

Performance evaluation of a parallel algorithm for a radiative transfer problem

Paulo Beleza Vasconcelos & Filomena Dias d'Almeida



PARA'04
WORKSHOP ON STATE-OF-THE-ART
IN SCIENTIFIC COMPUTING



Joint work with:

M. Ahues, A. Largillier, O. Titaud



B. Rutily



Description of the problem: astrophysics application

- Solve the radiative transfer equation in stellar atmospheres

$$T\varphi = z\varphi + f$$

Fredholm integral equation 2nd kind

- T integral operator defined on $X = L^1(I), I = [0, \tau^*]$

$$(Tx)(\tau) = \int_0^{\tau^*} g(|\tau - \tau'|) x(\tau') d\tau'$$

- τ is the optical depth of a stellar atmosphere
- τ^* optical thickness of the atmosphere
- z is on the resolvent set of T
- $f \in L^1(I)$ is the source term

Description of the problem: astrophysics application

- g is the (weakly singular) kernel defined by

$$g(\tau) := \frac{\varpi}{2} E_1(\tau) \quad 0 < \tau \leq \tau^*$$

- $\varpi \in]0,1[$ is the albedo and

$$E_1(\tau) := \int_1^\infty \frac{\exp(-\tau\mu)}{\mu} d\mu, \tau > 0$$

- is the first exponential-integral function and it belongs to the family $(E_\nu)_{\nu \geq 1}$ defined by

$$E_\nu(\tau) := \int_1^\infty \frac{\exp(-\tau\mu)}{\mu^\nu} d\mu, \tau \geq 0, \nu \geq 2$$

$$E'_{\nu+1}(\tau) = -E'_\nu(\tau)$$

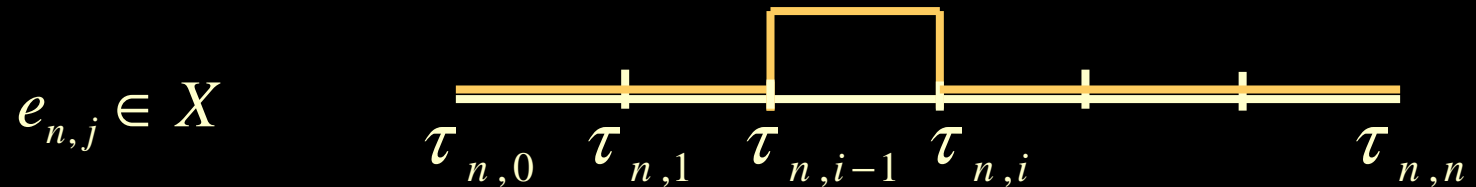
evaluated accordingly to
Abramowitz and Stegun

Projection methods

- Approximate $T\varphi = z\varphi + f$ by $T_n\varphi_n = z\varphi_n + f$

– consider a grid $0 = \tau_{n,0} < \tau_{n,1} < \dots < \tau_{n,n-1} < \tau_{n,n} = \tau^*$

– define $X_n = \text{span}(e_{n,j}, j = 1, \dots, n)$,



$$T_n x = \sum_{j=1}^n \langle x, \ell_{n,j} \rangle e_{n,j}, \quad \ell_{n,j} = T^* e_{n,j}^*$$

– where $e_{n,j}^*$ is the adjoint basis of $e_{n,j}$ in X^*

Approximate solution

- The solution of the approximate problem

$$T_n \varphi_n = z \varphi_n + f$$

- leads to the solution of a linear system with n eq's and n unknowns

$$(A_n - zI_n)x_n = b_n$$

– A_n is the restriction of T_n to X_n : $A_n = \left(\langle e_{n,j}, \ell_{n,i} \rangle \right)_{i,j=1}^n$

