Convergence rates for Kaczmarz type regularization methods

February 10, 2014

STEFAN KINDERMAN[1]
Institute for Industrial Mathematics, Johannes Kepler University, A-4040 Linz, Austria

ANTONIO LEITÃO
Department of Mathematics, Federal University of St. Catarina
P.O. Box 476, 88040-900 Florianópolis, Brazil

Abstract

This article is devoted to the convergence analysis of a special family of iterative regularization methods for solving systems of ill–posed operator equations in Hilbert spaces, namely Kaczmarz type methods. The analysis is focused on the Landweber–Kaczmarz (LK) explicit iteration an the iterated Tikhonov–Kaczmarz (iTK) implicit iteration. The corresponding symmetric versions of these iterative methods are also investigated (sLK and siTK). We prove convergence rates for the four methods above, extending and complementing the convergence analysis established originally in [24, 13, 12, 7].

MSC. Primary: 65J20, 65J15; Secondary: 47J06.

Keywords. Ill-posed systems; Landweber–Kaczmarz; Convergence rates; Regularization.

1 Introduction

Inverse problems under consideration

We consider ill-posed problems with a forward operator that has a block structure: Let

$$A_i : X \to Y_i$$

be linear, where X, Y_i are real Hilbert spaces. Whenever necessary, we shall denote by $X_{\mathbb{C}}$ the complexified version of a Hilbert space, i.e., the set of all $x_1 + ix_2$ with $x_1, x_2 \in X$.

Our goal is to solve the system of p equations

$$A_i x = y_i\,, \quad i = 0, \ldots, p-1\,, \tag{1}$$

where y_i are given (possibly noisy) data and the system is assumed to be ill-posed or ill-conditioned. In order to use a common framework, we define the operator $\mathbf{A}$ and the data vector $\mathbf{y}$ by

$$\mathbf{A} = \begin{pmatrix} A_0 \\ \vdots \\ A_{p-1} \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_0 \\ \vdots \\ y_{p-1} \end{pmatrix}.$$

[1]Corresponding author.

The operator $\mathbf{A}$ maps from X to the Hilbert space $\mathbf{Y} = (Y_0, \ldots, Y_{p-1})$. Thus, the equation to solve is now

$$\mathbf{A}x = \mathbf{y}\,. \tag{2}$$

In what follows, bold variables are used to denote block-structured ones.

We also represent the noisy data by $\mathbf{y}^\delta$, which satisfy the bound

$$\|y_i^\delta - y_i\| \le \delta_i\,, \quad i = 0, \ldots, p-1\,, \tag{3}$$

with noise level in $\mathbf{Y}$ satisfying

$$\delta^2 = \sum_{i=0}^{p-1} \delta_i^2\,. \tag{4}$$

In what follows we only require the overall noise level δ, instead of information on the particular δ_i.

Kaczmarz type methods

For ill-posed and ill-conditioned problems with a block structure, the class of Kaczmarz-type iteration is a useful iterative regularization method. The original Kaczmarz iteration [19] consists of a sequence of successive orthogonal projections (performed in a cyclic way), aiming to solve a system of linear equations in Hilbert spaces. This method was successfully applied to the inverse problem of computerized tomography [31] and was named *Algebraic Reconstruction Technique* (ART). We refer the reader to [32] for the application of the Kaczmarz method to other relevant inverse problems with bilinear structure. It is worth mentioning that the Kaczmarz iteration is closely related to the method of *adjoint fields* cited in the engineering literature [4]. For convergence analysis of the Kaczmarz method we refer the reader to [26, 27] (infinite dimensional spaces) and [28] (finite dimension). Acceleration of the Kaczmarz iteration for inconsistent linear systems is obtained in [18] by applying under-relaxation. Continuous and semicontinuous versions of Kaczmarz' method for the numerical resolution of linear algebraic equations arise from tomography and other areas of reconstruction from projections [30].

It is immediate to observe that Kaczmarz' strategy can be used in conjunction with any iterative method for solving ill-posed problems, e.g., gradient type methods (Landweber, Steepest descent [11]) or Newton type methods (Levenberg-Marquardt [20], IRGN [2], REGINN [36]). Essentially, one applies one iterative step of the chosen method to each of the equations of the system cyclically.

The investigation of Landweber–Kaczmarz methods for nonlinear ill-posed problems was initiated about ten years ago [24], where convergence of the iteration (without rates) was proven in case of exact data (the convergence proof for inexact data was incomplete). A complete convergence proof in the noisy data case (again without rates) was given in [14], where the authors introduced the *loping Landweber–Kaczmarz* iteration and changed the stop criteria in order to carry out the convergence proofs.

In what follows we give a brief overview on convergence analysis results for Kaczmarz-type methods (for both linear and nonlinear problems):

[2006] Iteratively-Regularized-Gauss-Newton–Kaczmarz [5]; convergence with rates;

[2007] Landweber–Kaczmarz [14, 13]; convergence without rates;

[2008] Steepest–Descent–Kaczmarz [8]; convergence without rates;

[2009] Expectation–Maximization–Kaczmarz [15]; convergence without rates;

[2009] Block–Landweber–Kaczmarz [12]; convergence without rates for linear systems;

[2010] Levenberg–Marquardt–Kaczmarz [3]; convergence without rates;

[2011] Iterated Tikhonov–Kaczmarz [7]; convergence without rates;

[2011] Parallel–Regularized–Newton–Kaczmarz [1]; convergence results with rates;

Notice that in [5] rates of convergence are obtained. In this article however, the assumptions on the nonlinearity of the operator equation (modeling the inverse problem) are by far the strongest. Convergence rate results can also be found in [1]. However, the method described there is not a cyclic (sequential) iteration, but it consists of solving in parallel all equations of the system and then computing a convex combination of a (regularized) Newton step for each subproblem.

While convergence results in the remaining articles are obtained using essentially the *tangential cone condition* [37, 17], the convergence proof in [5] require more delicate (stronger) assumptions as the *adjoint range invariance condition* [11] and a uniform bound on the convergence of the regularization operators [5, Sec. 3.1, assumption (3.5)].
Moreover, in order to derive rates of convergence, source conditions (smoothness assumptions on the solution) are also required.

Aim and scope

Differently from other iterative regularization methods such as Landweber iteration, CG, or iterated Tikhonov, a satisfactory convergence rate analysis for Kaczmarz-type iterations is not yet available, even in the simplest case of linear problems in Hilbert spaces. A possible explanation is the fact that Kaczmarz-type methods can be seen as nonsymmetric preconditioned versions of usual Richardson/Landweber type iterations (therefore, standard spectral theoretical approach cannot be used to derive rates).

The goal of this paper to close this gap and to establish a convergence rates analysis of the symmetric and nonsymmetric, implicit and explicit Landweber–Kaczmarz type iteration.

Our approach is based on the well-known formulation of these iterations as Gauss-Seidel preconditioned Landweber iteration [31, 10]. Moreover, we use the holomorphic functional calculus and functional calculus on the numerical range to obtain estimates for the approximation error and the propagated data error. In combination this leads to error estimates and convergence rates (using appropriate parameter choice rules), similar to the standard case for linear iterative regularization schemes. The methods of estimating the convergence rates in this paper can be found in the work of Plato [34] for sectorial operators (see also Nevanlinna [33]). These results, however, are for rather general operator equations, not necessarily ones coming from Kaczmarz-type iterations. The main difficulty concerning the use of these results is the problem of estimating the spectrum of the involved operators. In this work we use the numerical range, which is a spectral set, to derive some relevant inequalities. This allows to give computable conditions (see, e.g., (49), (50) or (55), (56)) for the convergence rates. It turns out that, for sufficiently small stepsizes, one always gets the standard Hölder convergence rates.

The paper is organized as follows. In Section 2 we define four Landweber–Kaczmarz type iterations. The first two are the classical method (here also referred to as the nonsymmetric LK method) and the iterated Tikhonov–Kaczmarz (iTK) method (its implicit version). Moreover, for each one of them we define their symmetric counterparts: sLK and siTK. We also compare them with the classical Landweber method and iterated Tikhonov method when applied to the full block system (i.e., when the Kaczmarz strategy is not applied). Furthermore in Section 2 we clarify the idea that the above mentioned Kaczmarz iterations can be seen as preconditioned version of the classical Landweber method or iterated Tikhonov method. In Section 3 we prove convergence rates for the nonsymmetric case including the classical Landweber–Kaczmarz method and the iterated Tikhonov–Kaczmarz (iTK) method. In Section 4 we discuss the obtained results. In Appendix A, for the sake of completeness of the presentation, we prove convergence rates for the symmetric iterations (sLK) and (siTK). This is done following the ideas in [11] to analyze iterative regularization methods in Hilbert spaces.

2 Kaczmarz and Block iteration methods

In this section we define the implicit and explicit Landweber–Kaczmarz iteration (symmetric and nonsymmetric versions of each method) applied to (2) and contrast them with the usual implicit or explicit Landweber iteration for block structured systems.

For generality we include p corresponding preconditioning operators M_i. In all of the following we assume the following condition:

$$M_i : Y_i \to Y_i \quad \text{symmetric bounded positive definite operators, } i = 0, \dots, p-1\,. \tag{5}$$

We collect the preconditioners into a block diagonal matrix $\mathbf{M}$

$$\mathbf{M} = \operatorname{diag}(M_i) \quad i = 0, \dots, p-1, \tag{6}$$

which is obviously symmetric bounded positive definite.

2.1 Nonsymmetric Kaczmarz-type iteration and block iterations

Let us first define the classical (nonsymmetric) *Landweber–Kaczmarz (LK) method* with preconditioning. The LK method defines a sequence of approximate solutions x_0, x_1, …, x_k, … to (2), which is based on the iteration

$$\begin{aligned} \bar{x}_{n+1} &= \bar{x}_n - A^*_{[n]} M_{[n]} (A_{[n]} \bar{x}_n - y^\delta_{[n]}) \qquad [n] := \operatorname{mod}(n,p) \\ x_k &:= \bar{x}_{kp} \qquad\qquad k = 0, 1, \dots \end{aligned} \tag{7}$$

starting at some initial element $\bar{x}_0$ and with M_i given as in (5). The approximate solutions to (1) are the iterates x_k. Hence, in order to compute x_{k+1} from x_k, one has to cycle through the equations (1) from top to bottom (i.e., $i = 0$ to $i = p-1$) performing Landweber-type steps. Commonly, the LK iteration is used with the trivial preconditioning $M_n = I$ or $M_{[n]} = \tau_{[n]} I$ with $\tau_{[n]} > 0$ being stepsize parameters.

This iteration can be compared with the one obtained by applying a standard Landweber iteration to the block system (1). This is called here *(block) Landweber method*, i.e., the sequence of approximate solutions to (2), x_0, x_1, …, x_k, … is defined by (compare with (7))

$$\begin{aligned} \bar{x}_{n+1} &= \bar{x}_n - A^*_{[n]} M_{[n]} (A_{[n]} x_{k-1} - y^\delta_{[n]})\,, \\ x_k &:= \bar{x}_{kp} \qquad\qquad k = 0, 1, \dots \end{aligned} \tag{8}$$

starting at some initial element $\bar{x}_0$. Equivalently, (8) can be written in the more common block form

$$x_k = x_{k-1} - \mathbf{A}^* \mathbf{M} (\mathbf{A} x_{k-1} - \mathbf{y}^\delta)\,, \tag{9}$$

with the block matrix $\mathbf{M}$ given as in (6). Once again, a common preconditioner for the block Landweber iteration is the choice $\mathbf{M} = \tau I$ with a positive stepsize parameter τ.

