Enhanced Reconstruction in Magnetic Particle Imaging by Whitening and Randomized SVD Approximation

Tobias Kluth* Bangti Jin†

March 15, 2019

Abstract

Magnetic particle imaging (MPI) is a medical imaging modality of recent origin, and it exploits the nonlinear magnetization phenomenon to recover a spatially dependent concentration of nanoparticles. In practice, image reconstruction in MPI is frequently carried out by standard Tikhonov regularization with nonnegativity constraint, which is then minimized by a Kaczmarz type method. In this work, we revisit two issues in the numerical reconstruction in MPI in the lens of inverse theory, i.e., the choice of fidelity and acceleration, and propose two algorithmic tricks, i.e., a whitening procedure to incorporate the noise statistics and accelerating Kaczmarz iteration via randomized SVD. The two tricks are straightforward to implement and easy to incorporate in existing reconstruction algorithms. Their significant potentials are illustrated by extensive numerical experiments on a publicly available dataset.

Keywords: magnetic particle imaging, reconstruction, randomized singular value decomposition

1 Introduction

Magnetic particle imaging (MPI) is a relatively new medical imaging modality [11]. It exploits the nonlinear magnetization behavior of (super-)paramagnetic nanoparticles in an applied magnetic field to reconstruct a spatially dependent concentration of nanoparticles. The experimental setup is as follows. A static magnetic field (selection field), given by a gradient field, generates a field free point or a field free line. Its superposition with a spatially homogeneous but time-dependent field (drive field) moves the field free point / line along a predefined trajectory defining the field-of-view. The most common trajectory is the so-called Lissajous curve. The change of the applied field induces a change of the nanoparticle magnetization, which can then be measured and used to recover the concentration of the nanoparticles.

MPI has a number of distinct features: high data acquisition speed, high sensitivity, potentially high spatial resolution and free from the need of harmful radiation. This makes MPI especially attractive for *in-vivo* applications, and the list of potential medical applications is long and growing. The potential for imaging blood flow was demonstrated in *in-vivo* experiments using a healthy mouse [46]. The feasibility of a circulating tracer for long-term monitoring was recently investigated [23]. The high temporal resolution of MPI is shown to be suitable for potential flow estimation [8], tracking medical instruments [15] and tracking and guiding instruments for angioplasty [40]. Further promising applications of MPI include cancer detection [48] and cancer treatment by hyperthermia [36].

Hence, the numerical reconstruction in MPI is of enormous practical importance, and has received much attention. In the literature, there are mainly two different groups of approaches, i.e., data-based v.s. model based, dependent of the description of the forward map. The data-based approach employs experimentally calibrated forward operators, whereas the model-based approach employs mathematical models to describe the physical process. Currently, the former delivers state of the art numerical reconstructions. The model based approach is predominantly based on the equilibrium model [24] or its variation, e.g., the x-space approach [13, 14], Chebyshev polynomials [39] and analytic inversion formulas [34, 6]. For example, x-space

---

[*] Center for Industrial Mathematics, University of Bremen, Bibliothekstr. 5, 28357 Bremen, Germany (`tkluth@math.uni-bremen.de`)

[†] Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK (`b.jin@ucl.ac.uk`, `bangti.jin@gmail.com`)

reconstruction [13, 14] views MPI as a linear shift-invariant imaging system with an analytic point spread function, and obtains fast image reconstruction via a gridding operation in the x-space. In both model-based and data-based approaches, MPI reconstruction techniques often boil down to solving a linear inverse problem using standard regularization techniques (see, e.g., the monographs [5, 42, 19]). The most popular idea is standard Tikhonov regularization with nonnegativity constraint, which is then minimized by Kaczmarz iteration [46, 29, 38]. The Kaczmarz method [22] is very attractive for large volume of data, due to its low operational complexity per iteration. This idea was also combined with preconditioning (row normalization) and row exclusion to improve the image reconstruction [29]. Recently, more advanced variational regularization techniques, e.g., nonnegative fused lasso penalty [43], total least-squares approach [26], approximation error modeling [1] and deep image prior [3], have been proposed and empirically evaluated. The total variation penalty allows recovering piecewise constant concentrations accurately. The approaches in [26, 1] allow incorporating model errors into the reconstruction process for enhanced imaging quality. We refer interested readers to the recent survey [27, Section 6] for an overview of other reconstruction methods. It is worth noting that all these reconstruction techniques can be very expensive for three-dimensional imaging problems, where the available datasets are of relatively large volume. Therefore, there is a significant demand in developing fast MPI image reconstruction algorithms (possibly with improved resolution).

There have been several important efforts [32, 41, 28, 31] in accelerating MPI reconstruction. One idea is to employ sparse approximations of the linear forward map in predefined basis sets, achieved by first applying discrete orthonormal transformations (e.g., Fourier / cosine or Chebyshev transforms) and then thresholding small elements. The sparse approximation enables reducing the computing time of iterative solvers (e.g., CGNE and LSQR) [32] or direct inversion techniques [41]. This idea simultaneously provides a memory-efficient sparse and approximate representation [32, 31, 41]. Alternatively, one can reduce the dimensionality of the forward map using a row selection technique, based on an SNR type quality measure [28] (see Section 2.2 for details). The speedup is achieved by dimension reduction in the data space.

In this work, we revisit two issues in MPI reconstruction, i.e., the choice of fidelity and acceleration, in the lens of inverse theory (see, e.g., [5, 42, 19]) and contribute to the development of robust, accurate and fast reconstruction techniques. First, we highlight the importance of noise covariance in the reconstruction algorithm, and propose a simple whitening procedure from the perspective of maximal likelihood estimation, leading to a least-squares fidelity for the whitened problem. Second, we propose a dimension reduction procedure in the data space to accelerate the benchmark MPI reconstruction algorithm using randomized singular value decomposition (SVD). It exploits the inherent ill-posed nature of the MPI imaging problem, that is, the system matrix admits a low-rank approximation, in order to reduce the effective number of equations. This step can be easily incorporated into any existing algorithms. Third and last, we present extensive numerical experiments on a publicly available dataset, i.e., the "shape" and "resolution" phantoms from `Open MPI dataset` (available at `https://www.tuhh.de/ibi/research/open-mpi-data.html`), to demonstrate the performance of the proposed algorithmic improvements. These represent the main contributions of the work. Our findings include that the whitening step can improve the reconstruction accuracy, and the randomized SVD can accelerate the benchmark algorithm by tens of times. Thus, these techniques may greatly facilitate fast and accurate MPI reconstruction.

The rest of the paper is organized as follows. In Section 2, we discuss the proper formulation of the MPI imaging problem, including system matrix calibration, frequency selection and whitening. In Section 3, we describe the classical reconstruction method based on Kaczmarz iteration and its acceleration via randomized SVD. Then in Section 4, we present extensive numerical results to illustrate the proposed algorithmic tricks. In Section 5, we present concluding remarks and additional discussions. In an appendix, we provide an error estimate of the approximate minimizer to the regularized functional with a low-rank approximation, to justify the acceleration procedure. In the supplementary materials, we provide alternative visualizations of the reconstructions in the inverted colormap, which can occasionally better display the background artifacts.

2 The MPI forward map

The accurate mathematical modeling of MPI is still in its infancy. Several mathematical models have been proposed; see the recent survey [24] for an overview. Nonetheless, state of art numerical reconstructions are achieved by experimentally calibrated forward maps, which we describe in Section 2.1 below. To give a self-contained description accessible to both practitioners and applied mathematicians, we give an abstract formalism of the MPI inverse problem below.

First, we briefly describe the mathematical modeling of the physical process, and point out related modeling challenges that have motivated data-based approaches. Let $\Omega \subset \mathbb{R}^3$ be the spatial domain occupied by the object of interest, and $c : \Omega \to \mathbb{R}_+$ be the concentration of the magnetic nanoparticles. Then the measured voltage signal $v_\ell : I := [0, T] \to \mathbb{R}$, for $\ell = 1, \ldots, L$, obtained at $L \in \mathbb{N}$ receive coils, $0 < T < \infty$, is given by

$$v_\ell(t) = \int_\Omega c(x) \underbrace{\int_I -a_\ell(t-t')\mu_0 p_\ell(x)^t \dot{\bar{m}}(x,t) \mathrm{d}t'}_{=s_\ell(x,t)} \mathrm{d}x + \underbrace{\int_I \int_{\mathbb{R}^3} -a_\ell(t-t')\mu_0 p_\ell(x)^t \dot{H}(x,t) \mathrm{d}x \mathrm{d}t'}_{=v_{\mathrm{E},\ell}(t)}, \tag{2.1}$$

where the superscript t denotes the transpose of a vector (or a matrix), and the notation $\cdot$ denotes taking derivative with respect to the time t. The relevant parameters in the model (2.1) are defined below

- $s_\ell : \Omega \times I \to \mathbb{R}$: system functions characterizing the magnetic behavior of the nanoparticles
- $\bar{m} : \Omega \times I \to \mathbb{R}^3$: mean magnetic moment of the nanoparticles
- $\mu_0 > 0$: magnetic permeability in vacuum
- $a_\ell : \bar{I} := [-T : T] \to \mathbb{R}$: analog filters in the signal acquisition chain
- $p_\ell : \mathbb{R}^3 \to \mathbb{R}^3$: sensitivity profiles of the receive coil units
- $H : \mathbb{R}^3 \times I \to \mathbb{R}^3$: applied magnetic field, which also induces a voltage in the receive coil
- $v_{\mathrm{E},\ell} : I \to \mathbb{R}$: direct feedthrough

The analog filters a_ℓ are employed to filter out the direct feedthrough $v_{\mathrm{E},\ell}$, and in practice, they are commonly taken to be band stop filters adapted to excitation frequencies of the drive field. However, the direct feedthrough $v_{\mathrm{E},\ell}$ is usually not perfectly removed by the analog filters a_ℓ. One big challenge in the modeling is that the analytic forms of the filters a_ℓ are rarely available. A second challenge in the modeling is the mean magnetic moment $\bar{m}$. One often assumes that the moment $\bar{m}$ is independent of the concentration c, and thus ignore possible particle-particle interactions (which is however present for high concentrations [33]). Under this and further assumptions, one popular way to relate the moment $\bar{m}$ to the applied magnetic field H is Langevin theory for paramagnetism, leading to the so-called equilibrium model [24]. These considerations lead to a simplified affine linear forward map $F : X \to Y^L$:

$$c \mapsto \left(\int_\Omega s_\ell(x,t) c(x) \ \mathrm{d}x + v_{\mathrm{E},\ell}(t) \right)_{\ell=1}^{L},$$

for suitable function spaces X and Y, e.g., $X = L^2(\Omega)$, $Y = L^2(I)$, and $\{s_\ell\}_{\ell=1}^L \subset L^2(\Omega \times I)$. The task in MPI is to recover the concentration c from the measured voltages $(v_\ell)_{\ell=1}^L \in Y^L$. However, due to the aforementioned practical complications, the precise kernels s_ℓ are usually unavailable, and instead they are calibrated experimentally for MPI image reconstruction.

2.1 System matrix calibration

First we describe the calibration process for obtaining the system matrix. Let $\Gamma \subset \mathbb{R}^3$ be a reference volume placed at the origin, which is often taken to be a small cube. Then one selects a set of calibration positions $\{x^{(i)}\}_{i=1}^m \subset \Omega$, which are often chosen such that the sets $\{x^{(i)} + \Gamma\}_{i=1}^m$ form a partition of the domain Ω, i.e., they are pairwise disjoint and $\Omega = \cup_{i=1}^m \{x^{(i)} + \Gamma\}$. Let χ_S denote the characteristic function of a set S. Then the set of piecewise constant functions $\{\chi_{x^{(i)}+\Gamma}\}_{i=1}^m$ forms an orthonormal basis (ONB) for a finite-dimensional subspace of $L^2(\Omega)$ (the space consisting of all square integrable functions), which can be used for approximating the concentration c in the domain Ω. In the experiment, a small sample is placed at these predefined grid points $\{x^{(i)}\}_{i=1}^m$, which is described as $c^{(i)} = c_0 \chi_{x^{(i)}+\Gamma}$ for some $c_0 > 0$ and represents one sample volume for calibration.

The measurements $\{v_\ell^{(i)} = F_\ell c^{(i)}\}_{i=1}^m$, $\ell = 1, \ldots, L$, are then used to characterize the discrete data-based forward operator via a discrete system matrix. Mathematically, this procedure can be formulated using the following map

$$\begin{aligned} Q_n : L^2(I)^L &\to \mathbb{R}^{n:=\sum_{i=1}^L 2|J_i|} \\ (v_\ell)_{\ell=1}^L &\mapsto \left[\begin{array}{c} \begin{pmatrix} \mathrm{Re}(\langle v_1, \psi_j\rangle)_{j\in J_1} \\ \mathrm{Im}(\langle v_1, \psi_j\rangle)_{j\in J_1} \end{pmatrix} \\ \hline \vdots \\ \hline \begin{pmatrix} \mathrm{Re}(\langle v_L, \psi_j\rangle)_{j\in J_L} \\ \mathrm{Im}(\langle v_L, \psi_j\rangle)_{j\in J_L} \end{pmatrix} \end{array}\right], \end{aligned} \tag{2.2}$$

where $\{\psi_j\}_{j\in\mathbb{N}} \subset L^2(I)$ is an ONB of $L^2(I)$, which is commonly taken to be the Fourier basis of time-periodic signals in $L^2(I)$, i.e. $\psi_j(t) = T^{-\frac{1}{2}}(-1)^j e^{i2\pi jt/T}$. The finite index sets $\{J_\ell\}_{\ell=1}^L \subset \mathbb{Z}$ serve as a preprocessing step prior to image reconstruction, to be described below. Note that the ordering of the indices for real and imaginary parts can be shuffled, and one may also employ complex measurements directly. Upon minor changes, the techniques to be developed below can still be applied.

The map Q_n in (2.2) consists of concatenating multiple receive coil signals, splitting real and imaginary parts (if necessary), index / frequency selection, and discretization via projection onto a finite subset of the ONB $\{\psi_j\}$ (indexed by J_ℓ). The system matrix S is then given by

$$S = \left[\; Q_n((v_\ell^{(1)})_\ell) \;\middle|\; \ldots \;\middle|\; Q_n((v_\ell^{(m)})_\ell) \;\right] \in \mathbb{R}^{n\times m}. \tag{2.3}$$

For the measured signals $\{v_\ell\}_{\ell=1}^L$, we build the measurement vector $v = Q_n((v_\ell)_{\ell=1}^L)$ analogously.

The background measurement $v^{(0)} = F\mathbf{0}$ (or more precisely, the mean over multiple measurements) is used to remove the influence of the direct feedthrough $v_{\mathrm{E},\ell}$. Then by subtracting the vector $v_0 = Q_n((v_\ell^{(0)})_{\ell=1}^L)$ and the rank-one matrix $S_0 = v_0\mathbf{1}_m^t$ ($\mathbf{1}_m \in \mathbb{R}^m$ is a vector with all its entries equal to unit), we obtain the following linear MPI reconstruction problem

$$Ax = y, \tag{2.4}$$

where $A \in \mathbb{R}^{n\times m}$, $y \in \mathbb{R}^n$ and $x \in \mathbb{R}^m$ are defined by

$$A = S - S_0, \quad y = v - v_0, \quad \text{and} \quad c = c_0 \sum_{i=1}^m x_i \chi_{x^{(i)}+\Gamma}.$$

In this construction, we have implicitly used a piecewise constant representation of the concentration c, with $c_0 x_i$ being the concentration c on the cell $x^{(i)} + \Gamma$. Problem (2.4) is the starting point of the proposed algorithmic tricks, e.g., whitening and low-rank approximation, which are clearly oblivious to the calibration process.

