Inverse Problems and Imaging doi:10.3934/ipi.2014.8.149
Volume 8, No. 1, 2014, 149–172

CONVERGENCE RATES FOR KACZMARZ-TYPE REGULARIZATION METHODS

Stefan Kindermann[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

(Communicated by Otmar Scherzer)

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 and 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 [22, 13, 12, 8].

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

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

where y_i are given (possibly noisy) data and the system is assumed to be ill-posed or ill-conditioned.

In what follows, bold variables are used to denote block-structured ones, i.e., objects of a (larger) product space. In order to use a common framework, we define the operator $\hat{\mathbf{A}}$ and the data vector $\hat{\mathbf{y}}$ by

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

2010 *Mathematics Subject Classification.* Primary: 65J20, 65J15; Secondary: 47J06.
Key words and phrases. Ill-posed systems, Landweber–Kaczmarz, convergence rates, regularization.
[1]Corresponding author.

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

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

For generality, we include p corresponding preconditioning operators M_i, for which we assume for the rest of this paper without further notice the following condition:

$$M_i : Y_i \to Y_i \quad \text{symmetric bounded positive definite operators}, \; i = 0, \ldots, p-1\,. \tag{3}$$

The preconditioned version of the system in (1) is

$$M_i^{\frac{1}{2}} A_i x = M_i^{\frac{1}{2}} y_i\,, \quad i = 0, \ldots, p-1\,, \tag{4}$$

which is clearly equivalent to (1) for positive definite M_i. We collect the preconditioners into a block diagonal matrix $\mathbf{M}$, i.e.,

$$\mathbf{M} = \operatorname{diag}(M_i) \quad i = 0, \ldots, p-1\,. \tag{5}$$

It will be convenient to work with the preconditioned operators and data, so let us define

$$\mathbf{A} := \mathbf{M}^{\frac{1}{2}} \hat{\mathbf{A}} = \begin{pmatrix} M_0^{\frac{1}{2}} A_0 \\ \vdots \\ M_{p-1}^{\frac{1}{2}} A_{p-1} \end{pmatrix}, \quad \mathbf{y} := \mathbf{M}^{\frac{1}{2}} \hat{\mathbf{y}} = \begin{pmatrix} M_0^{\frac{1}{2}} y_0 \\ \vdots \\ M_{p-1}^{\frac{1}{2}} y_{p-1} \end{pmatrix}.$$

We also represent the noisy data by $\hat{\mathbf{y}}^\delta$, respectively $\mathbf{y}^\delta = \mathbf{M}^{\frac{1}{2}} \hat{\mathbf{y}}^\delta$. As usual, a bound on the noise is assumed, i.e.,

$$\|y_i^\delta - y_i\| \leq \delta_i\,, \quad i = 0, \ldots, p-1\,. \tag{6}$$

The quantities δ_i are the *noise levels*, i.e., the amount of noise in the i-th equation. In our analysis, however, we only need to know the *overall noise level* δ,

$$\delta^2 = \sum_{i=0}^{p-1} \delta_i^2 = \|\hat{\mathbf{y}} - \hat{\mathbf{y}}^\delta\|^2\,, \tag{7}$$

instead of having information on each δ_i. We also introduce the noise level in the preconditioned version

$$\delta_M^2 := \sum_{i=0}^{p-1} \|M_i^{\frac{1}{2}}(y_i^\delta - y_i)\|^2 = \|\mathbf{y}^\delta - \mathbf{y}\|^2 \leq \|\mathbf{M}^{\frac{1}{2}}\|^2 \delta^2\,. \tag{8}$$

Kaczmarz-type methods. For ill-posed and ill-conditioned problems with a block structure, the class of Kaczmarz-type iterations is a useful iterative regularization method. The original Kaczmarz iteration [17] 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 [29] and was named *Algebraic Reconstruction Technique* (ART). We refer the reader to [30] 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 [24, 25] (infinite dimensional spaces) and [26] (finite dimension). Acceleration of the Kaczmarz iteration for inconsistent linear systems is obtained in [16] 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 [28].

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 [10]) or Newton-type methods (Levenberg-Marquardt [18], IRGN [2], REGINN [34]). 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 [22], 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 stopping criteria in order to carry out the convergence proofs.

In what follows we give a brief overview on the 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 [7]; convergence without rates;
[2009] Expectation–Maximization–Kaczmarz [11]; 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 [8]; 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* [35, 15], the convergence proof in [5] require more delicate (stronger) assumptions as the *adjoint range invariance condition* [10] 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 the iterated Tikhonov method, 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 (respectively Tikhonov) iteration [29, 9]. 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 [33] for sectorial operators (see also Nevanlinna [31]). 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 us to state computable conditions (see, e.g., (39) or (45)) sufficient for convergence rates, assuming a source condition. It turns out that, for sufficiently small stepsizes, one always obtains the standard Hölder convergence rates.

The paper is organized as follows. In Section 2, we define four 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 the iterated Tikhonov method when applied to the full block system (i.e., when the Kaczmarz strategy is not employed). 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. For the sake of completeness, we also present convergence rates for the symmetric iterations (sLK) and (siTK). These can be established following the ideas in [10] to analyze iterative regularization methods in Hilbert spaces. In Section 4 we discuss the obtained results.

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 (4) and contrast them with the usual implicit or explicit Landweber iteration for block structured systems.

For some of the iteration methods below, we have to relate the preconditioning operators to the operators A_i of the ill-posed problem appropriately. As a matter of fact, for the explicit Landweber-based methods, we require the following bound

$$\|A_i^* M_i A_i\| < 2 \qquad \forall i = 0, \dots p-1\,. \tag{9}$$

For bounded preconditioning operators, as the ones used in this paper, this condition can always be satisfied by introducing a stepsize $\tau > 0$, i.e., using scaled preconditioners τM_i instead of M_i. Hence, we also refer to (9) as stepsize constraint.

