A Recovery Algorithm based on the Kaczmarz Algorithm and ADMM Splitting with Application to Convex Optimization in Magnetic Particle Imaging

Marco Maass, Christine Droigk, Fabrice Katzberg, Philipp Koch, and Alfred Mertins
University of Lübeck
Institute for Signal Processing
D-23562 Lübeck, Germany
{marco.maass, c.droigk, f.katzberg, ph.koch, alfred.mertins}@uni-luebeck.de

***Abstract*—This work introduces a strategy for the extension of the standard Kaczmarz algorithm, which is popular for solving very large inverse problems, to priors other than the commonly used Tikhonov regularization. The proposed reformulation of the algorithm allows us to include more sophisticated priors while inheriting the row-wise operation structure of the classical Kaczmarz algorithm. The new method is developed with help of the alternating direction method of multipliers. The results show that also with suboptimal alternating direction method of multiplier steps, the proposed algorithm is able to solve convex optimization problems with very high accuracy. Especially, on the relative young preclinical medical imaging modality of magnetic particle imaging, the algorithm demonstrates high convergence rates. When the underlying matrix nearly shows mutually orthogonal rows, which is observed in the field of magnetic particle imaging, very high convergence rates can be expected.**

***Index Terms*—Kaczmarz method, alternating direction method of multipliers, convex optimization, inverse problems, medical imaging**

I. Introduction

Many recovery problems are modeled as linear inverse problems. Especially, such kind of problems occur in the field of medical imaging, such as computer tomography (CT) [1], magnetic resonance imaging (MRI) [2], and magnetic particle imaging (MPI) [3].

In contrast to the commonly known CT and MRI, magnetic particle imaging (MPI) is a relative young tracer-based medical imaging method to measure the spatial distribution of superparamagnetic iron-oxide nanoparticles (SPIONs) [3]. In MPI, two different magnetic fields are generated to image the spatial distribution of the SPIONs in one, two, or three dimensions. First, there is the so-called selection field, which induces the field free point (FFP), and secondly, one uses the so-called drive field, which moves the FFP inside an area of interest on a given trajectory and defines the field of view (FOV). The FFP is commonly moved on a periodic trajectory [3], [4]. The determination of the SPION distribution can be expressed as a linear inverse problem. Unfortunately, it is often ill-posed, because a high spatial resolution and a short acquisition time are sought at the same time.

Commonly, this problem is solved by a system-matrix based approach in a least squares error sense [4]. One widely used reconstruction method in MPI is the regularized algebraic reconstruction technique (ART) [1], also known as Kaczmarz algorithm (KA), because the method allows one to easily incorporate physically motivated constraints, like non-negative particle concentrations [5], [6]. However, the classical KA has a drawback, it works only for consistent linear systems. Therefore, commonly, a Tikhonov regularization is added to the least-squares problem, which allows for reformulating the least-squares problem as a consistent linear system [7]. Independent of the row-operating Kaczmarz algorithm with its fairly simple Tikhonov regularization, it has become common to use more sophisticated priors [8]–[10]. However, many of the sophisticated prior strategies come with the drawback that the MPI system matrix has to be applied also column-wise, such as [11], [12], which leads to an increased computational effort and makes these methods inefficient.

In the present work, a direct extension to the KA is derived that allows us to include sophisticated priors while operating only row-wise on the system matrix. The developed formulation of the reconstruction problem is based on the alternating direction method of multipliers (ADMM) [13], [14]. A related approach, which used another splitting approach to extend the classical Kaczmarz algorithm, can be found in [15]. Here, a linearized Bregman splitting was used to derive an ℓ_1-norm minimizing KA, called sparse KA [15]–[17]. For further details to linearized Bregman splitting see [18], [19]. However, these articles just deal with the consistent case of linear systems. By contrast, in this article we also deduce a formulation for the inconsistent case, based on the more general concept in [15].

II. Kaczmarz algorithm

The classical Kaczmarz algorithm [5] is generally formulated for consistent linear problems of the form

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{1}$$

where matrix $\boldsymbol{A} \in \mathbb{C}^{M\times N}$ is used to model a physical system, vector $\boldsymbol{b} \in \mathbb{C}^{M}$ contains the measurements, and $\boldsymbol{x} \in \mathbb{C}^{N}$ denotes an unknown vector to be recovered. The KA iteration for (1) is then given by

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \frac{b_i - \boldsymbol{a}_i \boldsymbol{x}^k}{\|\boldsymbol{a}_i\|_2^2} \boldsymbol{a}_i^H, \tag{2}$$

where $\boldsymbol{a}_i \in \mathbb{C}^{1\times N}$ denotes the i-th row of $\boldsymbol{A}$, and b_i is the i-th component of $\boldsymbol{b}$. One Kaczmarz iteration is a full run through all rows $\boldsymbol{a}_i$ with $i \in \{1, 2, \ldots, M\}$. The article uses random permutation to select the order of rows in each KA iteration.