It is worth noting that the calibration procedure is laborious, time consuming and highly problem dependent, and has limited spatial resolution. For example, it requires a recalibration whenever the experimental setting changes. Therefore, there is a huge demand in developing accurate model-based approaches or hybrid approaches for MPI image reconstruction. We refer interested readers to [24] for relevant mathematical models and [34, 6, 25] for preliminary mathematical analysis.

Remark 2.1. *Note that several different strategies have been proposed to perform the background subtraction for the system matrix in the literature [46, 45], which may require additional effort during the system matrix calibration, in view of costly robot movements.*

2.2 Frequency selection

In MPI there are two standard preprocessing approaches, i.e., band pass approach and SNR-type thresholding, and they are often combined via the index sets $\{J_\ell\}_{\ell=1}^L$. Let $J_{\mathrm{BP}} = \{j \in \mathbb{Z} | \; b_1 \le |j|/T \le b_2\}$ be the band pass indices for frequency band limits $0 \le b_1 < b_2 \le \infty$. The main purpose of band pass is to filter out the direct feedthrough $v_{\mathrm{E},\ell}$ (although not perfectly), and outside the frequency band J_{BP}, the signal is deemed to be too noisy and thus discarded. For the SNR-type thresholding, one standard quality measure is given by a ratio of mean absolute values from individual measurements $v_\ell^{(i)}$ (cf. Section 2.1) and a set of empty scanner measurements $\{v_{\ell,0}^{(k)}\}_{k=1}^K$ [7] obtained during the calibration process. Specifically, let $I_{\mathrm{SNR}} \subset \{1, \ldots, N\}$ be the index set of individual measurements with calibration positions being an element of the smallest cuboid region containing the trajectory of the field free point. Then we define

$$d_{\ell,j} = \frac{\frac{1}{|I_{\mathrm{SNR}}|} \sum_{i \in I_{\mathrm{SNR}}} |\langle v_\ell^{(i)} - \mu_\ell^{(i)}, \psi_j \rangle|}{\frac{1}{K} \sum_{k=1}^K |\langle v_{\ell,0}^{(k)} - \mu_\ell, \psi_j \rangle|}, \tag{2.5}$$

where $\mu_\ell = \frac{1}{K} \sum_{k=1}^K v_{\ell,0}^{(k)}$ is the mean measurement, and $\mu_\ell^{(i)} = \kappa_i v_{\ell,0}^{(k_i)} + (1 - \kappa_i) v_{\ell,0}^{(k_i+1)}$ is a convex combination of the previous (k_i-th) and following ($k_i + 1$-th) empty scanner measurement with respect to the i-th calibration scan. The parameters $\kappa_i \in [0, 1]$ are chosen equidistant for all calibration scans between two subsequent empty scanner measurements. Then for a given threshold $\tau \geq 0$, we define

$$J_\ell = \{j \in J_{\mathrm{BP}} | d_{\ell,j} \geq \tau\}, \quad \ell = 1, \ldots, L. \tag{2.6}$$

Remark 2.2. *The SNR-type thresholding was also used to obtain a dimensionality reduction in the system of linear equations in [28] to enable online reconstruction.*

2.3 Whitening

The calibration process leads to a linear inverse problem

$$Ax = y^\delta \quad \text{with } y^\delta = y^\dagger + \eta,$$

where η denotes the noise in the data, due to the imperfect data acquisition process. In practice, it is often assumed to follow a Gaussian distribution $N(\mu, C)$ with mean $\mu \in \mathbb{R}^n$ and covariance $C \in \mathbb{R}^{n \times n}$ (real symmetric positive semidefinite), upon invoking the central limit theorem (for repetitive measurements). These statistical parameters μ and C are then estimated from repetitive measurements. The mean μ is often approximately zero after background subtraction. The full covariance matrix C has a large number of parameters, and requires a large volume of data for a reliable estimation, which is not necessarily available in practice. One often imposes suitable structures on the covariance C, e.g., diagonal covariance, or uses more advanced options, e.g., sparse inverse covariance [9], to facilitate the estimation. In MPI experiments, the covariance C is often not a scalar multiple of the identity matrix (i.e., the noise components are not necessarily independent and identically distributed). Thus it is important to exploit the structure of the covariance C in image reconstruction, in the spirit of statistical inference. This can be achieved using a whitening matrix W such that $W(\eta - \mu)$ follows a zero mean Gaussian distribution with identity covariance. The whitening matrix W can be determined from the eigendecomposition (Q, Λ) of the covariance C (i.e., $C = Q\Lambda Q^t$) by $W = \Lambda^{-\frac{1}{2}} Q^t$. Alternatively one may employ the Cholesky decomposition $C = LL^t$ ($L \in \mathbb{R}^{n \times n}$ is lower triangular and $W = L^{-t}$) to whiten the noise. Then we arrive at the following linear problem

$$WAx = W(y^\delta - \mu). \tag{2.7}$$

The whitening step enables the use of the standard least-squares formulation in MPI reconstruction, in the spirit of the classical maximum likelihood approach, i.e.,

$$\|WAx - W(y^\delta - \mu)\|^2. \tag{2.8}$$

Conceptually, a large variance indicates that the corresponding measurement may be not so reliable, and thus may behave like an outlier within the dataset, for which an inadvertent use of the standard least-squares formulation may significantly compromise the reconstruction accuracy. Instead, it should be weighed down in the reconstruction step, which is precisely the role played by the whitening step. Clearly, the whitening in (2.8) is equivalent to the weighted least-squares $(Ax - (y^\delta - \mu))^t C^{-1} (Ax - (y^\delta - \mu))$, which corresponds to the maximum likelihood estimate for the data y^δ. This formulation also properly accounts for the noise statistics. However, the explicit whitening construction is advantageous for accelerating reconstruction via the randomized SVD described in Section 3.2 below.

Remark 2.3. *Weighting was also used in [29], with the weight w_k given by the energy of the kth row a_k of the system matrix A, i.e., $w_k = \|a_k\|^{-2}$. Thus the weighting in [29] represents a form of preconditioning, which differs from the covariance interpretation in* (2.7)*, despite the formal similarity. We refer also to [44] for a related experimental study on several different weighting schemes.*

For a calibrated system matrix A as in Section 2.1, the whitening process has to be adapted properly. Specifically, for an ONB $\{b_i\}_i \subset X$, we have

$$w_i = A^\dagger b_i + \eta_i,$$

where $A^\dagger : X \to Y$ denotes the (unknown) true forward map and η_i follow the same distribution as the noise η, i.e., $(\eta_i - \mu) \sim N(0, C)$. Then the (mean) corrected and noisy forward map A is given by

$$Ax = \sum_i \langle x, b_i \rangle (w_i - \mu) = A^\dagger x + \sum_i \langle x, b_i \rangle (\eta_i - \mu).$$

Thus, the noise term due to modeling error (in the forward map A) (relative to the exact one $A^\dagger$) is given by $\sum_i \langle x, b_i \rangle (\eta_i - \mu)$. The statistics of this term is actually dependent of the unknown concentration x: the mean is still zero, but the covariance is changed via a linear map dependent of x. In practical inversion, this error term is often lumped into the data error, and combined with the measurement error in the data y^δ, whose noise statistics are then x-dependent. This short discussion highlights the distinct role of modeling errors in data-based approaches. In the discussions below, we shall ignore model errors (for the acceleration step) and employ the whitening procedure described above. Clearly, more suitable approaches should employ alternatives, e.g., total least-squares approach [26] or approximation error modeling [1], which are, however, beyond the scope of this work.

3 Enhanced image reconstruction

Now we describe the common MPI reconstruction method, and its acceleration via randomized SVD.

3.1 The common approach

Currently, the most popular and successful idea in MPI reconstruction is based on the following constrained Tikhonov regularization with a quadratic penalty:

$$x^\dagger = \arg\min_{x \geq 0} \|Ax - y^\delta\|^2 + \alpha \|x\|^2, \tag{3.1}$$

where $\alpha > 0$ is the regularization parameter, controlling the tradeoff between the two terms [19]. The proper choice of the parameter α is notoriously challenging, and many rules have been proposed, e.g., L-curve criterion [17], discrepancy principle [35], balancing principle [19] and quasi-optimality criterion [10, 21], which however have not been extensively studied within MPI reconstruction. The nonnegativity constraint $x \geq 0$ is understood componentwise, and reflects the fact that the concentration x is nonnegative. The constraint is essential for obtaining physically meaningful reconstructions. The whitening approach in Section 2.3 may be implemented in the form (3.1) straightforwardly by penalizing the fidelity functional in equation (2.8), and thus all the discussions below adapt accordingly.

In practice, a variant of the popular Kaczmarz method [22], developed in [2], is often employed for solving the constrained optimization problem (3.1) in the MPI reconstruction. It has demonstrated excellent empirical performance [46, 29, 38], and has been implemented in commercial MPI scanners (included in ParaVision$^{\circledR}$ (Bruker BioSpin MRI GmbH, Germany) as reported in [8]). One distinct feature of the Kaczmarz method is that at each iteration, it operates only on one equation, instead of the whole linear system, and thus its computational complexity per iteration is nearly independent of the amount of data. This feature makes the algorithm especially attractive for problems with large datasets e.g., 3D MPI, and traditionally, it has been very successful within the computed tomography community [18, 37, 20]. The complete procedure of the variant in [2] is given in Algorithm 1. The algorithm often reaches the desired convergence within tens of sweeps through the equations, and thus its complexity is roughly proportional to the number n of rows in the matrix A. We shall employ it as the benchmark algorithm, and propose a preprocessing step to accelerate the computation.

Algorithm 1 Kaczmarz method for problem (3.1).

```
Input matrix $A \in \mathbb{R}^{n\times m}$, $y^\delta \in \mathbb{R}^n$, and $\alpha > 0$
Optional: initial value $x_0 \in \mathbb{R}^m$ (0 default), relaxation parameter $\omega \in (0, 2)$ (1 default);
Initialize $x = x_0$, $z = 0 \in \mathbb{R}^n$, $\bar{z} = 0 \in \mathbb{R}^m$;
for $k = 1, \ldots, K$ do
    $i = (k \bmod n) + 1$;    \\row index
    $\eta = -\omega \frac{\langle a_i, x\rangle + \sqrt{\alpha} z_i - y_i^\delta}{\|a_i\|^2 + \alpha}$;    \\$a_i$ is $i$-th row of $A$
    $z_i \leftarrow z_i + \eta\sqrt{\alpha}$;
    $x \leftarrow x + \eta a_i^t$;
    if $i = n$ or $k = K$ then
        $\bar{\eta} = -(\min(\bar{z}_j, \omega x_j))_{j=1,\ldots,m}$;
        $\bar{z} \leftarrow \bar{z} + \bar{\eta}$;
        $x \leftarrow x + \bar{\eta}$;    \\positivity constraint
    end if
end for
Return the approximation $x_K \leftarrow x$.
```

3.2 Acceleration by randomized SVD

Now we describe a simple acceleration method for Algorithm 1 based on randomized singular value decomposition (SVD). Recall that SVD of a matrix $A \in \mathbb{R}^{n\times m}$ is given by

$$A = U\Sigma V^t,$$

where $U = [u_1\ u_2\ \ldots\ u_n] \in \mathbb{R}^{n\times n}$ and $V = [v_1\ v_2\ \ldots\ v_m] \in \mathbb{R}^{m\times m}$ are column orthonormal matrices, $\Sigma \in \mathbb{R}^{n\times m}$ is a diagonal matrix, with the diagonal entries ordered in a nonincreasing manner: $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > \sigma_{r+1} = \ldots = \sigma_{\min(m,n)}$, where r is the rank of the matrix A. Traditional methods for computing SVD, e.g., Lanczos bidiagonalization, are not attractive for general dense matrices as arising in MPI. For example, the complexity of Golub-Reinsch algorithm for computing SVD is $4n^2m + 8m^2n + 9m^3$ (for $n \geq m$) [12, p. 254]. Thus, it can be prohibitively expensive for large-scale matrices. The randomized SVD (rSVD) provides an efficient way to construct a low-rank approximation by randomly mixing the columns of A [16]. The procedure is given in Algorithm 2 for the case $n \geq m$, and the case $n < m$ can be obtained by applying Algorithm 2 to the transposed matrix A^t.

Algorithm 2 rSVD for $A \in \mathbb{R}^{n\times m}$, $n \geq m$.

```
1: Input matrix $A \in \mathbb{R}^{n\times m}$, $n \geq m$, and target rank $k$;
2: Set parameters $p$ (default $p = 5$), and $q$ (default $q = 0$);
3: Sample a random matrix $\Omega = (\omega_{ij}) \in \mathbb{R}^{m\times(k+p)}$, with $\omega_{ij} \sim N(0, 1)$;
4: Compute the randomized matrix $Y = (AA^*)^q A\Omega$;
5: Find an orthonormal basis $Q$ of range($Y$);
6: Form the matrix $B = Q^*A$;
7: Compute the SVD of $B = WSV^*$;
8: Return the rank $k$ approximation $(\tilde{U}_k, \tilde{\Sigma}_k, \tilde{V}_k)$, cf. (3.2).
```

In Algorithm 2, Step 4 is to extract the column space $\mathcal{R}(A)$ of A, i.e., $\mathcal{R}(Y) \subset \mathcal{R}(A)$, and Step 5 is to find an orthonormal basis for $\mathcal{R}(Y)$, e.g., via QR decompositon or skinny SVD. The remaining steps can be regarded as one subspace iteration for computing SVD of the matrix QQ^tA. Since the involved matrices are of much smaller size, these SVDs can be carried out efficiently. The accuracy of $\mathcal{R}(Y)$ to $\mathcal{R}(A)$ is crucial to the success of the algorithm. A positive exponent q can improve the accuracy when the singular values of A decay slowly, and the oversampling parameter p is to improve the accuracy of the range probing. The low-rank approximation $\tilde{A}_k$ by rSVD is given by

$$\tilde{A}_k = \tilde{U}_k\tilde{\Sigma}_k\tilde{V}_k^t, \text{ with } \tilde{U}_k = (QW)_{:,1:k}, \quad \tilde{\Sigma}_k = S_{1:k,1:k}, \quad \tilde{V}_k = V_{:,1:k}, \tag{3.2}$$

where the notation $1 : k$ denotes taking the first k columns/rows of the matrix. The complexity of Algorithm 2 is around $4(q+1)kmn$, which is lower than computing SVD of A directly. Clearly, the efficiency of the approach relies crucially on the low-rank structure of A. In the context of MPI, it was rigorous justified for the equilibrium model in [25]: it is severely ill-posed for common experimental setups.