2.1. **Nonsymmetric Kaczmarz-type iterations 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,

$x_1, \ldots, x_k, \ldots$ to (1), respectively to (4), which is based on the iteration

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

starting at some initial element $\bar{x}_0$ and with M_i given as in (3). The approximate solutions to (1), respectively to (4), are the iterates x_k. Hence, in order to compute x_{k+1} from x_k, one has to cycle through the equations (4) 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), respectively to (4). This is called here *(block) Landweber method*, i.e., the sequence of approximate solutions to (4), x_0, $x_1, \ldots, x_k, \ldots$ is defined by (compare with (10))

$$\begin{aligned} \bar{x}_{n+1} &= \bar{x}_n - A^*_{[n]} M_{[n]} (A_{[n]} x_{k-1} - y^\delta_{[n]}) \,, \\ x_k &:= \bar{x}_{kp} \qquad\qquad k = 0, 1, \ldots \end{aligned} \tag{11}$$

starting at some initial element $\bar{x}_0$. Equivalently, (11) can be written in the more common block form

$$x_k = x_{k-1} - \mathbf{A}^*(\mathbf{A} x_{k-1} - \mathbf{y}^\delta) \,. \tag{12}$$

Once again, a common preconditioner for the block Landweber iteration is the choice $M_i = \tau I$ with a positive stepsize 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 (4). 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 (12) or (11) is here referred to as *(block) iterated Tikhonov regularization*, with iterations x_k given by:

$$\begin{aligned} \bar{x}_{n+1} &= \bar{x}_n - A^*_{[n]} M_{[n]} (A_{[n]} x_{n+1} - y^\delta_{[n]}) \,, \\ x_k &:= \bar{x}_{kp} \qquad\qquad k = 0, 1, \ldots \end{aligned} \tag{13}$$

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

$$x_k = x_{k-1} - \mathbf{A}^*(\mathbf{A} x_k - \mathbf{y}^\delta) \,. \tag{14}$$

This iteration is well-defined because $I + \mathbf{A}^*\mathbf{A}$ is invertible. Note that for computations, the expression (14) is rewritten as

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

Both the block Landweber iteration and the LK iteration have implicit counterparts. The implicit variant of the LK iteration (10) is the *iterated Tikhonov–Kaczmarz (iTK) method* and is defined by (compare with (10))

$$\begin{aligned} \bar{x}_{n+1} &= \bar{x}_n - A^*_{[n]} M_{[n]} (A_{[n]} \bar{x}_{n+1} - y^\delta_{[n]}) \,, \\ x_k &:= \bar{x}_{kp} \qquad\qquad k = 0, 1, \ldots \end{aligned} \tag{15}$$

starting from an arbitrary initial guess $\bar{x}_0$. This iteration is well-defined because 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 the type of a Tikhonov-regularization has to be solved in each step.

Remark 2.1. In view of the Landweber-Kaczmarz-type iteration, (10), the step-size condition (9) appears as a natural assumption in order to guarantee that the iteration operator $I - A_i^* M_i A_i$ is nonexpansive. Let us emphasize that we do not impose any specific bound on the norm of M_i itself (but only bounds on the norms of operators M_i in combination with A_i)! Actually, we expect that large parts of our analysis (e.g., the case of exact data) still holds for unbounded self-adjoint operators M_i as long as (9) is satisfied. However, allowing unbounded operators M_i leads in general to serious problems when noisy data are present (notice that M_i acts on y_i^δ, which does not have to be well-defined then). In this situation one could consider $M_i(y_i^\delta - y_i)$ (if defined) as individual noise level and demand this quantity to be bounded. To avoid the associated difficulties, we a-priori assume the boundedness of M_i in this paper.

As a second remark, let us mention, that our framework also covers the classical Kaczmarz iteration in the finite dimensional case. Using the preconditioners $M_i = (A_i A_i^*)^\dagger$ (cf. [29]), where $\dagger$ denotes the pseudo-inverse, these operators are bounded in the finite-dimensional case. Moreover, the operator $A_i^* M_i A_i$ is the projection onto the range of A_i^*. Since the norm of projection operators is clearly bounded by 1, condition (9) holds, so that our analysis is valid for the classical Kaczmarz case in finite-dimensional spaces.

2.2. **Symmetric Kaczmarz-type iterations.** Further variants of the Kaczmarz-type iterations are their symmetric versions [9]. In contrast to the block-iterations, the iterations LK, iTK, are not invariant if the ordering of the equations are reversed. For these 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. In other words, 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*

$$\bar{x}_{n+1} = \begin{cases} \bar{x}_n - A_{[n]}^* M_{[n]} (A_{[n]} \bar{x}_n - y_{[n]}^\delta) \\ \qquad\qquad \text{if } 0 \le \operatorname{mod}(n, 2p) \le p-1 \\ \bar{x}_n - A_{p-1-[n]}^* M_{p-1-[n]} (A_{p-1-[n]} \bar{x}_n - y_{p-1-[n]}^\delta) \\ \qquad\qquad \text{if } p \le \operatorname{mod}(n, 2p) \le 2p-1 \end{cases} \tag{16}$$
$$x_k := \bar{x}_{k2p} \qquad [n] := \operatorname{mod}(n, p)\,.$$

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

$$\bar{x}_{n+1} = \begin{cases} \bar{x}_n - A_{[n]}^* M_{[n]} (A_{[n]} \bar{x}_{n+1} - y_{[n]}^\delta) \\ \qquad\qquad \text{if } 0 \le \operatorname{mod}(n, 2p) \le p-1 \\ \bar{x}_n - A_{p-1-[n]}^* M_{p-1-[n]} (A_{p-1-[n]} \bar{x}_{n+1} - y_{p-1-[n]}^\delta) \\ \qquad\qquad \text{if } p \le \operatorname{mod}(n, 2p) \le 2p-1 \end{cases} \tag{17}$$
$$x_k := \bar{x}_{k2p} \qquad [n] := \operatorname{mod}(n, p)\,.$$

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 block Landweber iterations (respectively block iterated Tikhonov methods) preconditioned with a suitable block preconditioner. This has already been observed by Natterer [29], 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 [9]. In this section we extend their results to the iterated Tikhonov method.

Let us define the lower triangular operator $\mathbf{L} : \mathbf{Y} \to \mathbf{Y}$ as the part of $\mathbf{A}\mathbf{A}^*$ below the diagonal as follows

$$\mathbf{L} := \begin{pmatrix} 0 & & & 0 \\ M_1^{\frac{1}{2}} A_1 A_0^* M_0^{\frac{1}{2}} & 0 & & \\ \vdots & \ddots & 0 & \\ M_{p-1}^{\frac{1}{2}} A_{p-1} A_0^* M_0^{\frac{1}{2}} & \dots & M_{p-1}^{\frac{1}{2}} A_{p-1} A_{p-2}^* M_{p-2}^{\frac{1}{2}} & 0 \end{pmatrix}. \tag{18}$$

Then the following result holds true [9].