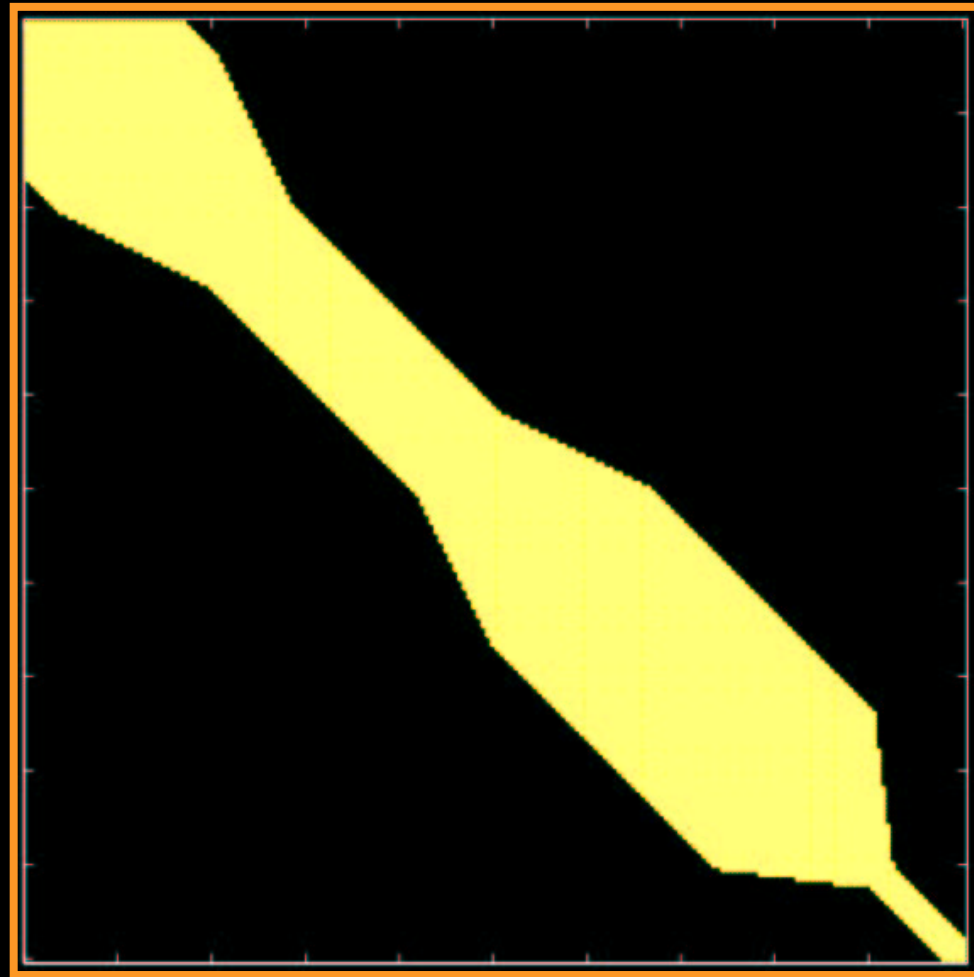
$$b_n = \left(\langle f, \ell_{n,i} \rangle \right)_{i=1}^n \quad x_n = \left(\langle \varphi_n, \ell_{n,i} \rangle \right)_{i=1}^n$$

- we recover φ_n from x_n by $\varphi_n = \frac{1}{z} \left(\sum_{j=1}^n x_n(j) e_{n,j} - f \right)$

Typical coefficient matrix

$$A_n - zI_n, \quad z = 1$$

band matrix
sparse matrix



Approximate solution

- How to solve $T_n \varphi_n = z \varphi_n + f$ when the associated coefficient matrix $A_n - zI_n$ has large dimension?
 - one can use preconditioned nonstationary iterative methods, or
 - an iterative refinement method (Newton-type method)

$$\begin{cases} \text{given } x^{(0)} \\ x^{(k+1)} = x^{(k)} - (T - zI)^{-1} (Tx^{(k)} - zx^{(k)} - f) \end{cases}$$

Iterative refinement methods

- Jacobian $R(z) = (T - zI)^{-1}$ can be approximated by
 - scheme 1: $R_n(z) := (T_n - zI)^{-1}$
 - Atkinson's algorithm
 - scheme 2: $\frac{1}{z} (TR_n(z) - I)$
 - Brakhage's algorithm
 - scheme 3: $\frac{1}{z} (R_n(z)T - I)$
 - new scheme (M.Ahues, A.Largillier)

Iterative refinement methods

Scheme 1:

$$\begin{cases} x^{(0)} = \varphi_n = R_n(z)f \\ x^{(k+1)} = x^{(0)} + R_n(z)(T_n x^{(k)} - Tx^{(k)}) \end{cases}$$

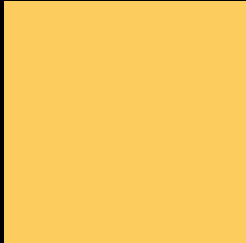


Scheme 2:

$$\begin{cases} x^{(0)} = \frac{1}{z}(R_n(z)T - I)f \\ x^{(k+1)} = x^{(0)} + \frac{1}{z}(R_n(z)T - I)(T_n x^{(k)} - Tx^{(k)}) \end{cases}$$

Scheme 3:

$$\begin{cases} x^{(0)} = \frac{1}{z}(T R_n(z) - I)f \\ x^{(k+1)} = x^{(0)} + \frac{1}{z}(T R_n(z) - I)(T_n x^{(k)} - Tx^{(k)}) \end{cases}$$

