FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG
INSTITUT FÜR INFORMATIK (MATHEMATISCHE MASCHINEN UND DATENVERARBEITUNG)

Lehrstuhl für Informatik 10 (Systemsimulation)

Algebraic multigrid for general inconsistent linear systems: Preliminary results

Constantin Popa

Lehrstuhlbericht 06-2

Algebraic multigrid for general inconsistent linear systems: Preliminary results

Constantin Popa*

Abstract

In this paper we analyse some basic concepts from classical Algebraic Multigrid theory - smoothing property and the correction step - for general rectangular systems of linear equations. We prove the smoothing property for the classical Kaczmarz algorithm in the consistent case, which generalizes a well known result by A. Brandt. We also prove a smoothing property for the Kaczmarz Extended algorithm in the inconsistent case. Some considerations and results are also described for a general correction scheme for inconsistent systems.

2000 MS Classification: 65F10, 65F20, 65N55

Key words and phrases: algebraic multigrid, smoothing property, Kaczmarz relaxation, inconsistent least squares problems, correction step.

1 Introduction

The basic concepts and results of Algebraic Multigrid (AMG, for short) have been first introduced by A. Brandt in the initial version of [2] (as an 1983 Technical Report at Weizmann Institute, Israel). Not far from this moment, J. Ruge and K. Stüben have developed and extended Brandt's ideas in [10]. According to this later paper we replay the following comments: *The algebraic multigrid approach is developed to solve* matrix equations *using the principles of usual multigrid methods. In contrast with "geometric" multigrid methods, the relaxation used in AMG is* fixed*....From a theoretical point of view, the process is best understood in the context of symmetric M-matrices, although, in practice, its use is not restricted to such cases.* According to both above mentioned comments and basic references, for designing an efficient AMG code we need to first state and try to obtain the "smoothing property" of the relaxation scheme used. This property is, in fact related only to the relaxation scheme we intend to use and not (yet) "connected" to the AMG elements: fine and coarse grids, interpolation and restriction operators, coarse grid matrices. The second basic AMG concept, which involves all the above mentioned elements, is the "approximation property". Both these properties establish clear quantitative relations between the errors and residuals before and after the respective AMG steps - relaxation or correction. For a better understanding we shall very briefly replay in what follows these facts. Let A be an $n \times n$ invertible matrix, $b \in \mathbb{R}^n$ and $x^* \in R^n$ the (unique) solution of the system

$$Ax^* = b. \tag{1}$$

We shall consider a 2-grid AMG algorithm, in which the correction step is made before the relaxation one. Let x be a current approximation of x^* and $\hat{x}, \bar{x}$ the approximations after the correction step, respectively the relaxation one. With these elements we define the corresponding errors and residuals by

$$e = x - x^*, \ \hat{e} = \hat{x} - x^*, \ \bar{e} = \bar{x} - x^*, \ \hat{r} = A\hat{e} = A\hat{x} - b. \tag{2}$$

Let $\| \cdot \|_1, \ \| \cdot \|_2$ some fixed vector norms on R^n.
Smoothing property: it exists a constant $\alpha \geq 0$ (independent on the dimension n or the eigenvalues of A) such that

$$\| \bar{e} \|_1^2 \leq \| \hat{e} \|_1^2 - \alpha \| \hat{r} \|_2^2 . \tag{3}$$

*"Ovidius" University of Constanta, Romania; e-mail: cpopa@univ-ovidius.ro. The paper was supported by the DAAD via a grant as a visiting professor at the Friedrich-Alexander Universität Erlangen- Nürnberg, Germany, in the period December 2005 - January 2006.

Approximation property: it exists a constant $\beta > 0$ (independent on the dimension n or the eigenvalues of A) such that

$$\| \hat{r} \|_2^2 \geq \frac{1}{\beta} \| \hat{e} \|_1^2 . \tag{4}$$

Beside the above two main concepts of AMG, we also need a specific property of the correction step (valid in some cases, as e.g. variational AMG, see [10]).
Correction error reduction: after the correction step we have

$$\| \hat{e} \|_1^2 \leq \| e \|_1^2 . \tag{5}$$

If all the above properties (3) - (5) hold, then $\alpha \leq \beta$ and the 2-grid AMG algorithm converges with an error reduction factor per iteration (with respect to the $\| \cdot \|_1$ norm) given by

$$\| \bar{e} \|_1^2 \leq \left(1 - \frac{\alpha}{\beta}\right) \| e \|_1^2 . \tag{6}$$

As Ruge and Stüben mentioned before, all these interesting theoretical results were obtained in the case of "symmetric M-matrices" (i.e. symmetric and positive definite with non-positive off-diagonal elements). In the present preliminary report we tried to adapt and extend, especially from a theoretical view point, some of the above described results and concepts to the general case of arbitrary (possible inconsistent) linear systems of equations. Such systems arise in many very important and actual real world problems, among which we can mention those coming from the field of computerized tomography image reconstruction from projections. In this respect, the report is organized as follows. In section 2 we replay the original definitions and results concerning the smoothing property for Gauss-Seidel and Kaczmarz relaxations obtained by A. Brandt in [2], together with extensions for SOR and ω-Kaczmarz methods that we obtained in a previous paper. In section 3 we prove the smoothing property for Kaczmarz relaxation in the case of arbitrary (rank-defficient), but consistent systems. This result generalizes Brandt's one for invertible matrices. In section 4 we prove a smoothing-like property for the Kaczmarz Extended algorithm which we introduced in [6] for arbitrary inconsistent linear least squares problems. In section 5 we analyse a very general correction-like step for inconsistent least squares problems and give examples (coming from the paper [9]) that satisfy the assumptions of this general construction. As mentioned in the title, the report is only a "preliminary" one. It must be continued with a deeper analysis and the design of a robust correction step together with some efficient implementation on computerized tomography image reconstruction problems.

2 The smoothing property - definition and classical results

Let B be an $n \times n$ symmetric and positive definite (SPD, for short) matrix and $c \in I\!R^n$ a given vector. By B_i, B_{ij} we shall denote the i-th row and (i,j)-th element of B. All the vectors that will appear will be column vectors. For the purposes of this section we consider the system of linear equations

$$Bx^* = c, \tag{7}$$

with $x^* = B^{-1}c$ its unique solution. Let

$$x^0 \in I\!R^n, x^{k+1} = Gx^k + g, \quad k \geq 0 \tag{8}$$

or componentwise

$$x_i^{k+1} = g_i + \sum_{j=1}^{n} G_{ij} x_j^k, \tag{9}$$

be a (convergent) relaxation scheme for (7) (i.e. a first order consistent iterative method, see [12] for details). Denoting by $\langle \cdot, \cdot \rangle, \| \cdot \|$ the Euclidean scalar product and norm, respectively we define the energy norms $\| \cdot \|_B$ and $\| \cdot \|_{D^{-1}}$

