2, \quad (13)$$

where Y is the reference images reconstructed from the full 400 projection views by using FBP, X is the images reconstructed from the sparse-view sinograms via FBP, and $F(X)$ stands for the images post-processed by FBPConvNet. The SART-TV and FLSQR are both iterative methods. In particular, FLSQR is the same algorithm as our inner iteration but without the outer restarted iterative scheme. As for quantitative measures, we used MSE, structural similarity index measure (SSIM) (Wang *et al* 2004), peak signal-to-noise ratio (PSNR), and computational time to evaluate the performance of each algorithm.

3. Results

3.1. Reconstructions of numerical simulation results

Reconstructed images of the Shepp–Logan phantom under various view-sampling conditions were compared in figure 2. The matrix size of the reconstruction image is 512×512 , and each pixel covers an area of $0.0765 \times 0.0765 \text{mm}^2$. For training the model of FBPConvNet, we generated a dataset consisting of 500 images representing ellipses with random intensity, size, and location. A fixed learning rate of 1×10^{-4} was employed. For all iterative reconstruction methods, we fine-tuned the regularization parameters to strike a balance between artifact removal and image resolution, following the examples outlined in the original literature. Specifically, in this case, the maximum number of iterations for the gradient descent step in the SART-TV method was set to 1000, with a regularization weight reduction factor of 2.5×10^{-4} , a fixed step size of 0.2 that directly influences image refinement, and a tolerance of 10^{-7} , consistent with other iterative algorithms. For the FLSQR method, the maximum number of iterations was set to 100. As for RGIRT, the inner iteration number was set to 1, while the outer iteration number was set to 300. The regularization parameters for these two methods were selected automatically by WGCV.

As observed in figure 2, all the reconstructed images from 39 views are significantly degraded, exhibiting the streaking artifacts indicated by the green arrow. In contrast, relatively reasonable reconstruction results are

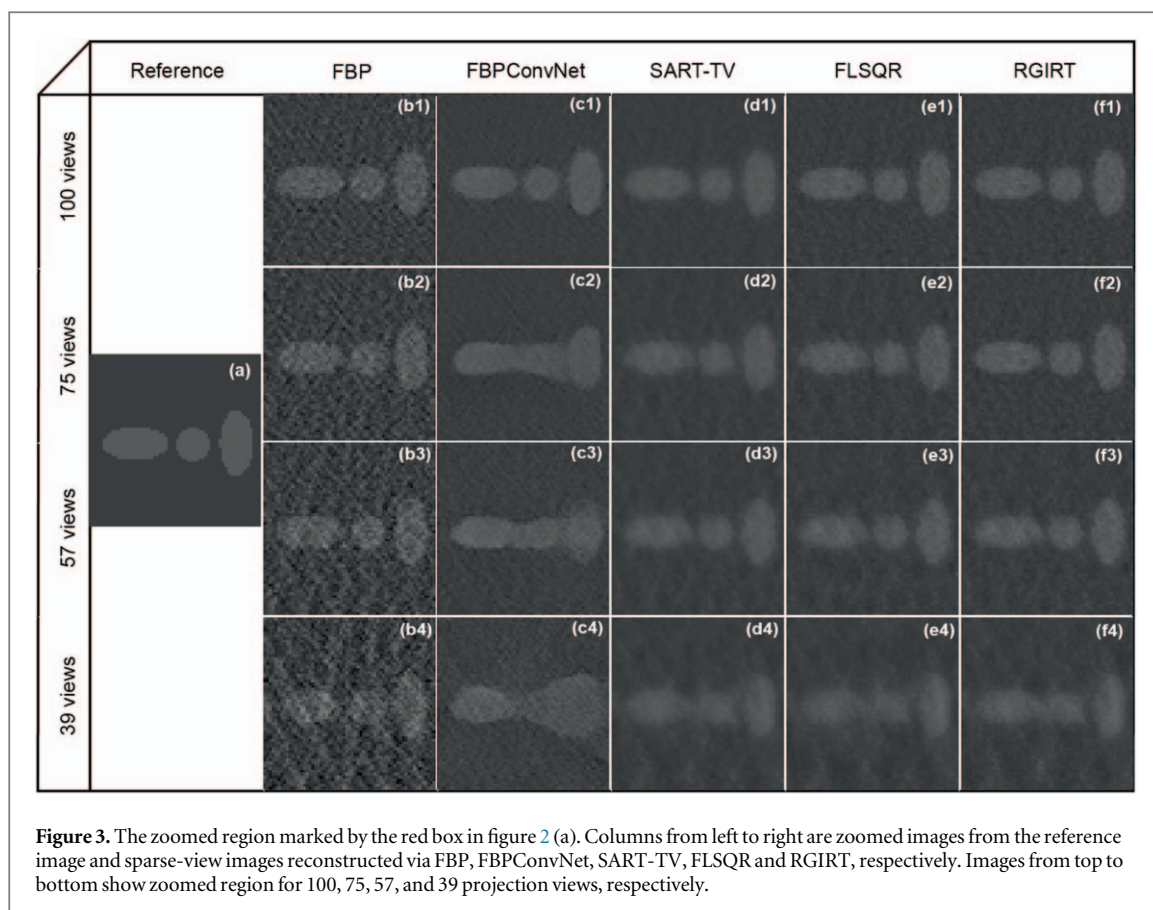
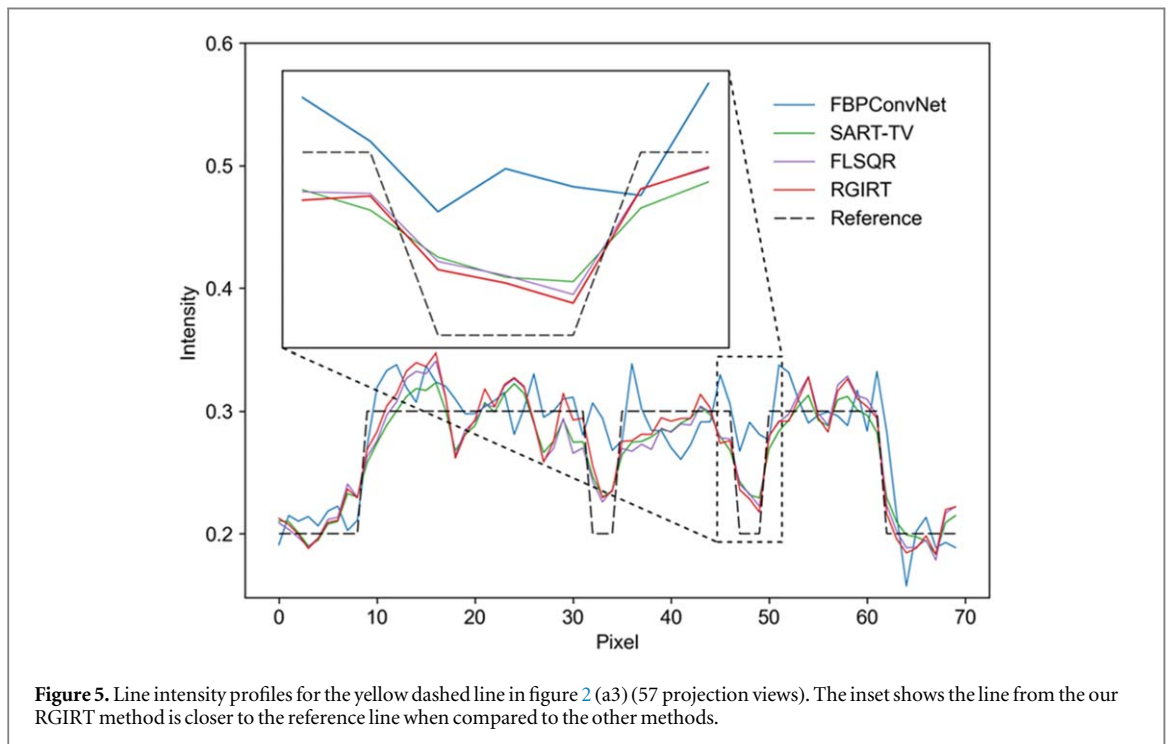
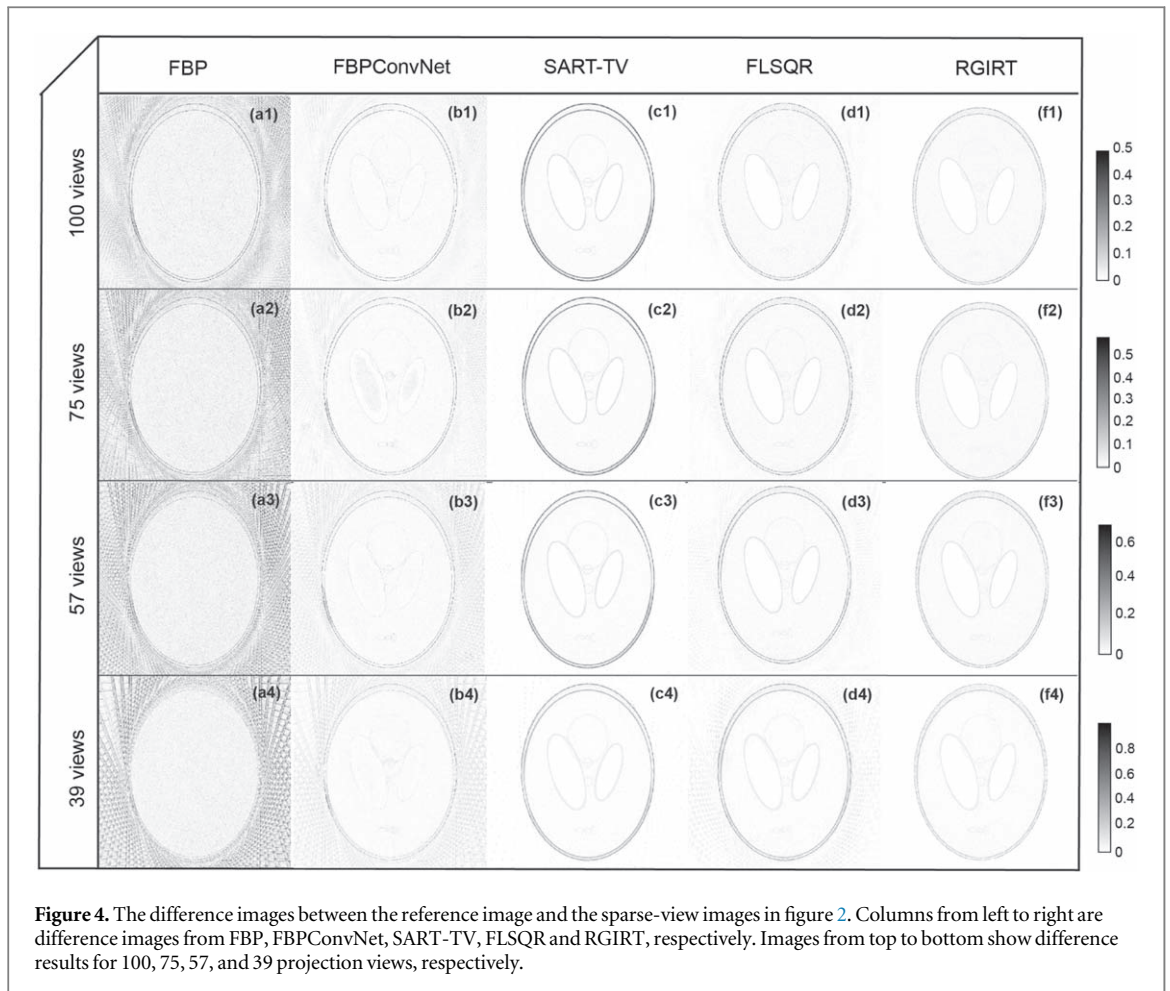