A. Block-Kaczmarz algorithm

The Kaczmarz methods has its origin in the projection onto convex set (POCS) algorithm [5], [20], where the projection in (2) can be interpreted as the orthogonal projection of the actual $\boldsymbol{x}^k$ onto the hyperplane $\mathcal{H}_i = \{\boldsymbol{x} \in \mathbb{C}^N \mid \boldsymbol{a}_i \boldsymbol{x} - b_i = 0\}$ spanned by $\boldsymbol{a}_i$. This interpretation allows one also to project a given $\boldsymbol{x}^k$ onto more than one row of $\boldsymbol{A}$. For this, let us define subsets $\boldsymbol{A_p}$ and $\boldsymbol{b_p}$ of $\boldsymbol{A}$ and $\boldsymbol{b}$ with index vector $\boldsymbol{p} = (p_1, p_2, \ldots, p_\nu)$, $p_j \in \{1, \ldots, M\}$, $\nu < M$:

$$\boldsymbol{A_p} = \left(\boldsymbol{a}_{p_1}^T, \boldsymbol{a}_{p_2}^T, \ldots, \boldsymbol{a}_{p_\nu}^T\right)^T,$$

$$\boldsymbol{b_p} = \left(b_{p_1}, b_{p_2}, \ldots, b_{p_\nu}\right)^T.$$

Hence, the orthogonal projection of $\boldsymbol{x}^k$ onto the linear subspace $\mathcal{H}_{\boldsymbol{p}} = \{\boldsymbol{x} \in \mathbb{C}^N \mid \boldsymbol{A_p x} - \boldsymbol{b_p} = \boldsymbol{0}\}$ is given by [21]

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \boldsymbol{A_p^+}(\boldsymbol{b_p} - \boldsymbol{A_p x}^k), \tag{3}$$

where $\boldsymbol{A_p^+}$ denotes the Moore-Penrose pseudoinverse of $\boldsymbol{A_p}$.

B. Inconsistent linear systems

For most problems, a consistent linear system as shown in (1) cannot be guaranteed due to noise and modeling errors. Therefore, the problem (1) is generally reformulated as a least squares (LS) one. This modified problem can be reformulated in a consistent one as $\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^H\boldsymbol{b}$, which is also known as normal equations. However, if N is large, the resulting matrix $\boldsymbol{A}^H\boldsymbol{A}$ becomes large and hard to manage. Especially, if no efficient way is known to calculate the rows of $\boldsymbol{A}^H\boldsymbol{A}$, one would prefer a row-operating algorithm like the classical KA that directly works on the rows of $\boldsymbol{A}$.

In [7] it was shown that the regularized LS problem

$$\underset{\boldsymbol{x}\in\mathbb{C}^N}{\arg\min} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \|\lambda \boldsymbol{x}\|_2^2 \tag{4}$$

with $\lambda > 0$ can be reformulated in the consistent problem

$$\underset{\boldsymbol{x}\in\mathbb{C}^N, \boldsymbol{v}\in\mathbb{C}^M}{\arg\min} \|\boldsymbol{x}\|_2^2 + \|\boldsymbol{v}\|_2^2 \quad \text{s.t.} \quad \begin{bmatrix}\boldsymbol{A} & \lambda \boldsymbol{I}_M\end{bmatrix} \begin{bmatrix}\boldsymbol{x} \\ \boldsymbol{v}\end{bmatrix} = \boldsymbol{b}, \tag{5}$$

where $\boldsymbol{v} = \frac{\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}}{\lambda}$ plays the role of the residuum.

III. Convex model

We consider the generalized convex optimization problem

$$\underset{\boldsymbol{x}\in\mathbb{D}}{\arg\min} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \beta \mathcal{R}(\boldsymbol{L}\boldsymbol{x}), \tag{6}$$

where $\boldsymbol{A} \in \mathbb{C}^{M\times N}$ models a physical system by a linear equation, $\boldsymbol{b} \in \mathbb{C}^M$ contains the measurements, $\boldsymbol{x} \in \mathbb{D} \subseteq \mathbb{C}^N$ denotes an unknown signal which is to be recovered and lies in a convex subset $\mathbb{D}$. $\boldsymbol{L} \in \mathbb{C}^{K\times N}$ is an arbitrary matrix, the function $\mathcal{R} : \mathbb{C}^K \to \mathbb{R}$ is a convex regularization term, and $\beta > 0$ denotes the weighting of the regularization.