$$\| z \|_B = \sqrt{\langle Bz, z \rangle}, \quad \| z \|_{D^{-1}} = \sqrt{\langle D^{-1}z, z \rangle}, \; z \in I\!R^n, \tag{10}$$

where

$$D = \mathrm{diag}(B) = \mathrm{diag}(B_{11}, \ldots, B_{nn}). \tag{11}$$

The smoothing property for the relaxation (8) was first introduced by A. Brandt in [2], in the context of algebraic multigrid algorithms. We shall briefly describe it in what follows. For this, let x be a given approximation of x^*, $\bar{x}$ the approximation of x^* after one step of the relaxation scheme (8) applied to x and $e, \bar{e}, r$ the corresponding errors and residual defined by

$$e = x - x^*, \quad r = Be = Bx - c, \quad \bar{e} = \bar{x} - x^*. \tag{12}$$

Definition 1 *We say that the relaxation scheme* (8) *satisfies the* **smoothing property** *(SP, for short) for the system* (7) *if it exists a constant* $\alpha > 0$ *(independent on* n *or the eigenvalues of* B*) such that*

$$\| \bar{e} \|_B^2 \leq \| e \|_B^2 - \alpha \| r \|_{D^{-1}}^2 . \tag{13}$$

In what follows we shall present classical results about the SP property for some well known relaxation schemes of the type (8).

Gauss-Seidel relaxation. Let $x^0 \in I\!R^n$; for $k = 0, 1, \ldots$ do

$$x_i^{k+1} = \frac{1}{B_{ii}} \left[c_i - \sum_{j<i} B_{ij} x_j^{k+1} - \sum_{j>i} B_{ij} x_j^k \right], \forall i = 1, \ldots, n. \tag{14}$$

SOR relaxation. Let $x^0 \in I\!R^n$; for $k = 0, 1, \ldots$ do

$$x_i^{k+1} = (1-\omega) x_i^k + \frac{\omega}{B_{ii}} \left[c_i - \sum_{j<i} B_{ij} x_j^{k+1} - \sum_{j>i} B_{ij} x_j^k \right], \forall i = 1, \ldots, n. \tag{15}$$

Kaczmarz relaxation. Let $x^0 \in I\!R^n$; for $k = 0, 1, \ldots$ do

$$\begin{cases} x^{k,0} = & x^k \\ x^{k,i} = & x^{k,i-1} - \frac{\langle x^{k,i-1}, B_i \rangle - c_i}{\langle B_i, B_i \rangle} B_i, \; i = 1, \ldots, n \\ x^{k+1} = & x^{k,n} \end{cases} \tag{16}$$

The following result is proved in [2].

Theorem 1 *The Gauss-Seidel relaxation* (14) *for the system* (7) *satisfies* (13) *with* $\alpha = \gamma_0$ *given by*

$$\gamma_0 = \frac{1}{(1 + \gamma_-(B))(1 + \gamma_+(B))} \tag{17}$$

and

$$\gamma_-(B) = \max_{1 \leq i \leq n} \sum_{j<i} \frac{|B_{ij}|}{B_{ii}}; \quad \gamma_+(B) = \max_{1 \leq i \leq n} \sum_{j>i} \frac{|B_{ij}|}{B_{ii}}. \tag{18}$$

In [5] we extended the above theorem for SOR relaxation (15) in the following form.

Theorem 2 *The SOR relaxation* (15) *for the system* (7) *satisfies* (13) *with* $\alpha = \delta_0$ *given by*

$$\delta_0 = \frac{\omega(2-\omega)}{(1 + \delta_-(B))(1 + \delta_+(B))} \tag{19}$$

and

$$\delta_-(B) = \max_{1 \leq i \leq n} \sum_{j<i} \frac{|B_{ij}|}{\sqrt{B_{ii} B_{jj}}}; \quad \delta_+(B) = \max_{1 \leq i \leq n} \sum_{j>i} \frac{|B_{ij}|}{\sqrt{B_{ii} B_{jj}}}. \tag{20}$$

Remark 1 *If the matrix* B *satisfies* $B_{ii} = B_{jj}, \forall i, j = 1, \ldots, n$ *(as happens in some finite differences approximations of boundary value problems), then the constants* $\delta_-(B), \delta_+(B)$ *from* (20) *are equal with* $\gamma_-(B), \gamma_+(B)$ *from* (18)*, respectively. Thus, in this case theorem 2 is an extension of theorem 1 for SOR relaxation.*

In order to derive an SP for Kaczmarz relaxation (16), we shall consider a general invertible matrix A (not neccessary SPD), $b \in I\!R^n$ a given vector and the linear system

$$Ax^* = b, \tag{21}$$

with x^* the unique solution. For a given approximation x of x^* from (21), let $\bar{x}$ be the approximation of x^* obtained after a Kaczmarz step (16) applied to x and $e = x - x^*$, $\bar{e} = \bar{x} - x^*$ and $r = Ae$ the corresponding errors and residual (see (12)).

Theorem 3 *With respect to the above definitons and notations, Kaczmarz relaxation* (16) *for the system* (21) *satisfies the following smoothing property (of the type* (13)*)*

$$\| \bar{e} \|^2 \leq \| e \|^2 - \tilde{\gamma}_0 \| \tilde{D}^{\frac{1}{2}} r \|^2 \tag{22}$$

where

$$\tilde{D} = \mathrm{diag}(\frac{1}{\| A_1 \|^2}, \dots, \frac{1}{\| A_n \|^2}), \tag{23}$$

$$\tilde{\gamma}_0 = \frac{1}{(1 + \tilde{\gamma}_-(A))(1 + \tilde{\gamma}_+(A))} \tag{24}$$

and

$$\tilde{\gamma}_-(A) = \max_{1 \leq i \leq n} \sum_{j<i} \frac{|\langle A_i, A_j \rangle|}{\| A_i \|^2}, \quad \tilde{\gamma}_+(A) = \max_{1 \leq i \leq n} \sum_{j>i} \frac{|\langle A_i, A_j \rangle|}{\| A_i \|^2}. \tag{25}$$

Proof. Let $g_i = (0, \dots, 1, \dots, 0)^T \in I\!R^n, i = 1, \dots, n$ be the cannonical basis. We first observe that, because of the symmetry of B, $Be_i = B_i$ then (14) can be written as follows: $x^0 \in I\!R^n$ given; for $k = 0, 1, \dots$ do

$$\begin{cases} x^{k,0} = & x^k \\ x^{k,i} = & x^{k,i-1} - \frac{\langle Bx^{k,i-1} - c, g_i \rangle}{\langle Bg_i, g_i \rangle} g_i, \ i = 1, \dots, n \\ x^{k+1} = & x^{k,n} \end{cases} \tag{26}$$