The block Landweber iteration can be seen as a sequence of explicit Euler steps for the gradient flow of the least squares functional for (2). For ill-posed operator equations, the implicit version of the Landweber iteration is usually called the iterated Tikhonov (iT) method. The implicit version of the block Landweber iteration (9) or (8) is here refered to as *(block) iterated Tikhonov regularization*, with iterations x_k given by:

$$\begin{aligned} \bar{x}_{n+1} &= \bar{x}_n - A^*_{[n]} M_{[n]} (A_{[n]} x_{n+1} - y^\delta_{[n]}) \,, \\ x_k &:= \bar{x}_{kp} \qquad\qquad k = 0, 1, \dots \end{aligned} \tag{10}$$

or, more commonly, written as block iteration in the form

$$x_k = x_{k-1} - \mathbf{A}^* \mathbf{M} \mathbf{A} (x_k - \mathbf{y}^\delta) \,, \tag{11}$$

with the operator $\mathbf{M}$ defined as in (6). This iteration is well-defined if $I + \mathbf{A}^*\mathbf{M}\mathbf{A}$, is invertible. Note that for computations, the expression (11) is used with

$$x_k = (I + \mathbf{A}^*\mathbf{M}\mathbf{A})^{-1} \left(x_{k-1} + \mathbf{A}^*\mathbf{M}\mathbf{y}^\delta \right) .$$

Both the block Landweber iteration and LK iteration have implicit counterparts. The implicit variant of the LK iteration (7) is the *iterated Tikhonov–Kaczmarz (iTK) method*, and is defined by (compare with (7))

$$\begin{aligned} \bar{x}_{n+1} &= \bar{x}_n - A^*_{[n]} M_{[n]} (A_{[n]} \bar{x}_{n+1} - y^\delta_{[n]}) \,, \\ x_k &:= \bar{x}_{kp} \qquad\qquad k = 0, 1, \dots \end{aligned} \tag{12}$$

starting from an arbitrary initial guess $\bar{x}_0$. This iteration is well-defined if all the operators $I + A^*_{[n]} M_{[n]} A_{[n]}$ are invertible (in each step of a cycle we have to solve a linear system involving this operator). Notice that, with the common choice $M_{[n]} = \frac{1}{\alpha} I$, a problem of Tikhonov-regularization type has to be solved in each step.

2.2 Symmetric Kaczmarz-type iterations

A further variant of the Kaczmarz type iterations are their symmetric versions [10]. Note that in contrast to the block-iterations, the iterations LK, iTK, are not invariant if the ordering of the equations are reversed. For this reasons (and since they are induced by a nonsymmetric block preconditioning) we call them the nonsymmetric Kaczmarz-type iterations. In what follows we define symmetric variants of LK and iTK.

At first a usual Kaczmarz cycle is performed, followed by another cycle, in which the order of the equations is reversed. I.e., in the second cycle the first iteration starts with A_{p-1}, y_{p-1}, followed by one with A_{p-2}, y_{p-2}. This yields the *symmetric Landweber–Kaczmarz (sLK) method*

$$\begin{aligned} \bar{x}_{n+1} &= \begin{cases} \bar{x}_n - A^*_{[n]} M_{[n]} (A_{[n]} \bar{x}_n - y^\delta_{[n]}) \\ \qquad\qquad \text{if } 0 \le \mathrm{mod}(n, 2p) \le p-1 \\ \bar{x}_n - A^*_{p-1-[n]} M_{p-1-[n]} (A_{p-1-[n]} \bar{x}_n - y^\delta_{p-1-[n]}) \\ \qquad\qquad \text{if } p \le \mathrm{mod}(n, 2p) \le 2p-1 \end{cases} \\ x_k &:= \bar{x}_{k2p} \qquad [n] := \mathrm{mod}(n, p) \,. \end{aligned} \tag{13}$$

For completeness of the presentation, we also define the symmetric variant of the iTK method, namely the *symmetric iterated Tikhonov Kaczmarz (siTK) method*, which is given by

$$\begin{aligned} \bar{x}_{n+1} &= \begin{cases} \bar{x}_n - A^*_{[n]} M_{[n]} (A_{[n]} \bar{x}_{n+1} - y^\delta_{[n]}) \\ \qquad\qquad \text{if } 0 \le \mathrm{mod}(n, 2p) \le p-1 \\ \bar{x}_n - A^*_{p-1-[n]} M_{p-1-[n]} (A_{p-1-[n]} \bar{x}_{n+1} - y^\delta_{p-1-[n]}) \\ \qquad\qquad \text{if } p \le \mathrm{mod}(n, 2p) \le 2p-1 \end{cases} \\ x_k &:= \bar{x}_{k2p} \qquad [n] := \mathrm{mod}(n, p) \,. \end{aligned} \tag{14}$$

It is easy to see that the symmetric versions double the computational amount per overall iterations. Moreover, comparing the LK method with the block Landweber iteration it is clear that the computational complexity is about the same, but the former is simpler since it does not requires one to store the old iterates x_k.

2.3 Kaczmarz iterations and Gauss-Seidel-preconditioning

The approach for a convergence analysis of Kaczmarz iterations is based on the fact that these methods can be expressed as ordinary Landweber (respectively iterated Tikhonov) iterations preconditioned with a suitable preconditioner. This has already been observed by Natterer [31], who showed that the classical Kaczmarz method with preconditioner $M_i = \omega(A_i A_i^*)^{-1}$ equals an SOR-method. For general preconditioning matrices M_i, the equivalence of the Landweber–Kaczmarz method to a Gauss-Seidel preconditioned Landweber iteration was shown by Elfving and Nikazad [10]. In this section we extend their results to the iterated Tikhonov method. We mostly stick here to the notation in [10].

Let us define the lower triangular operator $\mathbf{L} : \mathbf{Y} \to \mathbf{Y}$ and the block diagonal operator $\mathbf{D}$ as follows

$$\mathbf{L} := \begin{pmatrix} 0 & & & 0 \\ A_1 A_0^* & \ddots & & \\ \vdots & \ddots & \ddots & \\ A_{p-1}A_0^* & \dots & A_{p-1}A_{p-2}^* & 0 \end{pmatrix}, \quad \mathbf{D} := \operatorname{diag}(M_i^{-1}), i = 0, \dots, p-1\,. \tag{15}$$

Then the following result holds true [10]:

Theorem 1. *Let x_k be the iterates of the Landweber–Kaczmarz method (7) and let all M_i be invertible. Then the iteration (7) can be expressed as (nonsymmetric) preconditioned Landweber method of the form*

$$x_{k+1} = x_k - \mathbf{A}^* \mathbf{M}_B (\mathbf{A} x_k - \mathbf{y})\,, \tag{16}$$

with

$$\mathbf{M}_B = (\mathbf{D} + \mathbf{L})^{-1} \tag{17}$$

and $\mathbf{L}, \mathbf{D}$ *as in* (15).

Notice that, if all M_i are invertible, so is $\mathbf{M}_B$ since this is a lower triangular operator. However, except for nontrivial cases, the operator $\mathbf{M}_B$ is not symmetric. Thus, (16) cannot be seen as a symmetric preconditioned version of the classical block Landweber iteration (9).

A brief inspection of the proofs of this theorem shows that $(\mathbf{D} + \mathbf{L})^{-1}$ is not necessarily the only possible choice for $\mathbf{M}_B$. Actually, any operator $\mathbf{M}_B$ satisfying

$$\mathbf{A}^* \mathbf{M}_B (\mathbf{D} + \mathbf{L}) = \mathbf{A}^*$$

can be used as well in the iteration (16). A similar theorem, again due to Elfving and Nikazad [10], holds for the sLK method:

Theorem 2. *Let x_k be the iterates of the symmetric Landweber–Kaczmarz method (13) and let all M_i be invertible. Then the iteration (13) can expressed as a preconditioned Landweber iteration*

$$x_{k+1} = x_k - \mathbf{A}^* \mathbf{M}_{SB} (\mathbf{A} x_k - \mathbf{y})\,, \tag{18}$$

with

$$\mathbf{M}_{SB} = \mathbf{M}_B^* (2\mathbf{D} - \operatorname{diag}(A_i A_i^*)) \mathbf{M}_B\,, \tag{19}$$

and $\mathbf{L}$, $\mathbf{D}$ *as in* (15), *and* $\mathbf{M}_B$ *as in* (17).

In contrast to Theorem 1, we have here a symmetric preconditioning operator $\mathbf{M}_{SB}$, which also justifies the notion of symmetric/nonsymmetric iterations. The results in Theorem 2 carry over to the iterated Tikhonov case.

Theorem 3. *Let all M_i, $i = 0, \dots, p-1$ be invertible. The sequence $\{x_k\}$ generated by the iterated Tikhonov–Kaczmarz method (12) can be expressed as a (nonsymmetric) preconditioned block iterated Tikhonov iteration*

$$x_{k+1} = x_k - \mathbf{A}^*\mathbf{N}_B(\mathbf{A}x_{k+1} - \mathbf{y}) , \tag{20}$$

with

$$\mathbf{N}_B = ((\mathbf{D} - \mathbf{L})^{-1})^*. \tag{21}$$

Similarly, the iterates x_k of the symmetric iterated Tikhonov–Kaczmarz method (14) can be expressed as preconditioned iterated Tikhonov method

$$x_{k+1} = x_k - \mathbf{A}^*\mathbf{N}_{SB}(\mathbf{A}x_{k+1} - \mathbf{y}) , \tag{22}$$

with

$$\mathbf{N}_{SB} = \mathbf{N}_B \left(2\mathbf{D} + \operatorname{diag}(A_i A_i^*)\right) \mathbf{N}_B^*.$$

Proof. The iTK iterates satisfy

$$\bar{x}_n = \bar{x}_{n+1} + A_{[n]}^* M_{[n]}(A_{[n]}\bar{x}_{n+1} - y_{[n]}^\delta).$$

Define the permutation operator $\mathbf{P}$ that reverses the order of equations, i.e.,

$$\mathbf{P}\begin{pmatrix} z_0 \\ z_1 \\ \dots \\ z_{p-2} \\ z_{p-1} \end{pmatrix} = \begin{pmatrix} z_{p-1} \\ z_{p-2} \\ \dots \\ z_1 \\ z_0 \end{pmatrix}.$$

Moreover, define the vector $\omega = (\omega_0, \omega_1, \dots, \omega_p) := (\bar{x}_p, \bar{x}_{p-1}, \dots, \bar{x}_1, \bar{x}_0)$. Thus, ω_p can be expressed as the result of one cycle of a Landweber–Kaczmarz iteration (7) with initial element ω_0 and with $\bar{\mathbf{A}} = \mathbf{P}\mathbf{A}$, $\bar{\mathbf{y}} = \mathbf{P}\mathbf{y}$, $\bar{M}_{[n]} = -\left(\mathbf{P}\operatorname{diag}(M_i)\right)_i$ respectively replacing $\mathbf{A}$, $\mathbf{y}$, $\mathbf{M}$ in (7) (i.e., with the ordering of the equations reversed). Thus, according to Theorem 1, we can express

$$\omega_p = \omega_0 - \bar{\mathbf{A}}^*\bar{\mathbf{M}}_B[\bar{\mathbf{A}}\omega_0 - \bar{\mathbf{y}}] .$$

Going back to the original variables, this means that

$$\bar{x}_0 = \bar{x}_p - \mathbf{A}^*\mathbf{P}^*\bar{\mathbf{M}}_B\mathbf{P}(\mathbf{A}\bar{x}_p - y) .$$

Rearranging terms, and using the fact $\mathbf{P}^*\bar{\mathbf{M}}_B\mathbf{P} = \mathbf{N}_B$, we obtain the desired result in the first case. From this we conclude that, in the symmetric case,

$$(I + \bar{\mathbf{A}}^*\bar{\mathbf{N}}_B\bar{\mathbf{A}})x_{2p} = x_p + \bar{\mathbf{A}}^*\bar{\mathbf{N}}_B\bar{\mathbf{y}} ,$$

where $\bar{\mathbf{N}}_B$ is defined as in (21), but with the order of the operators reversed. It can be verified that $\mathbf{P}\bar{\mathbf{L}}\mathbf{P} = \mathbf{L}^*$, and hence $\mathbf{P}\bar{\mathbf{N}}\mathbf{P} = \mathbf{N}^*$. Therefore, with respect to the original variables, we obtain

$$(I + \mathbf{A}^*\bar{\mathbf{N}}_B^*\mathbf{A})x_{2p} = x_p + \mathbf{A}^*\bar{\mathbf{N}}_B^*\mathbf{y}.$$

Now, multiplication with $(I + \mathbf{A}^*\bar{\mathbf{N}}_B\mathbf{A})$ from the right, together with the identity

$$(I + \mathbf{A}^*\bar{\mathbf{N}}_B\mathbf{A})(I + \mathbf{A}^*\bar{\mathbf{N}}_B^*\mathbf{A}) = I + \mathbf{A}^*(\mathbf{N}_{SB})\mathbf{A}$$

yield the desired result in the symmetric case. □

Theorems 1–3 allows us to use the convergence theory of ordinary Landweber method and iterated Tikhonov method in Hilbert spaces to establish convergence rates. The symmetric case is considered in Appendix A, and is a rather straightforward application of the according theory [11]. The nonsymmetric case, however, is more demanding and will be treated in details in the following section.