Figure 3. The zoomed region marked by the red box in figure 2 (a). Columns from left to right are zoomed images from the reference image and sparse-view images reconstructed via FBP, FBPCovNet, SART-TV, FLSQR and RGIRT, respectively. Images from top to bottom show zoomed region for 100, 75, 57, and 39 projection views, respectively.

achievable using 57 views. From the results, FBP exhibits streaking artifacts across all sparse-view sampling conditions. FBPCovNet recovers sharper edges attributed to its ability to learn representative features from the training dataset. However, as the number of views for reconstruction decreases, the smaller ellipses marked in the red box in the FBPCovNet-reconstructed images become increasingly blurred. This blurring effect is primarily caused by the adoption of mean squared error (MSE) as the loss function, where a lower MSE corresponds to more blurred images. To further illustrate this, we selected and magnified the region outlined by the red box in figure 2 for each method, as shown in figure 3. While the FBPCovNet results merge the three ellipses within the red box for the images reconstructed from 79, 57, and 39 views, other iterative methods successfully separate them. SART-TV results exhibit blurred edges in the images reconstructed with 75 and 100 views. Although the FLSQR and RGIRT results appear similar in figure 2 due to the simple structure of the Shepp–Logan phantom, RGIRT demonstrates superior performance in the subsequent experiment involving more complex mouse cardiac micro-CT images.

To further demonstrate the advantages of RGIRT in terms of artifact reduction and preservation of image details, figure 4 presents the absolute difference images, which highlight the absolute discrepancies between the results obtained from each method and the reference image. In the FBPCovNet results, fewer differences are observed within the phantom but more outside of it. These outside residual streaking artifacts may be attributed to discrepancies between the training and test datasets, a challenge often encountered in deep learning-based methods known as dataset shift (Takahashi and Braga 2020). Specifically, for the Shepp–Logan phantom experiment, the training dataset consist of 500 images, each containing 20–40 randomly positioned ellipses. The length of the horizontal and vertical axes is randomly chosen between 32 and 128 pixels, and each ellipse rotates with a random angle. Although the Shepp–Logan phantom is comprised of ellipses, not all its features are randomly incorporated in the current training dataset. Additionally, significant differences of the three ellipses are visible, particularly in the images reconstructed by 75, 57, and 39 views sinogram, aligning with the findings depicted in figure 3. The SART-TV results exhibit noticeable differences along the edge of the phantom, though there are fewer artifacts outside of it. Moreover, in comparison, the FLSQR and RGIRT results exhibit less differences along the edge and less discrepancy in the region containing the three ellipses. Compared to the FLSQR results, our RGIRT results have fewer artifacts outside of the phantom, especially in the 57-view reconstructions, further demonstrate the robustness in cases of extreme sparse-view conditions. Among the evaluated methods, RGIRT exhibits the most superior performance across multiple metrics, indicating its higher overall competence in terms of artifact reduction, preservation of image details, and fidelity to the



reference image. Furthermore, figure 5 displays the comparison of line intensity profiles of various methods passing through the yellow dash line (see figure 2) in the 57-view case. It is evident that the line intensity profile from RGIRT closely resembles that of the reference CT image, especially the valley in the zoomed region.