Let's now write the system (21) in the form (A^T is the transpose of A)

$$(AA^T)y^* = b, \quad x^* = A^T y^*. \tag{27}$$

We apply the Gauss-Seidel relaxation (26) to (27), for $B = AA^T$, $c = b$ and the initial approximation y^0 and get

$$\begin{cases} y^{k,0} = & y^k \\ y^{k,i} = & y^{k,i-1} - \frac{\langle A^T y^{k,i-1}, A_i \rangle - b_i}{\| A_i \|^2} e_i, \ i = 1, \dots, n \\ y^{k+1} = & y^{k,n} \end{cases} \tag{28}$$

If we multiply from the left in (28) with A^T and replace the terms of the form $A^T y^{k,i}$ with $x^{k,i}$ we obtain exactly the Kaczmarz step (16) applied to the system (21). Thus, the Kaczmarz step (16) applied to the system (21) with an initial approximation of the form $x^0 = A^T y^0$, for some $y^0 \in I\!R^n$ is equivalent with the Gauss-Seidel step (26) applied to the system (27), with the initial approximation y^0 and setting $x^{k+1} = A^T y^{k+1}$. But, because the matrix AA^T is SPD, we can apply theorem 1 for the Gauss-Seidel iteration and get (see (13) and (10))

$$\langle (AA^T)\bar{f}, \bar{f} \rangle \leq \langle (AA^T)f, f \rangle - \tilde{\gamma}_0 \langle D^{-1}_{AA^T} AA^T f, AA^T f \rangle \tag{29}$$

with $\tilde{\gamma}_0$ computed as in (17)-(18) with AA^T instead of B and with $f, \bar{f}$ the corresponding errors with respect to the system (27). But, if $e, \bar{e}$ and r are the corresponding errors and residual, respectively for the Kaczmarz relaxation (as defined before), because of the considerations that we've made at the begining of the proof, we have

$$e = A^T f, \quad \bar{e} = A^T \bar{f}. \tag{30}$$

Then, by replacing (30) in (29) we get (22) with the elements from (23)-(25) and the proof is complete.

Remark 2 *A shorter proof of theorem 3 is given in* [2]. *We presented a detailed one here for the reasons of the results described in section* 2 *of the paper (see also remark* 3 *from below).*

A similar result as in theorem 3 can be proved using theorem 2, for the Kaczmarz iteration with relaxation parameter (for short, ω-Kaczmarz relaxation). This will be presented below.
ω-Kaczmarz relaxation. Let $x^0 \in I\!R^n$; for $k = 0, 1, \ldots$ do

$$\begin{cases} x^{k,0} = & x^k \\ x^{k,i} = & x^{k,i-1} - \omega \frac{\langle x^{k,i-1}, B_i \rangle - c_i}{\langle B_i, B_i \rangle} B_i, i = 1, \ldots, n \\ x^{k+1} = & x^{k,n} \end{cases} \tag{31}$$

Theorem 4 *With respect to the above definitons and notations, the ω-Kaczmarz relaxation* (31) *for the system* (21) *satisfies the following smoothing property (of the type* (13)*)*

$$\| \bar{e} \|^2 \leq \| e \|^2 - \tilde{\delta}_0 \| \tilde{D}^{\frac{1}{2}} r \|^2, \tag{32}$$

with $\tilde{D}$ *from* (23) *and*

$$\tilde{\delta}_0 = \frac{1}{(1 + \tilde{\delta}_-(A))(1 + \tilde{\delta}_+(A))}, \tag{33}$$

$$\tilde{\delta}_-(A) = \max_{1 \leq i \leq n} \sum_{j<i} \frac{|\langle A_i, A_j \rangle|}{\| A_i \| \| A_j \|}, \quad \tilde{\delta}_+(A) = \max_{1 \leq i \leq n} \sum_{j>i} \frac{|\langle A_i, A_j \rangle|}{\| A_i \| \| A_j \|}. \tag{34}$$

Proof. As in the proof of theorem 3 we show the above mentioned equivalence between the ω-Kaczmarz and the SOR relaxation (15) written in the form (26) (see also [4]).

Remark 3 *We have to observe that in* (22) *and* (32)*, for $\bar{e}$ and e we have the Euclidean norm, instead of the energy one from* (13). *Moreover, the proofs of the theorems* 3 *and* 4 *are based on theorems* 1 *and* 2*, i.e. the matrix $B = AA^T$ must be SPD; thus, the above theorems* 3 *and* 4 *cannot be extended (at least following the same way of proof) to arbitrary noninvertible systems like* (21)*, because in such a case, the matrix AA^T is no longer SPD. In the next section of the paper we shall prove such extensions, but using a direct proof (independent on the smoothing property of Gauss-Seidel or SOR relaxations for SPD matrices).*

3 Smoothing property of Kaczmarz relaxation for arbitrary consistent systems

Let A be an $m \times n$ matrix with $A_i \neq 0, \forall i = 1, \ldots m$ and $b \in I\!R^m$ such that the system

$$Ax = b \tag{35}$$

is consistent. We shall denote by $S(A; b), N(A), R(A)$ the solutions set for (35), null space and range of A, respectively. For a given vector subspace $E \subset I\!R^q$, $P_E(x)$ will be the orthogonal projection onto E of an element $x \in I\!R^q$ and $E^\perp$ will denote its orthogonal complement. If x_{LS} is the (unique) minimal norm solution of (35) it is well known that (see e.g. [1], [3])

$$x_{LS} \in N(A)^\perp = R(A^t), \quad S(A; b) = x_{LS} + N(A). \tag{36}$$

Thus, for a vector $z \in I\!R^n$ we shall denote by $s(z)$ the solution vector (see (36))

$$s(z) = P_{N(A)}(z) + x_{LS} \in S(A; b). \tag{37}$$

The Kaczmarz relaxation for (35) can be written as (see (16)): let $x^0 \in I\!R^n$ be given; for $k = 0, 1, 2 \ldots$ do

$$\begin{cases} x^{k,0} = & x^k \\ x^{k,i} = & x^{k,i-1} - \frac{\langle x^{k,i-1}, A_i \rangle - b_i}{\|A_i\|^2} A_i, i = 1, \ldots, m \\ x^{k+1} = & x^{k,m} \end{cases} \tag{38}$$

The following results are known (see for the proof [11], [3]).

Theorem 5 *For any $x^0 \in I\!R^n$ the sequence $(x^k)_{k \geq 0}$ generated by the algorithm* (38) *has the properties*

$$P_{N(A)}(x^k) = P_{N(A)}(x^0), \quad \forall k \geq 0 \tag{39}$$

and

$$\lim_{k \to \infty} x^k = P_{N(A)}(x^0) + x_{LS} = s(x^0). \tag{40}$$

Let now $x \in I\!R^n$ be a current approximation of $s(x^0)$ (generated by (38), for $k = 0, x = x^0$) and $\bar{x}$ the approximation after one step of (38) applied to x. Let $e, \bar{e}, r$ be the corresponding errors and residual, defined by (according to (37) and (39))

$$e = e - s(x^0), \quad \bar{e} = \bar{x} - s(x^0), \quad r = Ae = Ax_{LS} = b. \tag{41}$$

We are now in the position of proving the main result of this section.