With the rSVD $(\tilde{U}_k, \tilde{\Sigma}_k, \tilde{V}_k)$ at hand, we approximate problem (3.1) by

$$\begin{aligned} &\arg\min_{x\geq 0} \|Ax - y^\delta\|^2 + \alpha\|x\|^2 \\ \approx\; &\arg\min_{x\geq 0} \|\tilde{U}_k\tilde{\Sigma}_k\tilde{V}_k^t x - y^\delta\|^2 + \alpha\|x\|^2 \\ =\; &\arg\min_{x\geq 0} \|\tilde{\Sigma}_k\tilde{V}_k^t x - \tilde{U}_k^t y^\delta\|^2 + \alpha\|x\|^2. \end{aligned} \tag{3.3}$$

The number k of rows in problem (3.3) is much smaller than n, enabling a significant speedup of Algorithm 1. In essence, rSVD is a preprocessing step to extract essential information content in A, and can be viewed as a dimensionality reduction strategy in the data space; see Appendix A for error estimates on the minimizer due to the low-rank approximation. Note that this step does not alter the whole reconstruction procedure.

The low-rank approximation $(\tilde{U}_k, \tilde{\Sigma}_k\, \tilde{V}_k)$ (actually the factors $(\tilde{U}_k, \tilde{\Sigma}_k\tilde{V}_k^t)$ are sufficient) essentially underpins the acceleration in the formulation (3.3). In principle, any numerical method for constructing low-rank approximations can be employed, and rSVD given in Algorithm 2 represents a computationally attractive choice, which is more efficient than the standard SVD computation. Further, an approximate optimization problem can be constructed by means of sketching [47], which employs a random / deterministic matrix to approximately preserve the magnitude of the fidelity while reducing the dimension of the data space, which will not be explored in this work.

So far we have described an acceleration procedure for problem (3.1). It applies equally well to the whitened problem (2.7) derived in Section 2.3. This can be achieved simply by applying rSVD to the whitened matrix WA to construct an accurate low-rank approximation. The whitening matrix W may influence the spectral behavior of WA, which generally is different from that of A.

4 Numerical results and discussions

In this section, we present numerical results to illustrate the proposed algorithmic tricks, i.e., whitening and acceleration. The experimental setup is as follows. We employ a measured system matrix, where a band pass filter is applied (with $b_1 = 80$ kHz and $b_2 = 625$ kHz), which yields a system matrix $A \in \mathbb{R}^{n\times m}$ for the $L = 3$ receive channels. Background measurements (with the same band filter) are used to obtain a diagonal whitening operator $W \in \mathbb{R}^{n\times n}$ and thus also the whitened matrix $A_W = WA \in \mathbb{R}^{n\times m}$ (for the whitening approach). Let $(\tilde{U}_k, \tilde{\Sigma}_k, \tilde{V}_k)$ be the rSVD of A given by Algorithm 2, and analogously $(\tilde{U}_{W;k}, \tilde{\Sigma}_{W;k}, \tilde{V}_{W;k})$ for the rSVD of A_W. All forward maps are scaled to have a unit operator norm.

Below we compare the reconstructions by the proposed method with that by the standard Kaczmarz method (i.e., Algorithm 1) and the dimensionality reduction method proposed in [28]. Specifically, we consider the following four reconstruction methods:

- **[STD]**: The reconstructions x_{STD} and $x_{W;\mathrm{STD}}$ are respectively obtained by

  $$x_{\mathrm{STD}} = \arg\min_{x\geq 0} \|Ax - y^\delta\|^2 + \alpha\|x\|^2 \quad \text{and} \quad x_{W;\mathrm{STD}} = \arg\min_{x\geq 0} \|A_W x - Wy^\delta\|^2 + \alpha\|x\|^2,$$

  with Algorithm 1; cf. problem (3.1).

- **[SNR]**: For a given $k \in \mathbb{N}$, there exists a $\tau_k > 0$ such that a reduced system with k rows is obtained via the SNR-type frequency selection for $\tau = \tau_k$ (cf. Section 2.2) using Q_k to build the system matrix A_k and the measurement vector y_k^δ as proposed for online reconstruction in [28]. Then the diagonal whitening operator W_k is determined for the reduced system. The reconstructions x_{SNR} and $x_{W_k,\mathrm{SNR}}$ are respectively obtained by

  $$x_{\mathrm{SNR}} = \arg\min_{x\geq 0} \|A_k x - y_k^\delta\|^2 + \alpha\|x\|^2 \quad \text{and} \quad x_{W_k;\mathrm{SNR}} = \arg\min_{x\geq 0} \|A_{k;W_k} x - W_k y_k^\delta\|^2 + \alpha\|x\|^2,$$

  with Algorithm 1; cf. problem (3.1).

- **[rSVD1]**: The (rSVD) reconstructions x_{rSVD1} and $x_{W;\mathrm{rSVD1}}$ are respectively obtained by

  $$x_{\mathrm{rSVD1}} = \arg\min_{x\geq 0} \|\tilde{\Sigma}_k \tilde{V}_k^t x - \tilde{U}_k^t y^\delta\|^2 + \alpha\|x\|^2,$$
  $$x_{W;\mathrm{rSVD1}} = \arg\min_{x\geq 0} \|\tilde{\Sigma}_{W;k} \tilde{V}_{W;k}^t x - \tilde{U}_{W;k}^t W y^\delta\|^2 + \alpha\|x\|^2,$$

  with Algorithm 1 for given $k \in \mathbb{N}$; cf. problem (3.3).

- **[rSVD2]**: The reconstruction x_{rSVD2} is computed via

  $$x_{\mathrm{rSVD2}} = P_{\mathbb{R}_+^m} \tilde{V}_k \tilde{\Sigma}_k^{-1;\alpha} \tilde{U}_k^t y^\delta,$$

  with $\tilde{\Sigma}_k^{-1;\alpha} = \mathrm{diag}(\tilde{\Sigma}_{k;ii}/(\tilde{\Sigma}_{k;ii}^2 + \alpha^2)) \in \mathbb{R}^{k\times k}$, where the notation $P_{\mathbb{R}_+^m}$ denotes the projection into the set $\mathbb{R}_+^m$ which consists of vectors in $\mathbb{R}^m$ with all entries nonnegative. The reconstruction $x_{W;\mathrm{rSVD2}}$ is obtained similarly. These methods treat the nonnegativity constraint in an ad hoc manner, and can be used as rough approximations to x_{rSVD1} and $x_{\mathrm{W;rSVD1}}$, and can be viewed as close cousins of truncated SVD, in the form of pseudo-inverse from the randomized SVD using Tikhonov filters. Note that rSVD1 and rSVD2 are drastically different in nature: rSVD1 is an iterative reconstruction method (based on Kaczmarz iteration), whereas rSVD2 is a direct one.

The dimension reduction in the methods SNR, rSVD1 and rSVD2 may induce extra regularization besides the penalty $\|x\|^2$ (and nonnegativity constraint), in the spirit of the classical truncated SVD [5], when the target dimension k is small. This different source of regularization can significantly complicate the study (e.g., parameter tuning).

These reconstruction methods are evaluated on a publicly available 3D dataset, i.e., `open MPI dataset` (downloaded from `https://www.tuhh.de/ibi/research/open-mpi-data.html`, last accessed on January 19, 2019) provided in the MPI Data Format (MDF) [30]. The (measured) system matrix data $\{v_\ell^{(i)}\}_{i=1}^m$, $\ell = 1, 2, 3$, is obtained using a cuboid sample of size 2 mm × 2 mm × 1 mm. The calibration is carried out with Perimag® tracer with a concentration 100 mmol/l. The field-of-view has a size of 38 mm × 38 mm × 19 mm and the sample positions have a distance of 2 mm in x- and y-direction and 1 mm in z-direction, resulting in $19 \times 19 \times 19 = 6859$ voxels, which gives the number m of columns in the system matrix A. The entries of A are averaged over 1000 repetitions and empty scanner measurements are performed and averaged every 19 calibration scans. The measurements are averaged over 1000 repetitions of the excitation sequence, and with each phantom, an empty measurement with 1000 repetitions is provided, which are used for the background correction of the measurement and the system matrix A (cf. Section 2.1) and also for the approximation of the covariance C respectively the whitening matrix W (see Section 2.3). For the experiments presented below, the Kaczmarz method, i.e., Algorithm 1, is always run for 20 loops over the system, and numerically it is observed to be sufficient for ensuring convergence. Thus, there is no substantial regularizing effect arising from possibly early stopping of Kaczmarz iteration.

We validate the proposed methods on the "shape" and "resolution" phantoms in the dataset. The "shape" phantom is a cone defined by a 1 mm radius tip, an apex angle of 10 degree, and a height of 22 mm. The total volume is 683.9 μl. Perimag® tracer with a concentration of 50 mmol/l is used. See Fig. 1 for a schematic illustration of the phantom and the visualization structure of the 3D reconstructions below, where the pictures are partially adapted from `Open MPI dataset`. The "resolution" phantom consists of 5 tubes filled with Perimag® tracer with a concentration of 5 mmol/l. The 5 tubes have a common origin on one side of the phantom. The tubes extend in different angles from the origin within the x-y- and the y-z-planes. In z-direction, the angles in the y-z-plane are chosen smaller (10 deg and 15 deg) than in x-y-plane (20 deg and 30 deg); see Fig. 2 for the illustration. In all the reconstructions below, the concentration unit is mmol/l. See also the supplementary material for the reconstructions in an inverted colormap, which occasionally allow better display of the background artifacts.

4.1 The benefit of whitening

First, we illustrate the benefit of whitening in the standard reconstruction technique (i.e., STD). Due to the small number of repetitions of the empty measurements (1000 repetitions compared to 23482 indices in

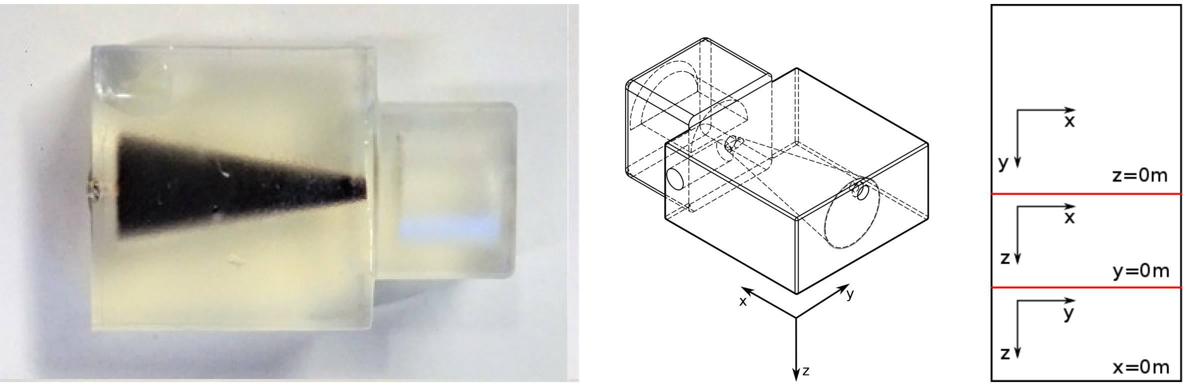


Figure 1: “Shape” phantom from the `open MPI dataset` (left, middle) and visualization structure for the 3D reconstructions (right).

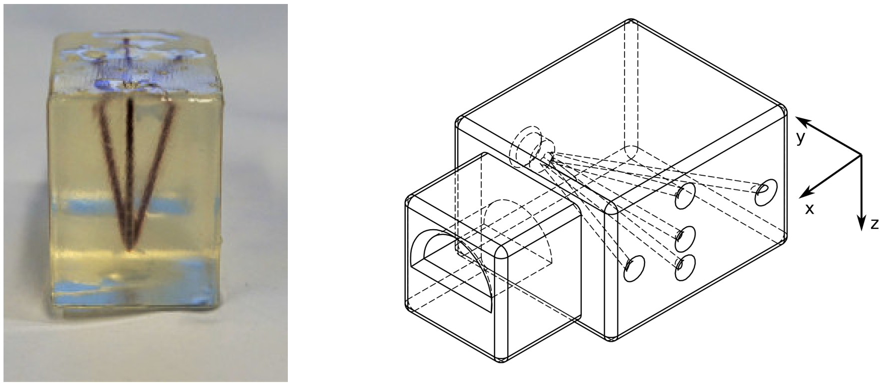


Figure 2: “Resolution” phantom from the `open MPI dataset`.

the band limits for each receive coil, when assuming the receive coils are independent), the covariance C is approximated by a diagonal one to ensure a reliable estimation. The estimated (diagonal) covariance C for the “shape” phantom is shown in Fig. 3. The magnitude of the noise variance is observed to vary dramatically with the frequency over the frequency band for both real and imaginary parts, and the behavior is similar for all three receive coils. The heteroscedastic nature of the noise necessitates the use of the whitening procedure in the reconstruction algorithm in Section 2.3.

The STD reconstructions for the non-whitened and whitened cases are shown in Figs. 4 and 5 for the “shape” and “resolution” phantoms, respectively. The results are presented for three different α values, including the cases of over, medium and under regularization, respectively. For the medium and small α values, the reconstructed phantoms for both non-whitened and whitened cases are of similar quality; see the middle and right columns of the figures. However, a closer inspection of Fig. 4 shows that the reconstruction in the non-whitened case suffers from pronounced background artifacts, whereas, in the whitened case, the artifacts can be greatly reduced even for much smaller α values. Meanwhile, for a large α value ($\alpha = 2^{-10}$) (the left column), the background artifacts largely disappear from the reconstructions in the non-whitened case but also the reconstructed cone tends to be overly smooth and to suffer from undesirable smearing effect, due to over-regularization introduced by the penalty; and these observations hold also for the whitened case. Thus, the whitening step makes the reconstruction algorithm more robust to the choice of the α value, which is highly desirable in practice, since its optimal choice is generally very challenging.

In summary, the STD reconstructions in Figs. 4 and 5 have similar quality in the whitened and non-whited cases, except some smaller background artifacts for the non-whitened approach.

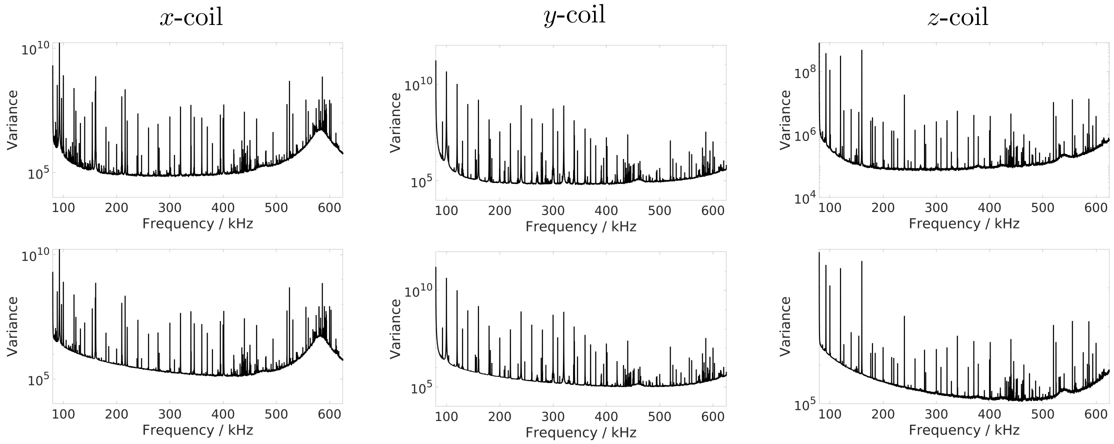


Figure 3: Variance structure of the diagonal covariance matrix C for the "shape" phantom. Visualized individually for each receive coil with respect to the frequency; real part (top), imaginary part (bottom).

