Block Preconditioning for Markov Chain Problems

Michele Benzi
Department of Mathematics and Computer Science
Emory University
Atlanta, GA
Acknowledgements

Thanks to:

- Bora Uçar (formerly of Emory, now at CERFACS)
- NSF (Computational Mathematics Program)
Outline

- Stochastic matrices and Markov chains
- Available solution methods
- Block preconditioning
- Spectral analysis
- Matrix partitioning
- Experimental results using MARCA models
- Conclusions


http://www.mathcs.emory.edu/~benzi
A homogeneous Discrete-Time Markov Chain (DTMC) is a stochastic process \( \{X_k\} \) described by the transition probability matrix \( P = [p_{ij}] \), where

\[
p_{ij} = P\{X_k = j \mid X_{k-1} = i\}
\]

is the probability of a transition from state \( i \) to state \( j \) at the \( k \)th step.

The transition matrix \( P \) is row-stochastic:

\[
p_{ij} \geq 0 \quad \forall i, j \quad \text{and} \quad \sum_j p_{ij} = 1 \quad \forall i.
\]

Here we assume that the number of possible states is finite, and equal to \( N \); hence, \( P \) is \( N \times N \).
Discrete-time Markov chains have countless applications throughout the natural and social sciences, in engineering (e.g., computer performance evaluation), in information retrieval (Google’s PageRank), operations research, linguistics, psychology, and so forth.

A Google search with key phrase “Markov chains” returns over 1.6 million hits (October 2007).

Markov chains were introduced by the Russian mathematician A. A. Markov in a 1906 paper in which he showed that independence of the random variables \(X_k\) is not necessary for the Law of Large Numbers and other limit (“ergodic”) theorems to hold.
Let $\pi^0 = [\pi^0_1, \pi^0_2, \ldots, \pi^0_N]$ be an initial probability distribution. Then the probability distribution at step $k$ is given by

$$\pi^k = \pi^{k-1}P = \pi^0P^k, \quad k = 1, 2, \ldots$$

For an ergodic Markov chain, there exists a steady-state probability distribution vector

$$\pi^\infty = \lim_{k \to \infty} \pi^k$$

independent of $\pi^0$. This is the unique probability distribution such that

$$\pi^\infty = \pi^\infty P.$$

Thus: $\pi^\infty$ is a left eigenvector of $P$ corresponding to the eigenvalue $\lambda = 1$. 
In order to describe the long-run behavior of the DTMC, we need to find the stationary distribution vector $\pi^{\infty}$. Hence, the computational problem is:

Find $\pi \in \mathbb{R}^{1 \times N}$ such that

$$\pi = \pi P, \quad \pi_i > 0, \quad \sum_{i=1}^{N} \pi_i = 1$$

or, equivalently,

$$(I - P^T)x = Ax = 0$$

where we have set $A = I - P^T$ and $x = \pi^T$.

$A$ is called the rate matrix, or the generator, associated with the DTMC.
For an ergodic chain, all states are reachable from any other state. Hence, $P$ is irreducible, and so is $A = I - P^T$.

Using Perron-Frobenius theory one can show that $A$ is a singular $M$-matrix with rank $N - 1$.

Further, there exists a unique null vector $x \in \mathbb{R}^N$ with

$$x_i > 0 \quad \text{and} \quad \sum_{i=1}^{N} x_i = 1.$$ 

In other words, there exists a unique positive vector $x = \pi^T$ such that

$$\text{span} \{x\} = \mathcal{N}(A).$$

It is this vector that we wish to compute.
Available Solution Methods

• Direct methods
  – Gaussian Elimination ($A = LU$)
  – GTH method (Grassmann-Taksar-Heyman 1985)
  – QR factorization (Golub & Meyer 1986)
  – DPM (B., 2004)

• Iterative methods
  – Classical: power, Gauss-Seidel, SOR, block variants
  – Multilevel, AMG-like: IAD
    (Iterative Aggregation/Disaggregation)
  – Preconditioned Krylov subspace methods

Direct methods are used for moderate-size problems, and as coarse solvers in multilevel methods, where “exact” solves are needed.
Here we assume that $P$ (therefore, $A$) is large and sparse.

However, we make the assumption that the nonzero entries of $A$ are explicitly available.

We consider preconditioned GMRES for solving the singular, homogeneous system $Ax = 0$.

There is a strong interest in developing effective preconditioners that can be efficiently implemented in parallel.