Theorem 1. *Let x_k be the iterates of the Landweber–Kaczmarz method* (10)*. Then the iteration* (10) *can be expressed as (nonsymmetric) block-preconditioned Landweber method of the form*

$$x_{k+1} = x_k - \mathbf{A}^* \mathbf{M}_B (\mathbf{A} x_k - \mathbf{y}^\delta)\,, \tag{19}$$

with

$$\mathbf{M}_B = (\mathbf{I} + \mathbf{L})^{-1}\,, \tag{20}$$

$\mathbf{L}$ *as in* (18)*, and* $\mathbf{I}$ *the identity operator.*

Notice that, since the M_i are bounded, the block operator $\mathbf{M}_B$ is invertible (this is a lower triangular operator). Its norm, as well as the norm of its inverse, can be bounded by constants depending on p and on $(\|M_i^{\frac{1}{2}} A_i\|)_{i=0}^{p-1}$. However, except for nontrivial cases, the operator $\mathbf{M}_B$ is not symmetric. Thus, (19) cannot be seen as a symmetric preconditioned version of the classical block Landweber iteration (12).

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

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

could be used as well in the iteration (19).

Here, we use a slightly different notation as in [9], where $\mathbf{L}$ was defined without the operators $M_i^{\frac{1}{2}}$, but instead $\mathbf{M}_B$ has the form $(\mathbf{D}+\mathbf{L})^{-1}$ with $\mathbf{D} = \mathbf{M}^{-1}$. Because we built in the M_i into $\mathbf{A}$ and $\mathbf{y}$, it is not difficult to see that (19) is equivalent to the statement in [9].

A similar theorem, again due to Elfving and Nikazad [9], holds for the sLK method:

Theorem 2. *Let x_k be the iterates of the symmetric Landweber–Kaczmarz method (16). Then the iteration (16) can expressed as a preconditioned Landweber iteration*

$$x_{k+1} = x_k - \mathbf{A}^* \mathbf{M}_{SB} (\mathbf{A} x_k - \mathbf{y}^\delta)\,, \tag{21}$$

with

$$\mathbf{M}_{SB} = \mathbf{M}_B^* (2\mathbf{I} - \operatorname{diag}(M_i^{\frac{1}{2}} A_i A_i^* M_i^{\frac{1}{2}}))\mathbf{M}_B\,, \tag{22}$$

and $\mathbf{L}, \mathbf{I}, \mathbf{M}_B$ as in Theorem 1.

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. *The sequence x_k generated by the iterated Tikhonov–Kaczmarz method (15) can be expressed as a (nonsymmetric) preconditioned block iterated Tikhonov iteration*

$$x_{k+1} = x_k - \mathbf{A}^* \mathbf{N}_B (\mathbf{A} x_{k+1} - \mathbf{y}^\delta)\,, \tag{23}$$

with

$$\mathbf{N}_B = ((\mathbf{I} - \mathbf{L})^{-1})^*\,. \tag{24}$$

Similarly, the iterates x_k of the symmetric iterated Tikhonov–Kaczmarz method (17) can be expressed as preconditioned iterated Tikhonov method

$$x_{k+1} = x_k - \mathbf{A}^* \mathbf{N}_{SB} (\mathbf{A} x_{k+1} - \mathbf{y}^\delta)\,, \tag{25}$$

with

$$\mathbf{N}_{SB} = \mathbf{N}_B \left(2\mathbf{I} + \operatorname{diag}(M_i^{\frac{1}{2}} A_i A_i^* M_i^{\frac{1}{2}})\right) \mathbf{N}_B^*.$$

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 (10) with initial element ω_0 and with $\bar{\mathbf{A}} = \mathbf{P}\mathbf{A}$, $\bar{\mathbf{y}}^\delta = \mathbf{P}\mathbf{y}^\delta$, respectively, replacing $\mathbf{A}$, $\mathbf{y}^\delta$, in (10) (i.e., with the ordering of the equations reversed) and with $-M_i$ replacing M_i. Thus, according to Theorem 1, we can express

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

with $\mathbf{M}_B$ defined via (20), (18) using $\bar{\mathbf{A}}$ instead of $\mathbf{A}$ and $-M_i$ instead of M_i. (Here we formally use $(-M_i)^{\frac{1}{2}}$ instead of M_i, with the rule $(-1)^{\frac{1}{2}}(-1)^{\frac{1}{2}} = -1$, i.e., $(-M_i)^{\frac{1}{2}} A_i A_j^* (-M_i)^{\frac{1}{2}} = -M_i^{\frac{1}{2}} A_i A_j^* M_i^{\frac{1}{2}}$; all this can be justified by rigorous calculations.)

Going back to the original variables, this means that

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

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,

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

where $\bar{\mathbf{N}}_B$ is defined as in (24), but with the order of the operators reversed. It can be verified that $\mathbf{P}\bar{\mathbf{L}}\mathbf{P} = \mathbf{L}^*$, and hence $\mathbf{P}\bar{\mathbf{N}}_B\mathbf{P} = \mathbf{N}_B^*$. Therefore, with respect to the original variables, we obtain

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

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

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

yield the desired result in the symmetric case. □

Theorems 1–3 allow us to use the convergence theory of the ordinary Landweber iteration and the iterated Tikhonov method in Hilbert spaces to establish convergence rates. The symmetric case is a rather straightforward application of the according theory [10]. The nonsymmetric case, however, is more demanding and will be treated in detail in the following section.

3. Analysis of the Kaczmarz-type methods.

3.1. Convergence rates for the nonsymmetric methods. In this section we present the convergence analysis and establish convergence rates for the Landweber–Kaczmarz method (10) and for the iterated Tikhonov–Kaczmarz method. According to Theorem 1 the former can be written as a Richardson-type iteration of the form (19). 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. Hence, the classical analysis based on self-adjoint operators cannot be applied.

For notational simplicity we define $G := \mathbf{A}^*\mathbf{M}_B\mathbf{A}$. The equivalence between (10) and (19) can be written in the form (see [9])

$$I - G = (I - A_{p-1}^*M_{p-1}A_{p-1})(I - A_{p-2}^*M_{p-2}A_{p-2})\ \dots\ (I - A_0^*M_0A_0)\,, \tag{26}$$

which immediately yields the following result:

Lemma 3.1. *If (9) holds, then G is an accretive operator, i.e., it satisfies*

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

Proof. By definition, for all $0 \leq i \leq p-1$, the operator $A_i^*M_iA_i$ is symmetric positive semidefinite with norm bounded by 2. Thus, $(I - A_i^*M_iA_i)$ is nonexpansive, and so is $I - G$. Consequently,

$$\mathrm{Re}(Gx, x)_{X_\mathbb{C}} = (x, x)_{X_\mathbb{C}} - \mathrm{Re}((I-G)x, x)_{X_\mathbb{C}} \geq \|x\|_{X_\mathbb{C}}^2 - \|(I-G)\|\,\|x\|_{X_\mathbb{C}}^2 \geq 0$$

concluding the proof. □

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

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

holds true [21, Chpt. 3, Th 3.2], i.e., G is a weakly sectorial operator [32, 33]. For such operators the fractional powers G^α, $\alpha > 0$, 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 (19) with noisy data $(y_i^\delta)_{i=0}^{p-1}$, while x_k denotes the iteration (19) with exact data $(A_i x^\dagger)_{i=0}^{p-1}$, i.e., x_k is the iteration with $\mathbf{y}^\delta$ replaced by $\mathbf{y} = \mathbf{A}x^\dagger$, where $x^\dagger$ denotes a solution to (1).