Theorem 6 *With respect to the above definitons and notations, Kaczmarz relaxation* (38) *for the system* (35) *satisfies the smoothing property* (22) – (25) *from theorem* 3.

Proof. Step 1. According to (38), the computation of $\bar{x}$ from x can be written as follows

$$\begin{cases} x^0 = & x \\ x^i = & x^{i-1} - \frac{\langle x^{i-1}, A_i \rangle - b_i}{\| A_i \|^2} A_i, i = 1, \dots, m \\ \bar{x} = & x^m \end{cases} \tag{42}$$

Let $\{f_1, f_2, \dots, f_m\}$ be the cannonical basis in $I\!R^m$ and e^i, r^i the errors and residuals defined by

$$e^i = x^i - s(x^0), \quad r^i = Ae^i, \ i = 1, \dots, m. \tag{43}$$

Then, $\forall i = 1, \dots, m$ we obtain (by also using the equality $A_i = A^T f_i$)

$$e^i = x^i - s(x^0) = x^{i-1} - \frac{\langle x^{i-1}, A_i \rangle - b_i}{\| A_i \|^2} A_i - s(x^0) =$$

$$e^{i-1} - \frac{\langle x^{i-1}, A^T f_i \rangle - \langle b, f_i \rangle}{\| A_i \|^2} A_i = e^{i-1} - \frac{\langle r^{i-1}, f_i \rangle}{\| A_i \|^2} A_i \tag{44}$$

and

$$r^i = Ae^i = r^{i-1} - \frac{\langle r^{i-1}, f_i \rangle}{\| A_i \|^2} AA_i. \tag{45}$$

From (44), by taking the Euclidean norm and using again the equality $A_i = A^T f_i$ and (43) we obtain

$$\| e^i \|^2 = \| e^{i-1} \|^2 - 2 \frac{\langle r^{i-1}, f_i \rangle \langle e^{i-1}, A_i \rangle}{\| A_i \|^2} + \frac{\langle r^{i-1}, f_i \rangle^2}{\| A_i \|^2} =$$

$$\| e^{i-1} \|^2 - \frac{\langle r^{i-1}, f_i \rangle^2}{\| A_i \|^2} = \| e^{i-1} \|^2 - \frac{(r_i^{i-1})^2}{\| A_i \|^2}, \ \forall i = 1, \dots, m, \tag{46}$$

where r_i^{i-1} is the i-th component of the residual vector r^{i-1}. But, because of (42) we have $x^0 = x, \bar{x} = x^m$, then $e^0 = e, \bar{e} = e^m$, thus by summing up in (46) following $i = 1, \dots, m$ we get

$$\| \bar{e} \|^2 = \| e \|^2 - \sum_{i=1}^m \frac{(r_i^{i-1})^2}{\| A_i \|^2}. \tag{47}$$

Step 2. Let now $r^* = (r_1^*, \dots, r_m^*)^T \in I\!R^m$ be the *dynamic residual* (as called in [2]), defined by

$$r_i^* = r_i^{i-1}, \ \forall i = 1, \dots, m, \tag{48}$$

i.e. r_i^* is the i-th component of the r^{i-1} residual (that is, the residual before the projection on the i-th equation of (42)). From (47) and (48) we then get

$$\| \bar{e} \|^2 = \| e \|^2 - \sum_{i=1}^m \frac{(r_i^*)^2}{\| A_i \|^2} = \| e \|^2 - \| \tilde{D}^{\frac{1}{2}} r^* \|^2, \tag{49}$$

with $\tilde{D}$ from (23). On the other hand, from the equation (45) we successively obtain

$$r^i = r^{i-1} - \frac{\langle r^{i-1}, f_i \rangle}{\parallel A_i \parallel^2} AA^T f_i = r^{i-1} - (\hat{A}\hat{A}^T) f_i \langle r^{i-1}, f_i \rangle =$$

$$r^{i-1} - (\hat{A}\hat{A}^T) f_i f_i^T r^{i-1} = r^{i-1} - BE_i r^{i-1}, \tag{50}$$

where

$$\hat{A} = \tilde{D}^{\frac{1}{2}} A, \quad B = \hat{A}\hat{A}^T, \quad E_i = f_i f_i^T. \tag{51}$$

Thus

$$\begin{cases} r^1 = & r^0 - BE_1 r^0 & = r - BE_1 r \\ r^2 = & r^1 - BE_2 r^1 & = r - (BE_1 r + BE_2 r^1) \\ \dots & \dots & \dots \\ r^{m-1} = & r^{m-2} - BE_{m-1} r^{m-2} & = r - (BE_1 r + BE_2 r^1 + \dots + BE_{m-1} r^{m-2}) \end{cases} \tag{52}$$

Then, from the definition of the matrices E_i in (51) and (48) we get $r_1^* = r_1^0 = r_1$ and for $i = 2, \dots, m$

$$r_i^* = r_i - (B_{i1} r_1^0 + B_{i2} r_2^1 + \dots + B_{i,i-1} r_i^{i-1}) \tag{53}$$

or in matrix form

$$r^* = \begin{bmatrix} r_1^0 \\ r_2^1 \\ \cdot \\ \cdot \\ \cdot \\ r_m^{m-1} \end{bmatrix} = r - \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ -B_{21} & 0 & 0 & \dots & 0 & 0 \\ -B_{31} & -B_{32} & 0 & \dots & 0 & 0 \\ .. & .. & .. & .. & .. & .. \\ -B_{m,1} & -B_{m,2} & -B_{m,3} & \dots & -B_{m,m-1} & 0 \end{bmatrix} \begin{bmatrix} r_1^0 \\ r_2^1 \\ \cdot \\ \cdot \\ \cdot \\ r_m^{m-1} \end{bmatrix} =$$

$$r - Lr^*, \tag{54}$$

where we denoted by L the corresponding strictly lower triangular matrix. Thus, from (54) it results (see also (49))

$$\tilde{D}^{\frac{1}{2}} r = \left(I + \tilde{D}^{\frac{1}{2}} L \tilde{D}^{-\frac{1}{2}} \right) \left(\tilde{D}^{\frac{1}{2}} r^* \right). \tag{55}$$

Step 3. If C is a square matrix and $\parallel C \parallel_2, \parallel C \parallel_\infty$ are, respectively its spectral and infinity norm (see e.g. [1]) we have

$$\parallel C \parallel_2^2 = \rho(C^T C) \leq \parallel C^T C \parallel_\infty \leq \parallel C^T \parallel_\infty \parallel C \parallel_\infty . \tag{56}$$