A main ingredient is the use of graph partitioning to reorder the matrix into a suitable block structure.
Block preconditioning

Assume that $A$ is an irreducible $M$-matrix that has been partitioned in the following block $2 \times 2$ form:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{11}$ is $n \times n$ and $A_{22}$ is $m \times m$, with $n + m = N$. Typically, $n > m$.

Then it is well known that:

1. $A_{11}$ is a nonsingular $M$-matrix.

2. The Schur complement $S := A_{22} - A_{21}A_{11}^{-1}A_{12}$ is an irreducible $M$-matrix. It is nonsingular iff $A$ is. If $A$ is singular, then $S$ has rank $m - 1$. 
Assume now that $A$ is **nonsingular**. Then $A$ has the block LU factorization

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ A_{21}A_{11}^{-1} & I_m \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & S \end{bmatrix}.$$
Assume now that $A$ is nonsingular. Then $A$ has the block LU factorization

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ A_{21}A_{11}^{-1} & I_m \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & S \end{bmatrix}.$$ 

Letting

$$P_T = \begin{bmatrix} A_{11} & A_{12} \\ 0 & S \end{bmatrix}$$
Assume now that $A$ is nonsingular. Then $A$ has the block LU factorization

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I_n & O \\ A_{21}A_{11}^{-1} & I_m \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ O & S \end{bmatrix}.$$ 

Letting $P_T = \begin{bmatrix} A_{11} & A_{12} \\ O & S \end{bmatrix}$

we have that $\sigma(AP_T^{-1}) = \sigma(P_T^{-1}A) = \{1\}$.

That is, the matrix $AP_T^{-1}$ (and hence $P_T^{-1}A$) has only the eigenvalue $\lambda = 1$ (of multiplicity $N$).
Furthermore, the preconditioned matrix $M = AP_T^{-1}$ satisfies $(M - I)^2 = O$. Hence, the minimum polynomial of $M$ has degree 2, which implies that GMRES will converge to the solution of $Ax = 0$ in at most two steps (in exact arithmetic).

Block triangular preconditioners of the form $P_T$ have been studied by Murphy, Golub and Wathen (SISC, 2000) and by others for solving saddle point problems where, typically,

$$A_{12} = A_{21}^T \quad \text{and} \quad A_{22} = O.$$

This approach has been extended to more general matrices by Ipsen (SISC, 2001). Here we study the application of block triangular preconditioning to Markov chains. A major difference is that now the $2 \times 2$ block structure is not given by the problem, but it must be imposed.
The “ideal” block triangular preconditioner $P_T$ is not practical. In practice one uses as preconditioner a block triangular matrix of the form

$$P_T = \begin{bmatrix}
\hat{A}_{11} & A_{12} \\
O & \hat{S}
\end{bmatrix}$$

where $\hat{A}_{11} \approx A_{11}$ and $\hat{S} \approx S$ are invertible approximations to $A_{11}$ and $S$.

Linear systems with $\hat{A}_{11}$ and $\hat{S}$ must be “easy” to solve. At the same time, the approximations must be good enough so as to retain fast convergence of GMRES.
Spectral analysis

For Markov chain problems, $A$ (and hence $S$) is singular. The preconditioned matrix, therefore, will have the simple eigenvalue $\lambda = 0$, and a cluster around $\lambda = 1$. The better the approximations $\tilde{A}_{11} \approx A_{11}$ and $\tilde{S} \approx S$, the tighter the cluster.

It is not easy to bound the eigenvalues of the preconditioned matrix, in general. Some simple results can be obtained by assuming that $\tilde{A}_{11} = A_{11}$ and that

$$\tilde{S} = A_{22} - A_{21} M_{11}^{-1} A_{21} \quad \text{where} \quad O \leq M_{11}^{-1} \leq A_{11}^{-1}.$$  

Note that the last inequality cannot be an equality.

The conditions on $\tilde{S}$ are satisfied if $M_{11}$ is obtained from $A_{11}$ by deletion of off-diagonal entries. For instance, $M_{11} = \text{diag}(A_{11})$ will do, unless $A_{11}$ is itself diagonal.
Theorem: Let $A$ be a singular, irreducible $M$-matrix partitioned in block $2 \times 2$ form. Let $P_T$ be a block triangular preconditioner with

$$\hat{A}_{11} = A_{11}, \quad \hat{S} = A_{22} - A_{21}M_{11}^{-1}A_{12}$$

where $M_{11} \neq A_{11}$ satisfies $O \leq M_{11}^{-1} \leq A_{11}^{-1}$.