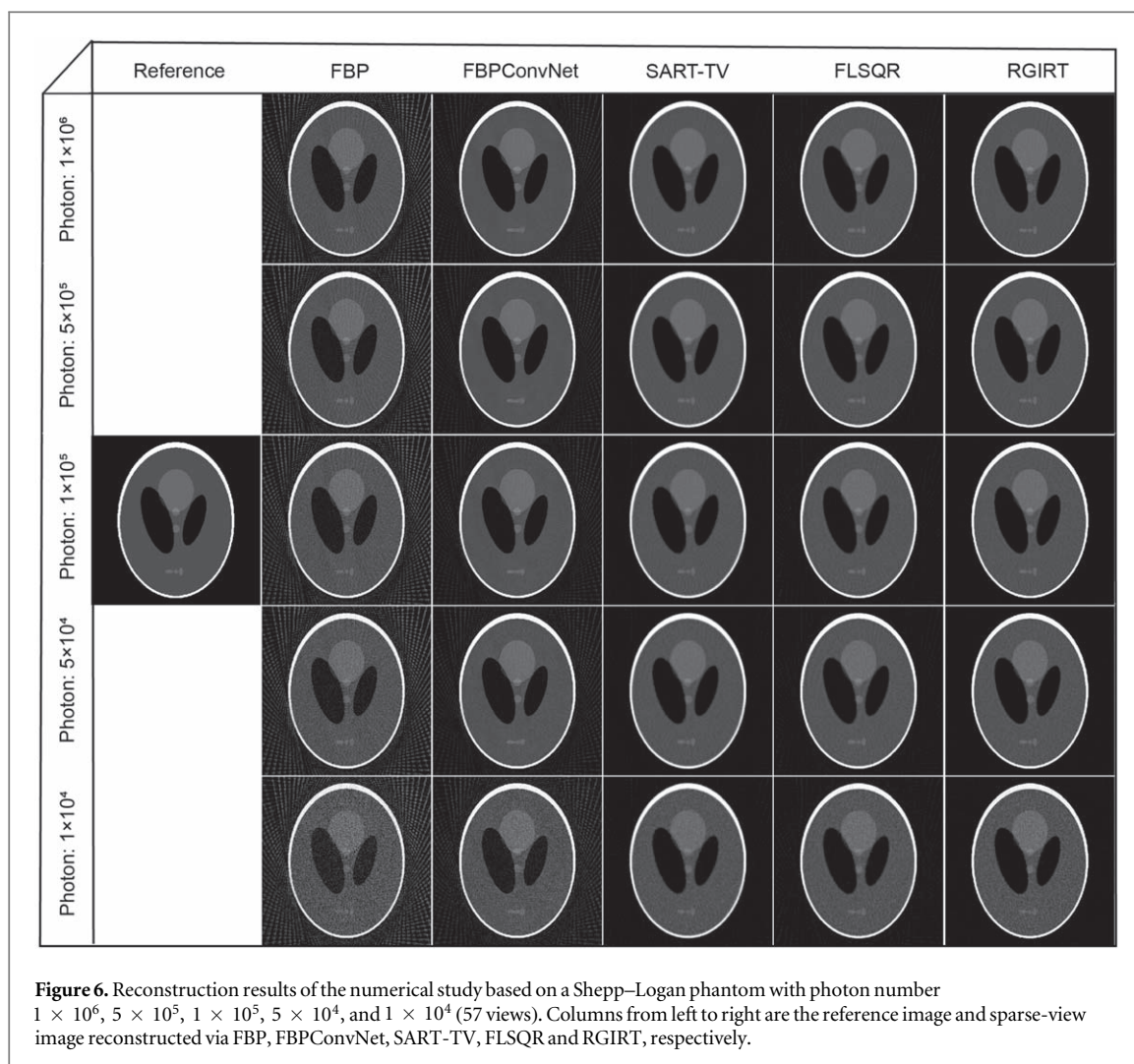
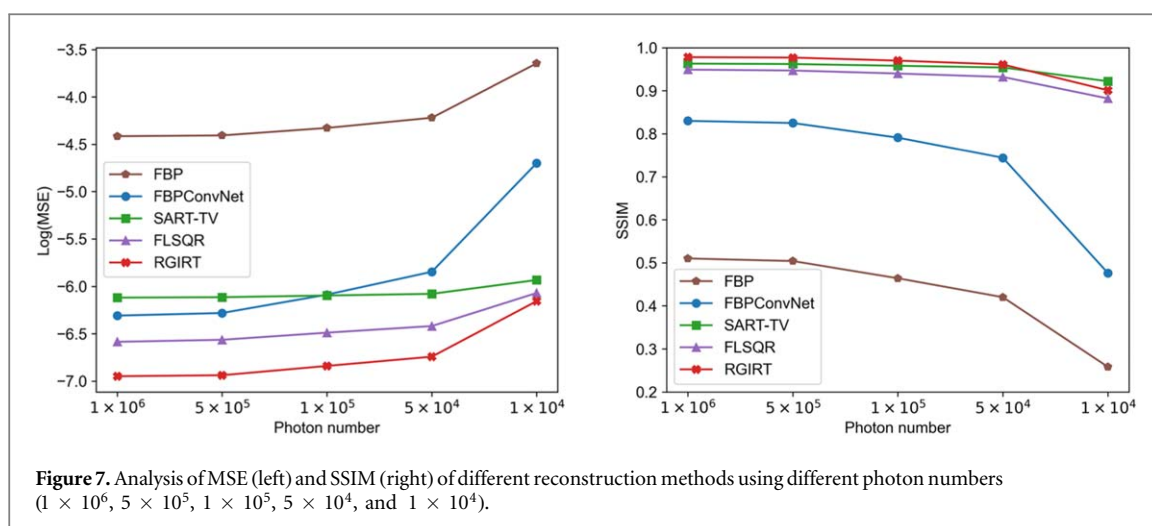


Figure 6. Reconstruction results of the numerical study based on a Shepp–Logan phantom with photon number 1×10^6 , 5×10^5 , 1×10^5 , 5×10^4 , and 1×10^4 (57 views). Columns from left to right are the reference image and sparse-view image reconstructed via FBP, FBPCovNet, SART-TV, FLSQR and RGIRT, respectively.

Table 1. Quantitative results for sparse-view Shepp–Logan phantom. The best results are marked in red, and the second-best results are marked in blue.

Method	MSE (10^{-2})				SSIM				PSNR			
	100	75	57	39	100	75	57	39	100	75	57	39
FBP	0.431	0.740	1.211	2.358	0.717	0.596	0.510	0.426	33.196	30.845	28.707	25.816
FBPCovNet	0.089	0.106	0.182	0.304	0.951	0.925	0.830	0.735	40.033	39.259	36.934	34.713
SART-TV	0.200	0.208	0.220	0.256	0.972	0.968	0.963	0.950	36.527	36.524	36.108	35.454
FLSQR	0.073	0.101	0.138	0.251	0.978	0.968	0.949	0.871	40.881	39.467	38.119	35.538
RGIRT	0.068	0.081	0.096	0.130	0.982	0.980	0.978	0.973	41.154	40.425	39.711	38.396

A quantitative analysis was conducted on the reconstructions of the sparse-view Shepp–Logan phantom and presented in table 1, with the best results highlighted in bold red and the second-best results in bold blue. Among all sparsity conditions, RGIRT demonstrated the smallest MSE and PSNR, except for the 100 views image, which closely resembled the performance of FLSQR. Furthermore, RGIRT ranked second highest in SSIM across all sparsity conditions, except for the 75 views image. While SART-TV outperformed RGIRT in SSIM, it exhibited significantly larger MSE. The results in table 1 demonstrate that RGIRT is capable of reducing artifacts and preserving image structure simultaneously, rather than excelling in one aspect while performing poorly in the other.



3.2. Robustness analysis to noise

In practical applications, the presence of photonic noise during imaging can significantly impact the quality of reconstructed images. To evaluate the robustness of our proposed method, we carried out further comparative experiments within the Shepp–Logan experiment, assessing the performance of our proposed method under various noise levels. Poisson noise with five different intensities was introduced to the simulated sinogram of 57 projection views, corresponding to 1×10^6 , 5×10^5 , 1×10^5 , 5×10^4 , 1×10^4 photons. The reconstruction results using FBP, FBPCovNet, SART-TV, FLSQR, and RGIRT at different intensity levels are shown in figure 6. Notably, while all methods are affected by noise at the intensity level of 1×10^4 , FBPCovNet exhibits a higher level of noise compared to the other methods. Reconstruction results (figure 6) and metrics (figure 7) show a good agreement that that RGIRT maintained its efficiency and accuracy both qualitatively and quantitatively, even in the presence of an increasing noise level. Although SART-TV also exhibits high noise tolerance (green curves in figure 7), its convergence is not guaranteed, resulting in a relatively poorer image quality when the stopping point is inappropriately selected. Another noteworthy observation is that, as the noise level increases, FBPCovNet experiences a sudden rise in MSE and a sharp decline in SSIM (figure 7), which emphasizes the susceptibility of FBPCovNet to noise interference, along with limited robustness and applicability.