Using (56) and taking the Euclidean norm in (55) we obtain

$$\parallel \tilde{D}^{\frac{1}{2}} r \parallel^2 \leq \parallel I + \tilde{D}^{\frac{1}{2}} L \tilde{D}^{-\frac{1}{2}} \parallel_\infty \parallel I + \tilde{D}^{-\frac{1}{2}} L^T \tilde{D}^{\frac{1}{2}} \parallel_\infty \parallel \tilde{D}^{\frac{1}{2}} r^* \parallel^2 =$$

$$\left(1 + \parallel \tilde{D}^{\frac{1}{2}} L \tilde{D}^{-\frac{1}{2}} \parallel_\infty \right) \left(1 + \parallel \tilde{D}^{-\frac{1}{2}} L^T \tilde{D}^{\frac{1}{2}} \parallel_\infty \right) \parallel \tilde{D}^{\frac{1}{2}} r^* \parallel^2 . \tag{57}$$

But, a simple computation gives us

$$1 + \parallel \tilde{D}^{\frac{1}{2}} L \tilde{D}^{-\frac{1}{2}} \parallel_\infty = \tilde{\gamma}_-(A), \quad 1 + \parallel \tilde{D}^{-\frac{1}{2}} L^T \tilde{D}^{\frac{1}{2}} \parallel_\infty = \tilde{\gamma}_+(A), \tag{58}$$

with $\tilde{\gamma}_-(A), \tilde{\gamma}_+(A)$ from (25). Then, using (49), (57) and (58) we get (22) and the proof is complete. For the ω-Kaczmarz relaxation a similar result can be proved.

Theorem 7 *With respect to the above definitons and notations, the ω-Kaczmarz relaxation for the system* (35) *satisfies the smoothing property* (22) $-$ (25) *from theorem* 3 *with the same constants $\tilde{\gamma}_-(A)$ and $\tilde{\gamma}_+(A)$ from* (25)*, but, instead of $\tilde{\gamma}_0$ from* (24)*, with $\tilde{\delta}_0$ given by*

$$\tilde{\delta}_0 = \frac{\omega(2-\omega)}{(1+\tilde{\gamma}_-(A))(1+\tilde{\gamma}_+(A))}. \tag{59}$$

Proof. The ω-Kaczmarz relaxation (31) for the system (35) can be written as

$$\begin{cases} x^0 = & x \\ x^i = & x^{i-1} - \omega \frac{\langle x^{i-1}, A_i \rangle - b_i}{\|A_i\|^2} A_i, i = 1, \dots, m \\ \bar{x} = & x^m \end{cases} \tag{60}$$

As in the proof of theorem 6 we get

$$e^i = e^{i-1} - \omega(2-\omega) \frac{\langle r^{i-1}, f_i \rangle}{\| A_i \|^2} A_i \tag{61}$$

and

$$r^i = Ae^i = r^{i-1} - \omega(2-\omega) \frac{\langle r^{i-1}, f_i \rangle}{\| A_i \|^2} AA_i. \tag{62}$$

Then we follow exactly the same way as in the above mentioned proof.

4 Smoothing property of Kaczmarz Extended relaxation for arbitrary inconsistent systems

Let A and b be as in section 2. Instead of the consistent system (35) we shall consider in this section the linear least squares formulation (inconsistent)

$$\| Ax - b \| = \min! \tag{63}$$

for which we shall denote by $LSS(A;b)$, x_{LS} the set of its solutions and the minimal norm one, respectively. Let b_A, b_A^* be defined by

$$b_A = P_{R(A)}(b), \quad b_A^* = P_{N(A^T)}(b). \tag{64}$$

Then (see e.g. [1])

$$b = b_A + b_A^*, \; LSS(A;b) = S(A;b_A) = x_{LS} + N(A). \tag{65}$$

Moreover, (see (37)) for any $z \in I\!R^n$ we have

$$s(z) = P_{N(A)}(z) + x_{LS} \in LSS(A;b) = S(A;b_A), \tag{66}$$

thus we define (as in section 2) the error and residual by

$$e = e(z) = z - s(z), \quad r = Ae = Az - b_A. \tag{67}$$

The Kaczmarz Extended algorithm for (63) (KE, for short), introduced in [6] (see also [7]) is the following.
KE relaxation. Let $x^0 \in R^n, y^0 = b$; for $k = 0, 1, \dots$ do

$$\begin{cases} y^{k+1} = \Phi(y^k) = & (\phi_1 \cdot \dots \cdot \phi_n)(y^k), \\ b^{k+1} = & b - y^{k+1}, \\ x^{k+1} = & \text{Kaczmarz}(b^{k+1}; x^k), \end{cases} \tag{68}$$

where

$$\phi_j(y) = y - \frac{\langle y, A^j \rangle}{\| A^j \|^2} A^j$$

and $A^j \neq 0, j = 1, \dots, n$ are the columns of A. In [8] we proved that the sequence $(x^k)_{k \geq 0}$ generated with the above KE algorithm satisfies the relation (39). Then, if $x = x^k$ (for some $k \geq 0$) is a current approximation, we shall define the error (see (67)) by

$$e = x - s(x^0). \tag{69}$$

Let $y = y^k$ be the corresponding element from the first step of (68) and $\bar{y} = y^{k+1}$, i.e.

$$\bar{y} = \Phi(y) = P_{N(A^T)}(y) + \tilde{\Phi}(y), \tag{70}$$

where (see for details [7])

$$\tilde{\Phi}(y) = \Phi P_{R(A)}(y) \in R(A). \tag{71}$$

From (70) and (71) we get that

$$P_{N(A^T)}(\bar{y}) = P_{N(A^T)}(y), \tag{72}$$

thus, recursively (see also (65))

$$P_{N(A^T)}(\bar{y}) = P_{N(A^T)}(b) = b^*_A. \tag{73}$$

Then, if $\bar{b} = b^{k+1}$ (see (68)), we get from (70) and (73)

$$\bar{b} = b - \bar{y} = b_A - \tilde{y}, \tag{74}$$

where

$$\tilde{y} = \tilde{\Phi}(y). \tag{75}$$

Remark 4 *Coming back at the original notation, from the above equalities we obtain*

$$\tilde{y} = \tilde{\Phi}(y) = \tilde{\Phi}^k(b). \tag{76}$$

Moreover, in [11] *it is proved that* $\| \tilde{\Phi} \|_2 < 1$, *thus*

$$\lim_{k \to \infty} \tilde{\Phi}^k(b) = 0. \tag{77}$$

Now we are able to prove the main result of this section.

Theorem 8 *With respect to the above definitions and notations, the KE relaxation* (68) *for the (inconsistent) problem* (63) *satisfies the following smoothing property*

$$\| \bar{e} \|^2 \leq \| e \|^2 - \frac{\tilde{\gamma}_0}{2} \| \tilde{D}^{\frac{1}{2}} r \|^2 + 2 \| \tilde{D}^{\frac{1}{2}} \tilde{y} \|^2 \tag{78}$$

with $\tilde{D}$ *and* $\tilde{\gamma}_0$ *from* (23) – (25).