First we estimate the propagated data error:

Lemma 3.2. *Let (9) hold. Moreover, let x_k^δ be the iteration (19) with noisy data and x_k the iteration (19) with exact data. Then we have the following estimate with a constant $C = C(\|\mathbf{A}\|, \|\mathbf{M}_B\|)$*

$$\|x_k^\delta - x_k\| \le Ck\delta_M\,. \tag{27}$$

If, additionally,

$$\sup_{k\in\mathbb{N}} \|(\mathbf{I} - \mathbf{M}_B\mathbf{A}\mathbf{A}^*)^k\| \le C_1 \tag{28}$$

holds with a constant C_1, then

$$\|x_k^\delta - x_k\| \le C\sqrt{k}\delta_M\,, \tag{29}$$

where the constant C depends on C_1, $\|\mathbf{M}_B\|, \|\mathbf{M}_B^{-1}\|$ but not on 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 $I - G$ is nonexpansive, (27) follows immediately with $C = \|\mathbf{A}^*\mathbf{M}_B\|$. Now assume that (28) holds true. Denoting by $g_L(x)$ the polynomial $g_L(x) = \sum_{j=0}^{k}(1-x)^j$, we can write

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

This inequality, together with (28), yields an estimate of the order $k\delta_M^2$, completing the proof. ☐

Our next step is to estimate the approximation error term $x_k - x^\dagger$. For this purpose we need the following lemma, which is proven in the Appendix A. Roughly speaking, it states that $I - G$ is a contraction for elements which are not in the null-space of G. The precise formulation follows:

Lemma 3.3. *Let (9) hold. Moreover, let $\eta > 0$ and $x \in X_{\mathbb{C}}$ with $\|x\|_{X_{\mathbb{C}}} = 1$ be given. If*

$$\operatorname{Re}(Gx, x)_{X_{\mathbb{C}}} \geq \eta\,, \tag{30}$$

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

$$\|(I - G)x\|_{X_{\mathbb{C}}}^2 \leq 1 - \gamma\,. \tag{31}$$

Remark 3.4. It follows from the proof in Appendix A that γ in (31) can be chosen as the largest number ϵ for which (59)–(61) fails to hold. In particular, for η sufficiently small (such that (61) implies (59) and (60)), Lemma 3.3 holds true with the following choice of γ:

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

We can now estimate the approximation error, assuming that a source condition is satisfied. As mentioned above, the usual approach via spectral theory is not possible here due to the lack of symmetry of the corresponding operators. As a replacement, we use functional calculus on the numerical range.

We define the numerical range of the operator G as

$$W(G) := \{(Gx, x)_{X_{\mathbb{C}}} \mid x \in X_{\mathbb{C}},\ \|x\|_{X_{\mathbb{C}}} = 1\}\,. \tag{32}$$

With this definition we are ready to state the main lemma needed to derive the approximation error estimates.

Lemma 3.5. *Let (9) hold and assume the existence of a constant $h > 0$ such that*

$$W(G) \subset \Sigma := \{\lambda \in \mathbb{C} \mid |\arg(\lambda)| \leq \tan^{-1}(h) < \tfrac{\pi}{2}\}\,. \tag{33}$$

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

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

holds true for all $\alpha > 0$.

Proof. It follows from (33), Lemma 3.3, and $|((I - G)x, x)| \leq \|x\|\|(I - G)x\|$ that, for any $\eta > 0$, there exists a constant $0 < \gamma < 1$ with

$$W(G) \subset (\Sigma \cap \{\lambda \in \mathbb{C} \mid 0 \leq \operatorname{Re}(\lambda) < \eta\}) \cup \{\lambda \in \mathbb{C} \mid \operatorname{Re}(\lambda) \geq \eta\,, \\ |1 - \lambda| < \sqrt{(1-\gamma)}\}\,.$$

We now fix $\eta = \cos(\psi)^2$, $\psi = \tan^{-1}(h)$ and take γ as the corresponding constant in (31). Using the functional calculus of Crouzeix [6] we conclude that

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

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

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

Hence,

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

which leads to the estimate

$$|(1-\lambda)^k\lambda^\alpha| \leq \left|1-|\lambda|\cos(\psi)\right|^{\frac{k}{2}}|\lambda|^\alpha \leq \frac{1}{\cos(\psi)^\alpha}\left|1-|\lambda|\cos(\psi)\right|^{\frac{k}{2}}|\lambda\cos(\psi)|^\alpha$$
$$\leq \frac{\alpha^\alpha}{\cos(\psi)^\alpha}(\frac{k}{2})^{-\alpha} \leq \frac{(4\alpha)^\alpha}{\cos(\psi)^\alpha}\frac{1}{(k+1)^\alpha}.$$

For the remaining 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 $C = \max\{\frac{(4\alpha)^\alpha}{\cos(\psi)^\alpha}, 2^\alpha C'\}$. □

Remark 3.6. For a moment we focus our attention to condition (33). Rewriting (33) in terms of the real and imaginary parts of $Gx = G(x+iy)$ and $G = \mathbf{A}^*\mathbf{M}_B\mathbf{A}$, one observes that an equivalent condition to (33) is the existence of an $h > 0$ such that

$$\begin{aligned}|(\mathbf{A}^*\mathbf{M}_B\mathbf{A} - \mathbf{A}^*\mathbf{M}_B^T\mathbf{A})x, y)| \leq & h\left((\mathbf{A}^*\mathbf{M}_B\mathbf{A}x, x) + (\mathbf{A}^*\mathbf{M}_B\mathbf{A}y, y)\right) \\ & \forall x, y, \in X, \|x\|^2 + \|y\|^2 = 1.\end{aligned} \tag{34}$$

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. Using the definition of $\mathbf{M}_B$ (see (20)), it follows that the above inequality holds if, for all z, $v \in Y$,

$$|(\mathbf{L}^T - \mathbf{L})z, v)| \leq h\left((z, z) + (v, v) + (\mathbf{L}z, z) + (\mathbf{L}v, v)\right). \tag{35}$$

This remark leads to the following lemma:

Lemma 3.7. *If q is such that*

$$\|\mathbf{L}\| \leq q < 1, \tag{36}$$

then (33) holds with $h = q/(1-q)$.

