A Strategy for the Extension of the Kaczmarz Algorithm to Different Priors

M. Maass[a*], C. Droigk[a], P. Koch[a], and A. Mertins[a]

[a] *Institute for Signal Processing, University of Lübeck, Lübeck, Germany*
[*] *Corresponding author, email: maass@isip.uni-luebeck.de*

Abstract: This work introduces a strategy for the extension of the standard weighted Kaczmarz algorithm, which is commonly used with Tikhonov regularization in system-matrix based magnetic particle imaging, to other priors. The proposed reformulation of the algorithm allows us to include more sophisticated priors while inheriting the fast convergence of the Kaczmarz iteration. The new method is developed with help of the alternating direction method of multipliers. The results show that also with a suboptimal alternating direction method of multiplier steps, the proposed algorithm solves the problem with very high convergence rate.

I. Introduction

Magnetic Particle Imaging (MPI) is based on the nonlinear magnetization characteristic of superparamagnetic iron-oxide nanoparticles (SPIOs). The superposition of different acceleration fields and a gradient field results in a field free point (FFP) that travels along a pre-defined trajectory. Commonly, for MPI scanners with a Lissajous FFP-trajectory, a system-matrix based approach is used for image reconstruction [1]. Therefore, often the Kaczmarz algorithm (KA) with an extended system is chosen [2] that solves a Tikhonov-regularized least squares problem. However, other approaches for the particle-distribution reconstruction with more sophisticated priors have also been proposed [3]. Typically, this results in extended reconstruction times. Thus, if a fast reconstruction is desired, still the KA is preferred in the community. Speculatively, the reason is that MPI is still in the preclinical field, and for a first fast reconstruction, fancy priors are not necessary. Another reason may be that the system matrix can be represented in frequency space with nearly orthogonal rows. This fact results in fast convergence for the KA. Additionally, a non-negativity constraint is easy to enforce. This leads to the question of whether it is possible to extend the KA to other priors. In [4], such an attempt has been made, but the method is inefficient, because row- and column-wise operations are needed. In the present work, a direct extension to the KA is derived that allows one to include sophisticated priors and that operates only row-wise. The developed formulation of the reconstruction problem is based on splitting methods, like the alternated direction method of multiplier (ADMM) [5].

II. Material and Methods

In MPI, the optimization problem is usually formulated as

$$\arg\min_{c\in\mathbb{R}_+^n} \|Sc-f\|_2^2+\lambda^2\|c\|_2^2, \tag{1}$$

where $S\in\mathbb{C}^{m\times n}$ denotes the system matrix in frequency space, $f\in\mathbb{C}^m$ are the measured frequency components, and $c\in\mathbb{R}_+^n$ is the unknown SPIO distribution. The problem (1) can be solved in a row-wise manner with help of the KA [6], where the trick lies in reformulating (1) into the equivalent problem

$$\arg\min_{c\in\mathbb{R}_+^n, v\in\mathbb{C}^m} \|c\|_2^2+\|v\|_2^2 \text{ s.t. } Sc+\lambda v=f, \tag{2}$$

which is consistent and can be solved by KA.

Now let us consider the use of a more general convex regularization function $\mathcal{R}(z)$:

$$\arg\min_{c\in\mathbb{R}_+^n} \|Sc-f\|_2^2+\beta\,\mathcal{R}(Lc), \tag{3}$$

where $L\in\mathbb{R}^{k\times n}$ is an arbitrary matrix. An example is the anisotropic total variation (A-TV), where $\mathcal{R}(z)=\|z\|_1$ and $L=\nabla\in\mathbb{R}^{dn\times n}$ is a discretized gradient operator with respect to d directions. The problem (3) can be rewritten as

$$\arg\min_{c\in\mathbb{R}_+^n} \|Sc-f\|_2^2+\beta\,\mathcal{R}(z) \text{ s.t. } Lc-z=0. \tag{4}$$

With help of the ADMM, Eq. (4) is split into

$$c^{k+1}=\arg\min_{c\in\mathbb{R}_+^n} \|Sc-f\|_2^2+\frac{\rho}{2}\|z^k-Lc+u^k\|_2^2, \tag{5}$$

$$\begin{aligned} z^{k+1} &= \arg\min_{z\in\mathbb{R}^d} \beta\,\mathcal{R}(z)+\frac{\rho}{2}\|z-Lc^{k+1}+u^k\|_2^2 \\ &= \operatorname{prox}_{\frac{\beta}{\rho}\mathcal{R}}(Lc^{k+1}-u^k), \end{aligned} \tag{6}$$

$$u^{k+1}=u^k+z^{k+1}-Lc^{k+1}. \tag{7}$$

The parameter $\rho>0$ is chosen adaptively by the strategy proposed in [5] (§3.4.1). To use the KA for solving (5), an additional damping parameter $\delta>0$ is introduced and the same trick as in (2) is used:

$$c^{k+1} = \arg\min_{c\in\mathbb{R}_+^n} \left\| \begin{pmatrix} S \\ \sqrt{\frac{\rho}{2}}\,L \end{pmatrix} c - \begin{pmatrix} f \\ \sqrt{\frac{\rho}{2}}(u^k + z^k) \end{pmatrix} \right\|_2^2 + \|\delta\, c\|_{2.}^2 \quad (8)$$

The entire matrix can either be split in a row-wise manner (KA I), or for the upper part with the system matrix S, a row-wise KA splitting is used, whereas for the lower part of the matrix $(\sqrt{\frac{\rho}{2}}L)$ a block KA splitting is applied (KA II).