3.3. Reconstructions of mouse cardiac micro-CT

The reconstructed images under various view-sampling conditions of the cardiac micro-CT are shown in figure 8. We can observe severe streaking artifacts (marked by the arrows) appearing in all the reconstructed images obtained from 39 views. As a comparison, the reconstructed results from 57 views have effectively eliminated the streaking artifacts. To train the FBPCovNet on mouse cardiac micro-CT images, the training dataset consisted of 300 images, with an additional 100 images allocated as the validation dataset. In this particular instance, RGIRT was configured with an inner iteration count of 3 and an outer iteration count of 20. The selection of the number of inner and outer iterations in contrast regions to be reconstructed. When the level of detail complexity is higher, we decrease the number of inner iterations and increase the number of outer iterations. The step size for SART-TV was set to 0.1. For all iterative methods, the tolerance was set to 10^{-7} . All other parameters used in FBPCovNet training process and the three iterative methods are the same as the Shepp–Logan experiment. The FBPCovNet, SART-TV, and FLSQR methods effectively suppress the streaking artifacts in 100 views. However, as the views become sparser, the different methods degrade differently. FBPCovNet tends to blur low-contrast regions such as the heart and enhance high-contrast regions such as bones. Particularly, in the case of 57 views, the FBPCovNet blurs the heart entirely, while other methods preserve its structure. SART-TV, on the other hand, tends to blur the entire image, including both high- and low-contrast regions. This blurring effect can be clearly observed in the results of 75, 57, and 39 views (figure 8). Compared with the FBP and FBPCovNet, RGIRT and FLSQR successfully preserve the geometries of imaging objects in both high- and low-contrast regions in a small number of views (e.g. 57 views and 75 views, figure 8). However, in the case of 57 views, FLSQR exhibits a slight degradation than RGIRT in high contrast regions, as clearly shown in the difference images in figure 8.

More detailed visual comparison is shown in figure 9, where the red box region of figure 8 (a) is enlarged to examine the thin interventricular septum more closely. The interventricular septum (indicated by the blue

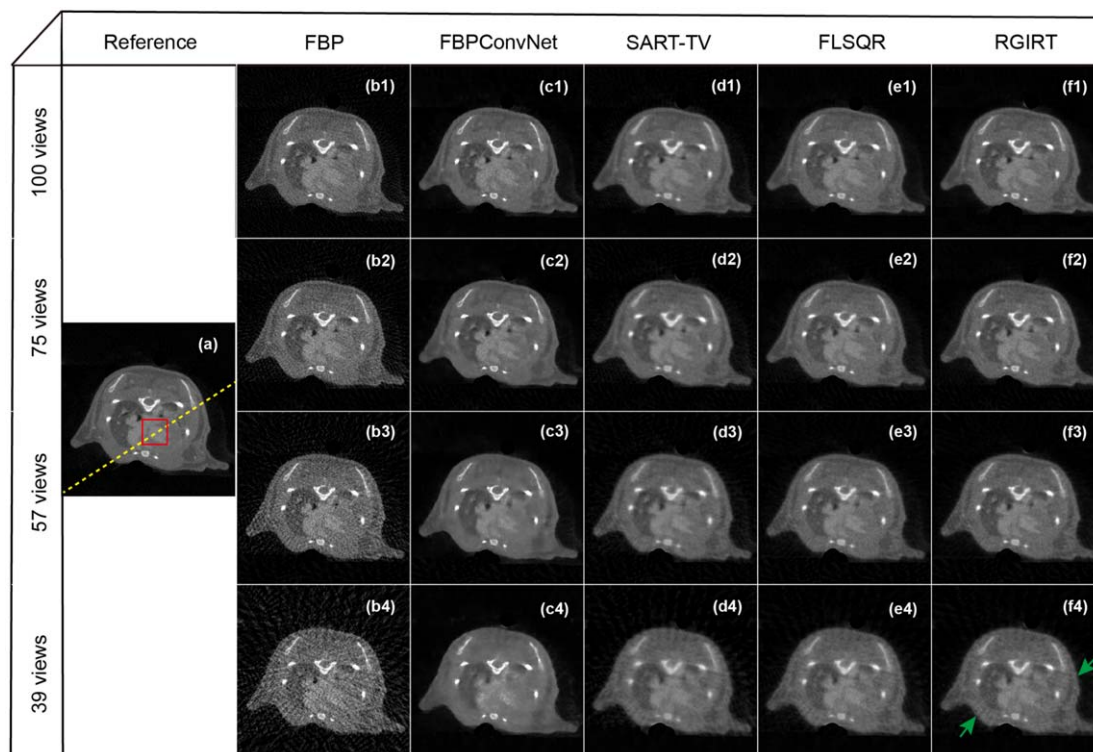


Figure 8. Reconstruction results for a mouse cardiac micro-CT dataset using different methods. Columns from left to right are the full-view reference image and sparse-view image reconstructed via FBP, FBPConvNet, SART-TV, FLSQR and RGIRT, respectively. Images from top to bottom show reconstruction results from 100, 75, 57, and 39 projection views, respectively.

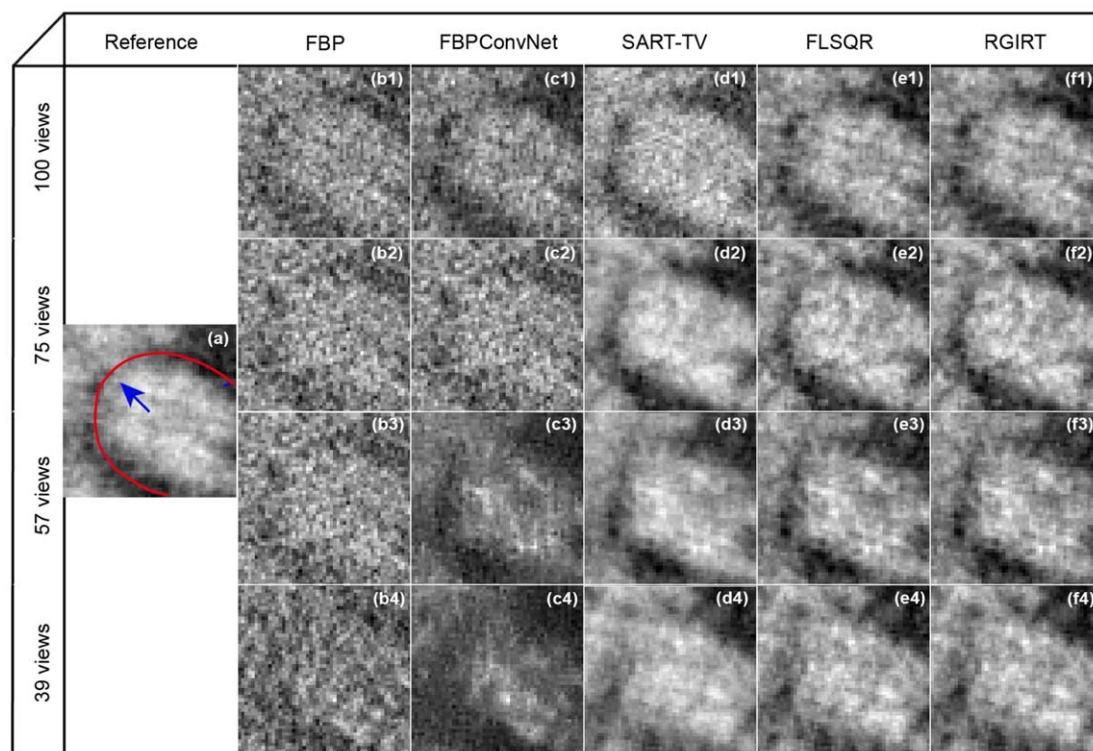
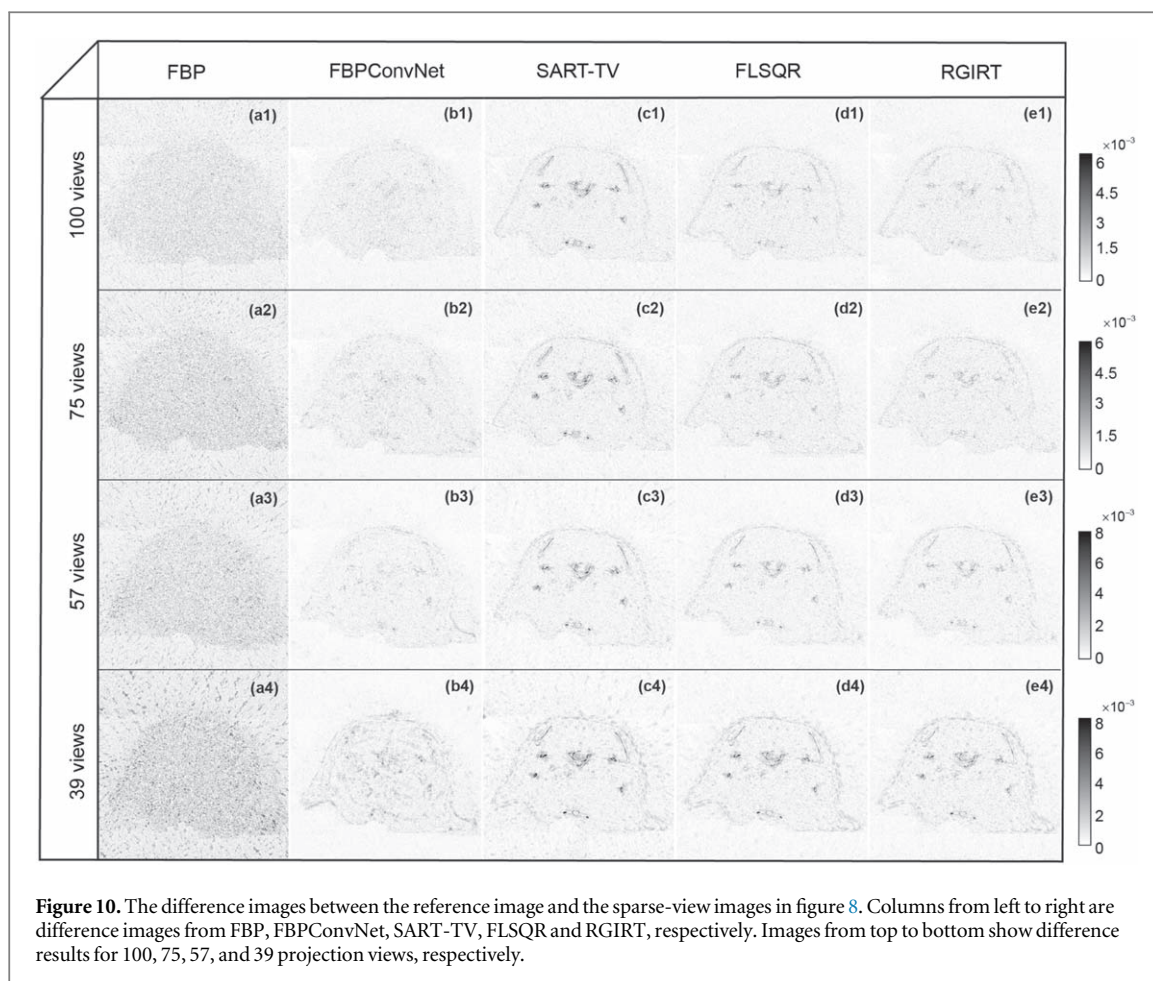


