SIAM J. SCI. COMPUT.
Vol. 29, No. 5, pp. 1982–1991

AN ALGEBRAIC MULTIGRID PRECONDITIONER FOR A CLASS OF SINGULAR M-MATRICES*

ELENA VIRNIK†

Abstract. We apply algebraic multigrid (AMG) as a preconditioner for solving large singular linear systems of the type $(I - T^T)x = 0$ with GMRES. Here, T is assumed to be the transition matrix of a Markov process. Although AMG and GMRES were originally designed for the solution of regular systems, with adequate adaptation their applicability can be extended to problems as described above.

Key words. algebraic multigrid, singular systems, Markov chains

AMS subject classifications. 65C20, 65N55, 65F10

DOI. 10.1137/060659272

1. Introduction.

This paper treats the application of an algebraic multigrid (AMG) method [12] as a preconditioner for solving large singular linear systems of the type

$$Ax := (I - T^T)x = 0 \tag{1}$$

using the generalized minimal residual (GMRES) method [13]. Here, T is the transition matrix of a Markov process. The applicability of GMRES to singular systems was examined, e.g., in [1, 4]. Our AMG strategy is based on the classical theory in [12] originally designed for systems involving regular, symmetric, positive definite M-matrices. The approach in [12] was generalized to symmetric positive definite and semidefinite matrices without the M-matrix property, e.g., in [3, 5, 6, 11]. A multigrid-like algorithm based on aggregation and disaggregation of Markov chains can be found in [9]. The main result of this paper is the extension of the AMG approach to a class of singular nonsymmetric M-matrices.

The theory of Markov chains [2, 14] represents an extremely important tool that has a broad variety of applications not only in the sciences, such as biology, physics, and chemistry, but also in business and economics. As an example, we consider a Markov chain model for blood circulation in the human body.

Example 1. Let us imagine the organs within a human body as states of a blood particle. Within one time step, it can advance from one organ to another with a certain probability. It is also possible that it stays within the organ and only continues its journey a few time steps later. On the other hand, an organ can also be considered as a set of substates, since the blood does not simply pass an organ but can also recirculate within it in some chaotic way. Hence, if we define a set of (sub)organs as states of the Markov chain and the corresponding transition probabilities, we get a large sparse (row) stochastic matrix. We are interested in the stationary distribution of the Markov chain, which would, e.g., represent the concentration of a dissolved substance after injection into an organ. For a more detailed description, see [16].

*Received by the editors May 8, 2006; accepted for publication (in revised form) January 18, 2007; published electronically September 28, 2007. This work was supported by the DFG Research Center MATHEON.

http://www.siam.org/journals/sisc/29-5/65927.html

†Technische Universität Berlin, Institut für Mathematik, Straße des 17. Juni 136, 10623 Berlin, Germany (virnik@math.tu-berlin.de).

2. Preliminaries. We call a vector $v \in \mathbb{R}^n$ *positive*, and we write $v > 0$ if all entries v_i are positive. A matrix $T \in \mathbb{R}^{n\times n}$, $T = (t_{ij})_{i,j=1,\dots,n}$, is called *positive (nonnegative)*, and we write $T > 0$ ($T \geq 0$) if all entries t_{ij} are positive (nonnegative). A matrix $T \in \mathbb{R}^{n\times n}$ is called *reducible* if there exists a permutation matrix $P \in \mathbb{R}^{n\times n}$, such that $PTP^T = \left[\begin{smallmatrix} T_{11} & 0 \\ T_{21} & T_{22} \end{smallmatrix}\right]$. Otherwise it is called *irreducible*.

A scalar $\lambda \in \mathbb{R}$ is called an *eigenvalue* of the matrix $T \in \mathbb{R}^{n\times n}$ if a vector $v \in \mathbb{R}^n,\ v \neq 0$ exists, such that $Tv = \lambda v$. The vector v is called a *(right) eigenvector* of T associated with λ. Accordingly, a vector $w \in \mathbb{R}^n,\ w \neq 0$, with $w^T T = \lambda w^T$ is called a *(left) eigenvector* of T. Let $T \in \mathbb{R}^{n\times n}$ have the eigenvalues $\lambda_i,\ i = 1, \dots, n$. Then we call $\rho(T) = \max_{1\leq i\leq n} |\lambda_i|$ the *spectral radius of* T.