3 Analysis of the nonsymmetric methods

In this section we present the convergence analysis and establish convergence rates for the Landweber–Kaczmarz method (7) and for the iterated Tikhonov–Kaczmarz method. According to Theorem 1 this iteration can be written as a Richardson-type iteration of the form (16). The main difficulty compared to the symmetric LK method is that the operator $\mathbf{A}^*\mathbf{M}_B\mathbf{A}$ is not symmetric, except in trivial cases, e.g., when all operators A_i commute. Hence, the classical analysis based on self-adjoint operators cannot be applied.

For notational simplicity we define $\mathbf{G} := \mathbf{A}^*\mathbf{M}_B\mathbf{A}$. The equivalence between (7) and (16) can be written in the form (see [10])

$$I - \mathbf{G} = (I - A_{p-1}^* M_{p-1} A_{p-1})(I - A_{p-2}^* M_{p-2} A_{p-2}) \ \ldots \ (I - A_0^* M_0 A_0)\,, \tag{23}$$

which immediately yields the following result:

Lemma 3.1. *If* $\max_i \|A_i^* M_i^{\frac{1}{2}}\| < \sqrt{2}$*, then* $\mathbf{G}$ *is an accreative operator, i.e., it satisfies*

$$\mathrm{Re}(\mathbf{G}x, x) \geq 0\,, \quad \forall x \in X_{\mathbb{C}}\,.$$

Proof. By definition, for all $0 \leq i \leq p-1$, the operator $A_i^* M_i A_i$ is symmetric positive semidefinite with norm bounded by 1. Thus, $(I - A_i^* M_i A_i)$ is non-expansive, and so is $I - \mathbf{G}$. Consequently,

$$\mathrm{Re}(\mathbf{G}x, x) = (x, x) - \mathrm{Re}((I - \mathbf{G})x, x) \geq \|x\|^2 - \|(I - \mathbf{G})\|\, \|x\|^2 \geq 0$$

concluding the proof. □

It follows from Lemma 3.1 that the spectrum of $\mathbf{G}$ is contained in the positive half space $\sigma(\mathbf{G}) \subset \{\lambda \in \mathcal{C} \,|\, \mathrm{Re}(\lambda) \geq 0\}$, and that the well-known resolvent estimate

$$\|(\mathbf{G} + tI)^{-1}\| \leq C\frac{1}{t} \quad \forall t > 0,$$

hold true [23, Chpt. 3, Th 3.2], i.e., $\mathbf{G}$ is a weakly sectorial operator [35, 34]. For such operators the fractional powers $\mathbf{G}^\alpha$, $\alpha > 0$, are well-defined by means of a Dunford-Schwartz-type integral.

For the remaining of this section, we adopt the (standard) notation: x_k^δ denotes the iteration (16) with noisy data $\mathbf{y} = (y_i^\delta)_{i=0}^{p-1}$, x_k denotes the iteration (16) with exact $\mathbf{y} = (A_i x^\dagger)_{i=0}^{p-1}$. First we estimate the propagated data error:

Lemma 3.2. *Let* x_k^δ *be the iteration* (16) *with noisy data and* x_k *the iteration* (16) *with exact data. Then we have the following estimate with a constant* $C = C(\mathbf{A}, \mathbf{M}_B)$

$$\|x_k^\delta - x_k\| \leq Ck\delta\,. \tag{24}$$

If additionally

$$\sup_{k\in\mathbb{N}} \|(I - \mathbf{M}_B\mathbf{A}\mathbf{A}^*)^k\| \leq C_1\,, \tag{25}$$

holds with a constant C_1*, then*

$$\|x_k^\delta - x_k\| \leq C\sqrt{k}\delta\,, \tag{26}$$

where the constant C *is independent of* k.

Proof. As for classical Landweber iteration we may write

$$x_k^\delta - x_k = \sum_{j=0}^{k} (I - \mathbf{A}^*\mathbf{M}_B\mathbf{A})^j \mathbf{A}^*\mathbf{M}_B(\mathbf{y} - \mathbf{y}^\delta)\,.$$

Thus, since $I - \mathbf{G}$ is nonexpansive, (24) follows immediately with $C = \|\mathbf{A}^*\mathbf{M}_B\|$. Now assume that (25) holds true. Denoting by $g_L(x)$ the polynomial $g_L(x) = \sum_{j=0}^k (1-x)^j$, we can write

$$\begin{aligned}
&\|x_k^\delta - x_k\|^2 \\
&= \Big(g_L(\mathbf{A}^*\mathbf{M}_B\mathbf{A})\mathbf{A}^*\mathbf{M}_B(\mathbf{y} - \mathbf{y}^\delta), g_L(\mathbf{A}^*\mathbf{M}_B\mathbf{A})\mathbf{A}^*\mathbf{M}_B(\mathbf{y} - \mathbf{y}^\delta) \Big) \\
&= \left(\mathbf{A}\mathbf{A}^* \sum_{j=0}^{k} (I - \mathbf{M}_B\mathbf{A}\mathbf{A}^*)^j \mathbf{M}_B(\mathbf{y} - \mathbf{y}^\delta), g_L(\mathbf{M}_B\mathbf{A}\mathbf{A}^*)\mathbf{M}_B(\mathbf{y} - \mathbf{y}^\delta) \right) \\
&= \Big(\mathbf{M}_B^{-1}\mathbf{M}_B\mathbf{A}\mathbf{A}^* g_L(\mathbf{M}_B\mathbf{A}\mathbf{A}^*)\mathbf{M}_B(\mathbf{y} - \mathbf{y}^\delta), g_L(\mathbf{M}_B\mathbf{A}\mathbf{A}^*)\mathbf{M}_B(\mathbf{y} - \mathbf{y}^\delta) \Big) \\
&\le \|\mathbf{M}_B^{-1}(I - (I - \mathbf{M}_B\mathbf{A}\mathbf{A}^*)^{k+1})\mathbf{M}_B(\mathbf{y} - \mathbf{y}^\delta)\| \|g_L(\mathbf{M}_B\mathbf{A}\mathbf{A}^*)\mathbf{M}_B(\mathbf{y} - \mathbf{y}^\delta)\| \\
&\le \|\mathbf{M}_B^{-1}\| \|(I - (I - \mathbf{M}_B\mathbf{A}\mathbf{A}^*)^{k+1}\| \|\mathbf{M}_B(\mathbf{y} - \mathbf{y}^\delta)\| \\
&\qquad\qquad \|g_L(\mathbf{M}_B\mathbf{A}\mathbf{A}^*)\mathbf{M}_B(\mathbf{y} - \mathbf{y}^\delta)\| \\
&\le \|\mathbf{M}_B^{-1}\|(1 + C_1) \left(\sum_{j=0}^{k} \|(I - \mathbf{M}_B\mathbf{A}\mathbf{A}^*)\|^j \right) \|\mathbf{M}_B(\mathbf{y}^\delta - \mathbf{y})\| \|\mathbf{y}^\delta - \mathbf{y}\|.
\end{aligned}$$

This estimate, together with (25), yields an estimate of the order $k\delta^2$, completing the proof. □

We now have to estimate the propagated error term $x_k - x^\dagger$. For this purpose we need an auxiliary lemma. Roughly speaking, it states that $I - \mathbf{G}$ is a contraction for elements which are not in the null-space of $\mathbf{G}$. The precise formulation follows:

Lemma 3.3. *Let* $\max_i \|A_i^* M_i^{\frac{1}{2}}\| < \sqrt{2}$. *Moreover, let* $\eta > 0$ *and* $x \in X_\mathbb{C}$ *with* $\|x\| = 1$ *be given. If*

$$\mathrm{Re}\,(\mathbf{G}x, x) \ge \eta\,, \tag{27}$$

then there exists a positive $\gamma < 1$ *depending on* η, p *and* $(\|A_i^* M_i^{\frac{1}{2}}\|)_{i=0}^{p-1}$ *such that*

$$\|(I - \mathbf{G})x\|^2 \le 1 - \gamma\,. \tag{28}$$

Proof. Denote by $E_{\lambda,i}$ the spectral family associated to $A_i^* M_i A_i$, $i = 0, \dots, p-1$. Moreover, for each $\xi > 0$ define the orthogonal projectors

$$P_{\xi,i} = \int_{\lambda \le \xi} dE_{\lambda,i} \qquad Q_{\xi,i} = I - P_{\xi,i} = \int_{\lambda > \xi} dE_{\lambda,i}\,.$$

Note that these orthogonal projectors have norm one and satisfy $\|P_{\xi,i}x\|^2 + \|Q_{\xi,i}x\|^2 = \|x\|^2$.

Let $0 < \xi < 1$ be such that $(1-\xi)^2 \ge (1 - \|A_i^* M_i A_i\|)^2$, for all $i = 0, \dots, p-1$. From our assumption $\max_i \|A_i^* M_i^{\frac{1}{2}}\| < \sqrt{2}$, such a ξ can be chosen out of an interval $(0, \xi_0)$.

Next we define $\theta := 1-(1-\xi)^2 = 2\xi - \xi^2 > \xi > 0$. Using spectral calculus, we obtain for our ξ

$$\begin{aligned}\|(I-A_i^*M_iA_i)x\|^2 &= \int_{\lambda\le\xi}(1-\lambda)^2 d\|E_{\lambda,i}x\|^2 + \int_{\lambda>\xi}(1-\lambda)^2 d\|E_{\lambda,i}x\|^2\\ &\le \|P_{\xi,i}x\|^2 + \max\{(1-\xi)^2,(1-\|A_i^*M_iA_i\|)^2\}\|Q_{\xi,i}x\|^2\\ &= \|P_{\xi,i}x\|^2 + (1-\xi)^2(\|x\|^2-\|P_{\xi,i}x\|^2) = (1-(1-\xi)^2)\|P_{\xi,i}x\|^2 + (1-\xi)^2\|x\|^2\\ &= \theta\|P_{\xi,i}x\|^2 + (1-\theta)\|x\|^2,\end{aligned} \tag{29}$$

and

$$\begin{aligned}\|A_i^*M_iA_ix\|^2 = \int_{\lambda\le\xi}\lambda^2 d\|E_{\lambda,i}x\|^2 + \int_{\lambda>\xi}\lambda^2 d\|E_{\lambda,i}x\|^2 &\le \xi^2\|P_{\xi,i}x\|^2 + \|M_i^{\frac12}A_i\|^4\|Q_{\xi,i}x\|^2\\ &\le \xi^2\|P_{\xi,i}x\|^2 + 4\left(\|x\|^2-\|P_{\xi,i}x\|^2\right).\end{aligned} \tag{30}$$

Now define for each $k\le p-1$ the operators

$$\mathbf{H}_k = \Pi_{i=0}^k(I-A_i^*M_iA_i) \quad\Leftrightarrow\quad \mathbf{H}_k = \mathbf{H}_{k-1} - A_k^*M_kA_k\mathbf{H}_{k-1}, \quad \mathbf{H}_0 = (I-A_0^*M_0A_0).$$

We know that $\|\mathbf{H}_k\|\le 1$. Moreover, from the recursion formula

$$\mathbf{H}_k - I = \mathbf{H}_{k-1} - I - A_k^*M_kA_k(\mathbf{H}_{k-1}-I) - A_k^*M_kA_k = (I-A_k^*M_kA_k)(\mathbf{H}_{n-1}-I) - A_k^*M_kA_k$$

we conclude (using induction) that, for any given x,

$$\|(\mathbf{H}_k-I)x\| \le \|(\mathbf{H}_{k-1}-I)x\| + \|A_k^*M_kA_kx\| \le \sum_{i=0}^k \|A_i^*M_iA_ix\|.$$