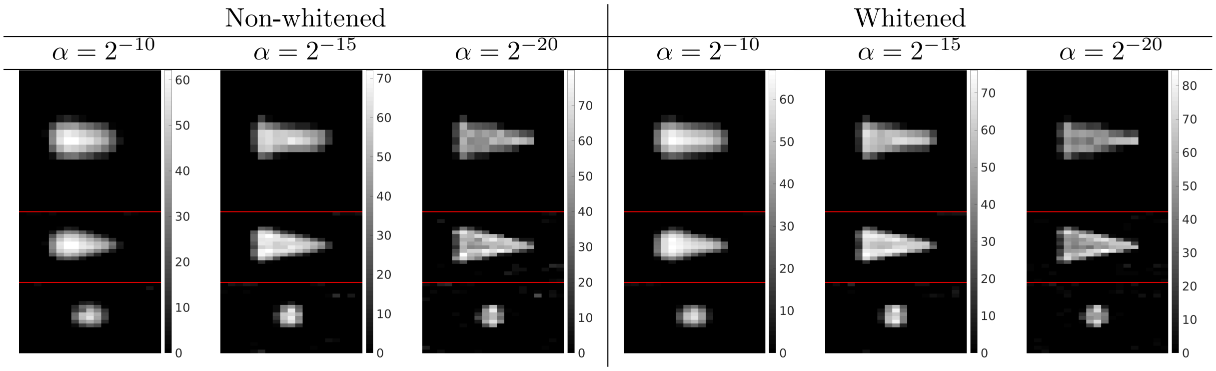


Figure 4: Non-whitened/whitened STD reconstructions of the "shape" phantom.

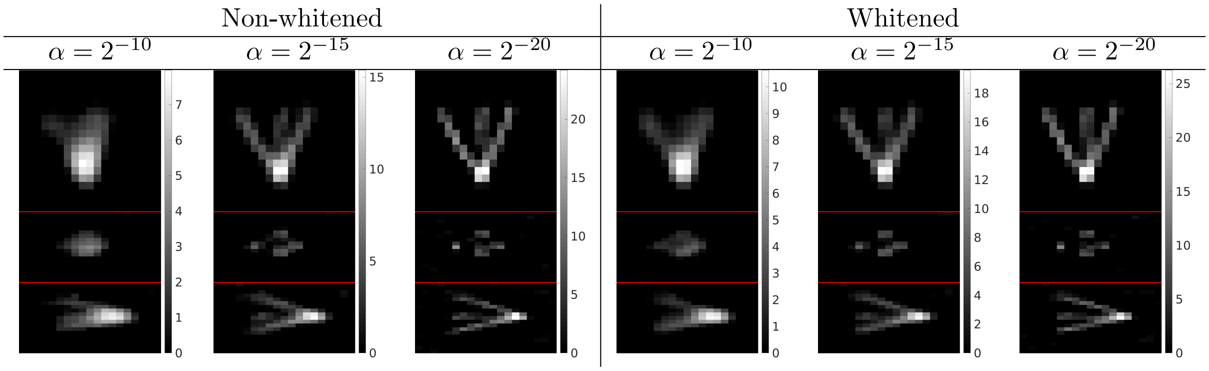
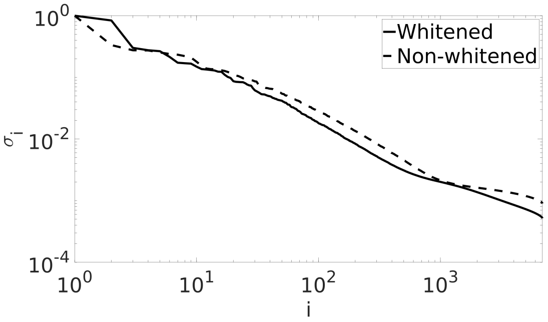


Figure 5: Non-whitened/whitened STD reconstruction of the "resolution" phantom.

4.2 Acceleration via randomized SVD

Now we illustrate randomized SVD for accelerating the Kaczmarz algorithm, and discuss its interplay with whitening. In Fig. 6, we plot the singular values (SVs) of the non-whitened and whitened system matrices for the "shape" phantom. The SVs decay algebraically with comparable decay rates for the whitened and non-whitened cases, indicating that the MPI inverse problem is mildly ill-posed. A useful quantitative measure of the low-rank approximation is the "energy" percentage $(\sum_{i=1}^{k}\sigma_i^2)/(\sum_{i=1}^{m}\sigma_i^2)$, which roughly corresponds to the optimal error bound on the rank-k approximation in the Frobenius norm. According to the table in



k	$(\sum_{i=1}^{k}\sigma_i^2)/(\sum_{i=1}^{m}\sigma_i^2)$ Whitened	Non-whitened
500	99.23	99.61
1000	99.46	99.73
1500	99.55	99.80
2000	99.61	99.85

Figure 6: Illustration of singular value decay (left) of the system $A \in \mathbb{R}^{n\times m}$ for the "shape" phantom. Table (right) including energy percentage $(\sum_{i=1}^{k}\sigma_i^2)/(\sum_{i=1}^{m}\sigma_i^2)$ for low rank approximations of $A \in \mathbb{R}^{n\times m}$; relevant for constructing randomized SVD approximation.

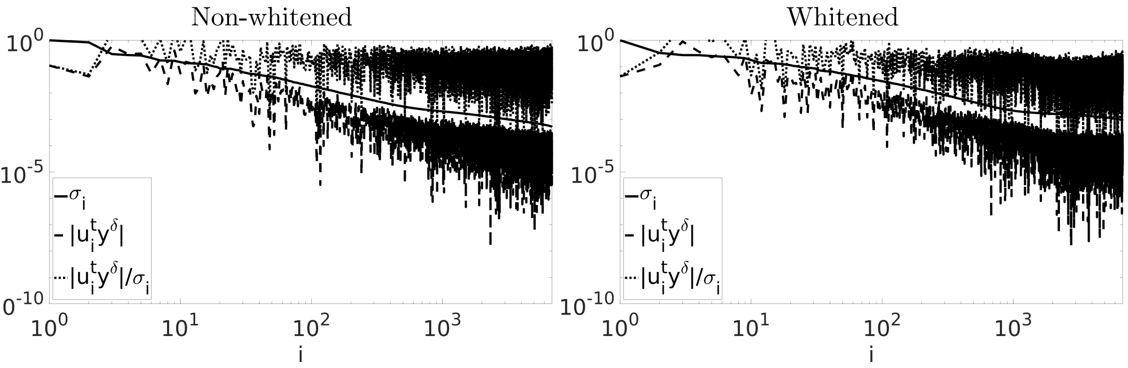


Figure 7: Picard plots for the "shape" phantom.

Fig. 6, five hundred SVs capture nearly all the energies for both whitened and non-whitened cases and thus can give an accurate low rank approximation; although in the non-whitened case, the same number of SVs can capture slightly more energy percentage than that for the whitened case. The decay behavior justifies the use of the rSVD approach for accelerating the algorithm in Section 3.2. In MPI, the SV decay was rigorously proved for simplified models in [6] and [25] for the one-dimensional and multi-dimensional cases, respectively. In Fig. 7, we plot the quantities $|u_i^t y^\delta|$ and $|u_i^t y^\delta|/\sigma_i$ for the "shape" phantom, which is often known as the discrete Picard plot and used for diagnosing the behavior of discrete inverse problems. The quantity $|u_i^t y^\delta|/\sigma_i$ tends to blow up as the index i increases, indicating the ill-posed nature of the inverse problem and necessitating regularization. Although not presented, similar behavior on the SV spectrum and Picard plot can be observed for the "resolution" phantom.

In view of the SV decay in Fig. 6 and the energy percentage in the table therein, two truncation numbers, i.e., $k = 500$ and $k = 1000$, are employed below for the accelerated reconstruction for both phantoms, which both can capture a sufficient percentage of energy. The numerical reconstructions for the "shape" and "resolution" phantoms are presented in Figs. 8–10 and Figs. 11–13, respectively, for three different α values, i.e., 2^{-20}, 2^{-15} and 2^{-10}, used in Figs. 4 and 5.

With the α value properly chosen (i.e., $\alpha = 2^{-15}$), rSVD1 can provide reasonable reconstructions in the whitened case for both k values (cf. Figs. 9 and 12), and the reconstructions are comparable with that by STD in Figs. 4 and 5. However, in the non-whitenend case, slight blurring appears in the reconstructed cone. With the choice $k = 1000$, SNR gives comparable reconstructions, but with $k = 500$, either significant distortions or smoothing appear in the reconstructions for all three α values, and thus the choice $k = 500$ in SNR seems insufficient to capture the essential information of the data. Thus, rSVD is more effective in compressing the data than SNR. Somewhat surprisingly, rSVD2, the simplest and fastest approach, can also provide reasonable reconstructions of comparable quality, even in the over-regularized case, but in the non-whitenend case, it gives reconstructions containing pronounced background artifacts, which, however,

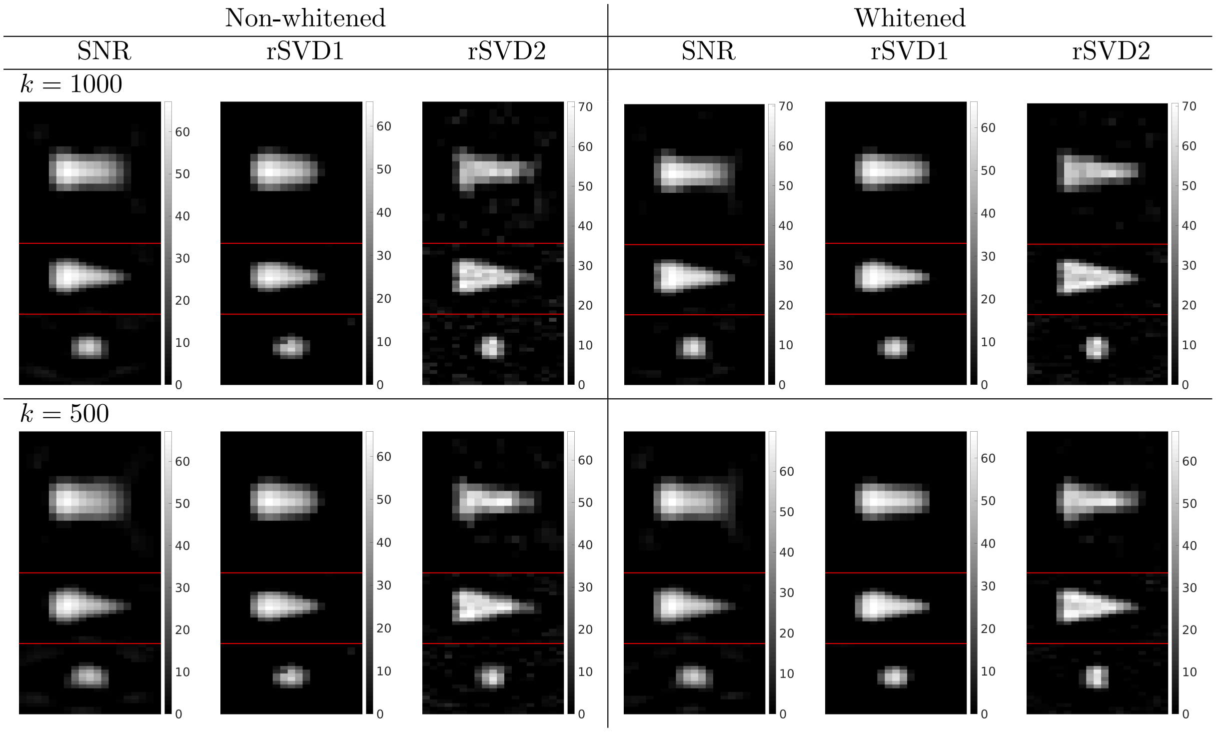


Figure 8: “Shape” phantom reconstructions for $\alpha = 2^{-10}$.

can be greatly reduced in the presence of whitening; see Figs. 9–10 and 12–13. This clearly shows the significant potential of the strategy whitening + rSVD2 for MPI reconstruction.

Next we focus on the role of whitening in rSVD acceleration. Whitening influences greatly both rSVD1 and rSVD2, especially when the α value is small: The whitened reconstructions are of better quality in the senses that the background artifacts are strongly reduced; see Figs. 10 and 13. Note that for small α, a small truncation number k can be very beneficial for improving reconstruction quality, due to its intrinsic regularizing effect (in a manner similar to the classical truncated SVD [5]). Further, comparing $k = 500$ for rSVD1 in Figs. 9 and 10 (and also Figs. 12 and 13), e.g., in the x-y-plane, the reconstruction shows that whitening may enable further dimension reduction while maintaining reconstruction quality (due to the change in the SV decay curve), concurring with the observation from Fig. 6. Among the three methods under analysis, rSVD2 benefits most from whitening, since for all three α values, the background artifacts disappear almost completely. While the precise mechanism of this phenomenon remains unclear, it may be attributed to the more robust SVD without too noisy singular vectors corresponding to large singular values. These observations indicate that whitening is advantageous in the reconstruction: it enables using smaller α values in rSVD1 and rSVD2 to obtain acceptable reconstructions without unnecessary smoothing the actual phantom because of using a sufficiently large α; see Figs. 8 and 11. However, SNR benefits little from whitening: it relies on the SNR-type quality measure for dimension reduction, already exploiting the background noise characteristic to some content. Thus, the dimensionality is already dramatically reduced, and the remaining rows of the reduced system have a large SNR-type quality measure and are only weakly influenced by the noise.