A process is called a *finite homogeneous Markov chain* if it has n states $s_1, \dots, s_n$ and the transition probability $P[s_i \rightsquigarrow s_j] =: t_{ij}$ is time-independent. The matrix $T = [t_{ij}]_{i,j=1,\dots,n}$ satisfies $t_{ij} \geq 0$ and $\sum_{j=1}^n t_{ij} = 1$ for $i, j = 1, \dots, n$, i.e., it is *(row) stochastic* and is called the *transition matrix* of a Markov chain. We denote by $x^k = (x_i^k)$ the *probability distribution vector*, where x_i^k is the probability that the system is in state s_i after k steps. We have $x_i^k \geq 0$ and $\sum_{i=1}^n x_i^k = 1$ for each k. A distribution vector x is said to be *stationary* if $x^T T = x^T$.

The well-known Perron–Frobenius theorem guarantees the existence and uniqueness of a stationary distribution.

THEOREM 1 (Perron–Frobenius theorem [2, p. 27]). *Let $T \geq 0$ be irreducible with spectral radius $\rho(T)$. Then $\rho(T)$ is a simple eigenvalue and T has a positive left and right eigenvector corresponding to $\rho(T)$.*

COROLLARY 2 (see [2, p. 28]). *A positive eigenvector x of a nonnegative matrix T corresponds to $\rho(T)$.*

COROLLARY 3. *Every finite homogeneous Markov chain has a stationary probability distribution vector. If the transition matrix T of the process is also irreducible, then the stationary probability distribution vector is unique.*

We define by $\mathcal{Z}^{n\times n} = \{A = [a_{ij}] \in \mathbb{R}^{n\times n} : a_{ij} \leq 0, i \neq j\}$ the set of all real matrices with nonpositive off-diagonal entries. Let $B \geq 0$ with spectral radius $\rho(B)$. A matrix A of the form $A = sI - B$, with $s > 0$ and $s \geq \rho(B)$, is called an *M-matrix*. If $s > \rho(B)$, then A is a *nonsingular* M-matrix, and if $s = \rho(B)$, then A is a *singular* M-matrix. Hence, in our case $A = (I - T^T)$ is a singular M-matrix.

3. AMG preconditioning. The GMRES method we use was introduced in [13]. Algorithm 1 shows a version with left preconditioning.

As we will see in the following, the application of AMG in our case requires the property that the matrix A have row sums zero. Since $A = (I - T^T)$ has column sums zero instead, for the construction of the preconditioner we use the matrix A^T and then use M^T for the iteration process.

3.1. Basic framework of AMG. The AMG method that is presented here follows the concept introduced in [12], which is motivated by geometric multigrid, where a sequence of grids is constructed from the underlying geometry with corresponding transfer operators between the grids. The main idea of geometric multigrid is to remove the smooth error, which cannot be eliminated by relaxation on the fine grid, by coarse grid correction. The solution process then as usual consists of presmoothing, transfer of residuals from fine to coarse grids, interpolation of corrections from coarse to fine levels, and optional postsmoothing. In contrast to geometric multigrid, the idea of AMG is to define an *artificial* sequence of systems of equations decreasing in

ALGORITHM 1: Preconditioned MGS-GMRES.

Input: $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, starting vector $x_0 \in \mathbb{R}^n$, preconditioner M
Output: $x_m \in \mathbb{R}^n$ solution approximate in the mth step

```
1  Compute r_0 = M^{-1}(b - Ax_0), β := ||r_0||_2 and v_1 := r_0/β
2  Define the (m+1) × m matrix H̄_m = {h_ij}_{1≤i≤m+1,1≤j≤m}. Set H̄_m = 0.
3  for j = 1, 2, ..., m do
4      Compute w_j := M^{-1} A v_j
5      for i = 1, ..., j do
6          h_ij := (w_j, v_i)
7          w_j := w_j - h_ij v_i
8      h_{j+1,j} = ||w_j||_2. If h_{j+1,j} = 0, set m := j and go to 12
9      v_{j+1} = w_j / h_{j+1,j}
10 Compute y_m the minimizer of ||βe_1 - H̄_m y||_2
11 Set x_m = x_0 + V_m y_m.
```

