

# Preconditioning for nonsymmetric systems

Andy Wathen  
Oxford University, UK



joint work with  
Jingkai Mao (Oxford University)  
Federico Danieli (Oxford University)

# Krylov subspace methods

For self-adjoint problems/symmetric matrices, iterative methods of choice exist: conjugate gradients for SPD, MINRES otherwise

but many possible methods for non-self-adjoint problems/nonsymmetric matrices: GMRES , BICGSTAB , QMR , IDR , . . .

# Krylov subspace methods

For self-adjoint problems/symmetric matrices, iterative methods of choice exist: conjugate gradients for SPD, MINRES otherwise

but many possible methods for non-self-adjoint problems/nonsymmetric matrices: GMRES , BICGSTAB , QMR , IDR , ...

For almost all need *preconditioning*

Preconditioner  $\mathbf{P}$  such that

$$\text{“}\mathbf{P}^{-1}\mathbf{B}\mathbf{x} = \mathbf{P}^{-1}\mathbf{b}\text{”}$$

has much faster convergence with the appropriate iterative method than  $\mathbf{B}\mathbf{x} = \mathbf{b}$ .

# Krylov subspace methods

Arises because of convergence guarantees:

- for symmetric matrices: descriptive convergence bounds based on eigenvalues  $\Rightarrow$  a priori estimates of iterations for acceptable convergence; good preconditioning ensures fast convergence.
- for nonsymmetric matrices: by contrast, to date there are no generally applicable *and descriptive* convergence bounds even for GMRES ; for any of the other nonsymmetric methods without a minimisation property, convergence theory is extremely limited  $\Rightarrow$  no good a priori way to identify what are the desired qualities of a preconditioner

A major theoretical difficulty, but heuristic ideas abound!

# Real Toeplitz matrices

can use *Pestana & W, 2015*: If  $\mathbf{B}$  is a real Toeplitz matrix then

$$\underbrace{\begin{bmatrix} a_0 & a_{-1} & \cdot & \cdot & a_{1-n} \\ a_1 & a_0 & a_{-1} & \cdot & \cdot \\ \cdot & a_1 & a_0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & a_{-1} \\ a_{n-1} & \cdot & \cdot & a_1 & a_0 \end{bmatrix}}_{\mathbf{B}} \quad \underbrace{\begin{bmatrix} 0 & 0 & \cdot & 0 & 1 \\ 0 & \cdot & 0 & 1 & 0 \\ \cdot & 0 & 1 & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 0 & \cdot & 0 & 0 \end{bmatrix}}_{\mathbf{Y}}$$

is the real *symmetric* (Hankel) matrix

$$\begin{bmatrix} a_{1-n} & \cdot & \cdot & a_{-1} & a_0 \\ \cdot & \cdot & a_{-1} & a_0 & a_1 \\ \cdot & \cdot & a_0 & a_1 & \cdot \\ a_{-1} & \cdot & \cdot & \cdot & \cdot \\ a_0 & a_1 & \cdot & \cdot & a_{n-1} \end{bmatrix}$$

so can use MINRES with positive definite preconditioner

Toeplitz  $\rightarrow$  Circulant approximation,  $C$ , and diagonalization  $C = F^* \Lambda F$  by the FFT ( $O(n \log n)$ ) is a well known approach.

To ensure positive definiteness, take the absolute value circulant  $|C| = F^* |\Lambda| F$

Theorem (*Pestana & W, 2015*)

$$|C|^{-1} B Y = J + R + E$$

where  $J$  is real symmetric and orthogonal with eigenvalues  $\pm 1$ ,  $R$  is of small rank and  $E$  is of small norm

$\Rightarrow$  guaranteed fast convergence because MINRES convergence only depends on eigenvalues which are clustered around  $\pm 1$  except for few outliers!

‘Krylov friendly’  $\rightarrow$  try GMRES simply with  $C$ ; generally converges in half the number of iterations

Natural to investigate how sensitive to exact Toeplitz structure.

Our simple test problem is the IVP:

$$y'(t) = ay(t) + g(t), \quad y(0) = y_0, \quad t \in [0, T]$$

approximated by the  $\theta$ -method:

$$(-1 - a(1 - \theta)h)y_n + (1 - a\theta h)y_{n+1} = hg(t_n)$$

and the BDF2 method

$$y_{n+2} - (4/3)y_{n+1} + (1/3)y_n = (2/3)hg(t_{n+2})$$

where for a regular mesh we have

$$t_n = nh, \quad n = 0, 1, \dots, N \quad \text{and} \quad Nh = T.$$

Both are written in monolithic (all-at-once) form yielding

$$Bx = b$$

where for example for the  $\theta$ -method with a regular grid we have the exact Toeplitz matrix

$$B = \begin{bmatrix} b & & & & & \\ c & b & & & & \\ & c & b & & & \\ & & \ddots & \ddots & & \\ & & & c & b & \\ & & & & & b \end{bmatrix},$$

$b = 1 - a\theta h$ ,  $c = -1 - a(1 - \theta)h$  and the vector  $x$  contains the solution at all time steps.