The problem in (6) can be rewritten as

$$\underset{\boldsymbol{x}\in\mathbb{C}^n}{\arg\min} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \beta \mathcal{R}(\boldsymbol{z}) + \chi_{\mathbb{D}}(\boldsymbol{x}), \text{ s.t. } \boldsymbol{L}\boldsymbol{x} - \boldsymbol{z} = \boldsymbol{0}, \tag{7}$$

where $\chi_{\mathbb{D}}(\boldsymbol{x})$ denotes the indicator function of the subspace $\mathbb{D}$. With help of ADMM [14], Eq. (7) can be split into

$$\boldsymbol{x}^{\ell+1} = \underset{\boldsymbol{x}\in\mathbb{C}^n}{\arg\min} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \frac{\rho}{2}\|\boldsymbol{z}^\ell - \boldsymbol{L}\boldsymbol{x} + \boldsymbol{u}^\ell\|_2^2 + \chi_{\mathbb{D}}(\boldsymbol{x}), \tag{8}$$

$$\begin{aligned}\boldsymbol{z}^{\ell+1} &= \underset{\boldsymbol{z}\in\mathbb{C}^K}{\arg\min}\, \beta\mathcal{R}(\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{z} - \boldsymbol{L}\boldsymbol{x}^{\ell+1} + \boldsymbol{u}^\ell\|_2^2 \qquad (9)\\ &= \operatorname{prox}_{\frac{\beta}{\rho}\mathcal{R}}(\boldsymbol{L}\boldsymbol{x}^{\ell+1} - \boldsymbol{u}^\ell)\end{aligned}$$

$$\boldsymbol{u}^{\ell+1} = \boldsymbol{u}^\ell + \boldsymbol{z}^{\ell+1} - \boldsymbol{L}\boldsymbol{x}^{\ell+1}, \tag{10}$$

where ℓ denotes the iteration index and $\rho > 0$ is the ADMM splitting parameter. The expression $\operatorname{prox}_{\frac{\beta}{\rho}\mathcal{R}}(\boldsymbol{L}\boldsymbol{x}^{\ell+1} - \boldsymbol{u}^\ell)$ denotes the proximity operator defined as

$$\operatorname{prox}_f(\boldsymbol{z}) = \underset{\boldsymbol{x}\in\mathbb{R}^N}{\arg\min}\, f(\boldsymbol{x}) + \frac{1}{2}\|\boldsymbol{z} - \boldsymbol{x}\|_2^2. \tag{11}$$

A run through (8), (9), and (10) is referred to as one full ADMM iteration.

To enable solving (8) with the KA, we introduce an additional damping parameter $\delta > 0$ and use the same trick as in (4) and (5). Therefore, we obtain the extended linear system

$$\underset{\substack{\boldsymbol{x}\in\mathbb{D}\\ \boldsymbol{v}\in\mathbb{C}^M\\ \boldsymbol{\epsilon}\in\mathbb{C}^K}}{\arg\min} \|\boldsymbol{x}\|_2^2 + \left\|\begin{pmatrix}\boldsymbol{v}\\ \boldsymbol{\epsilon}\end{pmatrix}\right\|_2^2 \text{ s.t. } \begin{pmatrix}\tilde{\boldsymbol{A}} & \delta \boldsymbol{I}_{M+K}\end{pmatrix} \begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{v}\\ \boldsymbol{\epsilon}\end{pmatrix} = \tilde{\boldsymbol{b}}, \tag{12}$$

where

$$\tilde{\boldsymbol{A}} = \begin{pmatrix}\boldsymbol{A}\\ \sqrt{\frac{\rho}{2}}\boldsymbol{L}\end{pmatrix} \quad \text{and} \quad \tilde{\boldsymbol{b}} = \begin{pmatrix}\boldsymbol{b}\\ \sqrt{\frac{\rho}{2}}\left(\boldsymbol{u}^\ell + \boldsymbol{z}^\ell\right)\end{pmatrix}. \tag{13}$$

A. Weighting strategy

With an adaptation strategy for δ, the objective function of the problem in (6) can be efficiently minimized. Here, we use the following rule: δ is increased by a factor of 1.01 if the objective function of (6) increases for three consecutive times after ADMM iterations, otherwise it is decreased by factor 0.99. After each ADMM iteration, the residual vectors $\boldsymbol{v}$ and $\boldsymbol{\epsilon}$ are set to $\boldsymbol{0}$. For the parameter $\rho > 0$ we use the strategy proposed in [22] (§3.4.1).