size,

$$A^m u^m = b^m, \tag{2}$$

where the superscript m denotes the mth iterate, directly from the underlying matrix. We have $A^1 = A$, and for $m = 1$ the system (2) is identical to (1). We call these equations "coarse grid" equations. The interpolation operator P and the restriction operator R define the transfer from finer to coarser grids and vice versa. More precisely, let $A^m \in \mathbb{R}^{n_m \times n_m}$; then

$$R^m : \mathbb{R}^{n_m} \to \mathbb{R}^{n_{m+1}},$$
$$P^m : \mathbb{R}^{n_{m+1}} \to \mathbb{R}^{n_m},$$

and the operator on the coarser grid is defined by

$$A^{m+1} = R^m A^m P^m \ \in \mathbb{R}^{n_{m+1}\times n_{m+1}}. \tag{3}$$

Thus, we do not need a geometry behind the problem.

The AMG method consists of two main parts: the setup phase and the solution phase. During the setup phase, the coarse grids and the corresponding operators are defined. We describe this in section 3.2. The solution phase consists of a multilevel iteration. Here, as an example, we illustrate a two-level iteration, which involves the combination of a smoothing process (e.g., Gauss–Seidel; see [7]), with a correction on a coarser grid:

$$x^{(k+1)} = x^{(k)} + B^{-1}(b^m - A^m x^{(k)}), \tag{4a}$$
$$x^{(k+2)} = x^{(k+1)} + P^m (A^{m+1})^{-1} R^m (b^m - A^m x^{(k+1)}). \tag{4b}$$

Here, the matrix $(A^{k+1})^{-1}$ is replaced by the recursive application of (4) to the solution of the coarse grid system. The number of recursive calls, which is the number of levels m, depends on the size and structure of the matrix. For our method, we use the V-cycle pattern (see, e.g., [8]), as shown in Figure 1. Other recursive patterns, e.g., the W-cycle that involves two recursive calls per cycle, are presented in [8].

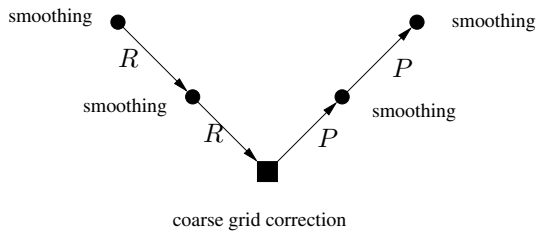


FIG. 1. *V-cycle pattern.*

3.2. Coarsening. The first main step of the coarsening process as introduced in [12] can be read as a permutation of the matrix $A \in \mathbb{R}^{n\times n}$ such that the fine grid nodes come first:

$$A \rightsquigarrow \pi^T A \pi = \left[\begin{array}{c|c} A_{FF} & A_{FC} \\ \hline A_{CF} & A_{CC} \end{array}\right].$$

We choose the fine grid nodes such that the block A_{FF} is close to a diagonal matrix. A further sparsification step is applied in which "small" off-diagonal entries in A_{FF} are added to the diagonal and the remaining off-diagonal entries in A_{FF} are redistributed onto the A_{FC} block:

$$A \rightsquigarrow \tilde{A} = \left[\begin{array}{c|c} \tilde{A}_{FF} & \tilde{A}_{FC} \\ \hline A_{CF} & A_{CC} \end{array}\right]. \tag{5}$$

It is shown in [12] that the "sparsification" process leaves the row sums of the matrix A unchanged.