Iterative refinement method

- In practice T is not used. The problem is restricted to X_m , $m \gg n$, considering a finer projection discretization of T , T_m ($m \times m$)
- T_m restricted to X_m : $A_m = \left(\langle e_{m,j}, \ell_{m,i} \rangle \right)_{i,j=1}^m$ 
- T_m restricted to X_n : $C = \left(\langle e_{m,j}, \ell_{n,i} \rangle \right)_{i,j=1}^{n,m}$ ($n \times m$)

- T_n restricted to X_m : $D = \left(\langle e_{n,j}, \ell_{m,i} \rangle \right)_{i,j=1}^{m,n}$ ($m \times n$)


Atkinson's scheme

given $A_n, A_m, C, D, x_n^{(0)}, x_m^{(0)}, z$
repeat until convergence

$$y_n = A_n x_n^{(k)} - C x_m^{(k)}$$

$$\text{solve } (A_n - zI) w_n = y_n$$

$$w_m = \frac{1}{z} \left(D(w_n - x_n^{(k)}) + A_m x_m^{(k)} \right)$$

$$x_n^{(k+1)} = x_n^{(0)} + w_n$$

$$x_m^{(k+1)} = x_m^{(0)} + w_m$$

$$k = k + 1$$

prolong. w_n

band block LU
or
sparse iterative methods

update
 x_n and x_m

Matrix coefficients

$$\begin{aligned} M_{\alpha,\beta}(i,j) &= \frac{\bar{\omega}}{2h_{\alpha,i}} \int_{\tau_{\alpha,i-1}}^{\tau_{\alpha,i}} \int_0^{\tau^*} E_1(|\tau - \tau'|) e_{\beta,j}(\tau') d\tau' d\tau \\ &= \frac{\bar{\omega}}{2h_{\alpha,i}} \int_{\tau_{\alpha,i-1}}^{\tau_{\alpha,i}} \int_{\tau_{\beta,j-1}}^{\tau_{\beta,j}} E_1(|\tau - \tau'|) d\tau' d\tau \end{aligned}$$

$$\begin{aligned} \alpha = \beta = n & : A_n = M_{n,n} \\ \alpha = \beta = m & : A_m = M_{m,m} \\ \alpha = n, \beta = m & : C = M_{n,m} \\ \alpha = m, \beta = n & : D = M_{m,n} \end{aligned}$$

$$\begin{aligned} h_{\alpha,i} &= \tau_{\alpha,i} - \tau_{\alpha,i-1}, \quad i = 1, \dots, \alpha \\ h_{\beta,j} &= \tau_{\beta,j} - \tau_{\beta,j-1}, \quad j = 1, \dots, \beta \end{aligned}$$

Matrix coefficients: A_n and A_m

grid $(\tau_{n,j})_{j=0}^n$ defined on $[0, \tau^*]$, for $i, j \in [1, n]$

$$A_n(i, j) = \frac{\varpi}{2h_{n,i}} \int_{\tau_{n,i-1}}^{\tau_{n,i}} \int_0^{\tau^*} E_1(|\tau - \tau'|) e_{n,j}(\tau') d\tau' d\tau$$

$$= \begin{cases} \frac{\varpi}{2h_{n,i}} [E_3(d_{n,i-1,j}) - E_3(d_{n,i-1,j-1}) + E_3(d_{n,i,j-1}) - E_3(d_{n,i,j})], & i \neq j \\ \varpi \left(1 + \frac{1}{h_{n,j}} \left[E_3 \left(h_{n,j} - \frac{1}{2} \right) \right] \right) & , i = j \end{cases}$$

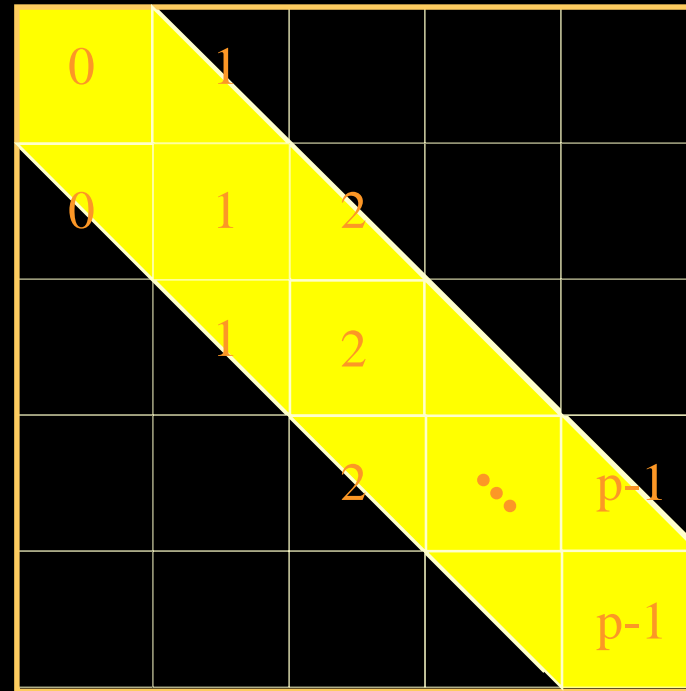
$$d_{n,i,j} = |\tau_{n,i} - \tau_{n,j}|, \quad i, j \in [0, n] \quad h_{n,j} = \tau_{n,j} - \tau_{n,j-1}, \quad j \in [1, n]$$