IV. Proposed algorithm

The proposed algorithm consists of the following steps:

1) Check the objective function for possible adaptation of δ as described in Section III-A.
2) Set $\boldsymbol{v} := \boldsymbol{0}$ and $\boldsymbol{\epsilon} := \boldsymbol{0}$.
3) Perform a fixed number of Kaczmarz iterations:
   - Kaczmarz iteration through the matrix $\boldsymbol{A}$:

$$\begin{aligned}\alpha &:= \frac{b_i - \boldsymbol{a}_i\boldsymbol{x} - \delta v_i}{\|\boldsymbol{a}_i\|_2^2 + \delta^2}\\ \boldsymbol{x} &:= \mathcal{P}_{\mathbb{D}}(\boldsymbol{x} + \alpha \boldsymbol{a}_i^H),\\ v_i &:= v_i + \alpha\delta\end{aligned} \tag{14}$$

where $\mathcal{P}_{\mathbb{D}}(\boldsymbol{x})$ denotes the orthogonal projection onto the convex set $\mathbb{D}$.

- Block-Kaczmarz iteration through $\boldsymbol{L}$:

$$\boldsymbol{w} := \boldsymbol{u} + \boldsymbol{z} - \boldsymbol{L}\boldsymbol{x} - \sqrt{\tfrac{2\delta^2}{\rho}}\boldsymbol{\epsilon}. \tag{15}$$

Depending on the dimensionality of $\boldsymbol{L}$, two different ways for the orthogonal projections are possible:

- If $K \leq N$

$$\begin{aligned} \boldsymbol{\gamma} &:= \left(\boldsymbol{L}\boldsymbol{L}^H + \frac{2\delta^2}{\rho}\boldsymbol{I}_K\right)^{-1}\boldsymbol{w} \\ \boldsymbol{\alpha} &:= \boldsymbol{x} + \boldsymbol{L}^H\boldsymbol{\gamma} \\ \boldsymbol{\epsilon} &:= \boldsymbol{\epsilon} + \sqrt{\tfrac{2\delta^2}{\rho}}\boldsymbol{\gamma}. \end{aligned} \tag{16}$$

- If $K > N$

$$\begin{aligned} \boldsymbol{s} &:= \left(\boldsymbol{L}^H\boldsymbol{L} + \frac{2\delta^2}{\rho}\boldsymbol{I}_N\right)^{-1}\boldsymbol{L}^H\boldsymbol{w} \\ \boldsymbol{\alpha} &:= \boldsymbol{x} + \boldsymbol{s} \\ \boldsymbol{\epsilon} &:= \sqrt{\tfrac{\rho}{2\delta^2}}\left(\boldsymbol{u} + \boldsymbol{z} - \boldsymbol{L}\boldsymbol{\alpha}\right). \end{aligned} \tag{17}$$

Orthogonal projection onto the convex subset $\mathbb{D}$:

$$\boldsymbol{x} := \mathcal{P}_{\mathbb{D}}(\boldsymbol{\alpha}). \tag{18}$$

4) Perform proximity update (9):

$$\boldsymbol{z} := \operatorname{prox}_{\frac{\beta}{\rho}\mathcal{R}}(\boldsymbol{L}\boldsymbol{x} - \boldsymbol{u}). \tag{19}$$

5) Update dual variable (10):

$$\boldsymbol{u} := \boldsymbol{u} + \boldsymbol{z} - \boldsymbol{L}\boldsymbol{x}. \tag{20}$$

6) Update ρ if necessary by the rules from [22] (§3.4.1).

V. Comparison to Methods in [15] and [16]

Following the formulation of the authors in [16], their method is able to optimize problems of the form

$$\begin{aligned} \arg\min_{\boldsymbol{x},\boldsymbol{z}} \frac{\beta}{2}\|\boldsymbol{z}\|_1 + \frac{1}{2}\left(\|\boldsymbol{x}\|_2^2 + \|\boldsymbol{z}\|_2^2\right) \quad \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{L}\boldsymbol{x} = \boldsymbol{z}. \end{aligned} \tag{21}$$

As in [15] conducted, a non-negativity constraint can be easily included. However, this formulation does not solve inconsistent problems of the form $\boldsymbol{A}\boldsymbol{x} \approx \boldsymbol{b}$. If the linear system is inconsistent, the sparse Kaczmarz ($\boldsymbol{L} = \boldsymbol{I}$) with random selection of the system matrix rows should converge in a least-squares sense [17]. However, we have observed that for some matrices $\boldsymbol{L}$ this is not the case.