To obtain the restriction and interpolation operators, we consider the incomplete block LU decomposition of the matrix $\tilde{A}$:

$$\tilde{A} = \left[\begin{array}{c|c} \tilde{A}_{FF} & \tilde{A}_{FC} \\ \hline A_{CF} & A_{CC} \end{array}\right] = \left[\begin{array}{c|c} I & 0 \\ \hline A_{CF}\tilde{A}_{FF}^{-1} & I \end{array}\right] \left[\begin{array}{c|c} \tilde{A}_{FF} & \tilde{A}_{FC} \\ \hline 0 & S \end{array}\right], \tag{6}$$

where $S = A_{CC} - A_{CF}\tilde{A}_{FF}^{-1}\tilde{A}_{FC}$ is the Schur complement with respect to $\tilde{A}_{FF}$. We take S to be the new coarse grid operator. This is the main difference from the classical approach in [12]. For the coarse grid operator we compute the exact Schur complement of an approximated matrix $\tilde{A}$, whereas in the classical approach an approximation to the Schur complement of the exact matrix A is computed. For other choices of restriction, interpolation, and coarse grid operators see, e.g., [5, 6, 12, 15]. Neither of these choices works here, since they do not preserve in our case important properties such as the singular M-matrix property and the row sums zero property; see section 3.3 for details.

The following lemma summarizes the main results for the coarsening process.

LEMMA 4. *Let A be sparsified and partitioned as in* (6) *and let N be the number of coarse grid nodes. Let S as in* (6) *be chosen as the new coarse grid operator. Then the following properties hold:*

1. *With the restriction and interpolation operators defined by*

$$R = \begin{bmatrix} -A_{CF}\tilde{A}_{FF}^{-1} & I \end{bmatrix} \in \mathbb{R}^{N\times n} \quad \textit{and} \quad P = \begin{bmatrix} -\tilde{A}_{FF}^{-1}\tilde{A}_{FC} \\ I \end{bmatrix} \in \mathbb{R}^{n\times N},$$

the new coarse grid operator S can be calculated from

$$S = R\tilde{A}P \in \mathbb{R}^{N\times N}. \tag{7}$$

2. *The interpolation operator P has all row sums equal to one.*
3. *The new coarse grid operator S has row sums equal to zero.*

Proof.

1. By straightforward calculation we get

$$\begin{aligned} R\tilde{A}P &= \begin{bmatrix} -A_{CF}\tilde{A}_{FF}^{-1} & I \end{bmatrix} \begin{bmatrix} \tilde{A}_{FF} & \tilde{A}_{FC} \\ A_{CF} & A_{CC} \end{bmatrix} \begin{bmatrix} -\tilde{A}_{FF}^{-1}\tilde{A}_{FC} \\ I \end{bmatrix} \\ &= \begin{bmatrix} -A_{CF}\tilde{A}_{FF}^{-1} & I \end{bmatrix} \begin{bmatrix} -\tilde{A}_{FC} + \tilde{A}_{FC} \\ -A_{CF}\tilde{A}_{FF}^{-1}\tilde{A}_{FC} + A_{CC} \end{bmatrix} \\ &= A_{CC} - A_{CF}\tilde{A}_{FF}^{-1}\tilde{A}_{FC} = S \in \mathbb{R}^{N\times N}. \end{aligned}$$

2. Let $\mathbb{1} = [1,\dots,1]^T$ represent the vector containing all entries equal to one. Hence, for all $i \in \{1,\dots,N\}$ we have

$$0 = e_i^T \tilde{A}\mathbb{1} = \sum_{k\in\mathcal{C}_i} \tilde{a}_{ik} + \tilde{a}_{ii},$$

and thus, as all $\tilde{a}_{ik} \le 0$ for $i \ne k$, we get

$$\frac{\sum_{k\in\mathcal{C}_i} |\tilde{a}_{ik}|}{\tilde{a}_{ii}} = 1,$$

where $\mathcal{C}_i$ is the set of interpolatory connections of i, i.e., we sum up the entries in the ith row of the $\tilde{A}_{FC}$ block. Hence, all row sums of the interpolation operator P are equal to one.