Proof. Step 1. The third step in (68) can be written as (see also (16))

$$\begin{cases} x^0 = & x \\ x^i = & x^{i-1} - \frac{\langle x^{i-1}, A_i \rangle - \bar{b}_i}{\| A_i \|^2} A_i, i = 1, \dots, m \\ \bar{x} = & x^m \end{cases} \tag{79}$$

Then, by defining the errors (see also (43))

$$e^i = x^i - s(x^0), \; i = 1, \dots, m, \quad \bar{e} = e^m, \tag{80}$$

we get

$$e^i = x^i - s(x^0) = x^{i-1} - \frac{\langle x^{i-1}, A_i \rangle - (b_A)_i + \tilde{y}_i}{\| A_i \|^2} A_i - s(x^0) =$$

$$e^{i-1} - \frac{\langle A x^{i-1} - b_A, f_i \rangle + \tilde{y}_i}{\| A_i \|^2} A_i = e^{i-1} - \frac{r_i^{i-1}}{\| A_i \|^2} A_i - \frac{\tilde{y}_i}{\| A_i \|^2} A_i. \tag{81}$$

From (81) we obtain (by also using the relation $A_i = A^T f_i$)

$$\| e^i \|^2 = \| e^{i-1} - \frac{r_i^{i-1}}{\| A_i \|^2} A_i \|^2 + \frac{|\tilde{y}_i|^2}{\| A_i \|^2} A_i - 2 \langle e^{i-1} - \frac{r_i^{i-1}}{\| A_i \|^2} A_i, \frac{\tilde{y}_i}{\| A_i \|^2} A_i \rangle =$$

$$\| e^{i-1} \|^2 - 2 \langle e^{i-1}, \frac{r_i^{i-1}}{\| A_i \|^2} A_i \rangle + \frac{(r_i^{i-1})^2}{\| A_i \|^2} + \frac{(\tilde{y}_i)^2}{\| A_i \|^2} - 2 \langle e^{i-1}, \frac{\tilde{y}_i}{\| A_i \|^2} A_i \rangle +$$

$$2 \langle \frac{r_i^{i-1}}{\| A_i \|^2} A_i, \frac{\tilde{y}_i}{\| A_i \|^2} A_i \rangle = \| e^{i-1} \|^2 - \frac{(r_i^{i-1})^2}{\| A_i \|^2} + \frac{(\tilde{y}_i)^2}{\| A_i \|^2}, \forall i = 1, \dots, m. \tag{82}$$

Then by adding the above relations (82), using (80) and the notations from section 2 we get

$$\parallel \bar{e} \parallel^2 = \parallel e \parallel^2 - \parallel \tilde{D}^{\frac{1}{2}} r^* \parallel^2 + \parallel \tilde{D}^{\frac{1}{2}} \tilde{y} \parallel^2 . \tag{83}$$

Step 2. From (81) it results (see again the notations in section 2)

$$r^i = Ae^i = Ae^{i-1} - \frac{r_i^{i-1}}{\parallel A_i \parallel^2} AA_i - \frac{\tilde{y}_i}{\parallel A_i \parallel^2} AA_i =$$

$$r^{i-1} - BE_i r^{i-1} - BE_i \tilde{y}, i = 1, \ldots, m, \tag{84}$$

thus

$$\begin{cases} r^1 = & r - BE_1 r - BE_1 \tilde{y} \\ r^2 = & r - (BE_1 r + BE_2 r^1) - (BE_1 + BE_2)\tilde{y} \\ \ldots & \ldots \\ r^{m-1} = & r - (BE_1 r + BE_2 r^1 + \ldots + BE_{m-1} r^{m-2}) - \\ & (BE_1 + BE_2 + \ldots + BE_{m-1})\tilde{y} \end{cases}$$

and by taking into account that (see section 2)

$$r^* = (r_1, r_2^1, \ldots, r_m^{m-1})^T$$

it results

$$\begin{cases} r_1^* = & r_1 \\ r_2^* = & r_2 - B_{21} r_1^* - B_{21} \tilde{y}_1 \\ r_3^* = & r_3 - (B_{31} r_1^* + B_{32} r_2^*) - (B_{31} \tilde{y}_1 + B_{32} \tilde{y}_2) \\ \ldots & \ldots \\ r_m^* = & r_m - (B_{m1} r_1^* + B_{m2} r_2^* + \ldots + B_{m,m-1} r_{m-1}^*) - \\ & (B_{m1} \tilde{y}_1 + B_{m2} \tilde{y}_2 + \ldots + B_{m,m-1} \tilde{y}_{m-1}) \end{cases} \tag{85}$$

Writing (85) in matrix form

$$\begin{bmatrix} r_1^* \\ r_2^* \\ . \\ . \\ . \\ r_m^* \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ . \\ . \\ . \\ r_m \end{bmatrix} - L \begin{bmatrix} r_1^* \\ r_2^* \\ . \\ . \\ . \\ r_m^* \end{bmatrix} - L \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ . \\ . \\ . \\ \tilde{y}_m \end{bmatrix}$$

with L from (54), we get

$$r^* = r - Lr^* - L\tilde{y}. \tag{86}$$

From (86) we obtain as in section 2

$$(I + L)r^* + L\tilde{y} = r \Leftrightarrow (I + \tilde{L})(\tilde{D}^{\frac{1}{2}} r^*) + \tilde{L}(\tilde{D}^{\frac{1}{2}} \tilde{y}) = \tilde{D}^{\frac{1}{2}} r, \tag{87}$$

where

$$\tilde{L} = \tilde{D}^{\frac{1}{2}} L \tilde{D}^{-\frac{1}{2}}. \tag{88}$$

Using (87) and (88) it follows that

$$\parallel \tilde{D}^{\frac{1}{2}} r \parallel^2 \leq \parallel (I + \tilde{L})(\tilde{D}^{\frac{1}{2}} r^*) + \tilde{L}(\tilde{D}^{\frac{1}{2}} \tilde{y}) \parallel^2 \leq$$

$$2 \left[\parallel I + \tilde{L} \parallel_2^2 \parallel \tilde{D}^{\frac{1}{2}} r^* \parallel^2 + \parallel \tilde{L} \parallel_2^2 \parallel \tilde{D}^{\frac{1}{2}} \tilde{y} \parallel^2 \right] \leq$$

$$2 \left[(1 + \parallel \tilde{L} \parallel_\infty)(1 + \parallel \tilde{L}^T \parallel_\infty) \parallel \tilde{D}^{\frac{1}{2}} r^* \parallel + \parallel \tilde{L} \parallel_\infty \parallel \tilde{L}^T \parallel_\infty \parallel \tilde{D}^{\frac{1}{2}} \tilde{y} \parallel^2 \right] \leq$$

$$2 \left[(1 + \parallel \tilde{L} \parallel_\infty)(1 + \parallel \tilde{L}^T \parallel_\infty) \right] \left(\parallel \tilde{D}^{\frac{1}{2}} r^* \parallel^2 + \parallel \tilde{D}^{\frac{1}{2}} \tilde{y} \parallel^2 \right). \tag{89}$$