Alternatively, one could try to reformulate (7) in a form that the splitting method in [15] can be applied in such a manner that the Kaczmarcz method can be used for the inconsistent linear system $\boldsymbol{A}\boldsymbol{x} \approx \boldsymbol{b}$. Let in the following the sparse prior be $\mathcal{R}(\boldsymbol{z}) = \|\boldsymbol{z}\|_1$ and let $\mathbb{D} = \mathbb{R}_+^N$. Then (7) reads

$$\arg\min_{\boldsymbol{x}\in\mathbb{R}_+^N} \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \frac{\beta}{2}\|\boldsymbol{z}\|_1 \quad \text{s.t} \quad \boldsymbol{L}\boldsymbol{x} = \boldsymbol{z}. \tag{22}$$

By adding the squared ℓ_2-norm of the variables $\boldsymbol{x}$ and $\boldsymbol{z}$, the splitting in [15] can be applied to

$$\begin{aligned} \arg\min_{\boldsymbol{x},\in\mathbb{R}_+^N\boldsymbol{z},\boldsymbol{v}} \frac{1}{2}\Big(\|\boldsymbol{v}\|_2^2 + \|\boldsymbol{x}\|_2^2 + \|\boldsymbol{z}\|_2^2\Big) + \frac{\beta}{2}\|\boldsymbol{z}\|_1 \\ \text{s.t} \quad \begin{pmatrix}\boldsymbol{A}, & -\delta\boldsymbol{I}_M\end{pmatrix}\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{v}\end{pmatrix} = \boldsymbol{b}, \quad \boldsymbol{L}\boldsymbol{x} = \boldsymbol{z}. \end{aligned} \tag{23}$$

This formulation allows one to perform the KA on the rows of $\begin{pmatrix}\boldsymbol{A}, & -\delta\boldsymbol{I}_M\end{pmatrix}$. For $\delta = 0$ we recover the algorithm from [16] with the problem in (21) with a non-negativity constraint.

The exact relationship between the algorithm in [16] and the proposed one can be verified by fixing the parameter δ and not resetting $\boldsymbol{\epsilon}$ and $\boldsymbol{v}$ in each iteration. Then, in the proposed algorithm, the problem in (7) with an additional regularization term $\delta^2\|\boldsymbol{x}\|_2^2$ will be optimized. When the new objective function is compared with the one in (23), it can be identified that the difference in both objective functions lies only in the additional quadratic regularization term regarding $\boldsymbol{z}$. The other obvious difference between both methods is that the parameter δ is steered depending on the development of the objective function in the proposed method.

VI. Experiments and Results

A. Reconstruction model and methods

For the regularization term $\mathcal{R}(\boldsymbol{L}\boldsymbol{x}) = \|\boldsymbol{L}\boldsymbol{x}\|_1$, we follow the proposal by Storath et al. in [8]. The authors used a weighted anisotropic total variation (TV) based regularization with an extended neighborhood relationship system. Additionally, they added an ℓ_1-norm regarding the particle distribution $\boldsymbol{x}$. This prior has shown good performances for the MPI task. For comparison purposes, we implemented the algorithm in [8], in the following called Storath method.

The entire reconstruction model reads

$$\arg\min_{\boldsymbol{x}\in\mathbb{R}_+^N} \|\boldsymbol{W}\left(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\right)\|_2^2 + \beta\sum_{s=1}^{S}\omega_s\|\nabla_s\boldsymbol{x}\|_1 + \alpha\|\boldsymbol{x}\|_1, \tag{24}$$

where ∇_s is the discretized gradient operator with respect to the neighborhood systems and ω_s denotes the corresponding weighting factor. The parameter α is set to $\frac{\beta}{4}$, which follows the empirically motivated suggestion in [8]. However, fine-tuning the parameters α and β without prior knowledge of the ground truth must be done quite carefully with respect to individual visual perception. The matrix $\boldsymbol{W}$ is a diagonal matrix that leads to normalized rows of the matrix $\boldsymbol{W}\boldsymbol{A}$.

The second implemented algorithm is the splitting methods by Condat [23], which allows for reconstruction without inversion of a matrix, called in the following the Condat method.