Proof. We start by proving (35). Notice that

$$\begin{aligned}|(\mathbf{L}^T - \mathbf{L})z, v)| &\leq |(z, \mathbf{L}v)| + |(\mathbf{L}z, v)| \leq \|z\|\|\mathbf{L}v\| + \|\mathbf{L}z\|\|v\|, \\ (\mathbf{L}z, z) + (\mathbf{L}v, v) &\geq -\|\mathbf{L}z\|\|z\| - \|\mathbf{L}v\|\|v\|.\end{aligned}$$

Thus, it suffices to prove

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

However, that this last inequality is a consequence of

$$\begin{aligned}&\|z\|\|\mathbf{L}v\| + \|\mathbf{L}z\|\|v\| + h\left(\|\mathbf{L}z\|\|z\| + \|\mathbf{L}v\|\|v\|\right) \\ &\qquad \leq q\left(2\|z\|\|v\| + h\|z\|^2 + h\|v\|^2\right) \leq q(1+h)\left(\|z\|^2 + \|v\|^2\right).\end{aligned}$$

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

It remains to investigate condition (28). For this purpose we rely on the following theorem [27, 23] (see also [19]):

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

$$\|(T-\lambda I)^{-1}\| \le C\frac{1}{|\lambda-1|}, \quad \forall\ |\lambda|>1\,,\ \lambda\in\mathbb{C}\,, \tag{37}$$

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

With the setting $T = \mathbf{I} - \mathbf{M}_B\mathbf{A}\mathbf{A}^*$, we thus obtain (28) if we can prove

$$\|(\mathbf{M}_B\mathbf{A}\mathbf{A}^* - \lambda\mathbf{I})^{-1}\| \le C\frac{1}{|\lambda|} \quad |\lambda-1|>1\,,\ \lambda\in\mathbb{C}\,. \tag{38}$$

The next result establishes a sufficient condition for (38).

Lemma 3.8. *If*

$$\|\mathbf{L}\| + \frac{1}{2}\|\mathbf{A}\mathbf{A}^*\| < 1\,, \tag{39}$$

then (38) is satisfied. Consequently, the propagated data error estimate (28) is also satisfied.

Proof. We prove (38). Define $\mathbf{S}(\lambda) := (\mathbf{A}\mathbf{A}^* - \lambda\mathbf{I})$ for $\lambda\in\mathbb{C}$. From the definition of $\mathbf{M}_B$, we obtain for each λ in (38)

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

as long as $|\lambda|\,\|\mathbf{S}^{-1}(\lambda)\|\,\|\mathbf{L}\| < 1$.

In order to prove this condition we use the estimate (for self-adjoint operators)

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

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

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

From the hypothesis we conclude that $\frac{2}{2-\|\mathbf{A}\mathbf{A}^*\|}\|\mathbf{L}\| < 1$. Thus, inequality (38) holds true. Equation (28) 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. The following theorem is then a collection of the previous results.

Theorem 5. *Let $\mathbf{L}, \mathbf{A}$ satisfy (39). Moreover, assume that the source condition*

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

is satisfied. Then, the iterates of the Landweber–Kaczmarz method (10) *satisfy the error estimate*

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

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

$$\|x_k - x^\dagger\| \sim \delta_M^{\frac{2\nu}{2\nu+1}} \le C\|\mathbf{M}^{\frac{1}{2}}\|\delta^{\frac{2\nu}{2\nu+1}}\,.$$

Notice that, by introducing a stepsize τ and making M_i sufficiently small, we can always achieve that the hypothesis in this theorem (except for the source condition) is satisfied.

3.2. **Analysis of the nonsymmetric iterated Tikhonov–Kaczmarz method.** In what follows, we derive convergence rates for the iTK method (15) (in the block form (23)) in the nonsymmetric case. First of all, from the equivalence between (15) and (23), it follows that

$$(I + \mathbf{A}^*\mathbf{N}_B\mathbf{A}) = (I + A_0^*M_0A_0)(I + A_1^*M_1A_1)\ldots(I + A_{p-1}^*M_{p-1}A_{p-1})\,.$$

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

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

We first investigate the propagated data error,

Lemma 3.9. *Let* x_k^δ *denote the iteration* (23) *with noisy data, and* x_k *the iteration* (23) *with exact data. Then we have the estimate with a constant* $C = C(\|\mathbf{A}\|, \|\mathbf{N}_B\|)$

$$\|x_k^\delta - x_k\| \le Ck\delta_M. \tag{40}$$

If, additionally,

$$\sup_{k\in\mathbb{N}} \|(\mathbf{I} + \mathbf{N}_B\mathbf{A}\mathbf{A}^*)^{-k}\| \le C_1 \tag{41}$$

holds with a constant C_1, *then*

$$\|x_k^\delta - x_k\| \le C\sqrt{k}\delta_M\,, \tag{42}$$

where the constant C *depends on* C_1, $\|\mathbf{N}_B\|$, $\|\mathbf{N}_B^{-1}\|$ *but not on* k.

Proof. We may express iteration (23) as

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

from which (40) immediately follows. Now, defining $g_k(\lambda) := \sum_{j=0}^{k-1}(1+\lambda)^{-j}$, we obtain the estimate (compare with the Landweber iteration, e.g., [10] Chap. 6)

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

Using (41) in the last inequality we obtain (42). □

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. *Assume the existence of an $h > 0$ such that the numerical range of $\mathbf{A}^*\mathbf{N}_B\mathbf{A}$ is contained in the sector*

$$W(\mathbf{A}^*\mathbf{N}_B\mathbf{A}) \subset \{\lambda \in \mathbb{C} \,|\, |\arg(\lambda)| \le \tan^{-1}(h) < \tfrac{\pi}{2}\}\,. \tag{43}$$

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 (43), 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 once again obtain

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

Furthermore, $|1+\lambda| \ge (1 + |\lambda|\cos(\psi))$, $\psi = \tan^{-1}(h)$. 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. □

Remark 3.11. As before, an equivalent condition to (43) is the existence of an $h > 0$ such that

$$\begin{aligned} \Big|\big([\mathbf{A}^*\mathbf{N}_B\mathbf{A} - \mathbf{A}^*\mathbf{N}_B^T\mathbf{A}]x, y\big)\Big| \;&\le h\big[(\mathbf{A}^*\mathbf{N}_B\mathbf{A}x, x) + (\mathbf{A}^*\mathbf{N}_B\mathbf{A}y, y)\big]\,, \\ &\forall\, x, y \in X,\ \|x\|^2 + \|y\|^2 = 1\,. \end{aligned} \tag{44}$$

In the next lemma we discuss a sufficient condition for (41), and (43) in Lemma 3.9, and Lemma 3.10 respectively.