repeat for m instead of n to obtain $A_m(i, j)$

Data distribution

- Distributed memory environment with p processors:
 $p=0, \dots, p-1$



Matrix coefficients: C and D

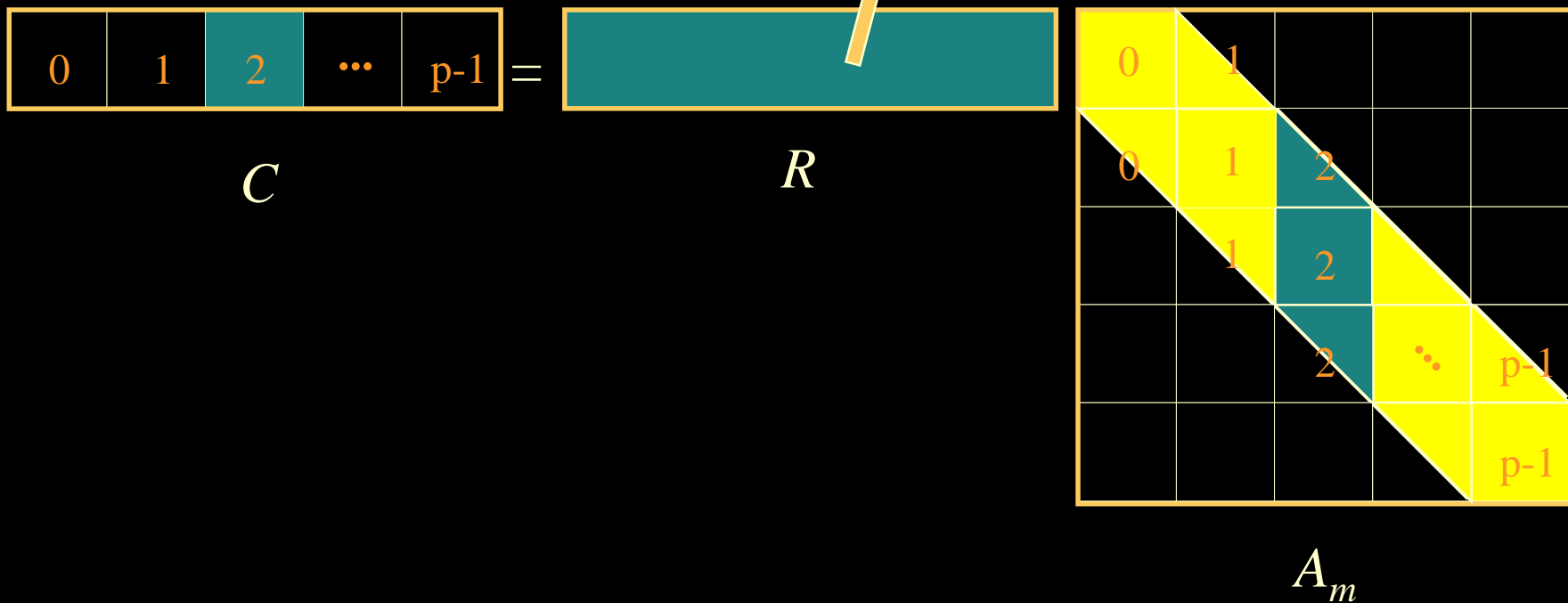
- We can compute matrices C and D :
 1. through the integral formulation (similar to A_n and A_m)
 2. or by doing some matrix computations using A_m

$$C = RA_m, \quad R(i, k) = \begin{cases} \frac{h_{m,k}}{h_{n,i}} & r \times (i-1) + 1 \leq k \leq r \times i \\ 0 & \text{otherwise} \end{cases}$$

$$D = A_m P, \quad P(k, j) = \begin{cases} 1 & r \times (j-1) + 1 \leq k \leq r \times j \\ 0 & \text{otherwise} \end{cases}$$