The computing times of the reconstruction methods are summarized in Table 1 for the “shape” phantom and Table 2 for the “resolution” phantom, where we have ignored the cost of preprocessing (e.g., frequency selection or rSVD) since it can be carried out offline. Interested readers can also find the computing times for rSVD in the last column of Table 1. All the computations are carried out on a server with 2×Intel ® Xeon® Broadwell-EP Series Processor E5-2687W v4, 3.00 GHz, 12-Core, and 1.5 TB DDR4 PC2666 main memory, and the mean and standard deviation are computed from 100 repetitions. The computing time for rSVD1 and SNR are more or less comparable, when using the same k value, due to similar complexity (more precisely, rSVD1 has slightly longer computing times due to the additional matrix-vector multiplication to

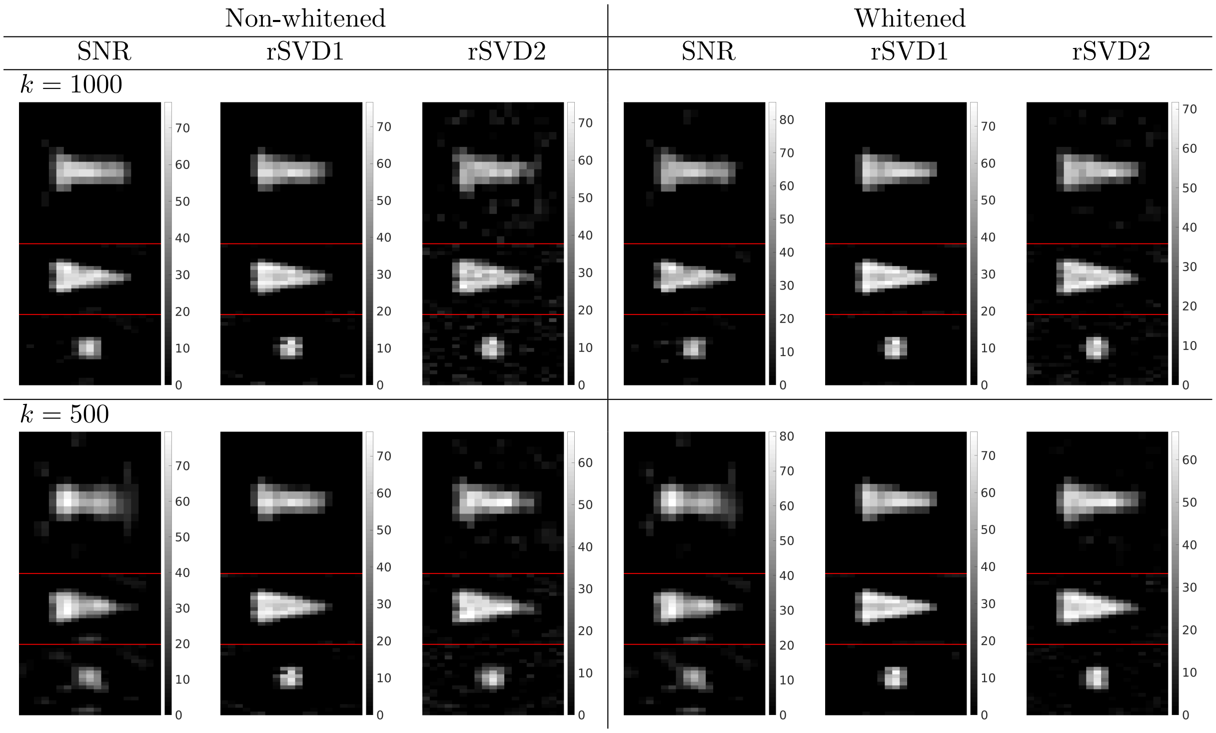


Figure 9: “Shape” phantom reconstructions for $\alpha = 2^{-15}$.

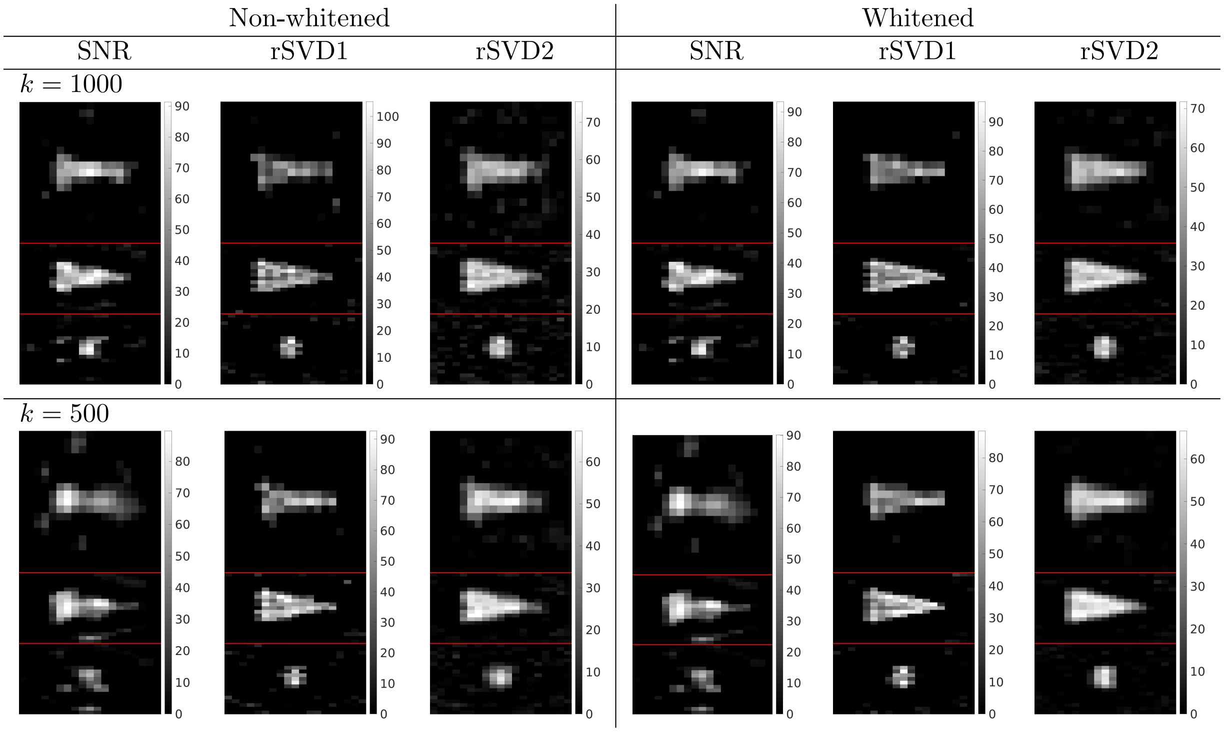


Figure 10: “Shape” phantom reconstructions for $\alpha = 2^{-20}$.

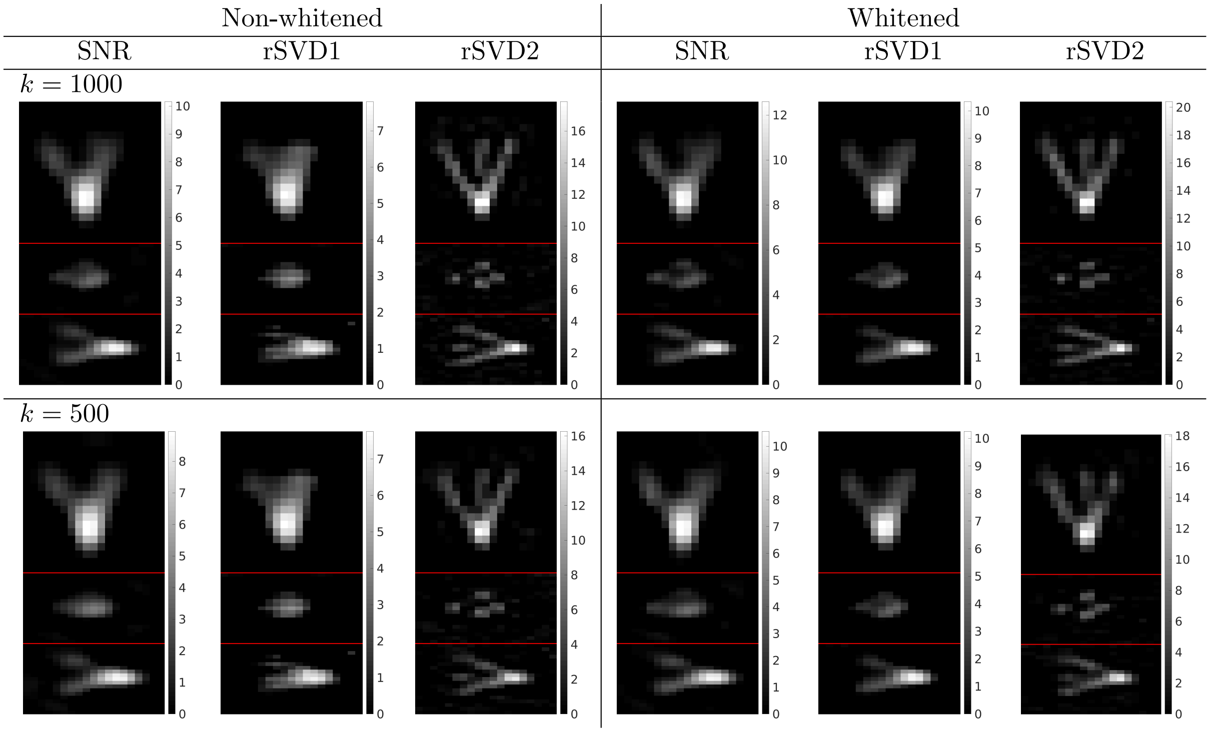


Figure 11: "Resolution" phantom reconstructions for $\alpha = 2^{-10}$.

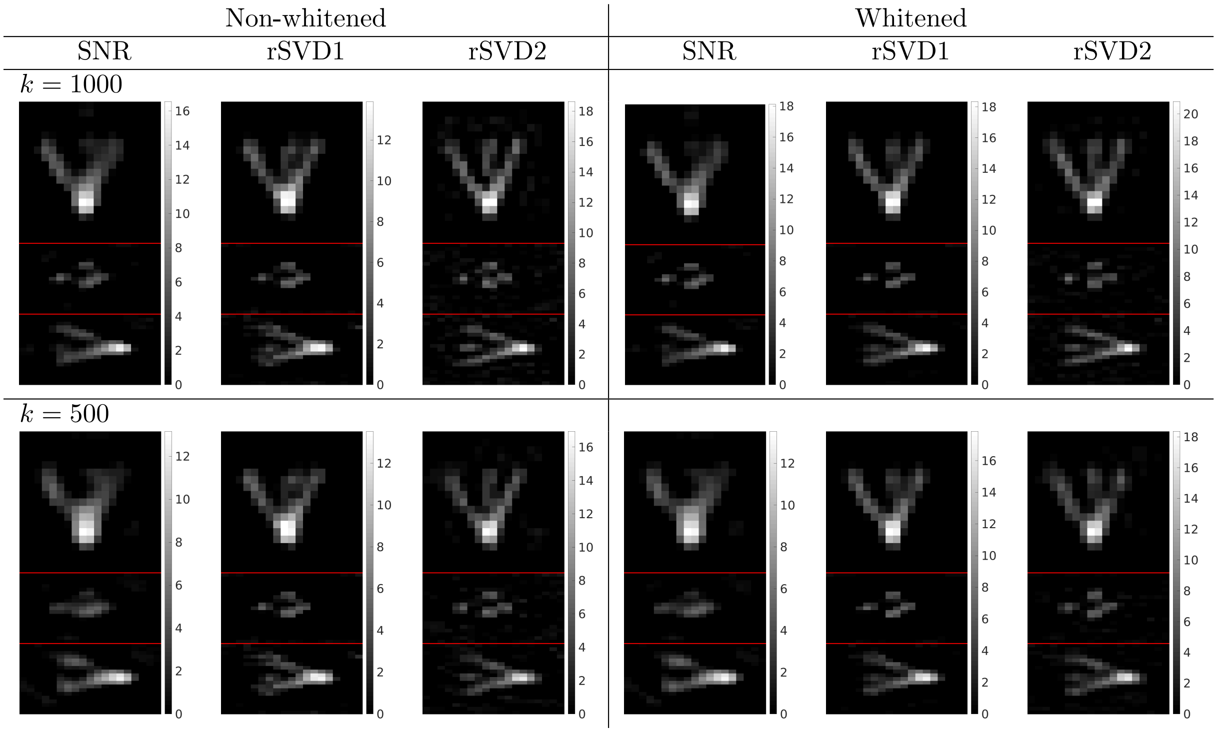


Figure 12: "Resolution" phantom reconstructions for $\alpha = 2^{-15}$.

project the measurement into the space spanned by the k singular vectors in $\tilde{U}_k$ respectively $\tilde{U}_{W;k}$). rSVD2 is the fastest method due to its non-iterative nature, even if one takes into account the 20 sweeps over the

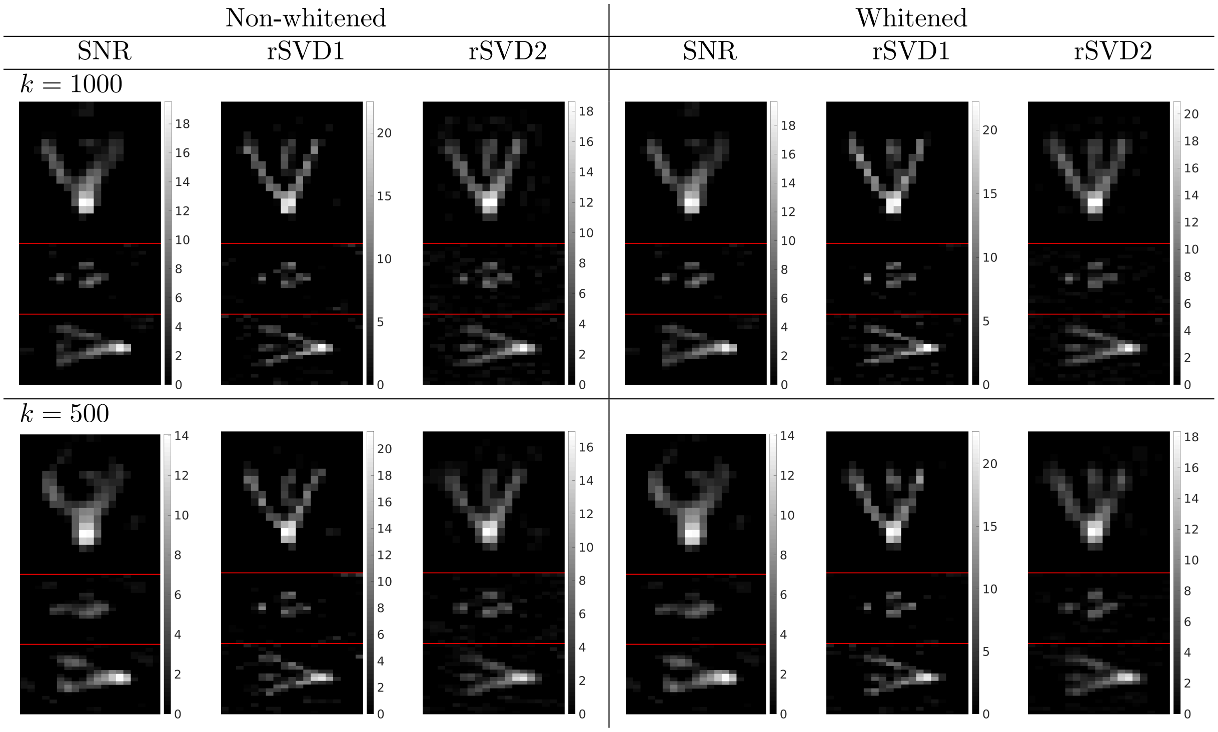


Figure 13: “Resolution” phantom reconstructions for $\alpha = 2^{-20}$.

k	SNR	rSVD1	rSVD2	rSVD
500	0.3792±0.0227	0.3880±0.0265	0.0115±0.0011	3.7237±0.1439
1000	0.7572±0.0391	0.7775±0.0403	0.0210±0.0022	7.1799±0.0959
1500	1.1489±0.0618	1.1743±0.0609	0.0300±0.0023	12.3714±0.1994
2000	1.5465±0.0910	1.5605±0.0712	0.0395±0.0022	18.6500±0.1881

Table 1: Computing times (in seconds) using `MATLAB` for the “shape” phantom. The computing time for STD is 53.5308±2.6653. STD for $20n$ iterations, SNR and rSVD1 for $20k$ iterations. The full SVD took 116.85s/123.64s (whitened/non-whitened); $n = 70446$ and $m = 6859$.

k	SNR	rSVD1	rSVD2
500	0.3761±0.0221	0.3868±0.0255	0.0118±0.0011
1000	0.7531±0.0328	0.7743±0.0337	0.0213±0.0036
1500	1.1336±0.0449	1.1533±0.0408	0.0308±0.0021
2000	1.5145±0.0630	1.5508±0.0666	0.0386±0.0026

Table 2: Computing times (in seconds) using `MATLAB` for the “resolution” phantom. The computing time for STD is 53.1284±2.1240. STD for $20n$ iterations, SNR and rSVD1 for $20k$ iterations.

corresponding reduced systems for SNR and rSVD1. Thus, all the acceleration approaches can significantly reduce the overall computational cost, with the speedup factor essentially determined by the size of the reduced system. STD is the most expensive one among all methods under consideration. It is worth noting that the computing time for rSVD increases nearly linearly with the rank k, concurring with the complexity in Section 3.2, and even after including the rSVD computation, the method rSVD1 is still far more efficient than STD. Note that the speedup is especially important in the multi-query context, since in practice one has to choose a proper value for the regularization parameter α, which inevitably requires solving a fair number of optimization problems. Further, it holds significant promise as a nearly online algorithm.