3. From 2. we conclude that

$$e_i^T S\mathbb{1} = e_i^T R\tilde{A}P\mathbb{1} = e_i^T R\tilde{A}\mathbb{1} = 0 \quad \text{for all } i,$$

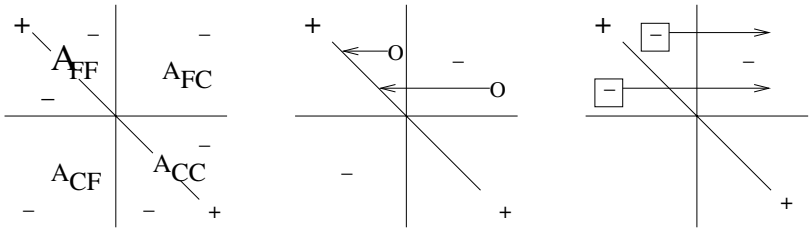
i.e., the zero row sums property is preserved for the new coarse grid operator S. ∎

3.3. Multilevel setup. For the setup of a recursive multilevel cycle, it is essential to ensure that the important properties of the finest grid operator carry over to all coarser grids. In our case it is on the one hand the row sums zero property and on the other hand the singular M-matrix property. These are important for an adequate interpolation of the smooth error that needs to be reduced by the coarse grid correction and can be written as

$$e_i \approx \sum_{j\ne i} \frac{|a_{ij}|}{a_{ii}} e_j, \quad \text{where } \sum_{j\ne i} \frac{|a_{ij}|}{a_{ii}} = 1.$$

Here, we can use only significant weights $|a_{ij}|/a_{ii}$ to interpolate e_i. Also the singular M-matrix property is important for the direct solution on the coarsest level; see section 3.5. In Lemma 4 we have shown that the coarsening process leaves the row sums zero. The following theorem states that the singular M-matrix property is preserved after the performance of one coarsening step.

THEOREM 5. *Let $A = \left[\begin{smallmatrix} A_{FF} & A_{FC} \\ A_{CF} & A_{CC} \end{smallmatrix}\right]$ be a singular M-matrix with zero row sums. Then $\tilde{A} = \left[\begin{smallmatrix} \tilde{A}_{FF} & \tilde{A}_{FC} \\ A_{CF} & A_{CC} \end{smallmatrix}\right]$ as in* (5) *and the corresponding new coarse grid operator $S = A_{CC} - A_{CF}\tilde{A}_{FF}^{-1}\tilde{A}_{FC}$ as in* (7) *are also singular M-matrices.*



(a) (b) (c)

FIG. 2.

Proof.

1. Show that $\tilde{A}$ is a singular M-matrix:
   A is a singular M-matrix. Therefore, it can be expressed in the form

$$A = sI - T, \quad \text{where } T \text{ is the transition matrix of a Markov process}$$
$$\rho(T) = 1 = s.$$

   In [12] it was shown that the coarsening process leaves the row sums invariant, i.e., in our case equal to zero. Hence, we can write $\tilde{A} = I - \tilde{T}$, where $\tilde{T}$ has row sums one. Furthermore, we have $\tilde{A} \in \mathcal{Z}^{n\times n}$, i.e., the off-diagonal entries in $\tilde{A}$ are nonpositive and the diagonal is nonnegative (as otherwise the row sums could not be zero). The diagonal entries are also at most one. One can see this by considering the following sparsification steps (see [12] for a detailed description of the "sparsification" process):
   - The initial matrix A has a sign and block structure, as illustrated in Figure 2(a). The sign "−" indicates that all entries in the corresponding section are nonpositive. Accordingly, "+" signifies that the entries in the corresponding section are nonnegative.
   - In the first step, "small" off-diagonal entries in each row are added to the diagonal (see Figure 2(b)). The diagonal entries become smaller but stay nonnegative.
   - Then the remaining off-diagonal entries of the A_{FF}-block are distributed onto the A_{FC}-block (see Figure 2(c)). The entries of the A_{FC}-block become smaller, but remain greater than -1, as otherwise the row sums could not stay zero.

   Thus, $\tilde{T}$ is in addition nonnegative and all entries are at most one. From this we conclude that $\tilde{T}$ is stochastic and thus $\rho(\tilde{T}) = 1 = s$. Consequently, $\tilde{A}$ is a singular M-matrix.
2. Show that S as in (7) is a singular M-matrix:
   In Lemma 4 we have shown that $S\mathbb{1} = 0$, i.e., S has zero row sums. From this and from $A_{CC} \in \mathcal{Z}^{N\times N}$ and $A_{CF}\tilde{A}_{FF}^{-1}\tilde{A}_{FC} \geq 0$, we can conclude that $S = A_{CC} - A_{CF}\tilde{A}_{FF}^{-1}\tilde{A}_{FC} \in \mathcal{Z}^{N\times N}$. Let S now have the representation

$$S = sI - \hat{T}, \qquad \hat{T} \geq 0, \quad s > 0.$$

   Then we get

$$S\mathbb{1} = (sI - \hat{T})\mathbb{1} = s\mathbb{1} - \hat{T}\mathbb{1} = 0$$
$$\Rightarrow \hat{T}\mathbb{1} = s\mathbb{1}.$$

