

3. COMPUTING THE PARTICULAR AND HOMOGENEOUS SOLUTIONS

We find the three solutions \mathbf{x}_p^R , \mathbf{x}_p^{UH} , and \mathbf{x}_p^{LH} by solving Eqs. (2)-(4). Exploiting overlapping calculations and elements with value 0 gives the following algorithm, with $13M$ binary floating point operations:

Forward elimination:

$$\begin{aligned} \omega_1 &= \frac{c_1}{b_1} & \omega_i &= \frac{c_i}{b_i - a_i \omega_{i-1}} & i &= 2, 3, \dots, M \\ \gamma_1 &= \frac{r_1}{b_1} & \gamma_i &= \frac{r_i - a_i \gamma_{i-1}}{b_i - a_i \omega_{i-1}} & i &= 2, 3, \dots, M \end{aligned}$$

Back substitution:

$$\begin{aligned} x_M^R &= \gamma_M & x_i^R &= \gamma_i - \omega_i x_{i+1}^R & i &= M-1, M-2, \dots, 1 \\ x_M^{LH} &= -\omega_M & x_i^{LH} &= -\omega_i x_{i+1}^{LH} & i &= M-1, M-2, \dots, 1 \\ \omega_M^{UH} &= \frac{a_M}{b_M} & \omega_i^{UH} &= \frac{a_i}{b_i - c_i \omega_{i+1}^{UH}} & i &= M-1, M-2, \dots, 1 \end{aligned}$$

Forward substitution:

$$x_1^{UH} = -\omega_1^{UH} \quad x_i^{UH} = -\omega_i^{UH} x_{i-1}^{UH} \quad i = 2, 3, \dots, M,$$

where the processor index p is implicitly present on all variables, and end elements a_1 and c_M are written in the appropriate positions in the a and c arrays. The sample code in ref. [1] implements this with no temporary storage arrays.

4. CONSTRUCTION AND SOLUTION OF THE REDUCED MATRIX

Once each processor has determined \mathbf{x}_p^R , \mathbf{x}_p^{UH} , and \mathbf{x}_p^{LH} , we construct and solve the reduced system of Eq. (6). We assume that the following functions are available for interprocessor communication:

- `Send (ToPid, data, n)`: When invoked by processor `FromPid`, the array `data` of length `n` is sent to processor `ToPid`. `Send()` is *nonblocking*.
- `Receive (FromPid, data, n)`: To complete data transmission, processor `ToPid` invokes `Receive()`. Upon execution, the array sent by processor `FromPid` is stored in the array `data` array of length `n`. `Receive()` is *blocking* (the processor waits for the data to be received before continuing).

Opening interprocessor communications is generally the most time-consuming step in the entire tridiagonal solution process, so it is important to minimize this. The following algorithm consumes a time of $T = (\log_2 P)t_c$ in opening communication channels (where t_c is the time to open one channel).

1. Each processor writes whatever data it has that is relevant to Eq. (6) in the array `OutData`.
2. The `OutData` arrays from each processor are concatenated as follows (Fig. 1):
 - (a) Each processor p sends its `OutData` array to processor $p - 1 \pmod{P}$, and receives a corresponding array from processor $p + 1 \pmod{P}$, as depicted in Fig. 1a. The incoming array is concatenated to the end of `OutData`.
 - (b) At the i^{th} step, repeat the first step, except sending to processor $p - 2^{i-1} \pmod{P}$, and receiving from processor $p + 2^{i-1} \pmod{P}$ (Fig. 1b,c), for $i = 1, 2, \dots$. After $\log_2 P$ iterations (or the next higher integer), each processor has the contents of the reduced matrix in the `OutData` array.
3. Each processor rearranges the contents of its `OutData` array into a local copy of the reduced tridiagonal system, and then solves. At this point, each processor has all the values in Eq. (5) stored locally.