Figure 9. The zoomed region marked by the red box in figure 8(a). Columns from left to right are zoomed images from the reference image and sparse-view images reconstructed via FBP, FBPConvNet, SART-TV, FLSQR and RGIRT, respectively. Images from top to bottom show zoomed region for 100, 75, 57, and 39 projection views, respectively. Blue arrow and red curve indicate a thin interventricular septum.

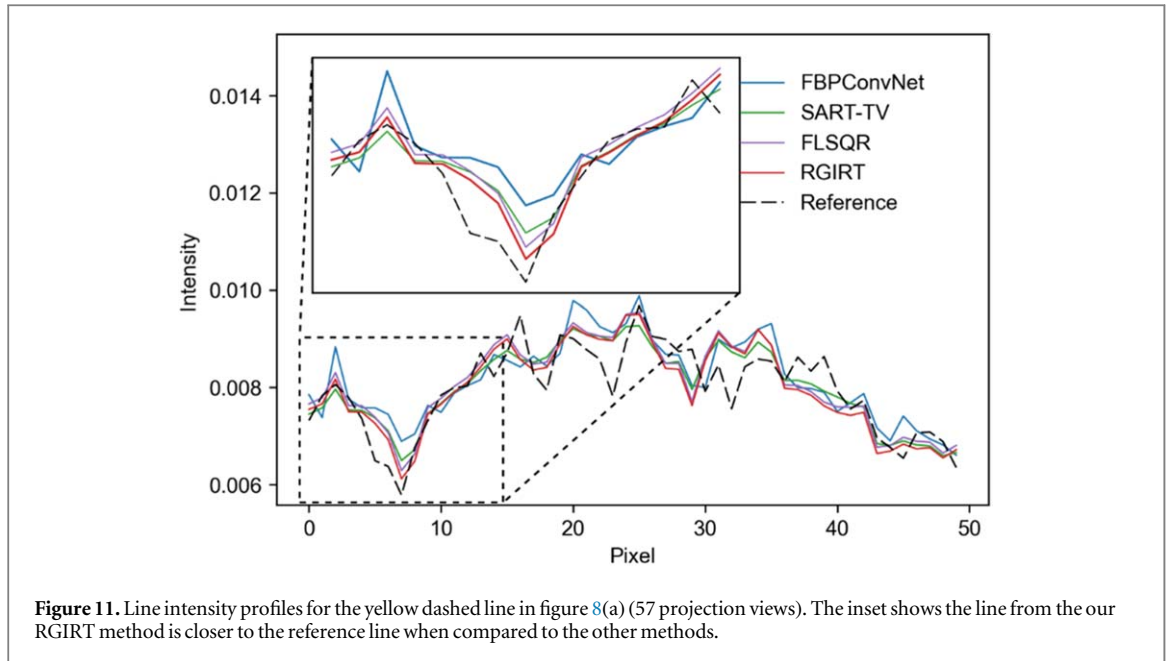


arrow and red curve in the reference image of figure 9) is not discernable at all from the sparse-view CT image reconstructed by FBP. For the images reconstructed by FBPCConvNet, it is somewhat visible from the 100-view case but barely visible from the remaining fewer-view (e.g. 75, 57 and 39 views) scenarios. The same interventricular septum is visible from the images reconstructed from SART-TV, FLSQR, and RGIRT, but a close look reveals that the interventricular septum reconstructed by SART-TV and FLSQR are a little bit blurry due to the over-smoothing from those two methods. The interventricular septum reconstructed by our RGIRT method seems most clear, especially when examining the small structure details. Moreover, the contrast between the interventricular septum and the neighboring iodine-filled ventricles appears clearer than that from the other methods, which indicates RGIRT is capable of preserving soft tissue contrast and suppressing noise. Judging from figure 9, as the sparsity changes from 100 views to 75, 57 and 39 views, the superiority of our RGIRT method becomes even more obvious.

The effectiveness of RGIRT is further demonstrated in the difference images shown in figure 10, which were calculated as the absolute difference between the reference image and the sparse-view images in figure 8. In these difference images, the darker the color, the larger the error. We can see that the difference image of RGIRT has the smallest overall difference, which confirms RGIRT's superiority in preserving the image details. Interestingly, comparing the difference images between FBPCConvNet and RGIRT, we found that RGIRT is better at preserving the soft tissues, while FBPCConvNet is better at preserving the bone structures. Because the heart is a soft-tissue organ without bone structure, our RGIRT method is indeed better suited for sparse-view cardiac CT reconstruction.

Figure 11 shows the comparison of line intensity profiles of various methods passing through the yellow dash line (see figure 8) in the 57-view case. It is clear that the line intensity profile from RGIRT resembles most closely to the one from the reference CT image, especially for the peak and valley in the zoomed region. This demonstrates the advantage of our method over the other methods on persevering edge and small features in the images.

We also carried out quantitative analysis for the sparse-view cardiac micro-CT reconstructions and the results are presented in table 2. As shown in table 2, RGIRT and FBPCConvNet clearly outperformed the other methods with smaller MSE and higher SSIM and PSNR. The FBPCConvNet method has the best MSE among many sparse-view cases, mainly because it uses MSE as the loss function. However, as the aforementioned



zoomed images (figure 9) clearly show, FBPCConvNet has a poor performance in reducing the overall noise and noise textures, which led to the loss of small image details. Thus, although the FBPCConvNet method provides good quantitative results, it cannot provide visual results as good as RGIRT. Compared to the other reference methods, the superiority of RGIRT is more obvious when it comes to fewer projection views. Therefore, we can conclude that RGIRT provides the most accurate reconstruction results for sparse-view CT.