The third algorithm is the fast gradient projection [24], which is FISTA (fast iterative-shrinkage thresholding algorithm) in combination with an additional TV-regularization. It should be noted that the algorithm in the article is described for "classical" TV-based problem, but it can be extended to the problem (24).

Furthermore, we used the conjugate gradient (CG) method instead of the KA to solve (8), called the CG method in the following. The rest of the algorithm is the same as the proposed one.

The fifth algorithm is the KA based algorithm by Lorenz et al. [15] with the implementation following [16] and an additional parameter $\delta \geq 0$. For $\delta = 0$ it is called unregularized Lorenz, otherwise it is called regularized Lorenz.

For the proposed method, the inner iterations for solving (8) were limited to just one KA iteration. The same was used for the CG method, because no change in the convergence behavior was observed when the iteration number was changed.

Computations were carried out on a server with 56 CPUs of the type Intel Xeon E5-2680 v4 at 2.40 GHz using Matlab, where no particular parallel processing steps were used.

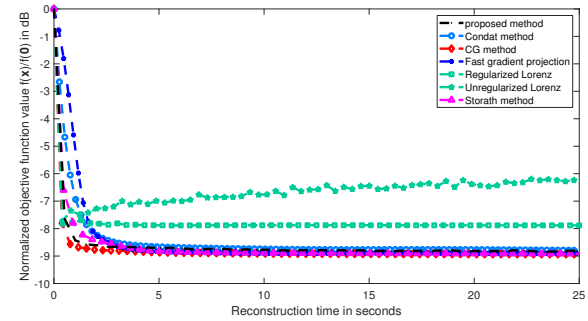
B. Simulation dataset

For the first test we simulated an MPI signal and the corresponding system matrix for a two-dimensional MPI scanner with a Lissajous FFP-trajectory. The simulations were performed with respect to the Langevin theory of paramagnetism [25]. The basis frequency of the scanner was set to $f_B = 2.5$ MHz, which was used as basis frequency for two drive fields working with the frequencies $f_x = \frac{f_B}{96}$ and $f_y = \frac{f_B}{93}$, respectively. This setup results in a periodic FFP-Lissajous trajectory with the frequency ratio of $f_x/f_y = 32/31$. The gradient field had a homogeneous field strength of 1 T/m. The resulting field of view, which is equivalent to the area of the FFP movement, had a size of 12.5×12.5 mm^2 and was discretized into 623×623 pixels to simulate the voltage signal. For reconstruction and to avoid so-called inverse crime, a second matrix was simulated with 256×256 pixels that was used for the reconstruction purpose. The tracer particle size was considered to be 30 nm.

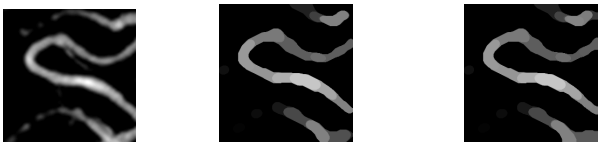
For the ground-truth particle distribution, we used a vessel angiography, which is shown in Fig. 1(b). The periodic voltage signal was corrupted with signal-independent additive white Gaussian noise, resulting in a signal-to-noise ratio (SNR) of 20 dB. Afterwards, we transformed the periodic MPI signal to the frequency domain and used only frequency components between 50 KHz and 2 MHz for the reconstruction. In addition, frequency components with an SNR lower than 1.5 were discarded, resulting in a highly underdetermined system matrix of the size 603×256^2.

C. Real-world dataset

An experiment with the real-world dataset from the Open-MPI database [26] was carried out. We decided to use the 3D MPI system matrix with 19^3 voxels spatial resolution and $3 \cdot 26929$ frequency components from three receive channels. We used the shape phantom to demonstrate that our algorithm converges to the same result as the algorithm in [24]. For further details on the dataset we refer the reader to [26]. In a first preprocessing step, the rows of the system matrix with an estimated SNR lower than 1.5 were discarded, resulting in a system matrix of size 20206×19^3. Due to the transfer function



(a)



(b) (c) (d)