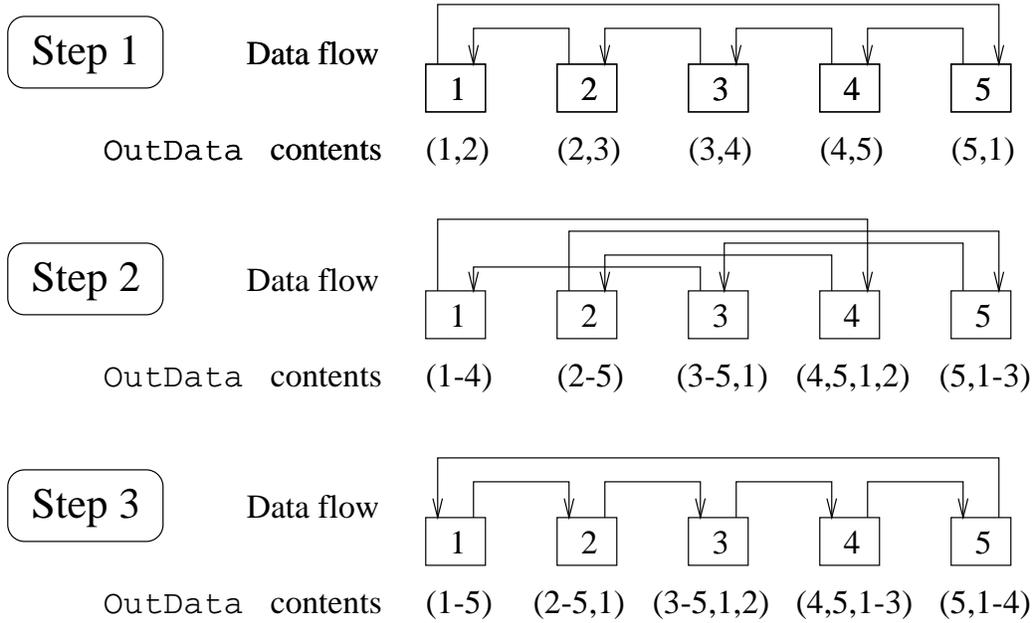


Figure 1: Illustration of the method to pass reduced matrix data between processors, shown for $P = 5$.

5. PERFORMANCE

The time consumption for this routine is as follows:

1. To calculate the three roots \mathbf{x}^R , \mathbf{x}^{UH} , and \mathbf{x}^{LH} requires $13M$ binary floating point operations by each processor, done in parallel.
2. To assemble the reduced matrix in each processor requires $\log_2 P$ steps where interprocessor communications are opened, and the i^{th} opening passes $8 \times 2^{i-1}$ real numbers.
3. Solution of the reduced system through LU decomposition requires $8(2P - 2)$ binary floating point operations by each processor, done in parallel.
4. Calculation of the final solution requires $4M$ binary floating point operations by each processor, done in parallel.

If t_b is the time of one binary floating point operation, t_c is the time required to open a communication channel (latency), and t_p is the time to pass one real number once communication is opened, then the time to execute this parallel routine is given by (optimally)

$$\begin{aligned}
 T_P &\simeq 13Mt_b + (\log_2 P)t_c + 8(P - 1)t_p + 8(2P - 2)t_b + 4Mt_b \\
 &\simeq (17M + 16P)t_b + (\log_2 P)t_c + 8Pt_p,
 \end{aligned} \tag{7}$$

for $P \gg 1$. For cases of present interest, T_P is dominated by $(\log_2 P)t_c$ and $17Mt_b$. The *parallel efficiency* is defined by $\epsilon_P \equiv \frac{T_S}{PT_P}$, where T_S is the execution time of a serial code which solves by LU decomposition. Serial LU decomposition solves an $N \times N$ system in a time $T_S = 8Nt_b$, so

$$\epsilon_P = \frac{8}{17 + 16P^2/N + (\log_2 P)Pt_c/Nt_b + 8P^2t_p/Nt_b}. \tag{8}$$

To test these claims empirically, we measured the execution times of working serial and parallel codes, and calculated ϵ_P both through its definition and through Eq. (8). Fig. 2 shows ϵ_P as a function of P for two cases, $N = 200$ and $N = 50,000$. We conclude from Fig. 2 that Eq. (8) (smooth lines) is reasonably accurate, both for the theoretical maximum efficiency (47%, achieved for small P and large N) and for the scaling with large P .