From (89) and (24) we then obtain

$$\parallel \tilde{D}^{\frac{1}{2}} r^* \parallel^2 \geq \frac{\tilde{\gamma}_0}{2} \parallel \tilde{D}^{\frac{1}{2}} r \parallel^2 - \parallel \tilde{D}^{\frac{1}{2}} \tilde{y} \parallel^2, \tag{90}$$

which together with (83) gives us (78) and the proof is complete.
A similar result can be derived for the Kaczmarz Extended algorithm with Relaxation Parameters (KERP, for short), introduced in [7].
KERP relaxation. Let $x^0 \in R^n, y^0 = b$; for $k = 0, 1, \dots$ do

$$\begin{cases} y^{k+1} = \Phi(\alpha; y^k) = & (\phi_1 \cdot \ldots \cdot \phi_n)(\alpha; y^k), \\ b^{k+1} = & b - y^{k+1}, \\ x^{k+1} = & \omega - \text{Kaczmarz}(b^{k+1}; x^k), \end{cases} \tag{91}$$

where

$$\phi_j(\alpha; y) = y - \alpha \frac{\langle y, A^j \rangle}{\| A^j \|^2} A^j.$$

Theorem 9 *With respect to the above definitions and notations, the KERP relaxation* (91) *for the (inconsistent) problem* (63) *satisfies the following smoothing property*

$$\| \bar{e} \|^2 \leq \| e \|^2 - \frac{\tilde{\delta}_0}{2} \| \tilde{D}^{\frac{1}{2}} r \|^2 + 2 \| \tilde{D}^{\frac{1}{2}} \tilde{y}(\alpha) \|^2 \tag{92}$$

with $\tilde{D}$ *and* $\tilde{\delta}_0$ *from* (23), (59) *and* (25), *and* $\tilde{y}(\alpha) = \Phi(\alpha; y)$.

Proof. We apply the above theorem 8 and theorem 4 for the ω-Kaczmarz relaxation.

Remark 5 *The result from theorem* 6 *is not a particular case of that one from theorem* 8. *Indeed, if* $b \in R(A)$ *for the problem* (63) *we have* $b_A^* = 0$, *thus* $\bar{y} = \tilde{y} \in R(A)$, *but we cannot drop out the term* $2 \| \tilde{D}^{\frac{1}{2}} \tilde{y} \|^2$ *in* (78). *Thus, in the consistent case theorems* 6 *and* 8 *provide two smoothing properties for two different algorithms: K and KE.*

5 Some considerations about the correction step

We shall refer to the general least squares problem (63). If A^+ is the Moore-Penrose pseudoinverse of A we replay the following well known properties (see e.g. [1])

$$x_{LS} = A^+ b = A^+ b_A, \quad N(A^+) = N(A^T), \quad R(A^+) = R(A^T). \tag{93}$$

Let $p < n$ be a fixed integer and the matrices A_p $(m \times p)$ and I_p^n $(n \times p)$ given such that I_p^n has full column rank. We consider the following correction process (see (93)).

$$\begin{cases} d = b - Ax \\ \| A_p v_p - d \| = \min! \Rightarrow \quad v_p = A_p^+ d = A_p^+ \left(P_{R(A_p)}(d) \right) \\ \bar{x} = x + I_p^n v_p \end{cases} \tag{94}$$

Remark 6 *The vector* v_p *from* (94) *satisfies the relation (see also* (66)*)*

$$A_p v_p = P_{R(A_p)}(d). \tag{95}$$

Assumption 1. *The matrices* A, A_p *and* I_p^n *satisfy the equality*

$$A_p = A I_p^n. \tag{96}$$

Based on this assumption we shall derive in what follows some properties of the correction process (94).

Proposition 1 *If* $x \in LSS(A; b)$, *then* $\bar{x} = x$.

Proof. From our hypothesis about x and by using (64) - (66), (93) and (96) it results

$$d = b - Ax = b - b_A = b_A^* \in N(A^T) \tag{97}$$

and

$$N(A_p^+) = N(A_p^T) = N((I_p^n)^T A^T) \supset N(A^T). \tag{98}$$

From (97) and (98) we obtain that $d = b_A^* \in N(A_p^+)$, thus

$$v_p = A_p^+ d = 0,$$

i.e. $\bar{x} = x$ and the proof is complete.

Proposition 2 *The correction process* (94) *is idempotent.*

Proof. Let $\bar{x}'$ the vector obtained after we aplly one more the correction step (94) to $\bar{x}$. We have (by also using (95) and (96))

$$\bar{d} = b - A\bar{x} = b - Ax - AI_p^n v_p =$$

$$d - A_p v_p = d - P_{R(A_p)}(d) = P_{N(A_p^T)}(d), \tag{99}$$

thus, following (93)

$$\bar{v}_p = A_p^+ \bar{d} = A_p^+ P_{N(A_p^T)}(d) = 0,$$

i.e. $\bar{x}' = \bar{x} + I_p^n \bar{v}_p = \bar{x}$ and the proof is complete.

Proposition 3 *Let $r, \bar{r}$ be the residuals before and after the correction step* (94)*, i.e. (see* (67)*)*

$$r = Ae = A(x - s(x)) = Ax - b_A, \quad \bar{r} = A\bar{e} = A(\bar{x} - s(\bar{x})) = A\bar{x} - b_A. \tag{100}$$

Then

$$A_p^T \bar{r} = 0, \tag{101}$$

$$\parallel \bar{r} \parallel \leq \parallel r \parallel . \tag{102}$$

Proof. By using (100), (94), (95) and (96) we successively obtain

$$\bar{r} = A\bar{x} - b_A = Ax + AI_p^n v_p - b_A = r + A_p v_p =$$

$$r + P_{R(A_p)}(d) = r + P_{R(A_p)}(-r + b_A^*) =$$

$$r - P_{R(A_p)}(r) + P_{R(A_p)}(b_A^*) = P_{N(A_p^T)}(r) + P_{R(A_p)}(b_A^*). \tag{103}$$

But, $b_A^* \in N(A_p^T) = R(A_p)^\perp$, thus $P_{R(A_p)}(b_A^*) = 0$, then from (103) it results

$$\bar{r} = P_{N(A_p^T)}(r) \Rightarrow A_p^T \bar{r} = 0.$$

For the inequality (102) we have

$$\parallel \bar{r} \parallel^2 = \parallel P_{N(A_p^T)}(r) \parallel^2 \leq \parallel P_{N(A_p^T)}(r) \parallel^2 + \parallel P_{R(A_p)}(r) \parallel^2 =$$

$$\parallel P_{N(A_p^T)}(r) + P_{R(A_p)}(r) \parallel^2 = \parallel r \parallel^2 .$$

Assumption 2. *The interpolation operator I_p^n and the coarse grid correction v_p satisfy*

$$(I - A^+ A) I_p^n v_p = 0, \tag{104}$$

where I is the $n \times n$ unit matrix.