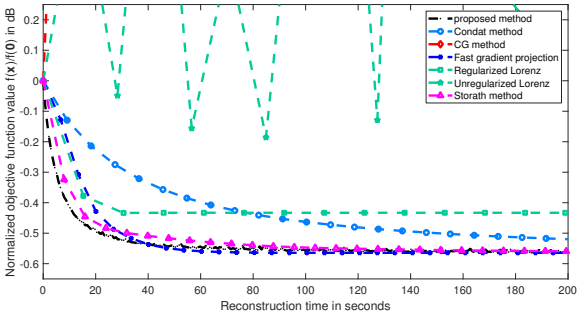
Fig. 1. Temporal progress and reconstruction results for phantom. (a) The temporal progress of the objective function as function of the reconstruction time. The objective function value is rescaled such that the value of $\boldsymbol{x} = \mathbf{0}$ is 0 dB in the figure. (b) The ground truth SPION distribution. (c) The reconstruction result by the algorithm by Storath et al. [8]. (d) The reconstruction result by the proposed algorithm.

of the scanner, also frequency components lower than 70 kHz were discarded. We used the empty measurements to estimate the scanner-background signal which was then subtracted from the system matrix and measurement vectors. This procedure has shown to increase the reconstruction quality in MPI [27]. The parameter β in (24) was set to 10^{-5}.

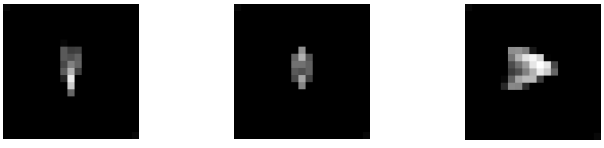
D. Results

Fig. 1(a) shows the objective function over the computation time for seven different algorithms for the simulated two-dimensional scenario in the first 25 seconds. We can clearly observe that the proposed method outperforms the Condat, fast gradient projection, and Storath algorithm. The proposed algorithm is slightly slower than the CG one. The unregularized Lorenz algorithm tends to diverge, whereas the regularized version converges fast on a relatively high-value for the objective function. Exemplarily, Fig. 1(b) shows the ground truth distribution, Fig. 1(c) the reconstruction with the Storath algorithm and Fig. 1(d) the reconstruction with the proposed algorithm after 2000 iterations. Because both methods optimize the same objective function, they nearly achieve the same result. Small derivations can be noticed in the upper left and lower right corners.

In Fig. 2(a) the temporal progress of the objective function as function of the reconstruction time is shown for the real-world 3D shape phantom. We show the reconstruction for 200 seconds. As in the first test example, the proposed method outperforms the other ones. After about 40 seconds the fast gradient projection method slightly surpassed the proposed ones. In contrast to the first test, the CG method did not work at all. This may be due to the ill-conditioning of the real-world system matrix, which itself includes a relatively

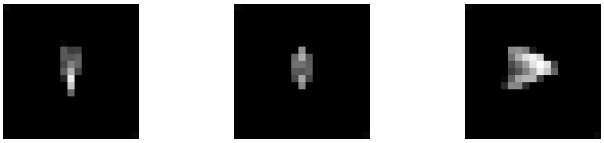


(a)



x-y plane, $z = 10$ x-z plane, $y = 8$ y-z, $x = 10$

(b)



x-y plane, $z = 10$ x-z plane, $y = 8$ y-z plane, $x = 10$

(c)

Fig. 2. Temporal progress and reconstruction results for shape phantom from OpenMPI database [26]. (a) The temporal progress of the objective function as function of the reconstruction time. The objective function value is rescaled such that the value of $\boldsymbol{x} = \mathbf{0}$ is 0 dB in the figure. (b) Different reconstruction planes for the fast gradient projection method [24]. (c) Different reconstruction planes for the proposed method.

high noise level. It can again be observed that the regularized Lorenz method is convergent, whereas the unregularized one shows a highly oscillating behavior. In Figs. 2(b) and (c) the reconstruction results are shown for the fast gradient projection algorithm from [24] and the proposed one. It can be observed that the results are nearly identical.

VII. Conclusions

We proposed a new formulation for Kaczmarz algorithm based on ADMM splitting for the purpose of MPI. Our new formulation allows for more sophisticated priors than the well-known Tikhonov regularization. We demonstrated the efficiency of the developed algorithm in comparison to other modern splitting methods for inverse imaging problems. We were able to show that the reconstruction time can be significantly decreased in the MPI problem setting. Even more, the formulation is quite general, which allows us to treat different inverse problems within the same framework. Still, there are also open questions that should be investigated in the future, like, for example, how fast the convergence is in cases that are not well suited for the classical KA.

References

[1] G. T. Herman, *Fundamentals of computerized tomography: Image reconstruction from projection*, ser. Advances in Computer Vision and Pattern Recognition. Springer-Verlag London, 2009, vol. 2.

[2] Z.-P. Liang and P. C. Lauterbur, *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*. IEEE Publications, 1999.