Then the spectrum of $P_T^{-1}A$ consists of:

- The simple eigenvalue $\lambda = 0$;
- The eigenvalue $\lambda = 1$ of multiplicity at least $n$;
- A cluster of at most $m - 1$ eigenvalues lying in the disk $D(1, 1) = \{ z \in \mathbb{C} ; |z - 1| < 1 \}$. The diameter of this cluster goes to zero as $\| S - \hat{S} \| \to 0$.

Moreover, the splitting $A = P_T - (P_T - A)$ is weak regular of the II kind; that is, $P_T^{-1} \geq O$ and $I - AP_T^{-1} \geq O$. 
Matrix partitioning

The spectral analysis of the preconditioned matrix suggests that the choice of blocks should be such that $n$, the size of the $(1,1)$ block $A_{11}$, is as large as possible, so as to maximize the number of eigenvalues at or near $\lambda = 1$.

Computational considerations, on the other hand, impose that the $A_{11}$ block should be easy to (approximately) invert. This is because we need to solve linear systems with $A_{11}$ and also because the inverse of $A_{11}$ appears in the definition of the Schur complement.

Also, in view of a possible parallel implementation, $A_{11}$ should be a block diagonal matrix.
Thus, for a given integer $p$, we would like to find a reordering (symmetric permutation) of $A$ into the block form

$$A = \begin{bmatrix} \frac{A_{11}}{A_{21}} & \frac{A_{12}}{A_{22}} \\ \frac{A_{21}}{A_{22}} & \frac{A_{22}}{A_{22}} \end{bmatrix} = \begin{bmatrix} A_1 & \cdots & B_1 \\ A_2 & \cdots & B_2 \\ \vdots & \ddots & \vdots \\ A_p & \cdots & B_p \\ C_1 & \cdots & C_p \end{bmatrix},$$

with the size of the block $A_{22} = A_S$ as small as possible.

With such a partitioning we have

$$A_{11} = \text{diag} (A_1, A_2, \ldots, A_p) \quad \text{and} \quad S = A_S - \sum_{i=1}^{p} C_i A_i^{-1} B_i.$$

Note that $S$ is well-defined since each $A_i$ is an invertible $M$-matrix.
Matrix partitioning (cont.)

Such a reordering of $A$ can be obtained using graph partitioning by vertex separator (GPVS) techniques.

Given an undirected graph $\mathcal{G} = (V, E)$ and an integer $p$, the $p$-way GPVS problem consists of finding a set of vertices $V_S$ of minimum size whose removal decomposes a graph into $p$ disconnected subgraphs $V_1, V_2, \ldots, V_p$ with balanced sizes. The problem is NP-hard.

We use standard graph partitioning software (METIS) applied to the undirected graph $\mathcal{G}$ associated with the symmetrized matrix $A + A^T$. Note that

$$|V| = N \quad \text{and} \quad |E| = n\text{nz}(A + A^T).$$
The matrix $A$ can be put into the $2 \times 2$ block structure

$$A = \begin{bmatrix}
A_1 & A_2 & \cdots & B_1 \\
& A_2 & \cdots & B_2 \\
& & \ddots & \vdots \\
C_1 & C_2 & \cdots & C_p & A_p & B_p \\
& & & & A_S 
\end{bmatrix}$$

by permuting the rows and columns associated with the vertices in $\bigcup_k V_k$ before the rows and columns associated with the vertices in $V_S$. That is, $V_S$ defines the rows and columns of the $(2,2)$ block $A_{22} \equiv A_S$.

Hence, the size of the separator set is $|V_S| = m$. 
Example: MARCA matrix ‘qnatm06’
We use this block structure to define block triangular preconditioners of the form

\[
P_T = \begin{bmatrix}
\tilde{A}_1 & B_1 \\
\tilde{A}_2 & B_2 \\
\vdots & \vdots \\
\tilde{A}_p & B_p
\end{bmatrix}
\]

where \( \tilde{A}_i \approx A_i \) for \( i = 1 : p \) and \( \hat{S} \) is a sparse approximation to the Schur complement. Taking just \( \hat{S} \approx \tilde{A}_S \) leads to a block Gauss–Seidel preconditioner with inexact block solves.

Lower triangular and block diagonal variants could also be used.
In our code, we use ILUTH (threshold-based incomplete LU factorization) to build the diagonal blocks $\tilde{A}_i$:

$$\tilde{A}_i = L_i U_i \approx A_i, \quad i = 1:p.$$ 

The approximate Schur complement $\tilde{S}$ is obtained in two stages. First we compute

$$\bar{S} = A_S - \sum_{i=1}^{p} B_i M_i^{-1} C_i$$

where $M_i^{-1} \approx A_i^{-1}$ for $i = 1:p$, then we use ILUTH to approximate $\bar{S}$.