With an adaptation strategy for δ, the objective function of the problem in (3) can be minimized. The δ is increased by a factor of 2.1 if the objective function of (3) becomes higher after one ADMM iteration, otherwise it is decreased by factor 0.9. Using this, the convergence rate is better than for the standard ADMM strategy, where the subproblem (5) is solved by a gradient descent (grad. descent) method based on FISTA [7]. For comparison purposes, also an exact solver of (5) (Matlab/*lsqnonneg)*, has been used. For the three inexact solvers (KA/grad. descent) for (5) an inner iteration is used, which was set in the tests to two iterations.

For evaluation, the test data have been simulated by the parameters in [8], but the FOV has been discretized to 50×50 pixels. The voltage signal has a signal-to-noise ratio of 20 dB. For the experiments, L and $\mathcal{R}(z)$ were chosen in such a way that the A-TV was optimized.

III. Results

In Fig. 1, the top row shows the phantom and reconstructions results. In the bottom left, the objective function is plotted vs. the number of ADMM iterations, where one iteration is defined by a full evaluation of the steps (5), (6), and (7). Of course, one iteration has different time consumptions for the different solvers. For comparison purposes, also a method in which (5) was solved in an exact manner is shown, which needs significantly more time than all other methods. Quite obviously, the KA solving strategies that use only two inner iterations for the approximation of problem (5) are (in terms of the objective function) nearly as good as the optimal ADMM strategy.

The gradient-based method has significant problems to follow the KAs. It should be noted that the situation becomes better if the number of inner iterations is chosen higher. However, this comes with an increase of the calculation times for one full iteration.

Interestingly, when we look at the mean squared error (MSE) in Fig. 1, bottom right, the optimal ADMM strategy starts to become a little bit worse, whereas the KA strategies always find an optimum.

IV. Discussion

The KAs clearly outperform the gradient descent method, because the system matrix rows are nearly orthogonal to each other. It should be mentioned that grad. descent can be accelerated if the system matrix rows are energy normalized. The energy normalization comes with the drawback that the noise floor is increased within the frequency components and a selection of frequency components becomes

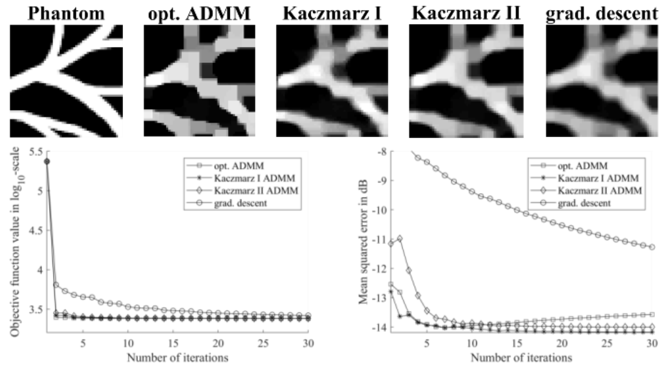

Figure 1: *Top: Reconstruction results after 100 iterations and two inner iterations for the approximation of (5). Bottom Left: The value of the objective function vs. the number of iterations for an anisotropic total variation problem solved by different solvers. Bottom Right: Mean squared error vs iterations.*

unavoidable. In contrast, KA is scale-invariant with respect to the rows of the matrix and the frequency selection is generally unnecessary, but can also help.

V. Conclusions

In this work, a strategy for the use of the Kaczmarz reconstruction in a row-wise manner with more sophisticated priors has been introduced and tested in the 2D total variation regularized setting. The method has been formulated in a generalized form and can be applied to different convex priors as well. If a closed-form solution for (6) is known or the result can be calculated, the algorithm can be efficiently implemented. The extension of the algorithm to non-convex priors is one of the next targets.

AUTHOR'S STATEMENT

This work was supported by the German Research Foundation under grant number ME 1170/7-1. Authors state no conflict of interest.

REFERENCES

[1] T. Knopp and T. M. Buzug. *Magnetic Particle Imaging: An Introduction to Imaging Principles and Scanner Instrumentation*. Springer, Berlin/Heidelberg, 2012. doi: 10.1007/978-3-642-04199-0.

[2] 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. doi: 10.1088/0031-9155/55/6/003.

[3] 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, doi:10.1109/TMI.2016.2593954.

[4] C. Popa and R. Zdunek. Penalized Least-Squares Image Reconstruction for Borehole Tomography. *Proc. of ALGORITMY*, pp. 260-269,2005.

[5] 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 Machine Learning*, 3(1):1-122, 2010, doi: 10.1561/2200000016.

[6] G. T. Herman, A. Lent, and H. Hurwitz. A Storage-Efficient Algorithm for Finding the Regularized Solution of Large, Inconsistent System of Equations. *J. Inst. Maths Applics*, 25(4):361-366, 1980.

[7] A. Beck and M. Teboulle. A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. *SIAM J. Imaging Sci.*, 2(1):183–202, 2009, doi:10.1137/080716542.

[8] M. Maass, M. Ahlborg, A. Bakenecker, F. Katzberg, H. Phan, T. M. Buzug, and A. Mertins. A Trajectory Study for Obtaining MPI System Matrices in a Compressed-Sensing Framework. *Int. J. Magn. Part. Imaging*, 3(2), 2017, doi:10.18416/ijmpi.2017.1706005.