In summary, randomized SVD can significantly accelerate the reconstruction algorithms, within which the whitening procedure is highly beneficial, while maintaining the overall accuracy.

5 Concluding remarks

In this work we have discussed two issues in the MPI reconstruction from the perspective of inverse theory. First, we propose to include a whitening strategy and to solve a generalized least squares problem that is adapted to the noise statistics. This step can significantly improve the robustness of the algorithm and reconstruction quality. Second, we propose a dimension reduction strategy in the data space via randomized SVD. The randomized SVD is computationally efficient and scales to very large matrices. This step can greatly reduce the number of equations, and arrive at an accurate reduced system (in the data space). The numerical results on a publicly available dataset show that by combining whitening and low-rank approximation, one can obtain reconstructions of similar quality compared to the benchmark approach, but at a much lower computational complexity, and meanwhile can improve the image quality when compared with alternative system reduction approaches like SNR. The experimental findings indicate that the algorithmic tricks can facilitate developing fast robust MPI reconstruction algorithms.

This study has several implications on MPI reconstruction. The low-rank approximation provides an alternative (and complementary) to the sparse approximation approaches for the forward map [32, 31, 41]. In contrast to these works, the dimension reduction based on randomized SVD does not rely on an *a priori* choice of a basis for system representation, and it is optimal with respect to the (weighted) Frobenius / spectral norm. In theory, the ill-posed nature of the MPI inverse problem [25] allows a memory-efficient representation (i.e., low-rank approximation with a small k) without significant loss of reconstruction quality, which is confirmed by the numerical results in Section 4. The proposed method also does not require an SNR-type quality measure, which is computed from the noisy measured system matrix data and empty scanner measurements [8]. The latter is utilized in the proposed method (with a sufficiently large number of repetitions) to obtain a reasonable approximation of the covariance matrix C for the whitening step. In contrast to the SNR-type quality measure, the noise characteristic is incorporated via a whitening strategy. The simple nature of the used background measurement correction does not require additional empty scanner measurements during the calibration, which may prolong the calibration due to expensive additional robot movements [46, 45]. Further, the optimal dimension reduction builds the basis for developing efficient online reconstructions. In this context, the proposed method is advantageous since it allows more robust and faster (due to possibly much more effective dimension reduction) image reconstruction when compared with the online reconstruction approach based on the SNR-type quality measure proposed in [28].

This work shows that the study of the MPI reconstruction from the perspective of inverse theory can be very fruitful, and it motivates further explorations along the line. The first is about choosing a suitable regularization parameter, which is a long standing issue and often done in an ad hoc manner in practice. A systematic investigation of the issue would greatly facilitate the development of automated reconstruction techniques. One important factor in the context of MPI is the presence of significant model errors (and possibly also the error due to low-rank approximation), which necessitates proper adaptation of existing rules, e.g., discrepancy principle, L-curve criterion and quasi-optimality criterion. The second is to extend the proposed low-rank trick to other penalties, e.g., total variation and sparsity. One obvious strategy is to reduce the computational complexity of the gradient evaluation in related iterative minimization algorithms (e.g., primal-dual algorithm and proximal gradient algorithm), using a low-rank approximation. Similarly, it is also of interest to explore the potentials of alternative dimension reduction strategies, e.g., random sketching. Third and last, it is important to analyze relevant theoretical issues, e.g., stability, uniqueness and resolution limit, of the MPI inverse problem, and their dependence on the experimental setup so as to guide the experimental design.

Acknowledgements

The authors are grateful to three anonymous referees for their constructive comments, which have led to an improvement in the quality of the work. T. Kluth acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - project number 281474342/GRK2224/1 “Pi^3 : Parameter

Identification - Analysis, Algorithms, Applications" and support by the project "MPI^2" (BMBF, project no. 05M16LBA).

References

[1] C. Brandt and A. Seppänen. Recovery from errors due to domain truncation in magnetic particle imaging: Approximation error modeling approach. *J. Math. Imag. Vis.*, 60(8):1196–1208, 2018.

[2] A. Dax. On row relaxation methods for large constrained least squares problems. *SIAM J. Sci. Comput.*, 14(3):570–584, 1993.

[3] S. Dittmer, T. Kluth, P. Maass, and D. O. Baguer. Regularization by architecture: A deep prior approach for inverse problems. Preprint, arXiv:1812.03889, 2018.

[4] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems.* SIAM, Philadelphia, PA, 1999.

[5] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems.* Kluwer, Dordrecht, 1996.

[6] W. Erb, A. Weinmann, M. Ahlborg, C. Brandt, G. Bringout, T. M. Buzug, J. Frikel, C. Kaethner, T. Knopp, T. März, M. Möddel, M. Storath, and A. Weber. Mathematical analysis of the 1D model and reconstruction schemes for magnetic particle imaging. *Inverse Problems*, 34(5):055012, 21 pp., 2018.

[7] J. Franke, U. Heinen, H. Lehr, A. Weber, F. Jaspard, W. Ruhm, M. Heidenreich, and V. Schulz. System characterization of a highly integrated preclinical hybrid MPI-MRI scanner. *IEEE Trans. Med. Imag.*, 35(9):1993–2004, 2016.

[8] J. Franke, R. Lacroix, H. Lehr, M. Heidenreich, U. Heinen, and V. Schulz. MPI flow analysis toolbox exploiting pulsed tracer information - an aneurysm phantom proof. *Int. J. Magnetic Particle Imag.*, 3(1):Article ID 1703020, 5 pp., 2017.

[9] J. Friedman, T. Hastie, and R. Tibshirani. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3):432–441, 2008.

[10] V. B. Glasko and Y. A. Kriksin. The principle of quasioptimality for linear ill-posed problems in Hilbert space. *Zh. Vychisl. Mat. i Mat. Fiz.*, 24(11):1603–1613, 1758, 1984.

[11] B. Gleich and J. Weizenecker. Tomographic imaging using the nonlinear response of magnetic particles. *Nature*, 435(7046):1214–1217, 2005.

[12] G. H. Golub and C. F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[13] P. W. Goodwill and S. M. Conolly. Multidimensional x-space magnetic particle imaging. *IEEE Trans. Med. Imag.*, 30(9):1581–1590, 2001.

[14] P. W. Goodwill, E. U. Saritas, L. R. Croft, T. N. Kim, K. M. Krishnan, D. V. Schaffer, and S. M. Conolly. X-space MPI: magnetic nanoparticles for safe medical imaging. *Adv. Mater.*, 24(28):3870–3877, 2012.

[15] J. Haegele, J. Rahmer, B. Gleich, J. Borgert, H. Wojtczyk, N. Panagiotopoulos, T. Buzug, J. Barkhausen, and F. Vogt. Magnetic particle imaging: visualization of instruments for cardiovascular intervention. *Radiology*, 265(3):933–938, 2012.

[16] N. Halko, P. G. Martinsson, and J. A. Tropp. Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Rev.*, 53(2):217–288, 2011.

[17] P. C. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev.*, 34(4):561–580, 1992.

[18] G. T. Herman, A. Lent, and P. H. Lutz. Relaxation method for image reconstruction. *Comm. ACM*, 21(2):152–158, 1978.

[19] K. Ito and B. Jin. *Inverse Problems: Tikhonov Theory and Algorithms.* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2015.

[20] Y. Jiao, B. Jin, and X. Lu. Preasymptotic convergence of randomized Kaczmarz method. *Inverse Problems*, 33(12):125012, 21 pp., 2017.

[21] B. Jin and D. A. Lorenz. Heuristic parameter-choice rules for convex variational regularization based on error estimates. *SIAM J. Numer. Anal.*, 48(3):1208–1229, 2010.

[22] S. Kaczmarz. Angenäherte auflösung von systemen linearer gleichungen. *Bull. Int. Acad. Polon. Sci. Lett. A*, 35:335–357, 1937. English transl.: *Int. J. Control* 57(6), 1269–1271, 1993.

[23] A. Khandhar, P. Keselman, S. Kemp, R. Ferguson, P. Goodwill, S. Conolly, and K. Krishnan. Evaluation of peg-coated iron oxide nanoparticles as blood pool tracers for preclinical magnetic particle imaging. *Nanoscale*, 9(3):1299–1306, 2017.

[24] T. Kluth. Mathematical models for magnetic particle imaging. *Inverse Problems*, 34(8):083001, 27 pp., 2018.

[25] T. Kluth, B. Jin, and G. Li. On the degree of ill-posedness of multi-dimensional magnetic particle imaging. *Inverse Problems*, 34(9):095006, 26 pp., 2018.

[26] T. Kluth and P. Maass. Model uncertainty in magnetic particle imaging: Nonlinear problem formulation and model-based sparse reconstruction. *Int. J. Magnetic Part. Imag.*, 3(2):1707004, 10 pp., 2017.

[27] T. Knopp, N. Gdaniec, and M. Möddel. Magnetic particle imaging: from proof of principle to preclinical applications. *Phys. Med. Biol.*, 67(14):R124–R178, 2017.

[28] T. Knopp and M. Hofmann. Online reconstruction of 3D magnetic particle imaging data. *Phys. Med. Biol.*, 61(11):N257–67, 2016.

[29] T. Knopp, J. Rahmer, T. F. Sattel, S. Biederer, J. Weizenecker, B. Gleich, J. Borgert, and T. M. Buzug. Weighted iterative reconstruction for magnetic particle imaging. *Phys. Med. Biol.*, 55(6):1577–1589, 2010.

[30] T. Knopp, T. Viereck, G. Bringout, M. Ahlborg, A. von Gladiss, C. Kaethner, A. Neumann, P. Vogel, J. Rahmer, and M. Möddel. MDF: Magnetic particle imaging data format. *Preprint, arXiv:1602.06072*, 2016.

[31] T. Knopp and A. Weber. Local system matrix compression for efficient reconstruction in magnetic particle imaging. *Adv. Math. Phys.*, 2015:article ID 472818, 7 pp., 2015.

[32] J. Lampe, C. Bassoy, J. Rahmer, J. Weizenecker, H. Voss, B. Gleich, and J. Borgert. Fast reconstruction in magnetic particle imaging. *Phys. Med. Biol.*, 57(4):1113–1134, 2012.

[33] N. Löwa, P. Radon, O. Kosch, and F. Wiekhorst. Concentration dependent MPI tracer performance. *Int. J. Magnetic Particle Imag.*, 2(1):601001, 5 pp., 2016.

[34] T. März and A. Weinmann. Model-based reconstruction for magnetic particle imaging in 2D and 3D. *Inverse Probl. Imaging*, 10(4):1087–1110, 2016.

[35] V. A. Morozov. On the solution of functional equations by the method of regularization. *Soviet Math. Dokl.*, 7:414–417, 1966.

[36] K. Murase, M. Aoki, N. Banura, K. Nishimoto, A. Mimura, T. Kuboyabu, and I. Yabata. Usefulness of magnetic particle imaging for predicting the therapeutic effect of magnetic hyperthermia. *Open J. Med. Imag.*, 5(02):85, 2015.

[37] F. Natterer. *The Mathematics of Computerized Tomography.* B. G. Teubner, Stuttgart; John Wiley & Sons, Ltd., Chichester, 1986.

[38] J. Rahmer, A. Halkola, B. Gleich, I. Schmale, and J. Borgert. First experimental evidence of the feasibility of multi-color magnetic particle imaging. *Phys. Med. Biol.*, 60(5):1775–1791, 2015.

[39] J. Rahmer, J. Weizenecker, B. Gleich, and J. Borgert. Signal encoding in magnetic particle imaging: properties of the system function. *BMC Med. Imag.*, 9(4):21 pp., 2009.

[40] J. Salamon, M. Hofmann, C. Jung, M. G. Kaul, F. Werner, K. Them, R. Reimer, P. Nielsen, A. vom Scheidt, G. Adam, T. Knopp, and H. Ittrich. Magnetic particle/magnetic resonance imaging: In-vitro mpi-guided real time catheter tracking and 4d angioplasty using a road map and blood pool tracer approach. *PloS One*, 11(6):e0156899–14, 2016.

[41] L. Schmiester, M. Möddel, W. Erb, and T. Knopp. Direct image reconstruction of Lissajous type magnetic particle imaging data using Chebyshev-based matrix compression. *IEEE Trans. Comput. Imag.*, 3(4):671–681, 2017.

[42] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski. *Regularization Methods in Banach Spaces.* Walter de Gruyter GmbH & Co. KG, Berlin, 2012.

[43] M. Storath, C. Brandt, M. Hofmann, T. Knopp, J. Salamon, A. Weber, and A. Weinmann. Edge preserving and noise reducing reconstruction for magnetic particle imaging. *IEEE Trans. Med. Imag.*, 36(1):74–85, 2017.

[44] P. Szwargulski, J. Rahmer, M. Ahlborg, C. Kaethner, and T. M. Buzug. Experimental evaluation of different weighting schemes in magnetic particle imaging reconstruction. *Curr. Direct. Biomed. Eng.*, 1(1):206–209, 2015.

[45] K. Them, M. G. Kaul, C. Jung, M. Hofmann, T. Mummert, F. Werner, and T. Knopp. Sensitivity enhancement in magnetic particle imaging by background subtraction. *IEEE Trans. Med. Imag.*, 35(3):893–900, 2016.

[46] J. Weizenecker, B. Gleich, J. Rahmer, H. Dahnke, and J. Borgert. Three-dimensional real-time in vivo magnetic particle imaging. *Phys. Med. Biol.*, 54(5):L1–L10, 2009.

[47] D. P. Woodruff. Sketching as a tool for numerical linear algebra. *Foundations and Trends® in Theoretical Computer Science*, 10(1–2):1–157, 2014.

[48] E. Y. Yu, M. Bishop, B. Zheng, R. M. Ferguson, A. P. Khandhar, S. J. Kemp, K. M. Krishnan, P. W. Goodwill, and S. M. Conolly. Magnetic particle imaging: A novel in vivo imaging platform for cancer detection. *Nano Letters*, 17(3):1648–1654, 2017.