As $\mathbb{1} > 0$ is a positive eigenvector of $\hat{T}$, it follows from Corollary 2 that $s = \rho(\hat{T})$. Hence, S is a singular M-matrix. □

From Theorem 5 we conclude that the singular M-matrix property is preserved on all coarser levels.

3.4. Integration of the side constraint into the solution process. For a badly chosen starting vector x_0 the GMRES method in Algorithm 1 may converge to the trivial solution, instead of the stationary distribution. In fact, for some problems we cannot choose the starting vector randomly but need to employ a specifically fixed starting vector as, e.g., in the case of the blood circulation model. Here, the initial distribution might represent an injection into a certain organ. Therefore, it makes sense to set $x_0 = [0, \ldots, 1, \ldots, 0]^T$, where the only nonzero entry $[x_0]_i = 1$ is positioned to represent the injection into the organ i. In this case the naive application of the preconditioned GMRES algorithm leads to a convergence to the trivial solution.

Hence, we have the side constraint $\mathbb{1}^T x = 1$ that we have to include into the solution process. The idea now is to shift the system such that the trivial solution is excluded as a possible solution. Consider the orthogonal projection matrix $Q := I - \frac{1}{n}\mathbb{1}\mathbb{1}^T$ with respect to $\mathbb{1}$, that is, $\mathbb{1}^T Q = 0$. Any vector x can be decomposed into

$$x := Qy + z$$

such that $z = \alpha\mathbb{1}$ for some constant $\alpha \in \mathbb{R}$. In particular, for a vector $\tilde{x}$ that fulfills $\mathbb{1}^T\tilde{x} = 1$, we obtain $z := \frac{1}{n}\mathbb{1}$.

Now we transform the initial problem as follows:

$$\begin{aligned} & Ax = 0 \\ \Leftrightarrow\ & A(Qy + z) = 0 \\ \Leftrightarrow\ & AQy = -\frac{1}{n}A\mathbb{1}. \end{aligned} \tag{8}$$

By setting $\hat{A} := AQ$ and $b := -\frac{1}{n}A\mathbb{1}$, we get a formulation of the linear system where the side constraint is embedded:

$$\hat{A}y = b. \tag{9}$$

Now, we can solve this system instead of (8) using AMG preconditioned GMRES. The system is still consistent since $b \in \mathrm{im}(A)$. As Q is orthogonal with respect to z, we have $\tilde{x} = Qy + z \neq 0$ for all y.

Note that in the implementation we do not compute the product $\hat{A} := AQ$ explicitly as Q is not sparse and matrix-matrix products are expensive, but rather successively calculate the corresponding matrix-vector products.

3.5. Direct solution on the coarsest level. On the coarsest level we solve the singular system

$$A^m x^m = b^m \tag{10}$$

directly using LU decomposition. From section 3.3 we know that the operator A^m on the coarsest level is still a singular M-matrix. In theory, the coarsening process does not necessarily guarantee irreducibility of A^m, although in our practical examples this is the case. Since the treatment of this problem is beyond the scope of this

paper, in the following discussion we will assume that A^m is irreducible. Then an LU decomposition of A^m with L nonsingular exists [10] and appears as follows:

$$A^m = LU = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \begin{bmatrix} & & \\ & \ddots & \\ & & 0 \end{bmatrix}.$$

To solve the system (10), we need only prove that it is consistent. That is, we need to ensure that L^{-1} applied to the right-hand side produces a zero in the last entry of the resulting vector. Since the entry in the lower right corner of U is zero, we have a consistent singular system. From the following lemma we conclude that whenever the right-hand side b^m is a zero vector the last entry of $L^{-1}b^m$ will be zero, and thus our system is consistent.