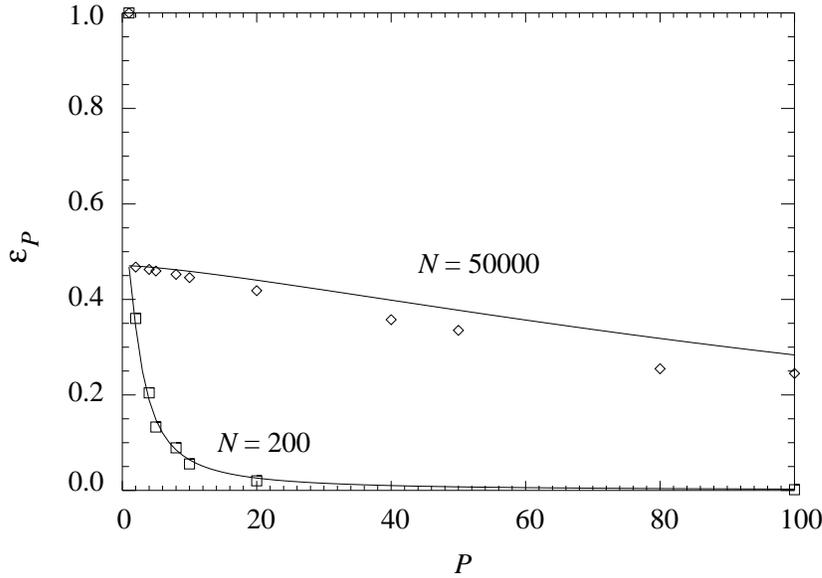


Figure 2: Results of scaling runs, comparing the parallel time with serial LU decomposition time. Here, ϵ_P is the parallel efficiency and P is the number of processors. The smooth lines represent Eq. (8), and the points are empirical results.

We made these timings on the BBN TC2000 machine at Lawrence Livermore National Laboratory, using 64-bit floating point arithmetic. This machine had 128 M88100 RISC processors, connected by a butterfly-switch network. To calculate the predictions of Eq. (7) we chose $t_c = 750\mu\text{sec}$, based on the average time of a send/receive pair measured in our code; based on other measurements, we chose the passage time of a single 64-bit real number as $t_p = 9\mu\text{sec}$; we chose $t_b = 1.4\mu\text{sec}$, based on our measured timing of 0.00218 sec for the serial algorithm on the $N = 200$ case.

6. PERIODIC TRIDIAGONAL SYSTEM

We have generalized our algorithm to a “periodic tridiagonal system.” This is a tridiagonal system with additional nonzero elements in the far upper and lower corners of the matrix, that is, Eq. (1) with Λ now of the form

$$\Lambda = \begin{pmatrix} B_1 & C_1 & & & & & A_1 \\ A_2 & B_2 & C_2 & & & & \\ & \cdot & \cdot & \cdot & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & & \\ C_N & & & & A_N & B_N & C_{N-1} \end{pmatrix}. \quad (9)$$

Solution proceeds almost precisely as before. First divide Equation (9) into tridiagonal subsystems, and solve for the particular, upper, and lower homogeneous solutions. The subsystems are again all tridiagonal, so no additional consideration need be given to this part. Then use these solutions to construct a reduced system analogous to Eq. (6). Here, however, the first and last subsystems acquire drives for the upper and lower homogeneous solutions, respectively, so the condition $\xi_1^{UH} = \xi_P^{UH} = 0$