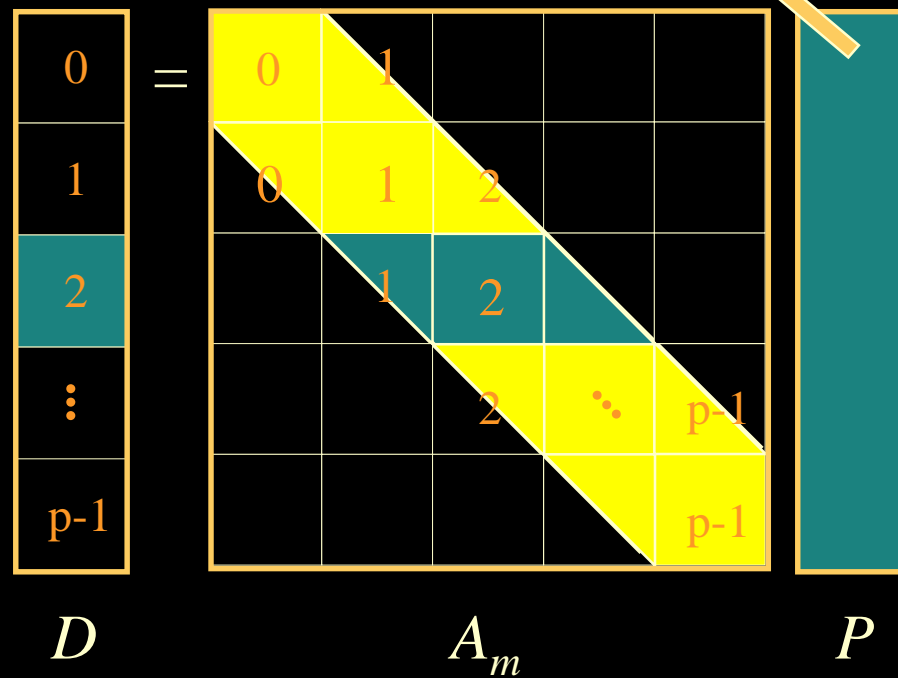
Matrix coefficients: C

$$C = RA_m, \quad R(i, k) = \begin{cases} h_{m,k} & r \times (i-1) + 1 \leq k \leq r \times i \\ h_{n,i} & \\ 0 & \text{otherwise} \end{cases}$$



Matrix coefficients: D

$$D = A_m P, \quad P(k, j) = \begin{cases} 1 & r \times (j-1) + 1 \leq k \leq r \times j \\ 0 & \text{otherwise} \end{cases}$$



Matrix coefficients: D

$$D(i, j) = \frac{\bar{\omega}}{2h_{m,i}} \int_{\tau_{m,i-1}}^{\tau_{m,i}} \int_{\tau_{n,j-1}}^{\tau_{n,j}} E_1(|\tau - \tau'|) d\tau' d\tau$$

$$= \begin{cases} \frac{\bar{\omega}}{2h_{m,i}} \left[-E_3(\tau_{n,j} - \tau_{m,i}) + E_3(\tau_{n,j} - \tau_{m,i-1}) + E_3(\tau_{n,j-1} - \tau_{m,i}) - \right. \\ \left. - E_3(\tau_{n,j-1} - \tau_{m,i-1}) \right] , & \tau_{m,i-1} \leq \tau_{m,i} \leq \tau_{n,j-1} \leq \tau_{n,j} \\ \frac{\bar{\omega}}{2h_{m,i}} \left[-E_3(\tau_{m,i} - \tau_{n,j}) + E_3(\tau_{m,i-1} - \tau_{n,j}) + E_3(\tau_{m,i} - \tau_{n,j-1}) + \right. \\ \left. - E_3(\tau_{m,i-1} - \tau_{n,j-1}) \right] , & \tau_{n,j-1} \leq \tau_{n,j} \leq \tau_{m,i-1} \leq \tau_{m,i} \\ \bar{\omega} \left[1 + \frac{1}{2h_{m,i}} \left(-E_3(\tau_{n,j} - \tau_{m,i}) + E_3(\tau_{n,j} - \tau_{m,i-1}) + E_3(\tau_{m,i} - \tau_{n,j-1}) - \right. \right. \\ \left. \left. - E_3(\tau_{m,i-1} - \tau_{n,j-1}) \right) \right] , & \tau_{n,j-1} \leq \tau_{m,i-1} \leq \tau_{m,i} \leq \tau_{n,j} \end{cases}$$