Proposition 4 *Under the assumption* (104) *we have*

$$P_{N(A)}(\bar{x}) = P_{N(A)}(x) \tag{105}$$

and

$$\bar{e} = e + I_p^n v_p. \tag{106}$$

Proof. According to [1] we have

$$P_{N(A)} = I - A^+A, \tag{107}$$

thus, from (94) and (104) we get

$$P_{N(A)}(\bar{x}) = (I - A^+A)x + (I - A^+A)I_p^n v_p = (I - A^+A)x = P_{N(A)}(x)$$

and the proof is complete. For the second equality (106) we have to observe that, according to (66) and (105) $s(\bar{x}) = s(x)$, thus

$$\bar{e} = \bar{x} - s(\bar{x}) = \bar{x} - s(x) = x + I_p^n v_p - s(x) = e + I_p^n v_p.$$

Remark 7 *The relation* (105) *is essential for developing a well defined 2-grid AMG algorithm for* (63)*. Indeed, if it doesn't hold, then the approximation $\bar{x}$ obtained after the correction step* (94) *will have a different component on $N(A)$ (compared with x) and this will destroy the consistency of the 2-grid method with the problem* (63)*. A stronger, but more practical assumption would be (see also section 5.2 from below)*

$$I_p^n(R(A_p^T)) \subset R(A^T). \tag{108}$$

5.1 An example satisfying Assumption 1

We consider the approach in [9]. Let A be the $m \times n$ matrix obtained by scanning the given image (we shall consider only the 2D case here). We shall suppose that

$$n = 4p \tag{109}$$

and let $P_1, \ldots, P_n$ be the pixels on the "fine grid". The "coarse grid" is obtained by considering the "bigger" pixels formed (each) by 4 adiacent ones from the fine grid, $P_1^H, \ldots, P_p^H$. For any $j \in \{1, \ldots, p\}$ we define the set of indices of fine grid pixels that form the coarse grid one P_j^H, i.e.

$$S(j) = \{j_1, j_2, j_3, j_4\} \tag{110}$$

such that

$$P_j^H = P_{j_1} \cup P_{j_2} \cup P_{j_3} \cup P_{j_4}. \tag{111}$$

We may suppose that

$$j_1 < j_2 < j_3 < j_4. \tag{112}$$

We construct the above coarse grid matrix A_p following the formulas

$$(A_p)_{ij} = \sum_{k \in S(j)} A_{ik}, \ \forall i = 1, \ldots, m, \ j = 1, \ldots, p \tag{113}$$

and the $n \times p$ interpolation operator I_p^n by

$$(I_p^n)_{ij} = \begin{cases} 1, & \text{if } i \ \in S(j) \\ 0, & \text{if } i \ \notin S(j), \end{cases} \tag{114}$$

$i = 1, \ldots, n, \ j = 1, \ldots, p.$

Proposition 5 *The above matrices A, A_p and I_p^n satisfy* (96)*. Moreover, the interpolation operator I_p^n has full column rank.*

Proof. We shall prove that $\forall i = 1, \ldots, m, \ \forall j = 1, \ldots, p$

$$(AI_p^n)_{ij} = (A_p)_{ij}. \tag{115}$$

Let $\{g_j\} \subset I\!R^p, \ \{f_i\} \subset I\!R^m$ be the cannonical basis. By the definiton of I_p^n, for a fixed $j \in \{1, \ldots, p\}$ the j-th column of I_p^n has 1's only on the positions $j_1, j_2, j_3, j_4 \in S(j)$ and 0's in rest. Then we have (by also using (113))

$$(AI_p^n)_{ij} = \langle AI_p^n g_j, f_i \rangle = \langle I_p^n g_j, A^T f_i \rangle = \langle \text{column } j \text{ of } I_p^n, \ \text{row } i \text{ of } A \rangle =$$

$$A_{i,j_1} + A_{i,j_2} + A_{i,j_3} + A_{i,j_4} = \sum_{k \in S(j)} A_{ik} = (A_p)_{ij}.$$

Concerning the second assertion, we first observe that from (110)-(111) we obtain

$$S(j) \cap S(j') = \emptyset, \ \forall j \neq j'. \tag{116}$$

From (116) it results that the j-th and j'-th columns of I_p^n are "structurally orthogonal", thus they are orthogonal and the proof is complete.

5.2 An example satisfying Assumption 2

We shall consider the matrix A from section 5.1 before. A sufficent condition for having (108) would be to define the interpolation operator I_p^n as

$$I_p^n = A^T E \tag{117}$$

where the $m \times p$ matrix E can be a "pick-up" one, i.e. only with 1's and 0's as entries. defined in the following way (see (110) - (111))

$$\begin{cases} 1, & \text{if the } i-th \text{ ray intersects at least} \\ & \text{one pixel } P_k \text{ with } k \in S(j) \\ 0, & \text{else.} \end{cases} \tag{118}$$

In this way we can get an (enough) sparse interpolation operator I_p^n. But, in order to keep the Assumption 1, we have to define the coarse grid matrix A_p as in (96), which gives us (see (117))

$$A_p = A I_p^n = A A^T E. \tag{119}$$

Then, beside the fact that its elements have to be computed as scalar products of the form $\langle A_i, A_j \rangle$, the sparsity structure can be different that one for A_p from section 5.1 (see (113)).

References

[1] Björk A., *Numerical methods for least squares problems*, SIAM Philadelphia, 1996.

[2] Brandt A., *Algebraic multigrid theory: The symmetric case*, Appl. Math. Comp., **19**(1986), 23-56.

[3] Censor Y., Stavros A. Z., *Parallel optimization: theory, algorithms and applications*, "Numer. Math. and Sci. Comp." Series, Oxford Univ. Press, New York, 1997.

[4] Natterer F., *The Mathematics of Computerized Tomography*, John Wiley and Sons, New York, 1986.

[5] Popa C., *On smoothing property of the SOR relaxation*, Studii si Cercetari Matematice, **41(5)**(1989), 399-406.

[6] Popa C., *Least-squares solution of overdetermined inconsistent linear systems using Kaczmarz's relaxation*; Intern. J. Comp. Math., **55**(1995), 79-89.

[7] Popa C., *Extensions of block-projections methods with relaxation parameters to inconsistent and rank-defficient least-squares problems*; *B I T*, **38(1)**(1998), 151-176.

[8] Popa C., *Characterization of the solutions set of inconsistent least-squares problems by an extended Kaczmarz algorithm*; Korean Journal on Comp. and Appl. Math., vol. **6(1)**(1999), 51-64.

[9] Prümmer, M. et al., *A full multigrid technique to accelerate an ART scheme for tomographic image reconstruction*, Proceedings of ASIM Conference, September 12-15 2005, Erlangen, Germany, SCS Publishing House e.V., 632-637.

[10] Ruge, J., Stüben, K., *Algebraic multigrid (AMG)*, Arbeitspapiere der GMD **210**(1986), Bonn.

[11] Tanabe K., *Projection Method for Solving a Singular System of Linear Equations and its Applications*, Numer. Math., **17**(1971), 203-214.

[12] Young D. M., *Iterative solution of large linear systems*, Academic Press, New York, 1971.