Since $\mathbf{G} = \mathbf{G}-I+I = I-\mathbf{H}_{p-1}$, we have

$$\|\mathbf{G}x\| \le \sum_{i=0}^{p-1}\|A_i^*M_iA_ix\|. \tag{31}$$

Thus, applying (29), we obtain the estimate

$$\begin{aligned}\|\mathbf{H}_kx\|^2 &= \|(I-A_k^*M_kA_k)\mathbf{H}_{k-1}x\| \le \theta\|P_{\xi,k}\mathbf{H}_{k-1}x\|^2 + (1-\theta)\|\mathbf{H}_{k-1}x\|^2\\ &\le (1-\theta)\|\mathbf{H}_{k-1}x\|^2 + \theta\left(\|P_{\xi,k}x\| + \|P_{\xi,k}(I-\mathbf{H}_{k-1})x\|\right)^2\\ &\le (1-\theta)\|\mathbf{H}_{k-1}x\|^2 + \theta\left(\|P_{\xi,k}x\| + \sum_{i=0}^{k-1}\|A_i^*M_iA_ix\|\right)^2\\ &\le (1-\theta)\|x\|^2 + \theta\left(\|P_{\xi,k}x\| + \sum_{i=0}^{k-1}\|A_i^*M_iA_ix\|\right)^2.\end{aligned} \tag{32}$$

Define now the sequence of numbers

$$D_0 = 5, \quad D_k = (9 + 8\sum_{i=0}^{k-1}\sqrt{D_i}), \qquad k = 1,\dots p-1.$$

We prove Lemma 3.3 by contradiction. Let us assume that the assertion does not hold true. Then we would be able to find some $\eta>0$ such that, for any $\epsilon>0$, there would exist an x with

$$|((I-\mathbf{G})x,x)|^2 \ge (1-\epsilon), \quad \mathrm{Re}(\mathbf{G}x,x)\ge\eta \quad \|x\|=1. \tag{33}$$

We now take $0 < \epsilon < 1$ small enough such that

$$\epsilon \leq \left(\frac{1}{1 + \sum_{i=0}^{p-1} \sqrt{D_i}} \right)^{2^p} \tag{34}$$

$$(1 - \sqrt{\epsilon})^2 \geq \max_{i=0,\dots,p-1} (1 - \|A_i^* M_i A_i\|)^2 \tag{35}$$

$$\epsilon < \left(\frac{\eta}{\sum_{i=0}^{p-1} \sqrt{D_i}} \right)^{2^{p+1}} . \tag{36}$$

Since $|((I - \mathbf{G})x, x)| \leq \|(I - \mathbf{G})x\|$ and

$$\|(I - \mathbf{G})x\| = \|\mathbf{H}_{p-1}x\| \leq \|\mathbf{H}_{p-2}x\| \leq \cdots \leq \|\mathbf{H}_0 x\| ,$$

it follows that for such ϵ we find an x as in (33) with

$$\|\mathbf{H}_k x\|^2 \geq 1 - \epsilon \qquad \forall k = 0, \dots p-1 .$$

By (35), the choice $\xi = \sqrt{\epsilon}$ can be used in (29) and (32). Therefore, we obtain with $\theta = 2\sqrt{\epsilon} - \epsilon > \sqrt{\epsilon}$ the inequality

$$1 - \epsilon \leq (1 - \theta) + \theta \left(\|P_{\sqrt{\epsilon},k} x\| + \sum_{i=0}^{k-1} \|A_i^* M_i A_i x\| \right)^2 , \qquad \forall k = 0, \dots p-1 ,$$

yielding the estimate

$$\|P_{\sqrt{\epsilon},k} x\| + \sum_{i=0}^{k-1} \|A_i^* A_i x\| \geq \sqrt{1 - \frac{\epsilon}{\theta}} \geq 1 - \frac{\epsilon}{\theta} \geq 1 - \sqrt{\epsilon} , \qquad \forall k = 0, \dots p-1 . \tag{37}$$

Using (30), we get

$$\|A_k^* M_k A_k x\|^2 \leq \epsilon \|P_{\sqrt{\epsilon},k} x\|^2 + 4(1 - \|P_{\sqrt{\epsilon},k} x\|^2) \leq \epsilon + 4(1 - \|P_{\sqrt{\epsilon},k} x\|^2) . \tag{38}$$

For $k = 0$ we obtain from (37) and (38)

$$\|P_{\sqrt{\epsilon},0} x\|^2 \geq 1 - \sqrt{\epsilon} , \qquad \|A_0^* M_0 A_0 x\|^2 \leq \epsilon + 4\sqrt{\epsilon} \leq D_0 \sqrt{\epsilon} .$$

We proceed by induction to show that

$$\|A_k^* M_k A_k x\|^2 \leq D_k \epsilon^{\frac{1}{2^{k+1}}} \quad \forall k = 0, \dots, p-1 . \tag{39}$$

Using (37) and the induction hypothesis for $k - 1, k \geq 1$, we find

$$\|P_{\sqrt{\epsilon},k} x\| \geq 1 - \sqrt{\epsilon} - \sum_{i=0}^{k-1} \epsilon^{\frac{1}{2^{i+2}}} \sqrt{D_i} \geq 1 - \epsilon^{\frac{1}{2^{k+1}}} \left(1 + \sum_{i=0}^{k-1} \sqrt{D_i} \right) .$$

Notice that, by (34), the right hand side in this inequality is positive. Hence, by (38) we obtain

$$\begin{aligned}
\|A_k^* M_k A_k x\|^2 &\leq \epsilon + 4 \left[1 - \left(1 - \epsilon^{\frac{1}{2^{k+1}}} \left(1 + \sum_{i=0}^{k-1} \sqrt{D_i} \right) \right)^2 \right] \\
&\leq \epsilon^{\frac{1}{2^{k+1}}} + 4 \left[2\epsilon^{\frac{1}{2^{k+1}}} \left(1 + \sum_{i=0}^{k-1} \sqrt{D_i} \right) - \left(\epsilon^{\frac{1}{2^{k+1}}} \left(1 + \sum_{i=0}^{k-1} \sqrt{D_i} \right) \right)^2 \right] \\
&\leq \epsilon^{\frac{1}{2^{k+1}}} \left(9 + 8 \sum_{i=0}^{k-1} \sqrt{D_i} \right) = \epsilon^{\frac{1}{2^{k+1}}} D_k ,
\end{aligned}$$

which verifies (39) for all $k = 0, \dots p-1$. Now we derive a contradiction to (33). Using (31) and (36) we obtain

$$\eta \le \mathrm{Re}(\mathbf{G}x, x) \le \|\mathbf{G}x\| \le \sum_{i=0}^{p-1} \|A_i^* M_i A_i x\| \le \sum_{i=0}^{p-1} \epsilon^{\frac{1}{2^{i+2}}} \sqrt{D_i} \le \epsilon^{\frac{1}{2^{p+1}}} \sum_{i=0}^{p-1} \sqrt{D_i} < \eta\,.$$

Hence, (33) cannot be true for such an ϵ and the proof is complete. □

Remark 3.4. *It follows from the above proof that γ in (28) can be taken as the largest number ϵ for which (34)–(36) fails to hold. In particular, for η sufficiently small (such that (36) implies (34) and (35)), Lemma 3.3 holds true with the following choice of γ:*

$$\gamma = \left(\frac{\eta}{\sum_{i=0}^{p-1} \sqrt{D_i}} \right)^{2^{p+1}}.$$

Lemma 3.5. *Let $\max_i \|A_i^* M_i^{\frac{1}{2}}\| < \sqrt{2}$ and assume the existence of a constant $h > 0$ such that*

$$\begin{aligned} |([\mathbf{G} - \mathbf{G}^T]x, y)| \ &\le h\big((\mathbf{G}x, x) + (\mathbf{G}y, y)\big)\,, \\ &\qquad \forall\ x, y \in X,\ \|x\|^2 + \|y\|^2 = 1\,. \end{aligned} \tag{40}$$

Then, there exists a constant C depending on α, h, p and $\left(\|A_i^* M_i^{\frac{1}{2}}\|\right)_{i=0}^{p-1}$ such that the inequality

$$\|(I - \mathbf{G})^k \mathbf{G}^\alpha\| \ \le \frac{C}{(k+1)^\alpha}$$

holds true for all $\alpha > 0$.

Proof. It follows from (40) that the numerical range $W(\mathbf{G}) := \{(\mathbf{G}x, x) \mid x \in X_{\mathbb{C}},\ \|x\| = 1\}$ of the operator $\mathbf{G}$ is contained in the sector $\Sigma := \{\lambda \in \mathbb{C} \mid |\arg(\lambda)| \le \psi < \frac{\pi}{2}\}$, $h = \tan(\psi)$. Moreover, it follows from Lemma 3.3 and $|((I - \mathbf{G})x, x)| \le \|x\| \|(I - \mathbf{G})x\|$ that, for any $\eta > 0$, there exists a constant $0 < \gamma < 1$ with

$$W(\mathbf{G}) \subset (\Sigma \cap \{\lambda \in \mathbb{C} \mid \mathrm{Re}\lambda < \eta\}) \cup \{\lambda \in \mathbb{C} \mid \mathrm{Re}\lambda \ge \eta\,,\ |1 - \lambda| < \sqrt{(1-\gamma)}\}\,.$$

We now fix $\eta = \cos(\psi)^2$ and take γ as the corresponding constant in (28). Using the functional calculus of Crouzeix [6] we conclude that

$$\|(I - \mathbf{G})^k \mathbf{G}^\alpha\|_{X_{\mathbb{C}}} \le C_C \sup_{\lambda \in \Sigma} |(1-\lambda)^k \lambda^\alpha|\,,$$

for some constant $C_C \le 11.08$. For the first part of the numerical range, where $\mathrm{Re}\lambda < \eta$, we have

$$|\lambda| \le \mathrm{Re}(\lambda)\sqrt{1 + h^2} \le \frac{\eta}{\cos\psi} \le \cos(\psi)\,.$$

Hence,

$$|(1-\lambda)|^2 \ = \ 1 + |\lambda|^2 - 2|\lambda|\cos(\arg(\lambda)) \ \le \ 1 - |\lambda|\cos(\psi)\,,$$

which leads to the estimate

$$\begin{aligned} |(1-\lambda)^k \lambda^\alpha| \le \ &\big|1 - |\lambda|\cos(\psi)\big|^{\frac{k}{2}} |\lambda|^\alpha \ \le \ \frac{1}{\cos(\psi)^\alpha} \big|1 - |\lambda|\cos(\psi)\big|^{\frac{k}{2}} |\lambda\cos(\psi)|^\alpha \\ &\le \ \frac{\alpha^\alpha}{\cos(\psi)^\alpha} (\frac{k}{2})^{-\alpha} \ \le \ \frac{(4\alpha)^\alpha}{\cos(\psi)^\alpha} \frac{1}{(k+1)^\alpha}\,. \end{aligned}$$

For the other part of the numerical range we have

$$|(1-\lambda)^k \lambda^\alpha| \leq 2^\alpha (1-\gamma)^{\frac{k}{2}}\,,\ \forall \mathrm{Re}\lambda \geq \eta\,.$$

Since the inequality $(1-\gamma)^{\frac{k}{2}} \leq C'(k+1)^{-\alpha}$, for all $k \geq 0$, holds true with some constant C', the lemma follows with $C = \max\{\frac{(4\alpha)^\alpha}{\cos(\psi)^\alpha}, 2^\alpha C'\}$. □

Before we proceed, the introduction of some notation is necessary. We define

$$\hat{\mathbf{L}} := \mathbf{D}^{-\frac{1}{2}}\mathbf{L}\mathbf{D}^{-\frac{1}{2}}, \tag{41}$$

i.e., $\hat{\mathbf{L}}$ is a lower triangular matrix with zero diagonal similar to $\mathbf{L}$, but with $M_i^{\frac{1}{2}} A_i A_j^* M_j^{\frac{1}{2}}$ replacing $A_i A_j^*$. Moreover, we also define the matrix $|\mathbf{L}| \in \mathbb{R}^{p\times p}$ as the lower triangular matrix with zero diagonal

$$|\mathbf{L}|_{i,j} = \begin{cases} 0 & j \geq i \\ \|\hat{\mathbf{L}}_{i,j}\| & \text{else} \end{cases}\,,\quad i,j = 0,\dots p-1\,. \tag{42}$$

It is worth noticing that the matrix entries in $|\mathbf{L}|$ start from 0. Notice also that, in what follows, $|\mathbf{L}|$ can be replaced by any other lower triangular matrix with zero diagonal and satisfying

$$\|\hat{\mathbf{L}}_{i,j}\| \leq |\mathbf{L}|_{i,j}\,,\quad i<j\,,\ i,j = 0,\dots p-1\,. \tag{43}$$

Consequently, only an upper bound on the norm of $\|\hat{\mathbf{L}}_{i,j}\|$ is needed.