Lemma 3.12. *If*

$$\|\mathbf{L}\| < 1\,, \tag{45}$$

then (41) and (43) hold true.

Proof. The proof of (43) follows the lines of the proof of Lemma 3.7. To prove (41), we use Theorem 4 with $T = (\mathbf{I} + \mathbf{N}_B\mathbf{A}\mathbf{A}^*)^{-1}$. Thus, for $\lambda \in \mathbb{C}$ with $|\lambda| > 1$, we estimate

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

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

$$\|(\mathbf{A}\mathbf{A}^* + s\mathbf{I} - s\mathbf{L}^*)^{-1}\| \;\le\; C\frac{1}{|s|}\,,\ \forall\ |s - 1| < 1\,.$$

Notice that

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

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

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

it follows that

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

establishing the desired bound. Consequently, inequality (41) follows from Theorem 4. □

Collecting the results, it follows from Lemma 3.12, Lemma 3.10 and Lemma 3.9 that we find the following convergence rates for the iterated Tikhonov–Kaczmarz method.

Theorem 6. *Let* $\mathbf{L}$ *satisfy* (45)*. 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 (13) *satisfies the estimate*

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

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

$$k \sim \delta_M^{\frac{-2}{2\nu+1}}$$

yields the convergence rate

$$\|x_k - x^\dagger\| \sim \delta_M^{\frac{2\nu}{2\nu+1}} \leq C \|\mathbf{M}^{\frac{1}{2}}\| \delta^{\frac{2\nu}{2\nu+1}}\,.$$

Remark 3.13. 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 probably 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 accretive operators [20] might be used.

Remark 3.14. The results of Theorems 5 and 6 use classical a-priori parameter choice rules $\alpha \sim \delta_M^{\frac{-2}{2\nu+1}}$. However, using the results of Plato and Hämarik [32], it is possible to establish convergence rates for a-posterior parameter choice rules as well, e.g., using the discrepancy principle.

The main conditions for deriving convergence rates results are (39) and (45). We now present some simpler sufficient conditions for these by estimating the norm of $\mathbf{L}$.

Define the matrix $|\mathbf{L}| \in \mathbb{R}^{p\times p}$ as the lower triangular matrix with zero diagonal

$$|\mathbf{L}|_{i,j} = \begin{cases} 0 & j \geq i \\ \|\mathbf{L}_{i,j}\| & \text{else} \end{cases}, \quad i,j = 0, \dots p-1\,. \tag{46}$$

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

$$\|\mathbf{L}_{i,j}\| \leq |\mathbf{L}|_{i,j}\,, \quad i < j\,,\ i,j = 0, \dots p-1\,. \tag{47}$$

Then we obtain

Lemma 3.15. *With $|\mathbf{L}|$ as in (46) or (47), we have*

$$\|\mathbf{L}\| \leq \sigma_{max}(|\mathbf{L}|)\,,$$

where σ_{max} denotes the largest singular value of its matrix argument.

Proof. For $z \in \mathbf{Y}$, the i-th component of $\mathbf{L}z$ can be bounded by

$$\|(\mathbf{L}z)_i\|_{Y_i} \leq \sum_{j=0}^{p-1} |\mathbf{L}|_{i,j} \|z_j\|\,.$$

Taking the sum of squares and noting that σ_{max} is the operator norm, yield the desired result. □

Consequently, we have

Corollary 1. *Let $|\mathbf{L}|$ be as in (46) or (47). The convergence rates result of Theorem 5 hold true if*

$$\sigma_{max}(|\mathbf{L}|) + \frac{1}{2}\sum_{i=0}^{p-1} \|A_i^* M_i A_i\| < 1\,. \tag{48}$$

The convergence rates result of Theorem 6 hold true if

$$\sigma_{max}(|\mathbf{L}|) < 1\,. \tag{49}$$

Remark 3.16. Considering the simplest preconditioner of the form $M_i = \tau I$, with $\tau > 0$ the stepsize parameter, we see that the Landweber-Kaczmarz-type iteration yields convergence rates for all τ sufficiently small. Indeed, for

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

where $|\mathbf{L}|$ is the lower triangular, zero diagonal matrix with $|\mathbf{L}|_{i,j} = \|A_i A_j^*\|$, Corollary 1 applies. The same holds using preconditioners M_i. Replacing them by τM_i, we always find an interval of possible stepsizes such that the hypothesis of Corollary 1 are satisfied.

3.3. **Convergence rates for the symmetric iterations sLK, siTK.** For completeness as well as for comparison reasons, we state the convergence rates results for the symmetric iterations. If we assume (9), then we can define

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

which yields

$$\mathbf{M}_{SB} = \mathbf{B}^* \mathbf{B} \,. \tag{51}$$

Hence, the iterates of the sLK method are equivalent to the Landweber iteration applied to the system

$$\mathbf{B}\mathbf{A} = \mathbf{B}\mathbf{y} \,, \tag{52}$$

where we also use the noise level $\delta_S = \|\mathbf{B}(\mathbf{y}^\delta - \mathbf{y})\|$. The classical theory for Landweber iteration immediately yields [10]:

Theorem 7. *Let (9) hold. Moreover, assume that a source condition $x^\dagger - x_0 = (\mathbf{A}^*\mathbf{M}_{SB}\mathbf{A})^\nu \omega$ hold. Then, for the iterations of the symmetric Landweber–Kaczmarz and the a-priori parameter choice $k \sim \delta_S^{-\frac{2}{(2\nu+1)}}$, we find the optimal order convergence rate*

$$\|x_k - x^\dagger\| \leq C \delta_s^{\frac{2\nu}{(2\nu+1)}} .$$

A similar result can be established for the siTK method. Define

$$\mathbf{B_T} := (2\mathbf{I} + \mathrm{diag}(M_i^{\frac{1}{2}} A_i A_i^* M_i^{\frac{1}{2}}))^{\frac{1}{2}} \mathbf{N}_B^* \,.$$

Moreover, consider the equation

$$\mathbf{B_T}\mathbf{A}x = \mathbf{B_T}\mathbf{y} \,, \tag{53}$$

and use the noise level $\delta_S = \|\mathbf{B_T}(\mathbf{y}^\delta - \mathbf{y})\|$. Then, the following result follows.

Theorem 8. *Let $x^\dagger - x_0$ satisfy the source condition $x^\dagger - x_0 = (\mathbf{A}^*\mathbf{N}_{SB}\mathbf{A})^\nu \omega$. Then, for the iterations of iterations of the symmetric iterated Tikhonov method and the a-priori parameter choice $k \sim \delta_S^{-\frac{2}{(2\nu+1)}}$, we find the optimal order convergence rate*

$$\|x_k - x^\dagger\| \leq C \delta_s^{\frac{2\nu}{(2\nu+1)}} .$$

Using estimates as above and Heinz' inequality (see [10]), we can even use more standard source conditions.