Atkinson's parallel scheme

given $A_n, x_n^{(0)}, x_m^{(0)}, z$
 repeat until convergence

receive $C[i] * x_m[i]$ from all P_i

$$y_n = A_n x_n^{(k)} - \sum_{i=1}^{p-1} C[i] * x_m[i]$$

$$\text{solve } (A_n - zI) w_n = y_n$$

compute $w_n - x_n^{(k)}$

receive $A_m[i] * x_m^{(k)}$

$$y_m = y_m - \sum_{i=1}^{p-1} A_m[i] * x_m^{(k)}$$

receive $D[i] * (w_n - x_n^{(k)})$ from all P_i

in location $w_m[(i-1) * n + 1 : i * n]$

$$w_m = \frac{1}{z} (w_n - y_m)$$

$$x_n^{(k+1)} = x_n^{(0)} + w_n; x_m^{(k+1)} = x_m^{(0)} + w_m$$

$k = k + 1$

given $A_m[i], C[i], D[i], z, x_n^{(0)}, x_m^{(0)}$
 repeat until convergence

compute $C[i] * x_m^{(k)}$

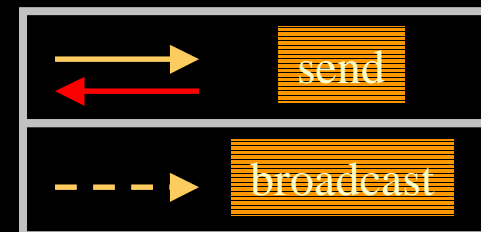
compute $A_m[i] * x_m^{(k)}$

receive $(w_n - x_n^{(k)})$ from P_0

compute $D[i] * (w_n - x_n^{(k)})$

receive $x_n^{(k+1)}$ and $x_m^{(k+1)}$ from P_0

$k = k + 1$

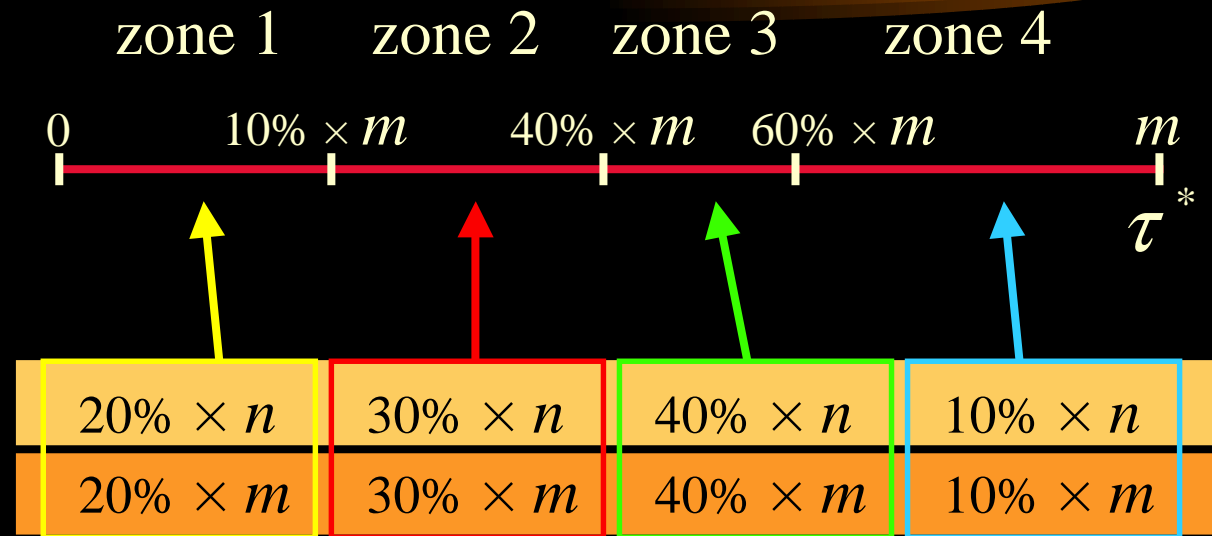


Problem specification

- Grid τ^* :

$$f(\tau) = \begin{cases} -1 & \text{if } 0 \leq \tau \leq \frac{\tau^*}{2} \\ 0 & \text{if } \frac{\tau^*}{2} < \tau \leq \tau^* \end{cases}$$

n. of points:



- parameters: $z = 1, \bar{\omega} = 0.75$; $tol : \epsilon \leq 10^{-12}$
- machine: *Beowulf cluster* with 20 PIII 450 MHz processors (128 Mb RAM) connected via fast Ethernet switch (100 Mbps)
- software: MPI, F77 & F95, LAPACK, Sparskit