Remark 3.6. *We should analyze (40) in more detail. In terms of $\mathbf{A}$ and $\mathbf{M}_B$ this condition reads*

$$\begin{aligned} |(\mathbf{A}^*\mathbf{M}_B\mathbf{A} - \mathbf{A}^*\mathbf{M}_B^T\mathbf{A})x,y)| \leq & h\,((\mathbf{A}^*\mathbf{M}_B\mathbf{A}x,x) + (\mathbf{A}^*\mathbf{M}_B\mathbf{A}y,y)) \\ & \forall x,y,\in X, \|x\|^2 + \|y\|^2 = 1\,. \end{aligned} \tag{44}$$

Substituting $z = \mathbf{M}_B\mathbf{A}x$ and $v = \mathbf{M}_B\mathbf{A}y$, this condition is satisfied if, for all z, $v \in Y$, the inequality

$$|(\mathbf{M}_B^{-T} - \mathbf{M}_B^{-1}z,v)| \leq h\left((\mathbf{M}_B^{-1}z,z) + (\mathbf{M}_B^{-1}v,v)\right)$$

holds true. Let $\hat{\mathbf{L}}$ be as in (41). Then, the above inequality holds if, for all z, $v \in Y$,

$$|(\hat{\mathbf{L}}^T - \hat{\mathbf{L}})z,v)| \leq h\left((z,z) + (v,v) + (\hat{\mathbf{L}}z,z) + (\hat{\mathbf{L}}v,v)\right)\,. \tag{45}$$

This remark leads to the following lemma:

Lemma 3.7. *If q is such that*

$$\|\hat{\mathbf{L}}\| \leq q < 1\,, \tag{46}$$

then (40) holds with $h = q/(1-q)$.

Proof. We start by proving (45). Notice that

$$|(\hat{\mathbf{L}}^T - \hat{\mathbf{L}})z,v)| \leq |(z,\hat{\mathbf{L}}v)| + |(\hat{\mathbf{L}}z,v)| \leq \|z\|\|\hat{\mathbf{L}}v\| + \|\hat{\mathbf{L}}z\|\|v\|\,,$$

$$(\hat{\mathbf{L}}z,z) + (\hat{\mathbf{L}}v,v) \geq -\|\hat{\mathbf{L}}z\|\|z\| - \|\hat{\mathbf{L}}v\|\|v\|\,.$$

Thus, it suffices to prove

$$\|z\|\|\hat{\mathbf{L}}v\| + \|\hat{\mathbf{L}}z\|\|v\| + h\left(\|\hat{\mathbf{L}}z\|\|z\| + \|\hat{\mathbf{L}}v\|\|v\|\right) \ \leq\ h\|z\|^2 + \|v\|^2\,.$$

However, that this last inequality is a consequence of

$$\|z\|\|\hat{\mathbf{L}}v\| + \|\hat{\mathbf{L}}z\|\|v\| + h\left(\|\hat{\mathbf{L}}z\|\|z\| + \|\hat{\mathbf{L}}v\|\|v\|\right) \\ \leq q\left(2\|z\|\|v\| + h\|z\|^2 + h\|v\|^2\right) \leq q(1+h)\left(\|z\|^2 + \|v\|^2\right) .$$

Indeed, the choice $h = q/(1-q)$ allow us to estimate the right hand side of the above expression by $h(\|z\|^2 + \|v\|^2)$. □

It remains to investigate condition (25). For this purpose we rely on the following theorem [29, 25] (see also [21]):

Theorem 4. *Let T be a bounded operator on a complex Banach space. If there is a constant C such that*

$$\|(T - \lambda I)^{-1}\| \leq C\frac{1}{|\lambda - 1|}, \quad \forall\ |\lambda| > 1,\ \lambda \in \mathbb{C}, \tag{47}$$

then $\sup_n \|T^n\| < \infty$.

For the setting $T = I - \mathbf{M}_B\mathbf{A}\mathbf{A}^*$, we thus obtain (25) if we can prove

$$\|(\mathbf{M}_B\mathbf{A}\mathbf{A}^* - \lambda I)^{-1}\| \leq C\frac{1}{|\lambda|} \quad |\lambda - 1| > 1,\ \lambda \in \mathbb{C}\,. \tag{48}$$

We have the following result

Lemma 3.8. *Let all M_i be symmetric and positive definite. Define $\hat{\mathbf{A}} := \mathbf{D}^{-\frac{1}{2}}\mathbf{A}$, and take $\hat{\mathbf{L}}$ as in (41). If*

$$\|\hat{\mathbf{L}}\| + \frac{1}{2}\|\hat{\mathbf{A}}\hat{\mathbf{A}}^*\| < 1\,, \tag{49}$$

then (48) is satisfied. Consequently, (25) is also satisfied.

Proof. We prove (48). Let $\hat{\mathbf{A}} = \mathbf{D}^{-\frac{1}{2}}\mathbf{A}$, $\hat{\mathbf{L}}$ be given as in (41). Moreover, define $\hat{\mathbf{S}}(\lambda) := (\hat{\mathbf{A}}\hat{\mathbf{A}}^* - \lambda I)$. From the definition of $\mathbf{M}_B$, we obtain for each λ in (48)

$$\begin{aligned}
\|(\mathbf{M}_B\mathbf{A}\mathbf{A}^* - \lambda I)^{-1}\| &\leq \|\mathbf{M}_B^{-1}\|\|(\mathbf{A}\mathbf{A}^* - \lambda(\mathbf{D} + \mathbf{L}))^{-1}\| \\
&= \|\mathbf{M}_B^{-1}\|\|(\hat{\mathbf{A}}\hat{\mathbf{A}}^* - \lambda(I + \hat{\mathbf{L}}))^{-1}\|\|\mathbf{D}^{-\frac{1}{2}}\|^2 \\
&\leq \|\mathbf{M}_B^{-1}\|\|\mathbf{D}^{-\frac{1}{2}}\|^2\|\left(I - \hat{\mathbf{S}}(\lambda)^{-1}\lambda\hat{\mathbf{L}}\right)^{-1}\hat{\mathbf{S}}(\lambda)^{-1}\| \\
&\leq C\|\hat{\mathbf{S}}(\lambda)^{-1}\|\frac{1}{1 - |\lambda|\|\mathbf{S}^{-1}\|\|\hat{\mathbf{L}}\|},
\end{aligned}$$

as long as $|\lambda|\,\|\mathbf{S}^{-1}\|\,\|\hat{\mathbf{L}}\| < 1$, where we used the estimate (for self-adjoint operators)

$$\|\mathbf{S}^{-1}(\lambda)\| \leq \sup_{t \in [0, \|\hat{\mathbf{A}}\hat{\mathbf{A}}^*\|]} \frac{1}{|\lambda - t|}\,.$$

It is worth noticing that the sup in the following expression is attained for $\lambda \to 2$

$$\sup_{\lambda, |\lambda - 1| > 1}\ \sup_{t \in [0, \|\hat{\mathbf{A}}\hat{\mathbf{A}}^*\|]} \frac{\lambda}{|\lambda - t|} = \frac{2}{2 - \|\hat{\mathbf{A}}\hat{\mathbf{A}}^*\|}\,.$$

From the hypothesis we conclude that $\frac{2}{2 - \|\hat{\mathbf{A}}\hat{\mathbf{A}}^*\|}\|\hat{\mathbf{L}}\| < 1$. Thus, inequality (48) holds true. Equation (25) follows now from Theorem 4. □

In the sequel we present the main result of this section, where convergence rates for the LK method are derived.

Theorem 5. *Let the lower triangular matrix* $|\mathbf{L}|$ *in* (42) *satisfy*

$$\sigma_{max}(|\mathbf{L}|) + \frac{1}{2}\sum_{i=0}^{p} \|A_i^* M_i^{\frac{1}{2}}\|^2 < 1 \tag{50}$$

(alternatively, let $\hat{\mathbf{L}}$ *satisfy* (49)*). Moreover, assume that the source condition*

$$x_0 - x^\dagger = (\mathbf{A}^*\mathbf{M}_B\mathbf{A})^\nu w\,, \qquad \text{for some } 0 < \nu < \infty$$

is satisfied. Then, the iterates of the Landweber–Kaczmarz method (7) *satisfy the error estimate*

$$\|x_k - x^\dagger\| \leq C_1 \frac{1}{k^\nu} + C_2\sqrt{k}\delta\,,$$

with some positive constants C_1*,* C_2*. In particular the* a-priori *choice rule* $k \sim \delta^{\frac{-2}{2\mu+1}}$ *yields the convergence rate*

$$\|x_k - x^\dagger\| \sim \delta^{\frac{2\mu}{2\mu+1}}\,.$$

If for some $\tau > 1$ *sufficiently large, the stopping index* k *is chosen according to the discrepancy principle (i.e., as the first index such that* $\|\mathbf{A}^*\mathbf{M}_B(\mathbf{A}x_k - \mathbf{y})\| \leq \tau\delta$*, this yields a parameter choice rule with the same rate.*

Proof. We already know that (50) implies (49), as well as (46) and $\max_i \|A_i^* M_i^{\frac{1}{2}}\| < \sqrt{2}$. The assertion is then a collection of the previous results. The convergence rates for the discrepancy principle is a consequence of the results of Plato and Hämarik [35]. □

It is worth noticing that, by choosing M_i sufficiently small, we can always achieve that the hypothesis in this theorem (except for the source condition) is satisfied.

3.1 Analysis of the nonsymmetric iterated Tikhonov–Kaczmarz method

In what follows we derive convergence rates for the iTK method (12) (in the block form (20)) in the nonsymmetric case. First of all, from the equivalence between (12) and (20), it follows that

$$(I + \mathbf{A}^*\mathbf{N}_B\mathbf{A}) = \Pi_{i=0}^{p-1}(I + A_i^* M_i A_i)\,.$$

In particular $(I + \mathbf{A}^*\mathbf{N}_B\mathbf{A})^{-1}$ is well defined and, as long as all M_i are symmetric positive semidefinite, we have

$$\|(I + \mathbf{A}^*\mathbf{N}_B\mathbf{A})^{-1}\| \leq 1\,.$$

We first investigate the propagated data error,

Lemma 3.9. *Let* x_k^δ *denote the iteration* (20) *with noisy data, and* x_k *the iteration* (20) *with exact data. Then we have the estimate with a constant* C

$$\|x_k^\delta - x_k\| \leq Ck\delta. \tag{51}$$

If, additionally,

$$\sup_{k\in\mathbb{N}} \|(I + \mathbf{N}_B\mathbf{A}\mathbf{A}^*)^{-k}\| \leq C_1 \tag{52}$$

holds with a constant C_1*, then*

$$\|x_k^\delta - x_k\| \leq C\sqrt{k}\delta\,, \tag{53}$$

where the constant $C = C(\mathbf{N}_B, \mathbf{A})$ *does not depend on* k*.*

Proof. We may express iteration (20) as

$$x_k^\delta - x_k = \sum_{j=0}^{k-1} (I + \mathbf{A}^*\mathbf{N}\mathbf{A})^{-j}\mathbf{A}^*\mathbf{N}(\mathbf{y} - \mathbf{y}^\delta)\,,$$

from which (51) follows immediately. Now, defining $g_k(\lambda) := \sum_{j=0}^{k-1}(1+\lambda)^{-j}$ we obtain the estimate (compare with the Landweber iteration)

$$\begin{aligned}
\|x_k^\delta - x_k\|^2 &= \Big(g_k(\mathbf{A}^*\mathbf{N}_B\mathbf{A})\mathbf{A}^*\mathbf{N}_B(\mathbf{y}-\mathbf{y}^\delta),\, g_k(\mathbf{A}^*\mathbf{N}_B\mathbf{A})\mathbf{A}^*\mathbf{N}_B(\mathbf{y}-\mathbf{y}^\delta)\Big) \\
&\quad \Big(\mathbf{N}_B^{-1}\mathbf{N}_B\mathbf{A}\mathbf{A}^* g_k(\mathbf{N}_B\mathbf{A}\mathbf{A}^*)\mathbf{N}_B(\mathbf{y}-\mathbf{y}^\delta),\, g_k(\mathbf{N}_B\mathbf{A}\mathbf{A}^*)\mathbf{N}_B(\mathbf{y}-\mathbf{y}^\delta)\Big) \\
&\le \|\mathbf{N}_B^{-1}\|\; \|I + \mathbf{N}_B\mathbf{A}\mathbf{A}^* - (I+\mathbf{N}_B\mathbf{A}\mathbf{A}^*)^{-k+1}\| \sum_{j=0}^{k-1} \|(I+\mathbf{N}_B\mathbf{A}\mathbf{A}^*)^{-k}\|\delta^2
\end{aligned}$$