3.4. Convergence analysis

For the three iterative methods (SART-TV, FLSQR, RGIRT), we analyzed their convergence behaviors from the following three aspects: whether it converges, the speed of convergence, and image error. This convergence analysis can be examined by plotting the MSE loss curves during iterations, which are shown in figure 12(a). This figure demonstrates the capability of the three iterative methods in stabilizing the semi-convergent behavior of the ill-posed problem in the 57-view mouse case. First, compared to FLSQR and RGIRT, SART-TV method has clear semi-convergent behavior, where MSE shows a trend of first decreasing and then slowly increasing as the number of iterations increases. As for the speed of convergence, it can be observed that RGIRT reaches the convergence point fastest with only 9 iterations. Furthermore, the image at the 9th iteration in figure 12(a) show that our RGIRT is able to obtain the most optimal result compared to the other methods. This is consistent with the performance of the MSE loss curve (see the zoomed area in figure 9). Combined with the fact that each iteration of RGIRT is to solve the dimensionality-reduced projected problem, the computing time of RGIRT spent on one iteration is the smallest among the three iterative methods. Thus, RGIRT is the fastest method to converge. In terms of image error, RGIRT produced the lowest MSE in all three iterative methods throughout the iterations. Therefore, among the three iterative methods studied in this work, RGIRT can provide the most accurate result with the fastest convergence speed.

When determining the maximum iteration number for inner and outer loops, it is essential to consider their respective roles—the inner iteration focuses on calculating solution space, while the outer iteration is dedicated to correcting the results obtained from the inner loop. An exploratory study concerning the selection of maximum iteration number for inner and outer loops is presented in figure 12(b). From the curve, we notice that the combination of fewer inner iterations with larger outer iterations often serves as a suitable parameter selection, which can improve the precision with high reconstruction speed. Theoretically speaking, FLSQR is a special case of RGIRT, wherein the number of outer iterations is set to 1 and no residual information is used for guidance.

3.5. Computational cost

Computational cost is an important factor for any reconstruction algorithm. All the algorithms in this work were implemented using MATLAB except the FBPCConvNet, which was implemented with PYTHON. Considering that the training process for a DL method is very time consuming, the pre-training time was not included in comparison. Although the algorithms were realized using different programming languages, the efficiency can be roughly compared on a same computer. The MSE, SSIM and reconstruction time from the

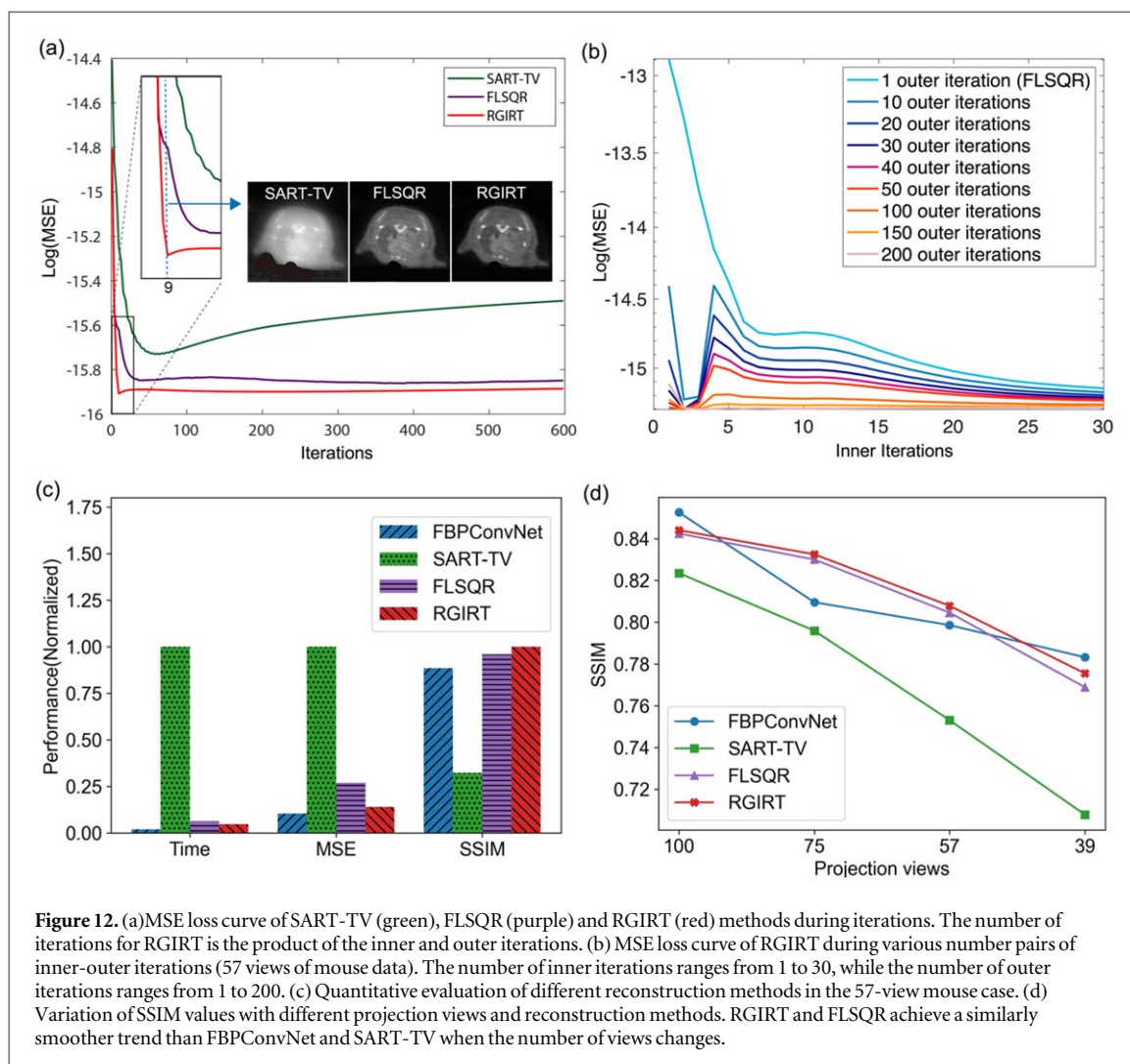


Table 2. Quantitative results for sparse-view mouse cardiac micro-CT image. The best results are marked in red, and the second-best results are marked in blue.

Method	MSE (10^{-7})				SSIM				PSNR			
	100	75	57	39	100	75	57	39	100	75	57	39
FBP	0.388	0.698	1.085	1.791	0.696	0.565	0.473	0.390	31.808	29.261	27.344	25.168
FBPCConvNet	0.134	0.182	0.215	0.305	0.862	0.809	0.798	0.783	36.403	35.080	34.371	32.854
SART-TV	0.202	0.249	0.335	0.481	0.823	0.795	0.753	0.707	34.626	33.721	32.441	30.875
FLSQR	0.162	0.186	0.273	0.346	0.842	0.830	0.804	0.768	35.592	34.993	33.942	32.301
RGIRT	0.153	0.176	0.220	0.318	0.844	0.832	0.807	0.775	35.843	35.229	34.272	32.674

different methods when reconstructing the images for the 57-view mouse case are shown in figure 12(c). All metrics in figure 12(c) were normalized to the largest value from various methods. For example, the SSIM of RGIRT is scaled to 1, and the SSIMs of other methods are normalized to the SSIM of RGIRT. The average reconstruction time for FBPCConvNet, SART-TV, FLSQR, and RGIRT is 1.57, 84.66, 5.33, and 3.89 s per CT slice, respectively. All the reconstructions were conducted on the same PC workstation (Intel Xeon Gold 6248 R CPU @ 3.00 GHz, 128 GB RAM and Nvidia Quadro RTX 4000 GPU card). Our RGIRT method clearly owns a considerable advantage in computational efficiency compared to the other iterative methods.