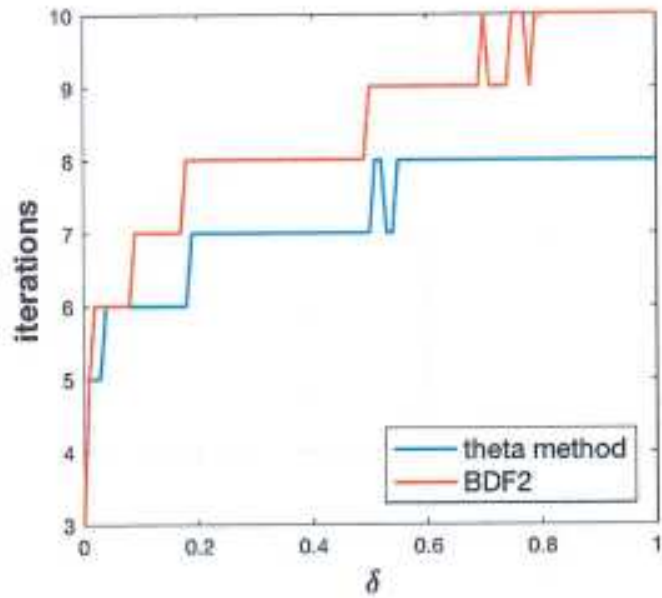
For an irregular grid we set

$$t_n = n h + \delta \mathit{Rand} h, \quad n = 1, 2, \dots, N - 1$$

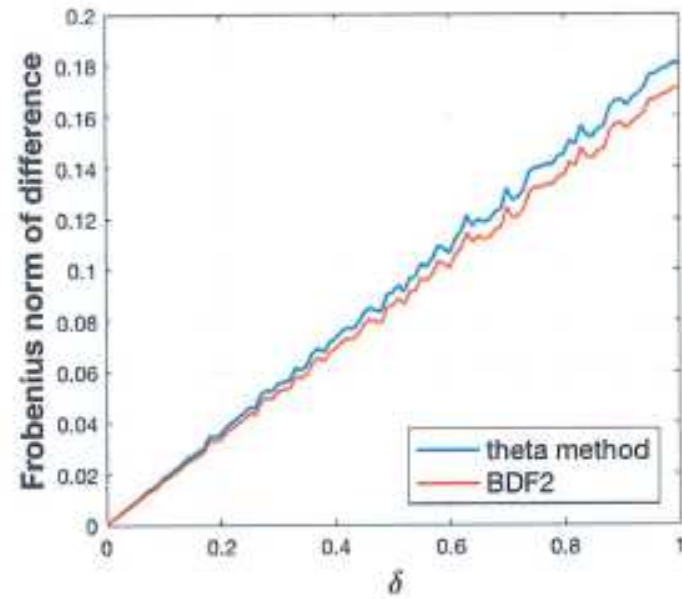
where  $\mathit{Rand}$  is, for each  $n$ , selected from a uniform random distribution on  $[-1/2, 1/2]$ .

$\delta = 0$  gives a regular grid and therefore a Toeplitz matrix, whereas increasingly irregular grids result for larger values of  $\delta$ . For increasing  $\delta$  the coefficient matrix is further from Toeplitz.

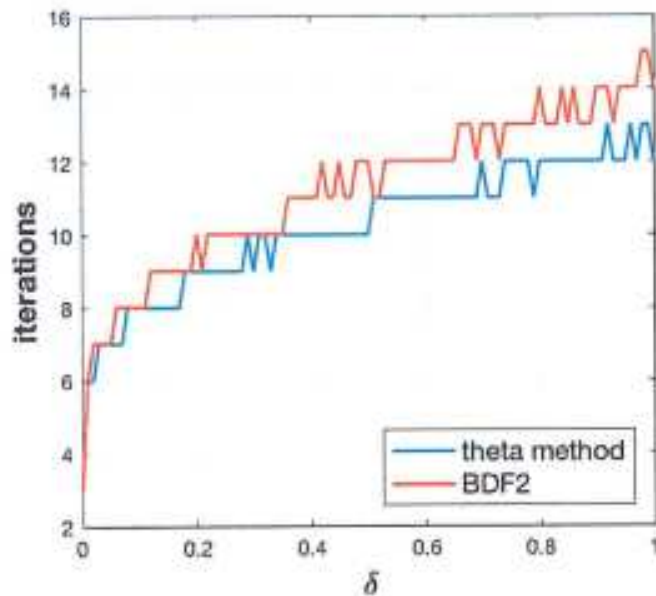
We precondition with the Strang circulant for the matrix obtained by averaging along each diagonal and solve for  $T = 10$  with  $h = 0.01$  and  $h = 0.1$ .