Using (52) in the last inequality we obtain (53). □

Next we investigate the approximation error. Notice that

$$x_k - x^\dagger = (I + \mathbf{A}^*\mathbf{N}_B\mathbf{A})^{-k}(x_0 - x^\dagger)\,,$$

Hence, if a source condition with the operator $(\mathbf{A}^*\mathbf{N}_B\mathbf{A})$ holds, we have to estimate the operator $(I + \mathbf{A}^*\mathbf{N}_B\mathbf{A})^{-k}(\mathbf{A}^*\mathbf{N}_B\mathbf{A})^\alpha$.

Lemma 3.10. *If there exists a $h > 0$ such that*

$$\begin{aligned}
\Big|\big([\mathbf{A}^*\mathbf{N}_B\mathbf{A} - \mathbf{A}^*\mathbf{N}_B^T\mathbf{A}]x, y\big)\Big| \;&\le h\big[(\mathbf{A}^*\mathbf{N}_B\mathbf{A}x, x) + (\mathbf{A}^*\mathbf{N}_B\mathbf{A}y, y)\big]\,, \\
&\forall\; x, y \in X,\; \|x\|^2 + \|y\|^2 = 1\,,
\end{aligned} \tag{54}$$

then $(\mathbf{A}^*\mathbf{N}_B\mathbf{A})^\alpha$ is well defined for all $\alpha \ge 0$ and there exists a constant C depending on α, h such that for all k

$$\|(I + \mathbf{A}^*\mathbf{N}_B\mathbf{A})^{-k}(\mathbf{A}^*\mathbf{N}_B\mathbf{A})^\alpha\| \le C\frac{1}{k^\alpha}\,.$$

Proof. Due to (54), the numerical range of $\mathbf{A}^*\mathbf{N}_B\mathbf{A}$ is contained in a sector analog to the one in the proof of Lemma 3.5. Using [6], we obtain once again

$$\|(I + \mathbf{A}^*\mathbf{N}_B\mathbf{A})^{-k}(\mathbf{A}^*\mathbf{N}_B\mathbf{A})^\alpha\| \le C_c \sup_{|\arg(\lambda)|<\psi=\arctan(h)<\frac{\pi}{2}} \frac{|\lambda|^\alpha}{|1+\lambda|^k}\,.$$

Furthermore, $|1+\lambda| \ge (1 + |\lambda|\cos(\psi))$. Thus, there exists a constant C' such that

$$\frac{|\lambda|^\alpha}{|1+\lambda|^k} \;\le\; \frac{1}{\cos(\psi)^\alpha}\frac{|\lambda\cos(\psi)|^\alpha}{|1+\lambda\cos(\psi)|^k} \;\le\; C'\frac{1}{\cos(\psi)^\alpha}\frac{1}{k^\alpha}\,, \quad \forall k \ge 1$$

(the last inequality follows from the convergence rate analysis of the standard iterated Tikhonov regularization), concluding the proof. □

The next lemma discuss a sufficient condition for the conditions (52) and (54) in Lemma 3.9 and Lemma 3.10 respectively.

Lemma 3.11. *If*

$$\|\hat{\mathbf{L}}\| < 1 \tag{55}$$

then (52) *and* (54) *hold true.*

Proof. The proof of (54) follows the lines of the proof of Lemma 3.7. To prove (52), we use Theorem 4 with $T = (I + \mathbf{N}_B \mathbf{A}\mathbf{A}^*)^{-1}$. Thus, for $\lambda \in \mathbb{C}$ with $|\lambda| > 1$, we estimate

$$\begin{aligned}
\|(T - \lambda I)^{-1}\| \leq & \ \|(I + \mathbf{N}_B \mathbf{A}\mathbf{A}^*)\| \, \|((1-\lambda)I - \lambda \mathbf{N}_B \mathbf{A}\mathbf{A}^*)^{-1}\| \\
\leq & \ \|(I + \mathbf{N}_B \mathbf{A}\mathbf{A}^*)\| \, \|\mathbf{N}_B\| \, |\lambda|^{-1} \|((1 - \frac{1}{\lambda})\mathbf{N}_B^{-1} + \mathbf{A}\mathbf{A}^*)^{-1}\| \\
\leq & \ \|(I + \mathbf{N}_B \mathbf{A}\mathbf{A}^*)\| \, \|\mathbf{N}_B\| \, \|\mathbf{D}^{-\frac{1}{2}}\|^2 |\lambda|^{-1} \|(\hat{\mathbf{A}}\hat{\mathbf{A}}^* + (1 - \frac{1}{\lambda})(I - \hat{\mathbf{L}}^*)\| \,.
\end{aligned}$$

Thus, setting $s := (1 - \frac{1}{\lambda})$, it is enough to prove that

$$\|(\hat{\mathbf{A}}\hat{\mathbf{A}}^* + sI - s\hat{\mathbf{L}}^*)^{-1}\| \ \leq \ C\frac{1}{|s|}\,, \ \forall \ |s-1| < 1\,.$$

Notice that

$$\begin{aligned}
\|(\hat{\mathbf{A}}\hat{\mathbf{A}}^* + sI - s\hat{\mathbf{L}}^*)^{-1}\| \leq & \ \|(\hat{\mathbf{A}}\hat{\mathbf{A}}^* + sI)^{-1}\| \, \|(I - (\hat{\mathbf{A}}\hat{\mathbf{A}}^* + sI)^{-1} s\hat{\mathbf{L}}^*)^{-1}\| \\
\leq & \ \frac{1}{|s|} \frac{1}{1 - \|(\hat{\mathbf{A}}\hat{\mathbf{A}}^* + sI)^{-1} s\hat{\mathbf{L}}^*\|}\,,
\end{aligned}$$

provided that $\|(\hat{\mathbf{A}}\hat{\mathbf{A}}^* + sI)^{-1} s\hat{\mathbf{L}}^*\| < 1$. However, due to the straightforward inequality

$$\|(\hat{\mathbf{A}}\hat{\mathbf{A}}^* + sI)^{-1} s\| \leq 1\,, \ \forall \ |s-1| < 1\,,$$

it follows that

$$\frac{1}{1 - \|(\hat{\mathbf{A}}\hat{\mathbf{A}}^* + sI)^{-1} s\hat{\mathbf{L}}^*\|} \leq \frac{1}{1 - \|\hat{\mathbf{L}}^*\|} = \frac{1}{1 - \|\hat{\mathbf{L}}\|}\,,$$

establishing the desired bound. Consequently, inequality (52) follows from Theorem 4. □

Our next step is to derive rates of convergence for the iterated Tikhonov–Kaczmarz method.

Theorem 6. *Let the lower triangular matrix* $|\mathbf{L}|$ *in* (42) *be such that*

$$\sigma_{max}(|\mathbf{L}|) < 1 \tag{56}$$

(alternatively, let $\hat{\mathbf{L}}$ *satisfy* (55)*). Moreover, assume the source condition*

$$x_0 - x^\dagger = (\mathbf{A}^* \mathbf{N}_B \mathbf{A})^\nu w\,, \quad \textit{for some} \ \ 0 < \nu < \infty\,.$$

Then, the sequence generated by the iterated Tikhonov–Kaczmarz method (10) *satisfies the estimate*

$$\|x_k - x^\dagger\| \ \leq \ C_1 \frac{1}{k^\nu} + C_2 \sqrt{k}\delta\,,$$

with some constants C_1, C_2. *In particular, the* a-priori *choice rule*

$$k \sim \delta^{\frac{-2}{2\mu+1}}$$

yields the convergence rate

$$\|x_k - x^\dagger\| \sim \delta^{\frac{2\mu}{2\mu+1}}\,.$$

If, for some (sufficiently large) $\tau > 1$, k *is chosen according to the discrepancy principle (i.e., the first index such that* $\|\mathbf{A}^* \mathbf{M}_B (\mathbf{A}x_k - \mathbf{y})\| \leq \tau\delta$*), this yields an* a-posteriori *parameter choice rule with the same rate.*

Proof. Since (56) implies (55), the results follow from Lemma 3.11, Lemma 3.10 and Lemma 3.9 respectively. The assertion concerning the discrepancy principle follows once again from [35]. □

Remark 3.12. *In Theorems 5 and 6 convergence rates are established under the source conditions* $\mathbf{x}_0 - x^\dagger \in R(\mathbf{A}^*\mathbf{M}_B\mathbf{A})^\nu$ *and* $\mathbf{x}_0 - x^\dagger \in R(\mathbf{A}^*\mathbf{N}_B\mathbf{A})^\nu$ *respectively. It would be interesting to replace these by the usual source conditions with ranges* $R(\mathbf{A}^*\mathbf{A})^\nu$*. It is not clear to us if this can be done under the same assumptions as in the above mentioned theorems. An equivalence between the source conditions can be shown if a norm equivalence*

$$d_1\|\mathbf{A}^*\mathbf{A}x\| \leq \|\mathbf{A}^*\mathbf{M}_B\mathbf{A}x\| \leq d_2\|\mathbf{A}^*\mathbf{A}x\|\,,$$

holds with some uniform constants d_1*,* d_2 *(analogously for* $\mathbf{A}^*\mathbf{N}_B\mathbf{A}$*). For this purpose, a generalization of the Kato-Heinz inequality to accreative operators [22] might be used.*

4 Conclusion

We have established convergence rates for the Landweber–Kaczmarz method and the iterated Tikhonov–Kaczmarz method, for the symmetric and nonsymmetric versions of each method. Since the only conditions for the convergence theorems are bounds on $A_i^*M_i$, it follows that for sufficiently small stepsizes (or appropriately scaled operators), standard convergence rates can always be established. In particular, we aimed to use bounds in our theorems (see, e.g., (50) or (56)), which are computable and can be used in numerical implementations.

As one would expect, if more information on the operators A_i is available, the weaker conditions (54), (52) (see also (48), (40)) can be proven directly.

Although, asymptotically, the Kaczmarz variants perform similar to their block iterations, they have some advantages, e.g., a simpler implementation and possibly a larger stepsize. However, even if they have similar convergence rates (as $\delta \to 0$), these two types of iterations can be quite different in practice. Depending on the distribution of the eigenvalues, on the structure of the exact solution, and on the noise level, it may happens that the Kaczmarz iterations perform better (specially at the first iterates).

If information on the distribution of the eigenvalue structure on the exact solution is available, the results in Section 3 can be used in order to estimate the decay rates of the error in appropriate subspaces. In particular, convergence of the nonsymmetric methods depends on how the eigenvalues are located in a sector of the positive complex half plane. The components that are close to zero and/or away from the real axis will contribute to a slow convergence.

The symmetric iterations have the advantage that they can be used even in cases where the conditions for the nonsymmetric iterations are not satisfied. However, as a drawback, one must pay the price of doubling the numerical computations.