For the approximate inverses $M_i^{-1} \approx A_i^{-1}$ we found simple diagonal approximations to be sufficient. Note that each $A_i$ is diagonally dominant (by columns).
Experimental results

- Solver: GMRES(50) with right preconditioning
- MATLAB 7.1.0 implementation
- 2.2 GHz dual core AMD Opteron (4GB main memory)
- Random $x_0$, stopping criterion: $\|Ax_k\| < 10^{-10}$
- Test matrices from MARCA (W. J. Stewart)

<table>
<thead>
<tr>
<th>Matrix</th>
<th>number of rows/cols</th>
<th>number of nonzeros</th>
<th>average</th>
<th>row</th>
<th>col</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>total</td>
<td>row/col</td>
<td>min</td>
<td>max</td>
</tr>
<tr>
<td>mutex09</td>
<td>65535 263950</td>
<td>1114079</td>
<td>17.0</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>mutex12</td>
<td>4031310</td>
<td>15.3</td>
<td>9</td>
<td>9</td>
<td>21</td>
</tr>
<tr>
<td>mutex09</td>
<td>62196 176851</td>
<td>420036</td>
<td>6.8</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>mutex12</td>
<td>1207051</td>
<td>6.8</td>
<td>2</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>mutex09</td>
<td>79220 130068</td>
<td>533120</td>
<td>6.7</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>mutex12</td>
<td>875896</td>
<td>6.7</td>
<td>3</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>qnatm06</td>
<td>13671 17081</td>
<td>67381</td>
<td>4.9</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>qnatm07</td>
<td>84211</td>
<td>4.9</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>tcomm16</td>
<td>66177 263169</td>
<td>263425</td>
<td>4.0</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>tcomm20</td>
<td>1050625</td>
<td>4.0</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>
The ‘mutex’ matrices behave differently from the others:
1. Preconditioning helps but is not essential;
2. Separator set is huge \( m \approx \frac{N}{5} \) already for \( p = 2 \).

Timings (secs.) for Block Gauss-Seidel, Product Splitting, and Block Triangular prec.

Block Gauss-Seidel with \( p = 8 \) is best for this problem.

<table>
<thead>
<tr>
<th>Matrix</th>
<th>GMRES Total time</th>
<th>Preconditioned GMRES Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Solve</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>BGS</td>
</tr>
<tr>
<td>mutex09</td>
<td>7.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3.60</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.52</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>1.45</td>
</tr>
<tr>
<td>mutex12</td>
<td>27.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>17.95</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5.49</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>3.37</td>
</tr>
</tbody>
</table>
- Experiments with ‘mutex’ matrices
- Average iteration counts for different values of $p$ and different preconditioners
- $BJ =$ Block Jacobi, $BD =$ Block Diagonal, $BGS =$ Block Gauss-Seidel, $PS =$ Product Splitting, $BT =$ Block Triangular

<table>
<thead>
<tr>
<th>Matrix</th>
<th>GMRES</th>
<th>Preconditioned GMRES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Preconditioners</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BJ</td>
</tr>
<tr>
<td>mutex09</td>
<td>97</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32</td>
</tr>
<tr>
<td>mutex12</td>
<td>91</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32</td>
</tr>
</tbody>
</table>
- Experiments with ‘ncd’ matrices
- BJ = Block Jacobi, BD = Block Diagonal, BGS = Block Gauss-Seidel, PS = Product Splitting, BT = Block Triangular
- Here ‘250’ means no convergence within 250 iterations

<table>
<thead>
<tr>
<th>Matrix</th>
<th>GMRES</th>
<th>Preconditioned GMRES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Preconditioners</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BJ</td>
</tr>
<tr>
<td>ncd07</td>
<td>250</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32</td>
</tr>
<tr>
<td>ncd10</td>
<td>250</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32</td>
</tr>
</tbody>
</table>
**Experimental results (cont.)**

- Iteration counts for ‘qnatm’ matrices
- BJ = Block Jacobi, BD = Block Diagonal, BGS = Block Gauss-Seidel, PS = Product Splitting, BT = Block Triangular
- Here ‘250’ means no convergence within 250 iterations

<table>
<thead>
<tr>
<th>Matrix</th>
<th>GMRES</th>
<th>Preconditioned GMRES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Preconditioners</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BJ</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td></td>
</tr>
<tr>
<td>qnatm06</td>
<td>250</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32</td>
</tr>
<tr>
<td>qnatm07</td>
<td>250</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32</td>
</tr>
</tbody>
</table>
Experimental results (cont.)