Proposition 3.1. *Let (9) hold, then there are constants m_1, m_2 such that*

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

Moreover, there are constants n_1, n_2 such that

$$n_1 \|\mathbf{y}\|_{\mathbf{Y}} \quad \leq \quad \|\mathbf{B_T}\mathbf{y}\| \ \leq \ n_2 \|\mathbf{y}\|_{\mathbf{Y}} \quad \forall y \in \mathbf{Y} \,.$$

In particular, for $0 \leq \nu \leq \frac{1}{2}$, in each of these cases the source conditions $x^\dagger - x_0 = (\mathbf{A}^*\mathbf{M}_{SB}\mathbf{A})^\nu \omega$ and $x^\dagger - x_0 = (\mathbf{A}^*\mathbf{A})^\nu \omega$, and $x^\dagger - x_0 = (\mathbf{A}^*\mathbf{N}_{SB}\mathbf{A})^\nu \omega$ and $x^\dagger - x_0 = (\mathbf{A}^*\mathbf{A})^\nu \omega$ are equivalent.

Thus, the results of Theorems 7 and 8 hold with the usual source condition $x^\dagger - x_0 = (\mathbf{A}^*\mathbf{A})^\nu \omega$ for $0 \le \nu \le \frac{1}{2}$ and, of course, with δ_S replaced by δ. Hence we find the exact same convergence rates under the same source condition as for the ordinary block Landweber iteration (or block iterated Tikhonov method).

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

$$\tau_{sLK} < \frac{2}{\max_i \|A_i^* A_i\|} .$$

This should be contrasted with the corresponding choice for the block-Landweber iteration, where τ has to be chosen such that $\tau\|\mathbf{A}\|^2 < 2$, i.e.,

$$\tau_L < \frac{2}{\sum_{i=0}^{p-1} \|A_i^* A_i\|} < \frac{2}{\max_i \|A_i^* A_i\|} .$$

A similar situation holds for the Landweber–Kaczmarz iteration. Although for convergence rates we required (50), for pure convergence the condition $\tau_{LK} < \frac{2}{\max_i \|A_i^* A_i\|}$ suffices [9] (at least in the finite dimensional case). Thus, besides the fact that the Kaczmarz-type iterations are easier to implement than the block Landweber ones, we may also choose a large stepsize.

4. Conclusion.

We have established convergence rates for the Landweber–Kaczmarz method and the iterated Tikhonov–Kaczmarz method; both the symmetric and the nonsymmetric versions of each method are considered. Since the only conditions for the convergence theorems are bounds on $A_i^* M_i A_i$, 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., (39), (45), (48), (49)), 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 (43), (41) (see also (38), (33)) can maybe be proven directly.

Although, asymptotically (as $\delta \to 0$), the Kaczmarz variants perform similarly to the corresponding block iterations, the former have some advantages, e.g., a simpler implementation and possibly a larger stepsize τ can be chosen. However, even if all iteration methods in this paper have similar convergence rates in δ, the Kaczmarz-type iterations can be quite different to the block variants in practice. For instance, in the first iterations, when the error is dominated by the approximation error, the Kaczmarz-type iterations may show a faster decay of the error when a larger stepsize can be used compared to the block iterations.

If detailed information on the structure of the exact solution is available, the results in Section 3 can be used in order to estimate the decay rates of the approximation error in appropriate subspaces. One may use Riesz projections of the exact solution onto parts of the spectrum of $\mathbf{A}^*\mathbf{M}_B\mathbf{A}$ (respectively $\mathbf{A}^*\mathbf{N}_B\mathbf{A}$) to analyze the convergence rates in more detail: The components of the exact solution in the subspace of the projections corresponding to points in the spectrum that are close to zero and/or away from the real axis will contribute to a slow convergence of the approximation error.

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 analysis results in Section 3 can be extended to general nonsymmetric preconditioned Landweber (and iterated Tikhonov) iterations, which are highly relevant in practical large scale applications.

Appendix A. **Proof of Lemma 3.3.** *Proof of Lemma 3.3.* 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

(54)
$$\begin{aligned}
&\|(I - A_i^* M_i A_i)x\|^2 = \int_{\lambda \le \xi} (1-\lambda)^2 d\|E_{\lambda,i}x\|^2 + \int_{\lambda > \xi} (1-\lambda)^2 d\|E_{\lambda,i}x\|^2 \\
&\le \|P_{\xi,i}x\|^2 + \max\{(1-\xi)^2, (1 - \|A_i^* M_i A_i\|)^2\} \|Q_{\xi,i}x\|^2 \\
&= \|P_{\xi,i}x\|^2 + (1-\xi)^2(\|x\|^2 - \|P_{\xi,i}x\|^2) = (1 - (1-\xi)^2)\|P_{\xi,i}x\|^2 + (1-\xi)^2\|x\|^2 \\
&= \theta\|P_{\xi,i}x\|^2 + (1-\theta)\|x\|^2,
\end{aligned}$$

and

$$\begin{aligned}
&\|A_i^* M_i A_i x\|^2 \\
&= \int_{\lambda \le \xi} \lambda^2 d\|E_{\lambda,i}x\|^2 + \int_{\lambda > \xi} \lambda^2 d\|E_{\lambda,i}x\|^2 \le \xi^2\|P_{\xi,i}x\|^2 + \|M_i^{\frac{1}{2}} A_i\|^4 \|Q_{\xi,i}x\|^2 \\
&\le \xi^2\|P_{\xi,i}x\|^2 + 4\left(\|x\|^2 - \|P_{\xi,i}x\|^2\right). \qquad (55)
\end{aligned}$$

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

$$H_k = \Pi_{i=0}^k (I - A_i^* M_i A_i) \quad \Leftrightarrow \quad \begin{aligned} H_k &= H_{k-1} - A_k^* M_k A_k H_{k-1}, \\ H_0 &= (I - A_0^* M_0 A_0)\,. \end{aligned}$$

We know that $\|H_k\| \le 1$. Moreover, from the recursion formula

$$\begin{aligned}
H_k - I =& H_{k-1} - I - A_k^* M_k A_k (H_{k-1} - I) - A_k^* M_k A_k \\
=& (I - A_k^* M_k A_k)(H_{n-1} - I) - A_k^* M_k A_k\,,
\end{aligned}$$

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

$$\|(H_k - I)x\| \le \|(H_{k-1} - I)x\| + \|A_k^* M_k A_k x\| \le \sum_{i=0}^{k} \|A_i^* M_i A_i x\|\,.$$