A Error estimate

Now we give an error estimate on the approximation $\tilde{x}_\alpha^\delta$ in a general setting. Let X be a Banach space, and Y be a Hilbert space, and $A : X \to Y$ be a compact linear operator. Consider the following inverse problem $Ax = y^\dagger$, where $x \in X$. Instead of the exact data $y^\dagger = Ax^\dagger$, corresponding to the exact solution $x^\dagger$, we have $y^\delta \in Y$ with an accuracy $\delta = \|y^\delta - y^\dagger\|$. Let $\tilde{A} : X \to Y$ be an approximate forward map with $\epsilon = \|\tilde{A} - A\|$. Then we aim at finding an approximate solution $\tilde{x}_\alpha^\delta$ by means of variational regularization [19]

$$\tilde{J}_\alpha(x) = \tfrac{1}{2}\|\tilde{A}x - y^\delta\|^2 + \alpha\psi(x),$$

where the functional $\psi : X \to \mathbb{R}_+ \cap \{0\}$ is a convex, proper and lower-semicontinuous functional. The common choice includes $\psi(x) = \frac{1}{2}\|x\|^2$, $\psi(x) = \|x\|_{\ell^1}$ and $\psi(x) = |x|_{\mathrm{TV}}$ etc. We denote by x_α^δ a minimizer to J_α and by $\tilde{x}_\alpha^\delta$ a corresponding minimizer to $\tilde{J}_\alpha$ with a noisy operator $\tilde{A}$. By $x^\dagger$, we denote a minimum-ψ solution of the equation $Ax = y^\dagger$: $x^\dagger = \arg\min_{x\in X: Ax=y^\dagger} \psi(x)$. It is easy to see that the derivation below remains valid for in the presence of nonnegativity constraint, so long as the minimum-ψ solution is feasible.

Let $\partial\psi(x)$ be the subdifferential of ψ at x [4]. For any $\xi \in \partial\psi(x)$, we define the Bregman distance from x to x' with respect to ξ by

$$d_\xi(x', x) = \psi(x') - \psi(x) - \langle \xi, x' - x\rangle.$$

Then we have the following error estimate on the approximation $\tilde{x}_\alpha^\delta$. It may serve as a guideline for determining the accuracy of the constructed approximation $\tilde{A}$: the model error $\epsilon = \|A - \tilde{A}\|$ should be comparable with data error δ in order not to compromise the reconstruction accuracy. The proof is standard [42, 19] and it is given only for completeness.

Theorem A.1. *Assume that the exact solution $x^\dagger$ fulfills the following source condition: there exists $w \in Y$ such that $A^*w \in \partial\psi(x^\dagger)$. Then for the minimizer $\tilde{x}_\alpha^\delta$ to the functional $\tilde{J}_\alpha$, there holds*

$$d_\xi(\tilde{x}_\alpha^\delta, x^\dagger) \leq \alpha^{-1}(\epsilon\|x^\dagger\| + \delta)^2 + \alpha\|w\|^2 + \epsilon\|w\|\|x^\dagger - \tilde{x}_\alpha^\delta\|.$$

Proof. By the minimizing property of $\tilde{x}_\alpha^\delta$, we obtain

$$\tfrac{1}{2}\|\tilde{A}\tilde{x}_\alpha^\delta - y^\delta\|^2 + \alpha d_\xi(\tilde{x}_\alpha^\delta, x^\dagger) \leq \tfrac{1}{2}\|\tilde{A}x^\dagger - y^\delta\|^2 - \alpha\langle \xi, \tilde{x}_\alpha^\delta - x^\dagger\rangle.$$

Under the source condition, rearranging the inequality yields

$$\tfrac{1}{2}\|\tilde{A}\tilde{x}_\alpha^\delta - y^\delta\|^2 + \alpha d_\xi(\tilde{x}_\alpha^\delta, x^\dagger) \leq \tfrac{1}{2}\|\tilde{A}x^\dagger - y^\delta\|^2 - \alpha\langle w, A(\tilde{x}_\alpha^\delta - x^\dagger)\rangle.$$

Next we rewrite the terms on the right hand side as

$$\begin{aligned}\|\tilde{A}x^\dagger - y^\delta\|^2 &= \|\tilde{A}\tilde{x}_\alpha^\delta - y^\delta\|^2 + 2\langle \tilde{A}\tilde{x}_\alpha^\delta - y^\delta, \tilde{A}(x^\dagger - \tilde{x}_\alpha^\delta)\rangle + \|\tilde{A}(x^\dagger - \tilde{x}_\alpha^\delta)\|^2,\\ \langle w, A(\tilde{x}_\alpha^\delta - x^\dagger)\rangle &= \langle w, A_k(\tilde{x}_\alpha^\delta - x^\dagger)\rangle + \langle w, (A - A_k)(\tilde{x}_\alpha^\delta - x^\dagger)\rangle.\end{aligned}$$

Combining the last three estimates yields

$$\begin{aligned}\tfrac{1}{2}\|\tilde{A}\tilde{x}_\alpha^\delta - y^\delta\|^2 + \alpha d_\xi(\tilde{x}_\alpha^\delta, x^\dagger) \leq& \tfrac{1}{2}\|\tilde{A}\tilde{x}_\alpha^\delta - y^\delta\|^2 + \langle \tilde{A}\tilde{x}_\alpha^\delta - y^\delta, \tilde{A}(x^\dagger - \tilde{x}_\alpha^\delta)\rangle + \tfrac{1}{2}\|\tilde{A}(x^\dagger - \tilde{x}_\alpha^\delta)\|^2\\ &+ \alpha\langle w, \tilde{A}(x^\dagger - \tilde{x}_\alpha^\delta)\rangle + \alpha\langle w, (A - \tilde{A})(x^\dagger - \tilde{x}_\alpha^\delta)\rangle\\ =& \tfrac{1}{2}\|\tilde{A}\tilde{x}_\alpha^\delta - y^\delta\|^2 + \langle \tilde{A}x^\dagger - y^\delta, \tilde{A}(x^\dagger - \tilde{x}_\alpha^\delta)\rangle - \tfrac{1}{2}\|\tilde{A}(x^\dagger - \tilde{x}_\alpha^\delta)\|^2\\ &+ \alpha\langle w, \tilde{A}(x^\dagger - \tilde{x}_\alpha^\delta)\rangle + \alpha\langle w, (A - \tilde{A})(x^\dagger - \tilde{x}_\alpha^\delta)\rangle.\end{aligned}$$

Collecting the terms, we obtain

$$\tfrac{1}{2}\|\tilde{A}(x^\dagger - \tilde{x}_\alpha^\delta)\|^2 + \alpha d_\xi(\tilde{x}_\alpha^\delta, x^\dagger) \leq \langle \tilde{A}x^\dagger - y^\delta, \tilde{A}(x^\dagger - \tilde{x}_\alpha^\delta)\rangle + \alpha\langle w, \tilde{A}(x^\dagger - \tilde{x}_\alpha^\delta)\rangle + \alpha\langle w, (A - \tilde{A})(x^\dagger - \tilde{x}_\alpha^\delta)\rangle$$

Thus, by means of Cauchy-Schwarz inequality and Young's inequality,

$$\alpha d_\xi(\tilde{x}_\alpha^\delta, x^\dagger) \leq \|\tilde{A}x^\dagger - y^\delta\|^2 + \alpha^2\|w\|^2 + \alpha\|w\|\|A - \tilde{A}\|\|x^\dagger - \tilde{x}_\alpha^\delta\|.$$

Meanwhile by the triangle inequality,

$$\begin{aligned}\|\tilde{A}x^\dagger - y^\delta\| &\leq \|(\tilde{A} - A)x^\dagger\| + \|Ax^\dagger - y^\delta\|\\ &\leq \|\tilde{A} - A\|\|x^\dagger\| + \|y^\dagger - y^\delta\| \leq \epsilon\|x^\dagger\| + \delta.\end{aligned}$$

Upon substituting the estimate, we obtain

$$d_\xi(\tilde{x}_\alpha^\delta, x^\dagger) \leq \alpha^{-1}(\epsilon\|x^\dagger\| + \delta)^2 + \alpha\|w\|^2 + \epsilon\|w\|\|x^\dagger - \tilde{x}_\alpha^\delta\|.$$

This completes the proof of the theorem. □

Remark A.1. *For the quadratic penalty $\psi(x) = \frac{1}{2}\|x\|^2$, the associated Bregman distance $d_\xi(x', x)$ is given by $d_\xi(x', x) = \frac{1}{2}\|x' - x\|^2$, and thus the error estimate in Theorem A.1 reduces to*

$$\|\tilde{x}_\alpha^\delta - x^\dagger\|^2 \leq 2\alpha^{-1}(\epsilon\|x^\dagger\| + \delta)^2 + 2\alpha\|w\|^2 + 2\epsilon\|w\|\|x^\dagger - \tilde{x}_\alpha^\delta\|,$$

which together with Young's inequality yields

$$\|\tilde{x}_\alpha^\delta - x^\dagger\|^2 \leq 4\alpha^{-1}(\epsilon\|x^\dagger\| + \delta)^2 + 4\alpha\|w\|^2 + 4\epsilon^2\|w\|^2.$$

The estimate shows that roughly one should choose $\epsilon := \|A - \tilde{A}\|$ such that $\epsilon\|x^\dagger\| \approx \delta$, in order to ensure that the overall accuracy is not compromised. This directly gives a guiding principle for constructing the low-rank approximation $\tilde{A}$ in Section 3.2. In particular, for a larger noise level δ, one may employ a less accurate low-rank approximation (i.e., smaller k), which can further reduce the computational complexity. Note that the last term in the estimate is generally of high order, and the first two terms essentially determines the accuracy.

B Supplements: Inverted colormap “shape” phantom

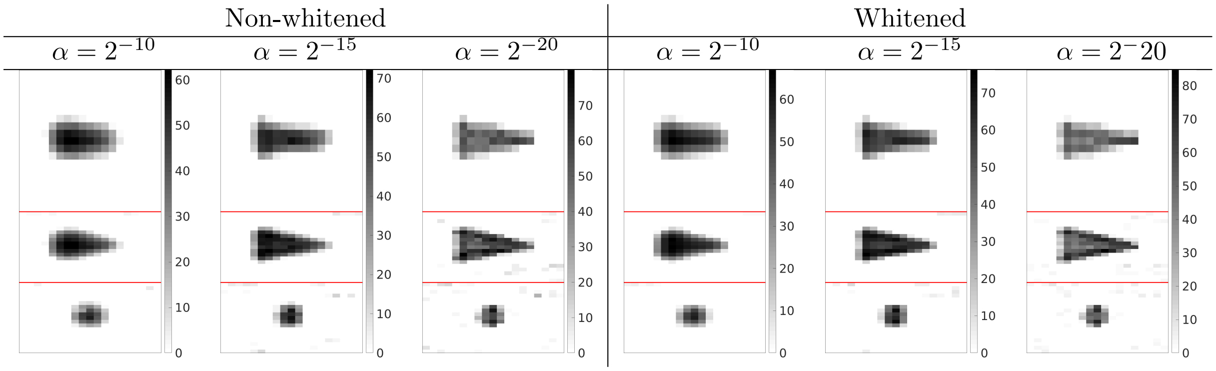


Figure 14: Non-whitened/whitened STD reconstructions of the “shape” phantom (Fig. 4 with inverted colormap).

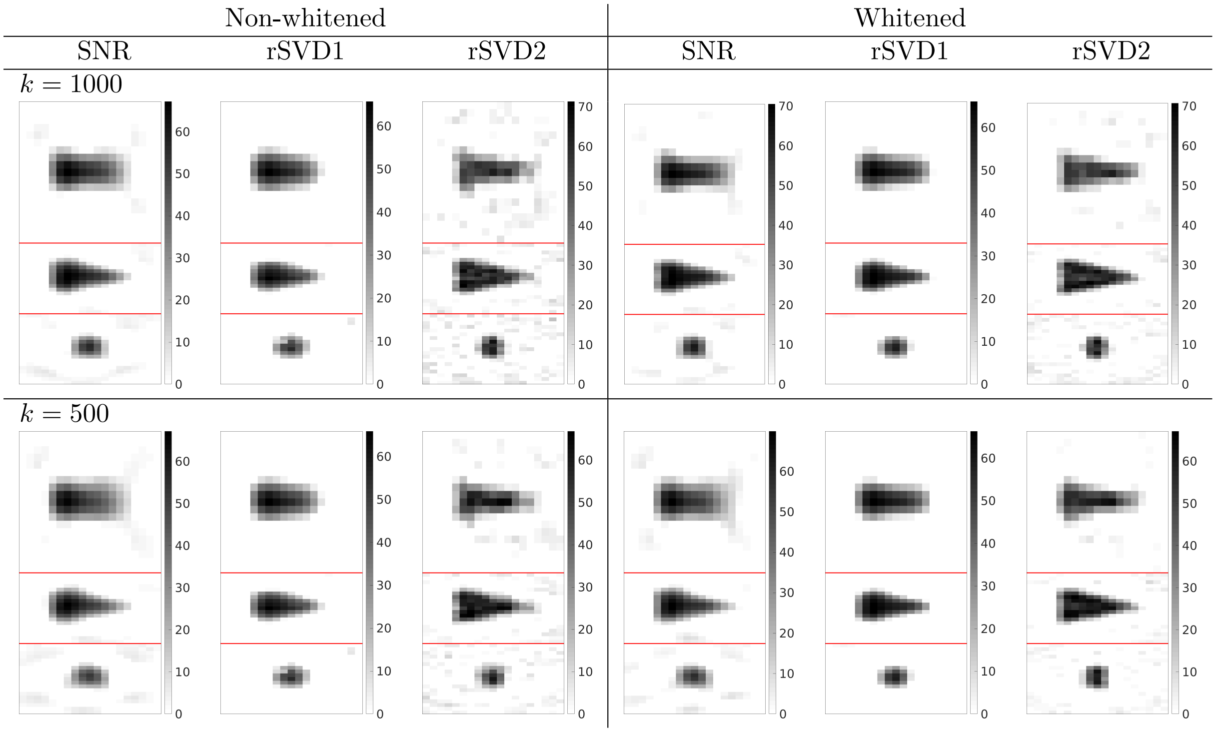


Figure 15: “Shape” phantom reconstructions for $\alpha = 2^{-10}$ (Fig. 8 with inverted colormap).

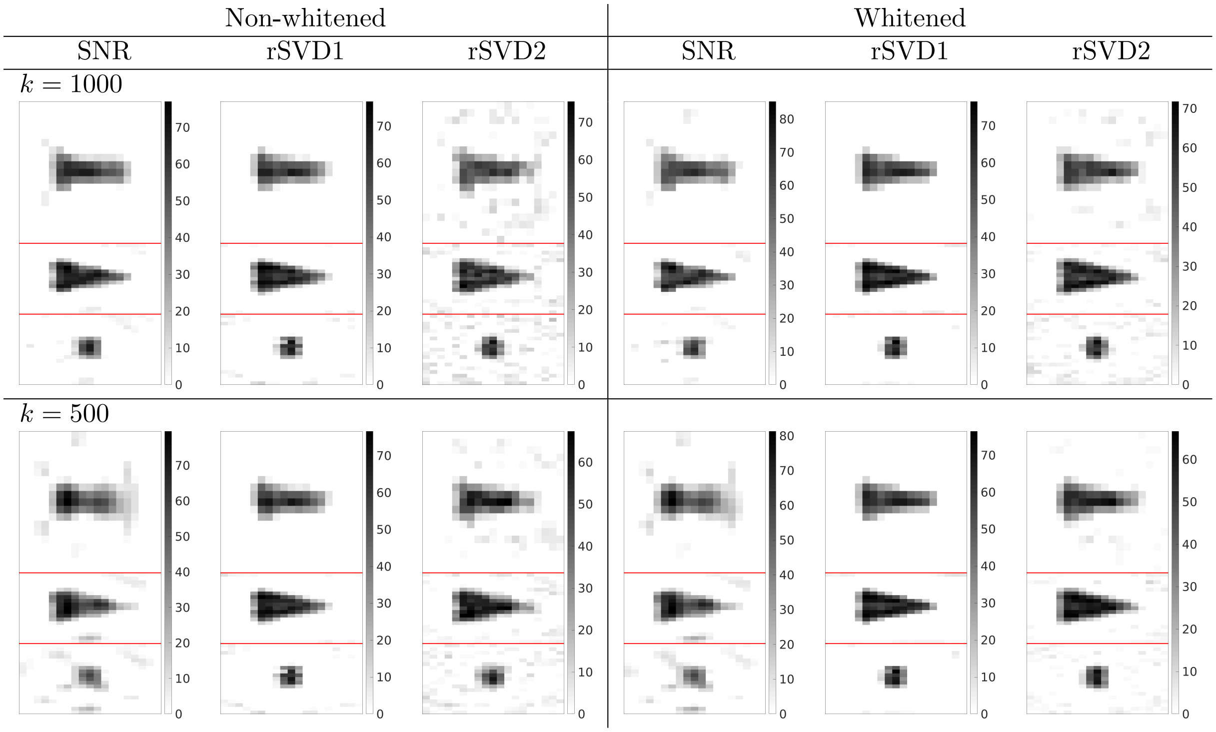


Figure 16: “Shape” phantom reconstructions for $\alpha = 2^{-15}$ (Fig. 9 with inverted colormap).