Nb. of iterations and elapsed times for the 3 iterative refinement methods

- $m = 10000, n = 1000$

	scheme 1	scheme 2	scheme 3
p	43 iterations	21 iterations	21 iterations
1	3472.7	3472.4	3472.2
2	1749.4	1749.0	1748.9
4	885.0	884.4	884.4
5	714.1	713.6	713.6
10	375.3	375.1	375.1

- $m = 100000, n = 5000$

20	14859.7	14855.0	14854.2
----	---------	---------	---------

Speedup and efficiency for the 3 iterative refinement methods

- $m = 10000, n = 1000$

$$S_p = \frac{t_1}{t_p}; \quad E_p = \frac{S_p}{p}; \quad t_p = \text{elapsed time using } p \text{ processors}$$

p	schemes 1-3	
	S_p	E_p
1	1	1.00
2	1.99	0.99
4	3.93	0.98
5	4.86	0.97
10	9.26	0.93

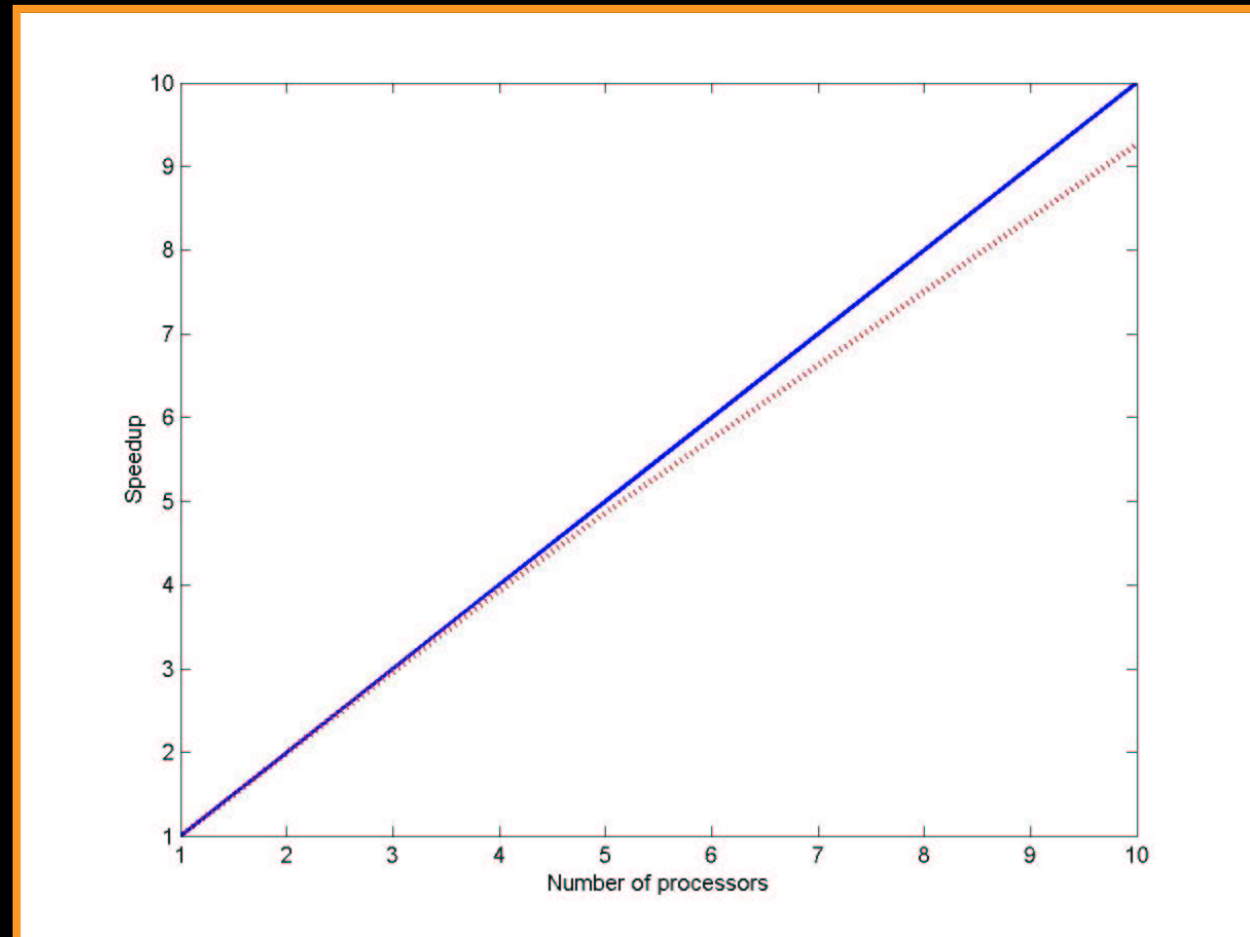
Speedup up to 10 processors

t_p = elapsed time using p processors

$$S_p = \frac{t_1}{t_p}$$

$$m = 10000$$

$$\frac{m}{n} = 10$$

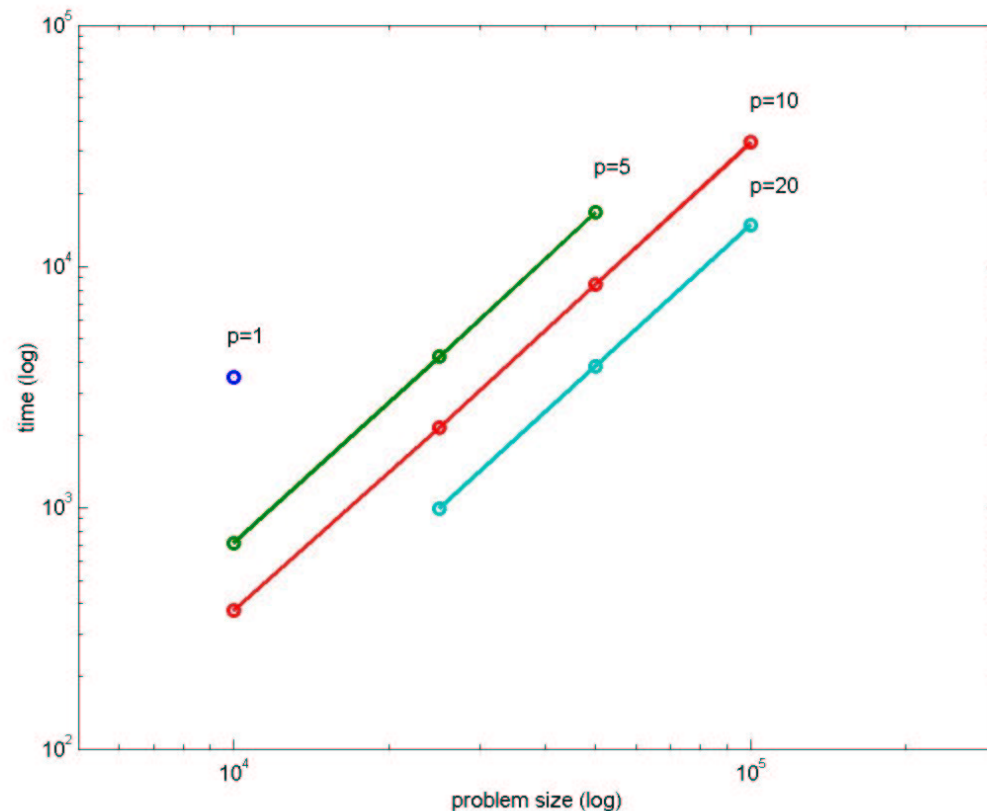


Scalability

- Isoefficiency or isogranularity (ScaLAPACK sense)
 - a constant memory use per node allows efficiency to be maintained

time for several
size problems
using up to 20
processors

$$\frac{m}{n} = 20$$



Conclusions



- We illustrate the parallel properties of this Fredholm integral equation
 - we showed it's scalability and efficiency in the solution of an Astrophysics problem
- Improvements on the computation of matrix D : **serial** vs. **parallel** approach
- For larger values of τ^* : \Rightarrow powerful parallel machine

Main references



- M. Ahues, A. Largillier and O. Titaud, The roles of a weak singularity and the grid uniformity in the relative error bounds, *Numer. Funct. Anal. And Optimiz.*, 22 7&8 789-814, 2001
- M. Ahues, F. D. d'Almeida, A. Largillier, O. Titaud and P. Vasconcelos, An L^1 refined projection approximate solution of the radiation transfer equation in stellar atmospheres, *J. Comput. Appl. Math.*, 140 13-26, 2002
- M. Ahues, F. D. d'Almeida, A. Largillier, O. Titaud and P. Vasconcelos, Iterative refinement schemes for an ill-conditioned transfer equation in astrophysics, *Algorithms for approximation IV*, Univ. of Huddersfield, 70-77, 2002
- I. W. Busbridge, *The Mathematics of Radiative Transfer*, Cambridge University Press, Cambridge, 1960
- B. Rutily, Multiple scattering theoretical and integral equations, *Integral methods in Science and Engineering: Analytic and Numerical Techniques*, Birkhäuser, 211-231, 2004