[3] B. Gleich and J. Weizenecker, "Tomographic imaging using the nonlinear response of magnetic particles," *Nature*, vol. 435, no. 7046, pp. 1214–1217, 2005.

[4] T. Knopp and T. M. Buzug, *Magnetic Particle Imaging: An Introduction to Imaging Principles and Scanner Instrumentation*. Berlin/Heidelberg: Springer, 2012.

[5] S. Kaczmarz, "Angenäherte Auflösung von Systemen linearer Gleichungen," *Bulletin International de l'Académie Polonaise des Sciences et des Lettres*, vol. A, pp. 355–357, 1937.

[6] T. Knopp, J. Rahmer, T. Sattel, S. Biederer, J. Weizenecker, B. Gleich, J. Borgert, and T. Buzug, "Weighted iterative reconstruction for magnetic particle imaging," *Phys. Med. Biol.*, vol. 55, no. 6, pp. 1577–1589, 2010.

[7] M. A. Saunders, "Solutions of sparse rectangular systems using LSQR and CRAIG," *BIT Numer. Math.*, vol. 35, no. 4, pp. 588–604, 1995.

[8] 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. Imaging*, vol. 36, no. 1, pp. 74–85, 2017.

[9] C. Bathke, T. Kluth, C. Brandt, and P. Maass, "Improved image reconstruction in magnetic particle imaging using structural a priori information," *Int. J. Mag. Part. Imag.*, vol. 3, no. 1, 2017.

[10] C. Droigk, M. Maass, and A. Mertins, "Multiresolution vessel detection in magnetic particle imaging using wavelets and a Gaussian mixture model," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, pp. 1913–1921, 2019.

[11] C. Popa and R. Zdunek, "Penalized least-squares image reconstruction for borehole tomography," in *Proc. ALGORITMY*, Vysoké Tatry, Slovakia, 2005, pp. 260–269.

[12] ——, "Regularized ART with gibbs priors for tomographic image reconstruction," in *Proc. of ASIM Conference*, 2005.

[13] N. Komodakis and J.-C. Pesquet, "Playing with duality," *IEEE Signal. Proc. Mag.*, vol. 32, no. 6, pp. 31–54, 2015.

[14] A. Beck, *First.-Order Methods in Optimization*, ser. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, 2017.

[15] D. A. Lorenz, F. Schöpfer, and S. Wenger, "The linearized Bregman method via split feasibility problems: Analysis and generalizations," *Linear Algebra Appl.*, vol. 7, no. 2, pp. 1237–1262, 2014.

[16] D. A. Lorenz, S. Wenger, F. Schöpfer, and M. Magnor, "A sparse Kaczmarz solver and a linearized Bregman method for online compressed sensing," in *Proc. IEEE Int. Conf. Image Proc.*, 2014, pp. 1347–1351.

[17] F. Schöpfer and D. A. Lorenz, "Linear convergence of the randomized sparse Kaczmarz method," *Math. Program.*, vol. 173, pp. 509–536, 2019.

[18] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing," *SIAM J. Imag. Sci.*, vol. 1, no. 1, pp. 143–168, 2008.

[19] W. Yin, "Analysis and generalizations of the linearized Bregman method," *SIAM J. Imag. Sci.*, vol. 3, no. 4, pp. 856–877, 2010.

[20] C. Byrne, "Iterative oblique projection onto convex sets and the split feasibility problem," *Inverse Probl.*, vol. 18, no. 2, pp. 441–453, 2002.

[21] T. Elfving, "Block-iterative methods for consistent and inconsistent linear equations," *Numer. Math. (Heidelb.)*, vol. 35, pp. 1–12, 1980.

[22] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Optimization*, vol. 3, no. 1, pp. 1–122, 2010.

[23] L. Condat, "A primal–dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms," *J. Optim. Theory. Appl.*, vol. 158, no. 2, pp. 460–479, 2013.

[24] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, 2009.

[25] J. Rahmer, J. Weizenecker, B. Gleich, and J. Borgert, "Signal encoding in magnetic particle imaging: Properties of the system function," *BMC Med. Imaging*, vol. 9, no. 4, p. 4, 2009.

[26] T. Knopp, P. Szwargulski, F. Griese, and M. Gräser, "OpenMPIData: An initiative for freely accessible magnetic particle imaging data," *Data in Brief*, vol. 28, p. 104971, 2020.

[27] 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. Imaging*, vol. 35, no. 3, pp. 893–900, 2016.