We conclude by mentioning that the convergence result analysis in Section 3 can be extended to general nonsymmetric preconditioned Landweber (and iterated Tikhonov) iterations, which is highly relevant in practical large scale applications.

Appendix: Analysis of the symmetric case

For the convenience of the reader we present in this appendix the convergence analysis for the symmetric Landweber–Kaczmarz (sLK) method and the symmetric iterated Tikhonov–Kaczmarz (siTK) method. As before, we consider the case of linear operator equations (1), under the assumption (3) on the data.

A.1 Convergence rates for the sLK method

By Theorem 2 the symmetric Landweber–Kaczmarz method is, roughly speaking, a block Landweber iteration with preconditioning $\mathbf{M}_{SB}$. This allows us to analyze the convergence and convergence rates using classical results for the (preconditioned) Landweber iteration (compare [9]).

If $\max_i \|A_i^* M_i^{\frac{1}{2}}\| \leq \sqrt{2}$ we define the operator

$$\mathbf{B} := (2\mathbf{D} - \mathrm{diag}(A_i A_i^*))^{\frac{1}{2}} \mathbf{M}_B \,.$$

Therefore, we have

$$\mathbf{M}_{SB} = \mathbf{B}^* \mathbf{B}, \tag{57}$$

and consequently Theorem 2 can be rephrased as follows:

Corollary 1. *If* $\max_i \|A_i^* M_i^{\frac{1}{2}}\| \leq \sqrt{2}$ *holds true, then the iterates of the sLK method are equivalent to the Landweber iteration applied to the system*

$$\mathbf{B}\mathbf{A} = \mathbf{B}\mathbf{y}, \tag{58}$$

i.e., $x_{k+1} = x_k - (\mathbf{B}\mathbf{A})^*(\mathbf{B}\mathbf{A}x_k - \mathbf{B}\mathbf{y}^\delta)$.

We denote by $x^\dagger$ a minimum norm solution of (58). Furthermore we impose a source condition of the form

$$x^\dagger - x_0 = \left[(\mathbf{B}\mathbf{A})^*(\mathbf{B}\mathbf{A})\right]^\mu \tilde{w} \quad \tilde{w} \in X. \tag{59}$$

The corresponding noise level is given by $\delta_S = \|\mathbf{B}(\mathbf{y}^\delta - \mathbf{y})\|$. From (61) below, we see that this modified noise level is always of the same magnitude as the one in (4).

Theorem 7. *Let* $\|\mathbf{B}\mathbf{A}\| < 2$*. If the data are exact, then the sLK iteration* x_k *converges to an* x_0*-minimum norm solution of* (58) *as* $k \to \infty$*.*

For noisy data, let the iteration be stopped either by the a-priori rule $k \sim \delta_S^{\frac{2\mu}{2\mu+1}}$*, or by the (a-posteriori) discrepancy principle, i.e., at the first index* $k = k(\delta_S, \mathbf{y}^\delta)$ *satisfying* $\|\mathbf{B}(\mathbf{A}x_k - \mathbf{y}^\delta)\| \leq \tau\delta_S$*, for some fixed* $\tau > 1$*. Moreover, assume that a source condition is satisfied for some* $\mu > 0$*. Then we have (optimal order) convergence rates* $\|x_k - x^\dagger\| \leq C\delta_S^{\frac{2\mu}{2\mu+1}}$*. In both cases the stopping index satisfies* $k = O(\delta_S^{-\frac{2}{2\mu+1}})$*.*

Proof. See [9, 11]. □

The convergence results in the above theorem can be stated in a more common form, i.e., using only bounds on $\mathbf{A}$, the noise level in (4), and a standard source condition.

At first we investigate the condition $\|\mathbf{B}\mathbf{A}\| < 2$. We have the following lemma.

Lemma A.1. *It* $\max_i \|A_i^* M_i^{\frac{1}{2}}\| < \sqrt{2}$ *hold true, then*

$$\|\mathbf{B}\mathbf{A}\| \leq 1 \,.$$

Proof. From Theorem 1 it follows that (compare with [10])

$$I - \mathbf{A}^* \mathbf{M}_{SB} \mathbf{A} = Q_B^* Q_B,$$

where

$$Q_B = (I - A_{p-1}^* M_{p-1} A_{p-1})(I - A_{p-2}^* M_{p-2} A_{p-2}) \ldots (I - A_0^* M_0 A_0) \,.$$

Notice that $\|Q_B\| \leq 1$ due to our assumptions. Consequently, $\|I - \mathbf{A}^* \mathbf{M}_{SB} \mathbf{A}\| \leq 1$. Since the operator $\mathbf{A}^* \mathbf{M}_{SB} \mathbf{A} = (\mathbf{B}\mathbf{A})^*(\mathbf{B}\mathbf{A})$ is symmetric and positive semidefinite, it's spectrum satisfies $\sigma(\mathbf{A}^* \mathbf{M}_{SB} \mathbf{A}) \subset (0, 1)$, and lemma follows. □

Next we relate the source condition (59) and the corresponding noise level to the more standard source condition

$$x^\dagger - x_0 = (\mathbf{A}^*\mathbf{A})^\mu w\,, \quad w \in X\,, \tag{60}$$

and the corresponding noise level δ. This can be done if M_B is an isomorphism.

Lemma A.2. *If* $\max_i \|A_i^* M_i^{\frac{1}{2}}\| < \sqrt{2}$*, then there exist positive constants* m_1*,* m_2 *such that*

$$m_1 \|\mathbf{y}\|_{\mathbf{Y}} \le \|\mathbf{B}\mathbf{y}\| \le m_2 \|\mathbf{y}\|_{\mathbf{Y}}\,, \quad \forall \mathbf{y} \in \mathbf{Y}\,. \tag{61}$$

These constants can be bounded by

$$m_2 \le \frac{\sqrt{2}\max_i \|M_i^{\frac{1}{2}}\|}{\sigma_{min}(I - |\mathbf{L}|)}\,, \qquad m_1 \ge \frac{\sqrt{(2 - \max_i \|A M_i^{\frac{1}{2}}\|^2)}}{\max_i \|M_i^{-\frac{1}{2}}\| \sigma_{max}(I + |\mathbf{L}|)}\,,$$

where σ_{max}*,* σ_{min} *denote the largest and smallest singular values of the matrix* $|\mathbf{L}| \in \mathbb{R}^{p \times p}$ *in* (42) *(see also* (43)*).*

Proof. Under the given assumptions it is clear that $\mathbf{B}$ is invertible. Moreover, from (57) it follows that

$$\mathbf{M}_{SB} = \mathbf{D}^{-\frac{1}{2}}(I + \hat{\mathbf{L}})^* \left(2I - \mathrm{diag}(M_i^{\frac{1}{2}} A_i A_i^* M_i^{\frac{1}{2}})\right)(I + \hat{\mathbf{L}})\mathbf{D}^{-\frac{1}{2}},$$

where $\hat{\mathbf{L}}$ is the operator defined in (41). Setting $z := (I + \hat{\mathbf{L}})^{-1}\mathbf{D}^{-\frac{1}{2}}\mathbf{y}$, equation (61) becomes equivalent to

$$m_1 \|\mathbf{D}^{\frac{1}{2}}(I + \hat{\mathbf{L}})z\|_{\mathbf{Y}} \;\le\; \| \left(2I - \mathrm{diag}(M_i^{\frac{1}{2}} A_i A_i^* M_i^{\frac{1}{2}})\right)^{\frac{1}{2}} z\| \;\le\; m_2 \|\mathbf{D}^{\frac{1}{2}}(I + \hat{\mathbf{L}})z\|\,.$$

However, from the assumption we have

$$\| \left(2I - \mathrm{diag}(M_i^{\frac{1}{2}} A_i A_i^* M_i^{\frac{1}{2}})\right)^{\frac{1}{2}} z\| \;\le\; \sqrt{2}\|z\|\,,$$

and also

$$\left((2I - \mathrm{diag}(M_i^{\frac{1}{2}} A_i A_i^* M_i^{\frac{1}{2}}))z, z\right) \ge \left(2 - \max_i(\|A_i^* M_i^{-\frac{1}{2}}\|^2\right)\|z\|^2\,.$$

On the other hand, we obtain for the operator $(I + \hat{\mathbf{L}})$ the estimate

$$\begin{aligned}\|(I + \hat{\mathbf{L}})z\|^2 &= \sum_{i=0}^{p-1} \| \sum_{j=0}^{p-1} (I + \hat{\mathbf{L}})_{i,j} z_j\|^2 \le \sum_{i=0}^{p-1} (\sum_{j=0}^{p-1} \|(I + \hat{\mathbf{L}})_{i,j}\| \|z_j\|)^2 \\ &\le \|I + |\mathbf{L}|\|_2^2 \|z\|^2 = \sigma_{max}(I + |\mathbf{L}|)^2 \|z\|^2\,.\end{aligned}$$

Let $x = (I + \hat{\mathbf{L}})^{-1}z$, i.e., $(I + \hat{\mathbf{L}})x = z$. Due to the triangular structure of $\hat{\mathbf{L}}$ it follows that

$$\|x_i\| \le \|z_i\| + \sum_{j=0}^{i-1} \|\hat{\mathbf{L}}_{i,j}\| \|x_j\|\,, \quad \forall i \in 0, \dots p-1\,.$$

Next, denoting by $|x| = (\|x_i\|)_{i=0}^{p-1}$, we obtain

$$(I - |\mathbf{L}|)|x| \le |z|\,,$$

where $\leq$ means componentwise inequality. Notice now that, if w is a componentwise positive vector then $(I - |\mathbf{L}|)^{-1}w$ is also componentwise positive. Indeed, this is shown by induction, since $v = (I - |\mathbf{L}|)^{-1}w$ satisfies the recursion formula

$$v_1 = w_1, v_i = w_i + \sum_{j=0}^{i-1} |\mathbf{L}|_{i,j} v_j \, .$$

Applying this results to $w = |z| - (I - |\mathbf{L}|)|x|$, we get

$$|x| \leq (I - |\mathbf{L}|)^{-1}|z| \, .$$

Consequently,

$$\|x\| \leq \|(I - |\mathbf{L}|)^{-1}|z|\|_{\mathbb{R}^p} \leq \|(I - |\mathbf{L}|)^{-1}\|_2 \|z\|_2 = \frac{1}{\sigma_{min}(I - |\mathbf{L}|)} \|z\|$$

or $\sigma_{min}(I - |\mathbf{L}|)\|x\| \leq \|(I + \hat{\mathbf{L}})x\|$. □

An immediate consequence of the above lemma is the equivalence of the source conditions in (59) and (60).

Corollary 2. *Let* $\max_i \|A_i^* M_i^{\frac{1}{2}}\| < \sqrt{2}$*, then for* $\nu \leq \frac{1}{2}$ *we have*

$$m_1^\nu \|(\mathbf{A}^*\mathbf{A})^\nu x\| \leq \| \left((\mathbf{BA})^*(\mathbf{BA}) \right)^\nu x\| \leq m_2^\nu \|(\mathbf{A}^*\mathbf{A})^\nu x\|, \tag{62}$$

with m_1, m_2 *being the constants in Lemma A.2. In particular, if (60) holds, then (59) also holds with* $\|\tilde{w}\| \leq \frac{1}{m_1^\nu}\|w\|$.

Proof. The first inequality is a consequence of Heinz' inequality (see [11]). The last inequality follows now from (62) when we set $\tilde{w} = ((\mathbf{BA})^*(\mathbf{BA}))^{-\nu} (\mathbf{A}^*\mathbf{A})^\nu w$. □

We are finally in position to state the main convergence rate result for the symmetric Landweber–Kaczmarz (sLK) method. The next theorem is actually a collection of the previous results, together with the standard estimates for the Landweber iteration. The last assertion on the discrepancy principle requires just a slight modification of the proofs in [11] (notice that we have $\delta_S \leq m_2\delta$).