- Iteration counts for ‘tcomm’ matrices
- BJ = Block Jacobi, BD = Block Diagonal, BGS = Block Gauss-Seidel, PS = Product Splitting, BT = Block Triangular
- Here ‘250’ means no convergence within 250 iterations

<table>
<thead>
<tr>
<th>Matrix</th>
<th>GMRES</th>
<th>Preconditioned GMRES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>BJ</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td></td>
</tr>
<tr>
<td>tcomm16</td>
<td>250</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32</td>
</tr>
<tr>
<td>tcomm20</td>
<td>250</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32</td>
</tr>
</tbody>
</table>
• Iteration counts for ‘twod’ matrices
• BJ = Block Jacobi, BD = Block Diagonal, BGS = Block Gauss-Seidel, PS = Product Splitting, BT = Block Triangular
• Here ‘250’ means no convergence within 250 iterations

<table>
<thead>
<tr>
<th>Matrix</th>
<th>GMRES</th>
<th>Preconditioned GMRES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>BJ</td>
</tr>
<tr>
<td>twod08</td>
<td>250</td>
<td>p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32</td>
</tr>
<tr>
<td>twod10</td>
<td>250</td>
<td>p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32</td>
</tr>
</tbody>
</table>
Experimental results (cont.)

- Timings for larger matrix of each remaining type

<table>
<thead>
<tr>
<th>Matrix</th>
<th>p</th>
<th>Total time</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Set-up</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BGS</td>
<td>PS</td>
<td>BT</td>
<td>BGS</td>
<td>PS</td>
</tr>
<tr>
<td>ncd10</td>
<td>2</td>
<td>1.18</td>
<td>1.93</td>
<td>1.68</td>
<td>19.40</td>
<td>5.39</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.04</td>
<td>1.99</td>
<td>1.59</td>
<td>15.85</td>
<td>6.16</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1.01</td>
<td>2.31</td>
<td>1.58</td>
<td>33.95</td>
<td>7.74</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>1.01</td>
<td>3.04</td>
<td>1.67</td>
<td>45.84</td>
<td>12.34</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>1.09</td>
<td>4.55</td>
<td>2.00</td>
<td>71.01</td>
<td>20.82</td>
</tr>
<tr>
<td>qnatm07</td>
<td>2</td>
<td>3.68</td>
<td>4.18</td>
<td>4.04</td>
<td>11.43</td>
<td>13.62</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2.34</td>
<td>2.98</td>
<td>2.71</td>
<td>11.69</td>
<td>15.12</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1.66</td>
<td>2.53</td>
<td>2.04</td>
<td>13.23</td>
<td>18.73</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>1.34</td>
<td>2.75</td>
<td>1.84</td>
<td>18.07</td>
<td>26.62</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>1.30</td>
<td>3.72</td>
<td>1.96</td>
<td>23.85</td>
<td>46.35</td>
</tr>
<tr>
<td>tcomm20</td>
<td>2</td>
<td>0.10</td>
<td>0.13</td>
<td>0.11</td>
<td>0.43</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.63</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.10</td>
<td>0.18</td>
<td>0.10</td>
<td>0.93</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>0.10</td>
<td>0.20</td>
<td>0.10</td>
<td>1.98</td>
<td>2.47</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>0.10</td>
<td>0.30</td>
<td>0.10</td>
<td>5.01</td>
<td>8.30</td>
</tr>
<tr>
<td>twod10</td>
<td>2</td>
<td>7.04</td>
<td>7.92</td>
<td>7.64</td>
<td>10.46</td>
<td>10.90</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>6.65</td>
<td>7.55</td>
<td>7.15</td>
<td>12.85</td>
<td>14.32</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>4.95</td>
<td>6.47</td>
<td>5.64</td>
<td>14.41</td>
<td>18.83</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>4.09</td>
<td>6.43</td>
<td>4.89</td>
<td>16.78</td>
<td>30.60</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>3.02</td>
<td>7.17</td>
<td>4.18</td>
<td>18.27</td>
<td>53.05</td>
</tr>
</tbody>
</table>
Conclusions

• Block triangular preconditioning is promising

• Results are fairly stable with respect to $p$

• Comparisons on MARCA models suggest the method is often superior to other techniques with similar complexity and storage requirements

• Future work:
  1. Parallel implementation
  2. What to do when the separator set is huge?
  3. Multilevel version?