Since $G = G - I + I = I - H_{p-1}$, we have

$$\|Gx\| \le \sum_{i=0}^{p-1} \|A_i^* M_i A_i x\|\,. \qquad (56)$$

Thus, applying (54), we obtain the estimate

$$\begin{aligned}
\|H_k x\|^2 &= \|(I - A_k^* M_k A_k) H_{k-1} x\|^2 \le \theta \|P_{\xi,k} H_{k-1} x\|^2 + (1-\theta)\|H_{k-1}x\|^2 \\
&\le (1-\theta)\|H_{k-1}x\|^2 + \theta \left(\|P_{\xi,k} x\| + \|P_{\xi,k}(I - H_{k-1})x\|\right)^2 \\
&\le (1-\theta)\|H_{k-1}x\|^2 + \theta \left(\|P_{\xi,k} x\| + \sum_{i=0}^{k-1} \|A_i^* M_i A_i x\|\right)^2 \qquad (57)\\
&\le (1-\theta)\|x\|^2 + \theta \left(\|P_{\xi,k} x\| + \sum_{i=0}^{k-1} \|A_i^* M_i A_i x\|\right)^2 .
\end{aligned}$$

Next, define 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

$$|((I-G)x, x)_{X_{\mathbb{C}}}|^2 \ge (1-\epsilon), \quad \mathrm{Re}(Gx, x)_{X_{\mathbb{C}}} \ge \eta \quad \|x\|_{X_{\mathbb{C}}} = 1\,. \qquad (58)$$

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

$$\epsilon \le \left(\frac{1}{1 + \sum_{i=0}^{p-1} \sqrt{D_i}}\right)^{2^p} \qquad (59)$$

$$(1-\sqrt{\epsilon})^2 \ge \max_{i=0,\dots,p-1} (1 - \|A_i^* M_i A_i\|)^2 \qquad (60)$$

$$\epsilon < \left(\frac{\eta}{\sum_{i=0}^{p-1} \sqrt{D_i}}\right)^{2^{p+1}} . \qquad (61)$$

Since $|((I-G)x, x)_{X_{\mathbb{C}}}| \le \|(I-G)x\|$ and

$$\|(I-G)x\| = \|H_{p-1}x\| \le \|H_{p-2}x\| \le \cdots \le \|H_0 x\|\,,$$

it follows that for such ϵ we can find an x as in (58) with

$$\|H_k x\|^2 \ge 1 - \epsilon \qquad \forall k = 0, \dots p-1\,.$$

By (60), the choice $\xi = \sqrt{\epsilon}$ can be used in (54) and (57). Therefore, we obtain with $\theta = 2\sqrt{\epsilon} - \epsilon > \sqrt{\epsilon}$ the inequality

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

yielding the estimate

$$\|P_{\sqrt{\epsilon},k} x\| + \sum_{i=0}^{k-1} \|A_i^* A_i x\| \ge \sqrt{1 - \frac{\epsilon}{\theta}} \ge 1 - \frac{\epsilon}{\theta} \ge 1 - \sqrt{\epsilon}\,, \qquad \forall k = 0, \dots p-1\,. \qquad (62)$$

Using (55), we get

$$\|A_k^* M_k A_k x\|^2 \le \epsilon \|P_{\sqrt{\epsilon},k} x\|^2 + 4(1 - \|P_{\sqrt{\epsilon},k} x\|^2) \le \epsilon + 4(1 - \|P_{\sqrt{\epsilon},k} x\|^2)\,. \qquad (63)$$

For $k = 0$, we obtain from (62) and (63)

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

We proceed by induction to show that

$$\|A_k^* M_k A_k x\|^2 \le D_k \epsilon^{\frac{1}{2^{k+1}}} \quad \forall k = 0, \dots, p-1\,. \tag{64}$$

Using (62) and the induction hypothesis for $k-1, k \ge 1$, we find

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

Notice that, due to (59), the right hand side in this inequality is positive. Hence, by (63) we obtain

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

which verifies (64) for all $k = 0, \dots p-1$. Now we derive a contradiction to (58). Using (56) and (61) we obtain

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

Hence, (58) cannot be true for such an ϵ and the proof is complete. □

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 309767/2013–0.

REFERENCES

[1] P. K. Anh and C. V. Chung, Parallel regularized Newton method for nonlinear ill-posed equations, *Numer. Algorithms*, **58** (2011), 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), 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), 153–182.

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

[7] 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), 596–607.

[8] 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), 1–17.

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

[10] H. W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems*, vol. 375 of Mathematics and its Applications, Kluwer Academic Publishers Group, Dordrecht, 1996.

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

[12] M. Haltmeier, Convergence analysis of a block iterative version of the loping Landweber-Kaczmarz iteration, *Nonlinear Anal.*, **71** (2009), 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), 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), 289–298.

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

[16] M. Hanke and W. Niethammer, On the acceleration of Kaczmarz's method for inconsistent linear systems, *Linear Algebra Appl.*, **130** (1990), 83–98, Linear algebra in image reconstruction from projections.

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

[18] B. Kaltenbacher, A. Neubauer and O. Scherzer, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, Walter de Gruyter GmbH & Co. KG, Berlin, 2008.

[19] N. Kalton, S. Montgomery-Smith, K. Oleszkiewicz and Y. Tomilov, Power-bounded operators and related norm estimates, *J. London Math. Soc. (2)*, **70** (2004), 463–478.

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

[21] T. Kato, *Perturbation Theory for Linear Operators*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the 1980 edition.

[22] R. Kowar and O. Scherzer, Convergence analysis of a Landweber-Kaczmarz method for solving nonlinear ill-posed problems, in *Ill-posed and inverse problems*, VSP, Zeist, (2002), 253–270.

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

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

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

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

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

[28] M. Z. Nashed, Continuous and semicontinuous analogues of iterative methods of Cimmino and Kaczmarz with applications to the inverse Radon transform, in *Mathematical aspects of computerized tomography (Oberwolfach, 1980)*, vol. 8 of Lecture Notes in Med. Inform., Springer, Berlin, (1981), 160–178.

[29] F. Natterer, *The Mathematics of Computerized Tomography*, SIAM, Philadelphia, PA, 2001.

[30] 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.

[31] O. Nevanlinna, *Convergence of Iterations for Linear Equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1993.

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

[33] R. Plato, *Iterative and Other Methods for Linear Ill-Posed Equations*, Habilitation thesis, Department of Mathematics, TU Berlin, 1995.

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

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

Received August 2012; revised November 2013.

E-mail address: kindermann@indmath.uni-linz.ac.at
E-mail address: acgleitao@gmail.com