Theorem 8. *Let* $\max_i \|A_i^* M_i^{\frac{1}{2}}\| < \sqrt{2}$*, and let* $x^\dagger - x_0$ *satisfy (60) with* $\nu \leq \frac{1}{2}$*. Moreover, let* $x^\dagger$ *be a least squares solution of (58). Then, for the iterations of the symmetric Landweber–Kaczmarz, the following estimate holds true:*

$$\|\mathbf{x}_k - x^\dagger\| \leq \sqrt{k} m_2 \delta + \frac{1}{(k+1)^\nu} \frac{1}{m_1^\nu} \|w\| \, .$$

In particular, the a-priori parameter choices $k \sim \delta^{-\frac{2\nu}{(2\nu+1)}}$ *or* $k \sim \delta_S^{-\frac{2\nu}{(2\nu+1)}}$ *yield the order optimal rate. The same estimate holds if* k *is chosen as the according to the discrepancy principle in Theorem 7.*
If k *is chosen as the first index satisfying* $\|\mathbf{A}x_k - \mathbf{y}^\delta\| \leq \tau m_2 \delta$ *with* $\tau > 1$*, the symmetric Landweber–Kaczmarz iteration converges with the order optimal rate.*

Comparing the sLK iteration with the usual block Landweber iteration we notice one difference: If we set the simple preconditioners $M_i = \tau I$, then for the symmetric Landweber–Kaczmarz iteration $\tau = \tau_{sLK}$ can be chosen as $\tau_{sLK} < 2\big(\max_i \|A_i\|^2\big)^{-1}$. This should be contrasted with the corresponding choice for the block-Landweber iteration, where τ has to be chosen such that $\tau\|\mathbf{A}\|^2 \le 2$, i.e.,

$$\tau \;<\; 2\Big(\sum_{i=0}^{p} \|A_i\|^2\Big)^{-1} \;<\; 2\big(\max_i \|A_i\|^2\big)^{-1}.$$

Thus, besides the fact that the sLK iteration is easier to implement, we may also choose a large stepsize. This has the effect that the sLK iteration will damp errors corresponding to the larger singular values of $\mathbf{A}$ in a more efficient way. Asymptotically this difference is not relevant, since both iterations yield to the same order of convergence.

Note that the stepsize in the sLK iteration can be selected rather independently of the number of blocks p. However, the constants m_1 and m_2 will in general depend on p via the singular values of $I \pm |\mathbf{L}|$. The following (conservative) upper bound can be derived from (43)

$$\|L_{i,j}\| \le \max_{i,j} \|A_i\|\|A_j\| =: \eta =: |\mathbf{L}|_{i,j}\,,\;\; \forall\, 1 \le j < i\,. \tag{63}$$

In this case, we can select $|\mathbf{L}|$ as the lower triangular matrix with constant entries η. A direct estimate shows that $\sigma_{max}(I + |\mathbf{L}|)$ grows linearly with p, while $\frac{1}{\sigma_{min}(I-|\mathbf{L}|)}$ grows exponentially.

A.2 Convergence rates for the siTK method

The convergence theory for the symmetric iterated Tikhonov–Kaczmarz method can be established using spectral theory for the operator $\mathbf{A}^*\mathbf{N}_{SB}\mathbf{A}$ and the well-known filter function

$$g_k(\lambda) = \sum_{i=1}^{k} \frac{1}{(1+\lambda)^i}\,.$$

In contrast to Landweber-type iterations, the spectrum of the forward operator does not have to be in the interval $[0, 2)$. Moreover, similarly as in Lemma A.2, it can be shown the existence of constants n_1, n_2 such that

$$n_1\|z\| \le \|(\mathbf{D} + \mathrm{diag}(A_iA_i^*))^{\frac{1}{2}}\mathbf{N}_B^* z\| \le n_2\|z\|\,,\;\; \forall z \in \mathbf{Y}\,.$$

Consequently, the source conditions $x^\dagger - x_0 \in R(\mathbf{A}^*\mathbf{A})^\nu)$ and $x^\dagger - x_0 \in R(\mathbf{A}^*\mathbf{N}_{SB}\mathbf{A})^\nu)$ are equivalent.

We collect the main results in the next theorem, which follow from standard results on the iterated Tikhonov regularization [16, 11].

Theorem 9. *Let* $\max_i \|A_i^* M_i^{\frac{1}{2}}\| < \sqrt{2}$*, and let* $x^\dagger - x_0$ *satisfy the source condition* (60) *with* $\nu \le \frac{1}{2}$*. Moreover, let* $x^\dagger$ *be a least squares solution of* (58)*. The iterations of the symmetric iterated Tikhonov method satisfy the estimate*

$$\|x_k - x^\dagger\| \;\le\; C_1\sqrt{k}\delta + \frac{1}{(k+1)^\nu} C_2\|w\|\,,$$

where C_1, C_2 *depends only on* p *and* n_1, n_2*. In particular, the a-priori parameter choice rules* $k \sim \delta^{-\frac{2\nu}{(2\nu+1)}}$ *or* $k \sim \delta_S^{-\frac{2\nu}{(2\nu+1)}}$ *yields order optimal rate. The same holds if* k *is chosen according to the discrepancy principle in Theorem 7. If* k *is chosen as the first index satisfying* $\|\mathbf{A}x_k - \mathbf{y}^\delta\| \le \tau n_2 \delta$*, the symmetric iterated Tikhonov method converges with order optimal rate.*

Acknowledgments

The authors would like to thank two anonymous referees for the valuable comments on the original version of the manuscript. Particularly what concerns the proof of Lemma 3.3.
The work of A.L. is partially supported by the Brazilian National Research Council CNPq, grant 303098/2009–0.

References

[1] P. K. Anh and C. V. Chung, *Parallel regularized Newton method for nonlinear ill-posed equations*, Numer. Algorithms, 58 (2011), pp. 379–398.

[2] A. B. Bakushinsky, M. Y. Kokurin, and A. Smirnova, *Iterative methods for ill-posed problems*, Walter de Gruyter GmbH & Co. KG, Berlin, 2011.

[3] J. Baumeister, B. Kaltenbacher, and A. Leitão, *On Levenberg-Marquardt-Kaczmarz iterative methods for solving systems of nonlinear ill-posed equations.*, Inverse Probl. Imaging, 4 (2010), pp. 335–350.

[4] H. Bui and Q. Nguyen, *Thermomechanical Couplings in Solids*, North-Holland Publishing Co., Amsterdan, 1987.

[5] M. Burger and B. Kaltenbacher, *Regularizing Newton-Kaczmarz methods for nonlinear ill-posed problems.*, SIAM J. Numer. Anal., 44 (2006), pp. 153–182.

[6] M. Crouzeix, *Numerical range and functional calculus in Hilbert space.*, J. Funct. Anal., 244 (2007), pp. 668–690.

[7] A. De Cezaro, J. Baumeister, and A. Leitão, *Modified iterated Tikhonov methods for solving systems of nonlinear ill-posed equations*, Inverse Probl. Imaging, 5 (2011), pp. 1–17.

[8] A. De Cezaro, M. Haltmeier, A. Leitão, and O. Scherzer, *On steepest-descent-Kaczmarz methods for regularizing systems of nonlinear ill-posed equations.*, Appl. Math. Comput., 202 (2008), pp. 596–607.

[9] H. Egger, *$\mathcal{Y}$-scale regularization.*, SIAM J. Numer. Anal., 46 (2008), pp. 419–436.

[10] T. Elfving and T. Nikazad, *Properties of a class of block-iterative methods.*, Inverse Problems, 25 (2009), pp. Article ID 115011, 13 p.

[11] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems.*, Kluwer, Dordrecht, 1996.

[12] M. Haltmeier, *Convergence analysis of a block iterative version of the loping Landweber-Kaczmarz iteration*, Nonlinear Anal., 71 (2009), pp. e2912–e2919.

[13] M. Haltmeier, R. Kowar, A. Leitão, and O. Scherzer, *Kaczmarz methods for regularizing nonlinear ill-posed equations. II. Applications*, Inverse Probl. Imaging, 1 (2007), pp. 507–523.

[14] M. Haltmeier, A. Leitão, and O. Scherzer, *Kaczmarz methods for regularizing nonlinear ill-posed equations. I: Convergence analysis.*, Inverse Probl. Imaging, 1 (2007), pp. 289–298.

[15] M. Haltmeier, A. Leitão, and E. Resmerita, *On regularization methods of EM-Kaczmarz type*, Inverse Problems, 25 (2009), pp. 075008, 17.

[16] M. Hanke and C. Groetsch, *Nonstationary iterated Tikhonov regularization.*, J. Optimization Theory Appl., 98 (1998), pp. 37–53.

[17] M. Hanke, A. Neubauer, and O. Scherzer, *A convergence analysis of the Landweber iteration for nonlinear ill-posed problems*, Numer. Math., 72 (1995), pp. 21–37.

[18] M. Hanke and W. Niethammer, *On the acceleration of Kaczmarz's method for inconsistent linear systems.*, Linear Algebra Appl., 130 (1990), pp. 83–98.

[19] S. Kaczmarz, *Angenäherte Auflösung von Systemen linearer Gleichungen.*, Bull. Int. Acad. Polon. Sci. A, 1937 (1937), pp. 355–357.

[20] B. Kaltenbacher, A. Neubauer, and O. Scherzer, *Iterative regularization methods for nonlinear ill-posed problems*, Walter de Gruyter GmbH & Co. KG, Berlin, 2008.

[21] N. Kalton, S. Montgomery-Smith, K. Oleszkiewicz, and Y. Tomilov, *Power-bounded operators and related norm estimates.*, J. Lond. Math. Soc., II. Ser., 70 (2004), pp. 463–478.

[22] T. Kato, *A generalization of the Heinz inequality.*, Proc. Japan Acad., 37 (1961), pp. 305–308.

[23] ——, *Perturbation theory for linear operators.*, Springer, Berlin-Heidelberg-New York, 1980.

[24] R. Kowar and O. Scherzer, *Convergence analysis of a Landweber-Kaczmarz method for solving nonlinear ill-posed problems.* Romanov, V. G. (ed.) et al., Ill-posed and inverse problems. Dedicated to Academician Mikhail Mikhailovich Lavrentiev on the occasion of his 70th birthday. Utrecht (2002)., 2002.

[25] Y. Lyubich, *Spectral localization, power boundedness and invariant subspaces under Ritt's type condition.*, Stud. Math., 134 (1999), pp. 153–167.

[26] S. McCormick, *An iterative procedure for the solution of constrained nonlinear equations with application to optimization problems*, Numer. Math., 23 (1975), pp. 371–385.

[27] ——, *The methods of Kaczmarz and row orthogonalization for solving linear equations and least squares problems in Hilbert space*, Indiana Univ. Math. J., 26 (1977), pp. 1137–1150.

[28] K.-H. Meyn, *Solution of underdetermined nonlinear equations by stationary iteration methods*, Numer. Math., 42 (1983), pp. 161–172.

[29] B. Nagy and J. Zemánek, *A resolvent condition implying power boundedness.*, Stud. Math., 134 (1999), pp. 143–151.

[30] M. Nashed, *Continuous and semicontinuous analogues of iterative methods of Cimmino and Kaczmarz with applications to the inverse Radon transform.* Mathematical aspects of computerized tomography, Proc., Oberwolfach 1980, Lect. Notes Med. Inf. 8, 160-178 (1981)., 1981.

[31] F. Natterer, *The mathematics of computerized tomography*, SIAM, Philadelphia, PA, 2001.

[32] F. Natterer and F. Wübbeling, *Mathematical methods in image reconstruction*, SIAM Monographs on Mathematical Modeling and Computation, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.

[33] O. Nevanlinna, *Convergence of iterations for linear equations.*, Birkhäuser, Basel, 1993.

[34] R. Plato, *Iterative and parametric methods for linear ill-posed equations*, Habilitationschrift, Fachbereich Mathematik, TU Berlin, 1995.

[35] R. Plato and U. Hämarik, *On pseudo-optimal parameter choices and stopping rules for regularization methods in Banach spaces.*, Numer. Funct. Anal. Optimization, 17 (1996), pp. 181–195.

[36] A. Rieder, *On the regularization of nonlinear ill-posed problems via inexact Newton iterations*, Inverse Problems, 15 (1999), pp. 309–327.

[37] O. Scherzer, *Convergence criteria of iterative methods based on Landweber iteration for solving nonlinear problems*, J. Math. Anal. Appl., 194 (1995), pp. 911–933.