4. Discussion and conclusion

Although DL-based sparse-view CT reconstruction methods have gained increasing popularity in recent years, the interpretability and generalizability of trained networks remain unresolved issues in clinical applications (Antun *et al* 2020). A feasible strategy for resolving these issues is to combine deep learning networks with model-based iterative reconstruction methods (Zhang *et al* 2020, Xiang *et al* 2021, Su *et al* 2022). The focus of this study is to develop a robust iterative reconstruction algorithm for sparse-view CT, which is beneficial for the future research into sparse-view CT reconstruction methods that leverage the synergistic powers of data-driven and model-driven methods. The proposed RGIRT algorithm is based on an inner-outer iteration framework, where FLSQR is implemented in the inner iteration to build up the solution space and a restarted iterative scheme is applied in the outer iteration as a correction. In this study, we utilized data from 100, 75, 57, and 39 projection views for reconstruction and conducted comparative experiments involving four other algorithms (FBP, FBPCovNet, SART-TV and FLSQR) to explore the performance and limitations of RGIRT. Given sufficient projection views (e.g. 100 views), all reconstruction algorithms, including FBP, can yield outstanding reconstruction results. In contrast, in a situation with considerable data scarcity (e.g. 39 views), noticeable artifacts emerge, indicating that 39 views exceed the practical limitations of all considered algorithms. Consequently, we selected reconstructions from 57 views for further evaluation, in which the performance differences (e.g. accuracy, stability) among various algorithms become particularly noticeable. It is important to mention that all the presented reconstruction methods including RGIRT cannot fully eliminate artifacts under extremely sparse-view conditions, e.g. 57 and 39 views. One has to be careful in selecting an appropriate number of projection angles according to specific scenarios and image quality requirements.

The key strength of RGIRT lies in its exceptional robustness, primarily manifesting in the following aspects. Firstly, compared with other methods, RGIRT can deliver more accurate reconstruction results, especially under sparse-view conditions, since its inner-outer framework incorporates the residual-guided solution correction during the iterative process. We observe that the performance of all methods deteriorates as the number of views decreases. FBPCovNet tends to blur the details in low-contrast regions as shown in cardiac data of 57 and 39 views (figure 9) and numerical data of 75, 57, and 39 views (figure 3). SART-TV displaces a large error in the edges of high-contrast regions, which is clearly visible in both cardiac data and numerical data for each case of view numbers (figures 4 and 10). Compared with these two methods, our method maintains a relatively smaller error in strong edges and depicts the low-contrast regions more clearly even under some extremely sparse-view conditions, e.g. 57 views, for both *in vivo* and numerical experiments. Furthermore, the variation of SSIM with decreasing numbers of projection views indicates that RGIRT is distinguished by the lowest gradient of change, providing greater stability and robustness in more ill-posed cases with fewer projection views compared to other methods (figure 12(d)).

Secondly, the convergence of RGIRT is validated through both theoretical and experimental results. RGIRT employs a restarted iterative scheme in the outer loop to iteratively minimize the residual and refine the solution using the result obtained from the inner loop, thereby ensuring convergence. Classic iterative methods such as SART-TV have semi-convergent behaviour (figure 12(a)). Similarly, it has been shown that during network training practical DL networks fail to converge to optimum solutions (Kawaguchi and Sun 2021), due to the challenge of solving highly nonlinear optimization problems. We believe the convergence guarantee of our RGIRT framework would be valuable in designing hybrid learning methods that combine RGIRT and deep learning techniques, particularly in improving algorithmic stability and robustness, a claim that can be validated through proof-of-concept. This would certainly improve the explainability and generalizability of hybrid DL methods. Furthermore, RGIRT features a faster convergence speed, primarily attributed to its adoption of the efficient Golub–Kahan bidiagonalization method, and the use of residual-guided correction expedites the approximation of the optimal solution.

Thirdly, RGIRT outperforms other methods in terms of MSE and SSIM across different noise levels (figure 7). FBPCovNet occurs a sharp decrease in reconstruction accuracy when introducing Poisson noise with 1×10^4 intensity levels. For RGIRT, because of the correction in the outer loop, the inner solver doesn't require high degree of accuracy, which enhances the robustness of our algorithm against various levels of noise. Moreover, we can also reduce the number of iteration for the inner solver to save runtime memory. In figure 12(b), We investigated the performance of RGIRT across varying numbers of inner iterations and observed that a high number of outer iterations combined with a relatively smaller number of inner iterations can yield precise results.

Our work is subject to following limitations. Firstly, RGIRT is currently designed for a certain type of regularization, i.e. ℓ_1 -norm regularization. In fact, other sparsity-encouraging prior knowledge can be adopted, such as ℓ_0 -norm regularized dictionary learning method (Wu *et al* 2018). It is obvious that more complex and advanced priors may improve the reconstructed image quality, although they are more computationally demanding. In practice, one would balance reconstruction performance and computational cost. Secondly, the inner-outer double iteration scheme of RGIRT could cause some computational complexity, especially when the convergence speed is slow for some special CT geometries. In this study we mainly focused on the short-scan

in fan-beam geometry, since for a cone-beam CT scanner the short-scan is advantageous in terms of decreased mechanical manipulation, increased scanning flexibility, and shortened acquisition time. Given the prevalent use of DL-based methods nowadays, we expect that our method can be hybridized with an DL-based method. This may involve a neural network to pinpoint an optimal mix of the regularization parameter and penalty function, followed by employing our algorithm for iterative reconstruction. Such framework will maintain the algorithmic convergence while reducing the uncertainties inherent in DL-methods.

In conclusion, we introduce a robust sparse-view CT reconstruction algorithm named RGIRT, which features an inner-outer dual iteration framework. The inner iteration efficiently yields an intermediate result, while the outer iteration minimizes the residual and refines the solution. Both numerical phantoms and real experimental Micro-CT data demonstrate the robustness and accuracy of RGIRT. Meanwhile, theoretical outcomes confirm the convergence of our approach.

Acknowledgments

This work was partially supported by National Natural Science Foundation of China under Grants 12101406 (JJ), 62105205 (WR), 62273238 (GC), Shanghai Science and Technology Innovation Program under Grant 21YF1429100 (JJ), and start-up funding of ShanghaiTech University (WR, GC).

Data availability statement

The data cannot be made publicly available upon publication because they are owned by a third party and the terms of use prevent public distribution. The data that support the findings of this study are available upon reasonable request from the authors.

Ethics declarations

All animal operation protocols have been approved by the guidelines of the Institutional Animal Care and Use Committee, ShanghaiTech University.

Competing interests

The authors declare no competing interests.

Appendix

Proof of convergence of the residual

Proof:

The residual at $(\ell + 1)$ th outer iterations is

$$\begin{aligned} &= \mathbf{b} - \mathbf{A}\mathbf{x}^{(\ell+1)} \\ &= \mathbf{b} - \mathbf{A}(\mathbf{x}^{(\ell)} + \mathbf{s}) \\ \mathbf{r}^{(\ell+1)} &= (\mathbf{b} - \mathbf{A}\mathbf{x}^{(\ell)}) - \mathbf{A}\mathbf{s} \\ &= \mathbf{r}^{(\ell)} - \mathbf{A}\mathbf{s} \end{aligned} \tag{A14}$$

where \mathbf{s} denotes the solution of the inner problem at ℓ th outer iterations formulated as (4) in the main text and that $\mathbf{0}$ (zero vector) is in the solution space of \mathbf{s} ,

$$\|\mathbf{A}\mathbf{s} - \mathbf{r}^{(\ell)}\|_2^2 + \lambda\|\mathbf{D}(\mathbf{s})\mathbf{s}\|_2 \leq \|\mathbf{A}\mathbf{0} - \mathbf{r}^{(\ell)}\|_2^2 + \lambda\|\mathbf{D}(\mathbf{0})\mathbf{0}\|_2 \tag{A15}$$

$$\|\mathbf{A}\mathbf{s} - \mathbf{r}^{(\ell)}\|_2^2 + \lambda\|\mathbf{D}(\mathbf{s})\mathbf{s}\|_2 \leq \|\mathbf{r}^{(\ell)}\|_2^2. \tag{A16}$$

Meanwhile, due to $\lambda > 0$,

$$\|\mathbf{A}\mathbf{s} - \mathbf{r}^{(\ell)}\|_2^2 \leq \|\mathbf{r}^{(\ell)}\|_2^2. \tag{A17}$$