LEMMA 6. *Let A be a singular M-matrix with column sums zero and let $A = LU$ be an LU decomposition of A with L nonsingular. Then*

$$e_n^T L^{-1} = \mathbb{1}^T.$$

Proof. $A = LU \Leftrightarrow L^{-1}A = U$. If we apply the nth unit vector from the left, we get

$$e_n^T L^{-1} A = e_n^T U = [0, \ldots, 0].$$

Thus, it must hold that $e_n^T L^{-1} = \alpha \mathbb{1}^T$ for some scalar α. Yet, as L^{-1} has ones on the diagonal, we get $\alpha = 1$ and hence $e_n^T L^{-1} = \mathbb{1}^T$. ∎

By construction, the AMG ensures that on any level j, the system (10) satisfies $\mathbb{1}^T b^j = 0$ since this holds for the initial right-hand side in (9), and the new right-hand sides on coarser grids are obtained via restriction of the residual, i.e.,

$$b^{j+1} = R^j(b^j - A^j x^k).$$

Note that Lemma 4 is applied to A^T. Thus, R^j corresponds to $(P^j)^T$ in Lemma 4.

3.6. Numerical tests. For numerical tests, we used MATLAB Version 7.1 run on a PC with an Intel Pentium 4 CPU 3.20GHz processor. The relative machine precision was `eps` $= 2.2204 \times 10^{-16}$. As convergence criterion for GMRES we used the tolerance $t = \|r_0\|\sqrt{\texttt{eps}}$, where r_0 is the initial residual, i.e., GMRES has converged when the residual norm falls under the tolerance t. For the AMG iteration we used the V-cycle with one presmoothing and one postsmoothing sweep; see section 3.1. The time measurement was conducted via the commands `tic` and `toc` and represents an upper bound for the required time. It can be further reduced by a more efficient implementation of the AMG algorithm.

In section 1, we have discussed the Markov chain model of blood circulation in the human body. Table 1 presents some results of the AMG preconditioned GMRES algorithm applied to transition matrices of Markov chains that arise from the blood circulation problem; see Example 1. Here, we consider the case where the blood particles can recirculate in the neighboring organs before passing on. Hence, the considered matrices have about five nonzero entries per row.

In the first column of Table 1 we have the size n of the problem. The second column states the number nnz of nonzero elements in the matrix. The column labeled Grid shows the time required by AMG for the grid construction. In $\#level$ the

TABLE 1
Performance of AMG preconditioned GMRES.

		AMG+GMRES				$\sum$AMG
n	nnz	Grid	$\#level$	Time	Iters.	Time
1125	5571	1.94	7	0.2	12	2.14
2598	12936	4.28	8	0.39	13	**4.67**
4952	24706	8.27	9	0.85	13	**9.12**
9544	47666	16.97	10	2.69	12	**19.66**

TABLE 2
Comparison with other methods.

		AMGsa		$\sum$AMGsa	GS		GS+GMRES	
n	nnz	Time	Iters.	Time	Time	Iters.	Time	Iters.
1125	5571	0.20	23	2.14	0.97	1061	**0.75**	144
2598	12936	0.68	27	5.96	8.49	4214	5.25	297
4952	24706	1.73	27	10.0	26.25	5547	18.01	428
9544	47666	4.08	20	21.05	–	–	162.60	955

number of constructed grids is given. The following Time and Iters. columns represent the computation time of AMG preconditioned GMRES and the corresponding number of iterations. The last column summarizes the total time required by AMG preconditioned GMRES.

In Table 2 we give the results of alternative methods for the same problems as in Table 1. The first two columns as in Table 1 contain the sizes and the numbers of nonzeros of the problems. The following three columns contain the results for AMG as a stand-alone function, where the total required time is given in the third column. The next two columns labeled GS display the time and number of iterations required by Gauss–Seidel iteration. The last two columns present the results for GMRES with Gauss–Seidel as preconditioner. The top times are marked in bold.