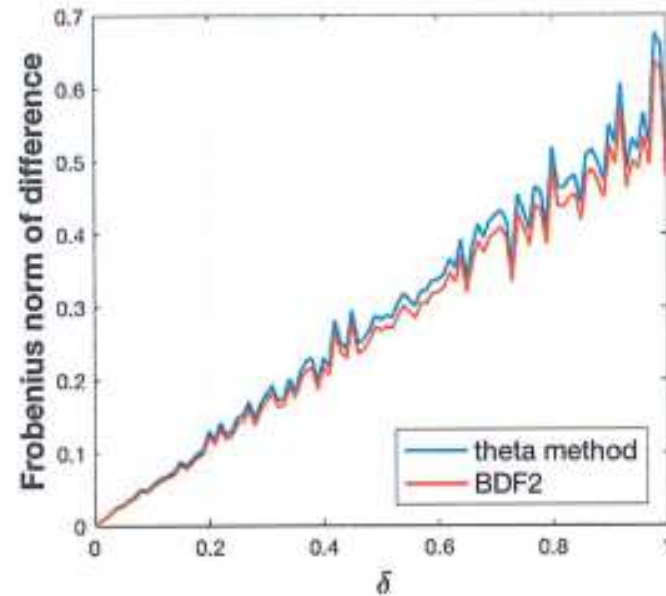
(a) Number of iterations of GMRES ( $h = 0.01$ ).



(b) Frobenius norm of  $D$  ( $h = 0.01$ ).



(c) Number of iterations of GMRES ( $h = 0.1$ ).



(d) Frobenius norm of  $D$  ( $h = 0.1$ ).

For a genuinely irregular grid—100 grid points generated randomly in  $[0, T]$ —100 iterations are required as might be expected for a Krylov method.

In a further experiment we randomly generate matrices of dimension 100 with differing distance from a circulant:

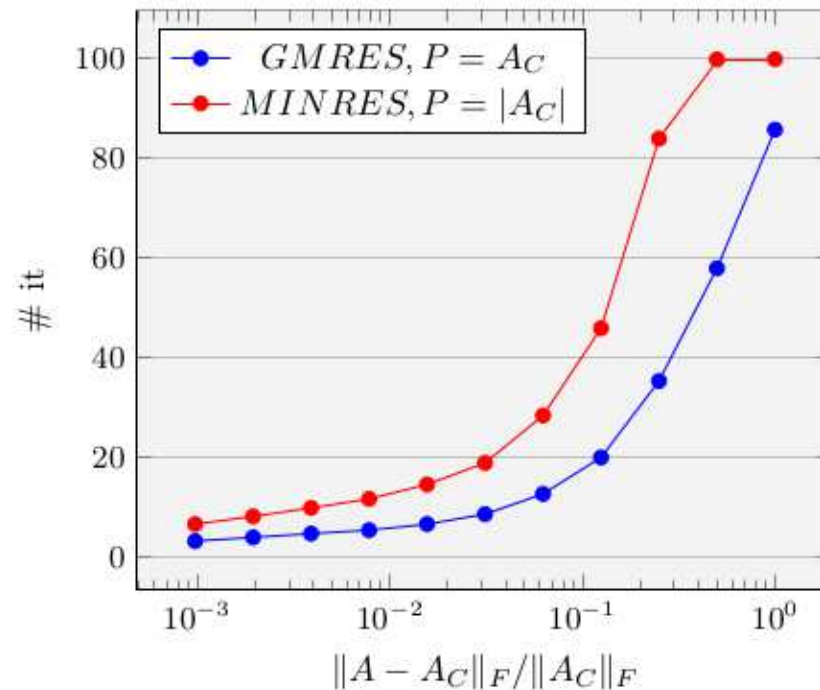


Figure 4: Average number of GMRES (blue) and MINRES (red) iterations to convergence applied to 20 dense symmetric randomly generated (uniform distributions) matrices  $A$  of size 100. As a preconditioner, the closest Circulant  $A_C$  is used for GMRES, while its absolute value  $|A_C|$  for MINRES. Different values of the relative norm of the non-Circulant component of  $A$ ,  $\|A - A_C\|_F / \|A_C\|_F$  are considered. Convergence reached with residual norm  $\leq 10^{-6}$ .

# Summary

The property of circulant matrices as excellent preconditioners for Toeplitz matrices to some extent applies for perturbations of Toeplitz matrices even in the nonsymmetric case.

# Summary

The property of circulant matrices as excellent preconditioners for Toeplitz matrices to some extent applies for perturbations of Toeplitz matrices even in the nonsymmetric case.

Reference:

Pestana, J. & Wathen, A.J., 2015,  
'A preconditioned MINRES method for nonsymmetric  
Toeplitz matrices',

SIAM J. Matrix Anal. Appl. **36**, pp. 273–288.