Combining the relationship of (14) and (17), we can obtain

$$\|\mathbf{r}^{(\ell+1)}\|_2 \leq \|\mathbf{r}^{(\ell)}\|_2. \tag{A18}$$

ORCID iDs

Jianru Zhang  <https://orcid.org/0009-0000-9348-0143>

Wuwei Ren  <https://orcid.org/0000-0002-0940-7590>

Jiahua Jiang  <https://orcid.org/0000-0002-4000-8846>

References

- Andersen A H and Kak A C 1984 Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm *Ultrason. Imaging* **6** 81–94
- Antun V et al 2020 On instabilities of deep learning in image reconstruction and the potential costs of AI *Proc. Natl. Acad. Sci. USA* **117** 30088–95
- Baker A H, Jessup E R and Manteuffel T 2005 A technique for accelerating the convergence of restarted GMRES *SIAM J. Matrix Anal. Appl.* **26** 962–84
- Beck A and Teboulle M 2009 A fast iterative shrinkage-thresholding algorithm for linear inverse problems *SIAM J. Imag. Sci.* **2** 183–202
- Beister M, Kolditz D and Kalender W A 2012 Iterative reconstruction methods in x-ray CT *Phys. Med.* **28** 94–108
- Brenner D J and Hall E J 2007 Computed tomography—an increasing source of radiation exposure *New Engl. J. Med.* **357** 2277–84
- Cao G et al 2010 Prospective-gated cardiac micro-CT imaging of free-breathing mice using carbon nanotube field emission x-ray *Med. Phys.* **37** 5306–12
- Chung J and Gazzola S 2019 Flexible Krylov methods for l_p regularization *SIAM J. Sci. Comput.* **41** S149–71
- Chung J, Nagy J G and O’leary D P 2008 A weighted GCV method for Lanczos hybrid regularization *Electron. Trans. Numer. Anal.* **28** 149–67 (<https://nist.gov/publications/weighted-gcv-method-lanczos-hybrid-regularization>)
- Colton D, Piana M and Potthast R 1997 A simple method using morozov’s discrepancy principle for solving inverse scattering problems *Inverse Prob.* **13** 1477–93
- Demmel J W 1997 *Applied Numerical Linear Algebra* (SIAM)
- Eiermann M, Ernst O G and Schneider O 2000 Analysis of acceleration strategies for restarted minimal residual methods *J. Comput. Appl. Math.* **123** 261–92
- Elfving T, Hansen P C and Nikazad T 2014 Semi-convergence properties of Kaczmarz’s method *Inverse Prob.* **30** 055007
- Elfving T, Hansen P C and Nikazad T 2017 Convergence analysis for column-action methods in image reconstruction *Numer. Algorithms* **74** 905–24
- Ergül A, Malas T and Gürel L 2010 Solutions of large-scale electromagnetics problems using an iterative inner-outer scheme with ordinary and approximate multilevel fast multipole algorithms *Prog. Electromagn. Res.* **106** 203–23
- Fessler J A 1994 Penalized weighted least-squares image reconstruction for positron emission tomography *IEEE Trans. Med. Imaging* **13** 290–300
- Golub G H, Hansen P C and O’Leary D P 1999 Tikhonov regularization and total least squares *SIAM J. Matrix Anal. Appl.* **21** 185–94
- Gordon R, Bender R and Herman G T 1970 Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography *J. Theor. Biol.* **29** 471–81
- Gorodnitsky I F and Rao B D 1992 A new iterative weighted norm minimization algorithm and its applications [1992] *IEEE Sixth SP Workshop on Statistical Signal and Array Processing*
- Holschneider M 1991 Inverse Radon transforms through inverse wavelet transforms *Inverse Prob.* **7** 853–61
- Hudson H M and Larkin R S 1994 Accelerated image reconstruction using ordered subsets of projection data *IEEE Trans. Med. Imaging* **13** 601–9
- Jin K H et al 2017 Deep convolutional neural network for inverse problems in imaging *IEEE Trans. Image Process.* **26** 4509–22
- Kawaguchi K and Sun Q 2021 A recipe for global convergence guarantee in deep neural networks *Proc. of the AAAI Conf. on Artificial Intelligence*
- Kim D et al 2013 Accelerating ordered subsets image reconstruction for x-ray CT using spatially nonuniform optimization transfer *IEEE Trans. Med. Imaging* **32** 1965–78
- Kim H et al 2016 Non-local total-variation (NLTV) minimization combined with reweighted L1-norm for compressed sensing CT reconstruction *Phys. Med. Biol.* **61** 6878–91
- Kim K et al 2014 Sparse-view spectral CT reconstruction using spectral patch-based low-rank penalty *IEEE Trans. Med. Imaging* **34** 748–60
- Kingma D P and Ba J 2015 Adam: a method for stochastic optimization. *Proceedings of the International Conference on Learning Representations*
- Kudo H, Suzuki T and Rashed E A 2013 Image reconstruction for sparse-view CT and interior CT—introduction to compressed sensing and differentiated backprojection *Quant. Imaging Med. Surg.* **3** 147–61
- Lee H et al 2018 Deep-neural-network-based sinogram synthesis for sparse-view CT image reconstruction *IEEE Trans. Radiat. Plasma Med. Sci.* **3** 109–19
- Li R et al 2022 DDPTransformer: dual-domain with parallel transformer network for sparse view CT image reconstruction *IEEE Trans. Computational Imaging* **8** 1101–16
- Liu Y et al 2013 Total variation-stokes strategy for sparse-view x-ray CT image reconstruction *IEEE Trans. Med. Imaging* **33** 749–63
- Lu Y et al 2012 TV-based image reconstruction of multiple objects in a fixed source-detector geometry *J Xray Sci Technol* **20** 277–89
- Magreñán Á A and Argyros I 2018 *A Contemporary Study of Iterative Methods: Convergence, Dynamics and Applications*. (Academic)
- Niu S et al 2014 Sparse-view x-ray CT reconstruction via total generalized variation regularization *Phys. Med. Biol.* **59** 2997–3017
- Rodriguez P and Wohlberg B 2008 An efficient algorithm for sparse representations with l_p data fidelity term *Proc. of 4th IEEE Andean Technical Conf. (ANDESCON)*
- Saad Y 1993 A flexible inner-outer preconditioned GMRES algorithm *SIAM J. Sci. Comput.* **14** 461–9
- Scherzer O 1995 Convergence criteria of iterative methods based on Landweber iteration for solving nonlinear problems *J. Math. Anal. Appl.* **194** 911–33
- Shepp L A and Vardi Y 1982 Maximum likelihood reconstruction for emission tomography *IEEE Trans. Med. Imaging* **1** 113–22
- Singh S et al 2010 Abdominal CT: comparison of adaptive statistical iterative and filtered back projection reconstruction techniques *Radiology* **257** 373–83
- Su T et al 2022 Generalized deep iterative reconstruction for sparse-view CT imaging *Phys. Med. Biol.* **67** 025005

- Takahashi C C and Braga A P 2020 A review of off-line mode dataset shifts *IEEE Comput. Intell. Mag.* **15** 16–27
- van Aarle W *et al* 2016 Fast and flexible x-ray tomography using the ASTRA toolbox *Opt. Express* **24** 25129–47
- Vogel C R 2002 *Computational Methods for Inverse Problems* (SIAM)
- Wang G, Ye J C and De Man B 2020 Deep learning for tomographic image reconstruction *Nat. Mach. Intell.* **2** 737–48
- Wang G, Yu H and De Man B 2008 An outlook on x-ray CT research and development *Med. Phys.* **35** 1051–64
- Wang Z *et al* 2004 Image quality assessment: from error visibility to structural similarity *IEEE Trans. Image Process.* **13** 600–12
- Wu W *et al* 2018 Low-dose spectral CT reconstruction using image gradient $l(0)$ -norm and tensor dictionary *Appl Math Model* **63** 538–57
- Xiang J, Dong Y and Yang Y 2021 FISTA-Net: learning a fast iterative shrinkage thresholding network for inverse problems in imaging *IEEE Trans. Med. Imaging* **40** 1329–39
- Zhang H *et al* 2020 MetaInv-Net: meta inversion network for sparse view CT image reconstruction *IEEE Trans. Med. Imaging* **40** 621–34
- Zhang Z *et al* 2018 A sparse-view ct reconstruction method based on combination of densenet and deconvolution *IEEE Trans. Med. Imaging* **37** 1407–17