As we can see, although for the smallest problem the Gauss–Seidel iteration and the Gauss–Seidel preconditioned GMRES outperform the AMG method, for larger problems AMG preconditioning becomes very effective. The computational effort of AMG preconditioned GMRES grows linearly with the problem size, whereas the number of iterations stays constant. The number of iterations for the Gauss–Seidel preconditioned GMRES, in contrast, grows linearly with the problem size. The effectiveness of AMG preconditioning is due to the fact that the overall computation time of the AMG preconditioned GMRES is dominated by the time required to construct the "algebraic grid." The computational effort required for grid construction, although it also grows linearly, grows much slower with the matrix size than the computational effort for Gauss–Seidel preconditioned GMRES. The AMG algorithm as a stand-alone function also works quite well, although the computational effort is slightly smaller in the AMG preconditioned GMRES due to fewer iterations. This also supports the suitability of the AMG method as a preconditioner for this type of problem.

4. Conclusions. In this paper, the application of the AMG method as a preconditioner for solving large singular linear systems of the type $(I - T^T)x = 0$ with GMRES was examined. In doing so, we concentrated on the case in which T is the transition matrix of a Markov process. The AMG method [12] was originally devel-

oped for regular, symmetric, positive definite M-matrices. In our case, although the matrix $(I - T^T)$ is singular and nonsymmetric, it is a singular M-matrix that possesses a number of useful properties. It turns out that with the adaptation discussed in sections 3.3–3.5, the method becomes applicable to singular matrices such as those in the blood circulation model. Numerical experiments illustrate that preconditioning with the AMG method leads to significant acceleration of the convergence speed.

Acknowledgments. The author thanks M. Bollhöfer for suggesting the problem and for the fruitful discussions throughout the project, as well as V. Mehrmann and D. B. Szyld for reading the manuscript with great care.

REFERENCES

[1] V. A. Barker, *Numerical solution of sparse singular systems of equations arising from ergodic Markov chains*, Comm. Statist. Stochastic Models, 5 (1989), pp. 335–381.
[2] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Classics Appl. Math. 9, SIAM, Philadelphia, 1994 (revised reprint of the 1979 original).
[3] M. Brezina, R. Falgout, S. MacLachlan, T. Manteuffel, S. McCormick, and J. Ruge, *Adaptive algebraic multigrid*, SIAM J. Sci. Comput., 27 (2006), pp. 1261–1286.
[4] P. N. Brown and H. F. Walker, *GMRES on (nearly) singular systems*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 37–51.
[5] Q. Chang and Z. Huang, *Efficient algebraic multigrid algorithms and their convergence*, SIAM J. Sci. Comput., 24 (2002), pp. 597–618.
[6] Q. Chang, Y. S. Wong, and H. Fu, *On the algebraic multigrid method*, J. Comput. Phys., 125 (1996), pp. 279–292.
[7] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, 1983.
[8] W. Hackbusch, *Multigrid Methods and Applications*, Springer-Verlag, Berlin, 1985.
[9] G. Horton and S. T. Leutenegger, *A multi-level solution algorithm for steady-state Markov chains*, in Proceedings of the 1994 ACM SIGMETRICS Conference on Measurement and Modeling of Computer Systems, ACM Press, New York, 1994, pp. 191–200.
[10] I W. Kuo, *A note on factorizations of singular M-matrices*, Linear Algebra and Appl., 16 (1977), pp. 217–220.
[11] S. MacLachlan, T. Manteuffel, and S. McCormick, *Adaptive reduction-based AMG*, Numer. Linear Algebra Appl., 13 (2006), pp. 599–620.
[12] J. W. Ruge and K. Stüben, *Algebraic multigrid*, in Multigrid Methods, S. F. McCormick, ed., Frontiers Appl. Math. 3, SIAM, Philadelphia, 1987, pp. 73–130.
[13] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
[14] W. J. Stewart, *Introduction to the Numerical Solution of Markov Chains*, Princeton University Press, Princeton, NJ, 1994.
[15] U. Trottenberg, C. W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, San Diego, CA, 2001.
[16] E. Virnik, *Algebraic Multigrid Preconditioners for Computing Stationary Distributions of Markov Processes*, diplomarbeit, Technische Universität Berlin, Berlin, 2004; available online from http://www.math.tu-berlin.de/preprints/abstracts/Report-10-2005.rdf.html.