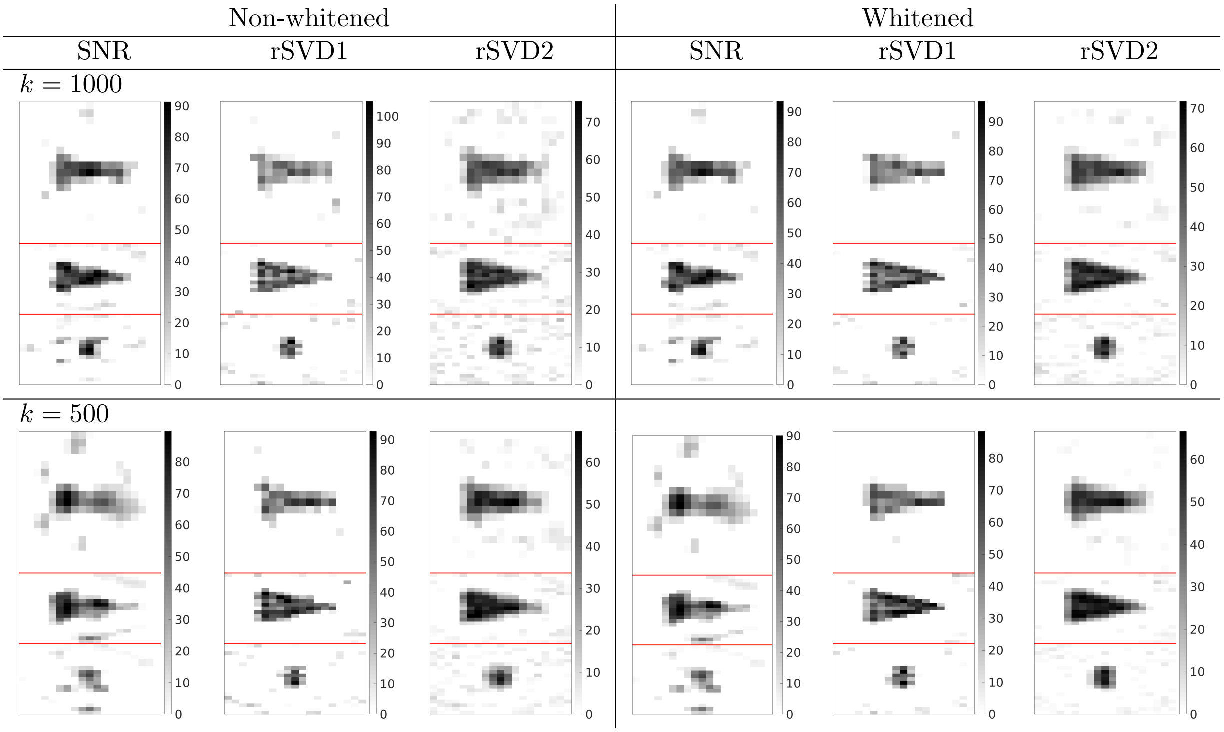


Figure 17: “Shape” phantom reconstructions for $\alpha = 2^{-20}$ (Fig. 10 with inverted colormap).

C Supplements: Inverted colormap “resolution” phantom

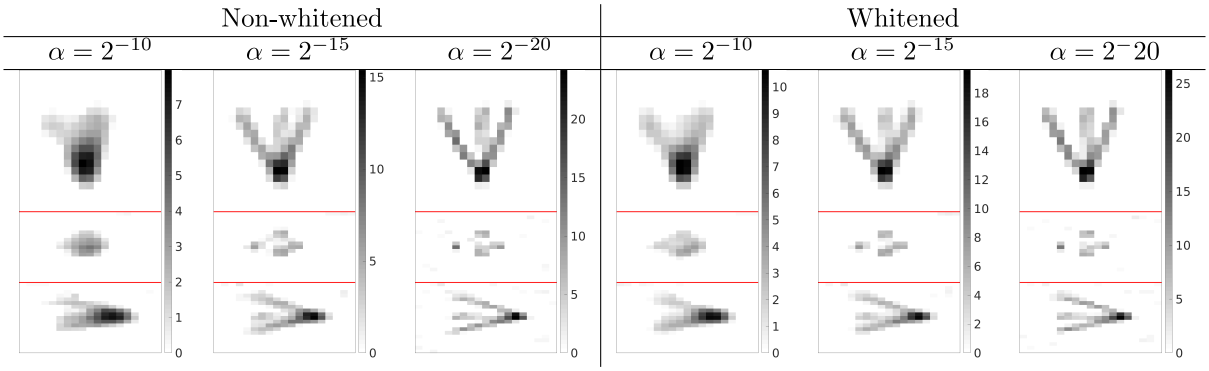


Figure 18: Non-whitened/whitened STD reconstructions of the “resolution” phantom (Fig. 5 with inverted colormap).

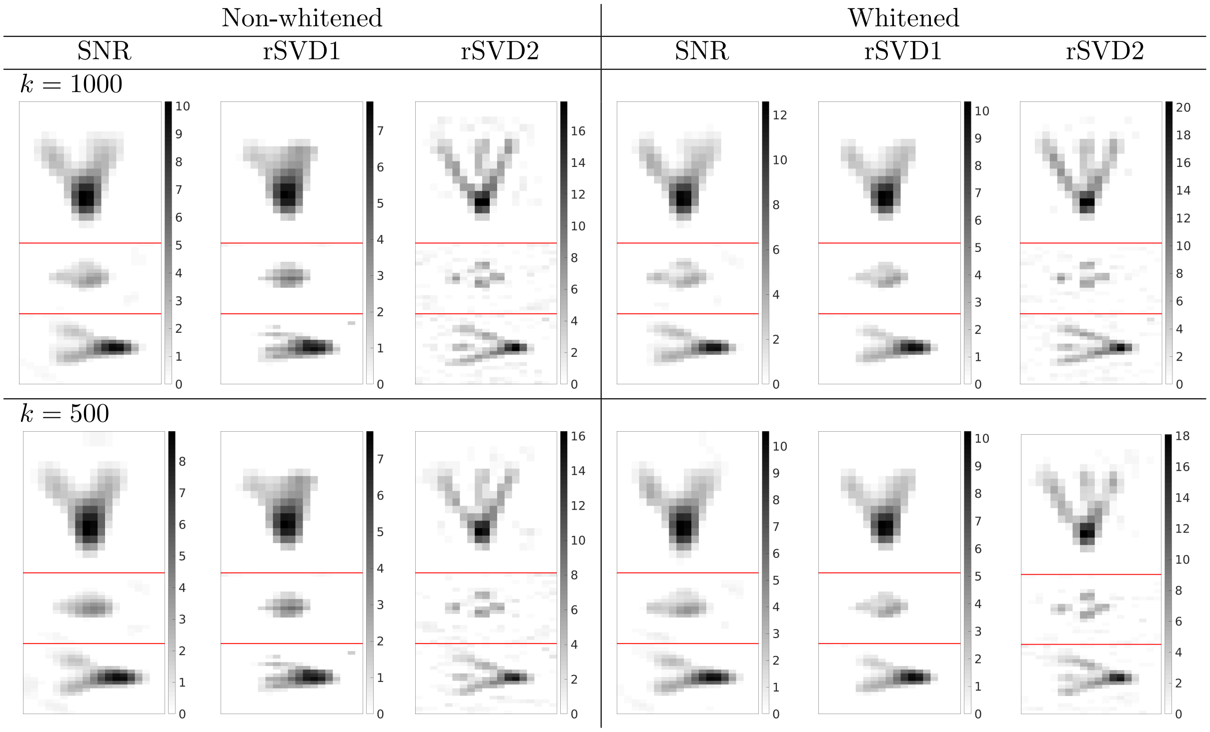


Figure 19: “Resolution” phantom reconstructions for $\alpha = 2^{-10}$ (Fig. 11 with inverted colormap).

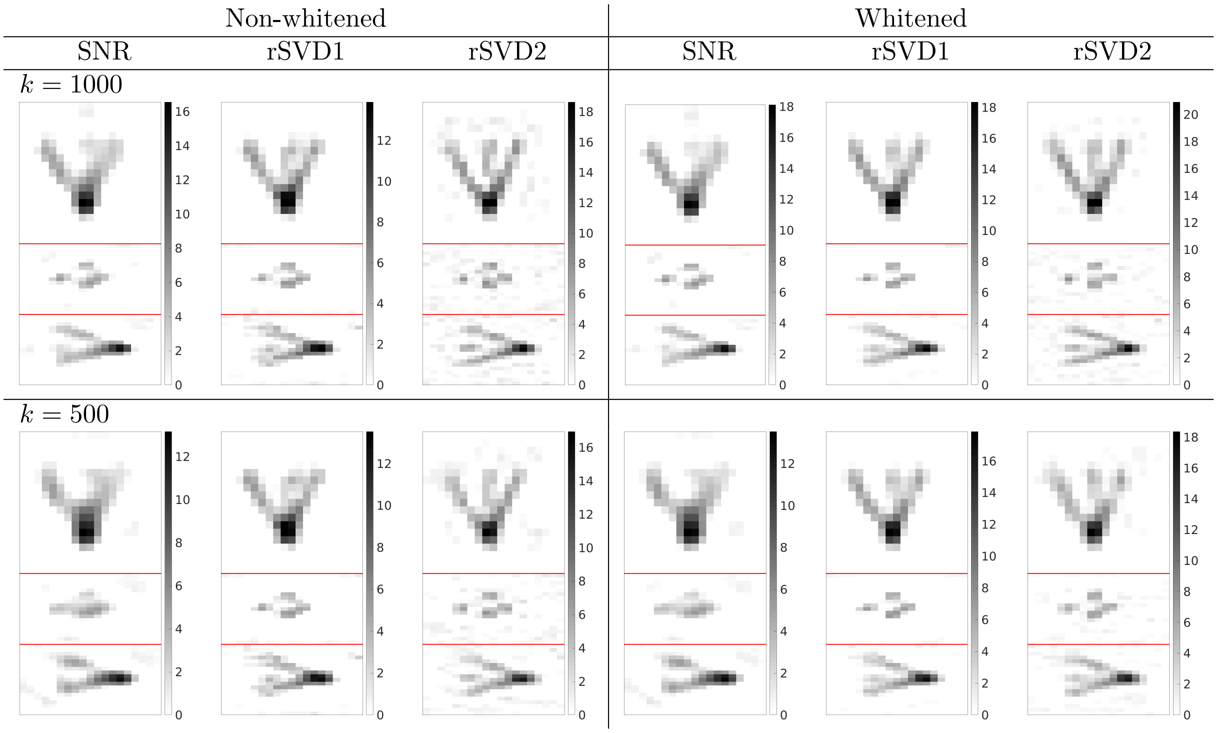


Figure 20: “Resolution” phantom reconstructions for $\alpha = 2^{-15}$ (Fig. 12 with inverted colormap).

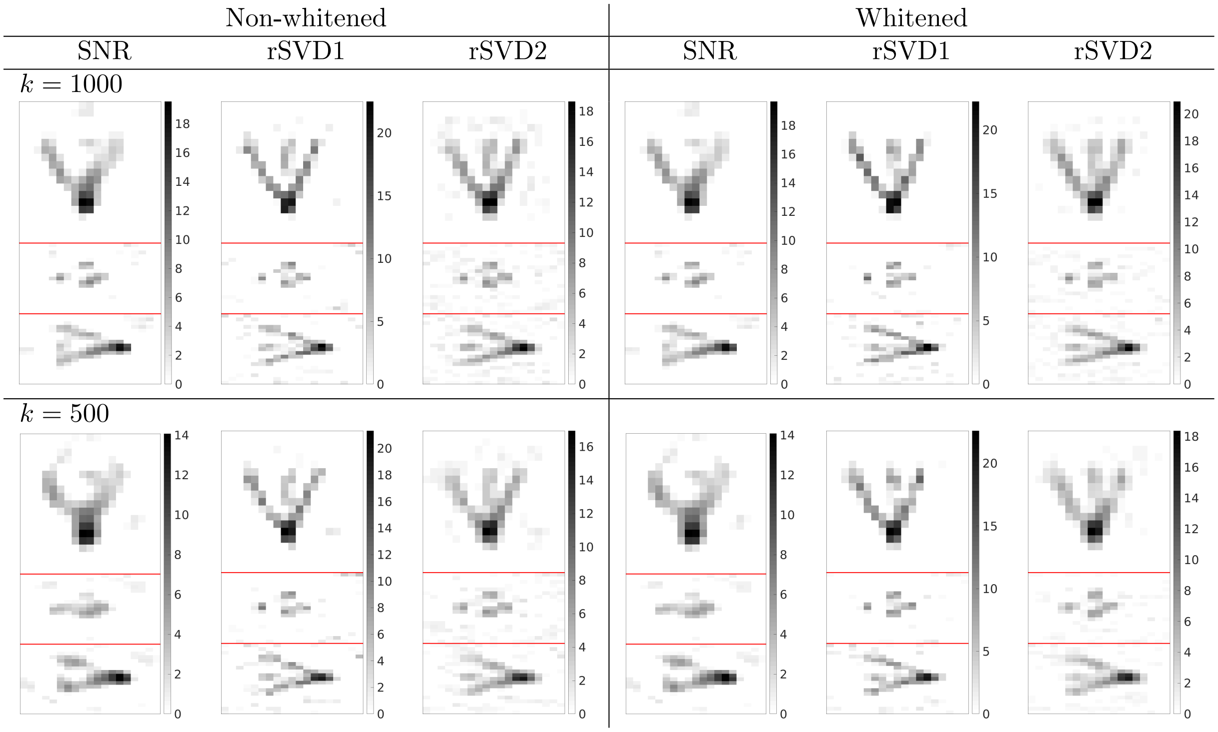


Figure 21: “Resolution” phantom reconstructions for $\alpha = 2^{-20}$ (Fig. 13 with inverted colormap).