Evaluation of Linear Solvers for an Astrophysics Problem

P.B. Vasconcelos

U. PORTO

CMUP
Centro de **Matemática**
Universidade do Porto

O.A. Marques

Berkeley Lab
75 Years of World-Class Science 1931-2006

Description of the astrophysics problem

- Solve radiative transfer equation in stellar atmospheres

$$\boxed{T\varphi = z\varphi + f}$$

Fredholm integral equation 2^{nd} kind

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

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

  - 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 astrophysics problem

- g is the kernel defined by $g(\tau) := \dfrac{\varpi}{2} E_1(\tau),\ 0 < \tau \le \tau^*$
  - $\varpi \in]0,1[$ is the albedo and
  - E_1 is the first exponential-integral function and it belongs to the family

$$E_\nu(\tau) := \int_1^\infty \frac{\exp(-\tau\mu)}{\mu^\nu} d\mu \ ,\tau > 0 \ ,\nu \ge 1$$

$$E'_{\nu+1}(\tau) = -E_\nu(\tau);\ E_\nu(0) = \frac{1}{\nu - 1}, \nu > 1$$

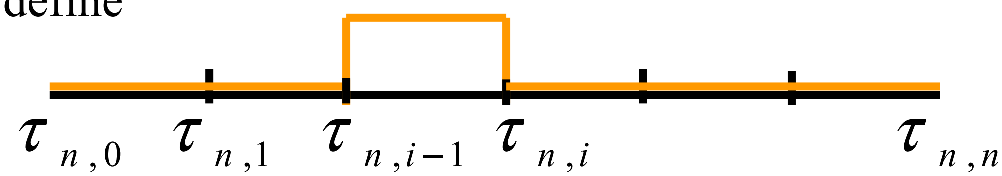
  - g is weakly singular in the sense that

$$\lim_{\tau \to 0^+} g(\tau) = +\infty;\ g \in C^0(]0,\tau^*]) \cap X;\ \sup_{\tau \in [0,\tau^*]} \int_0^{\tau^*} g(|\tau - \tau'|)\, d\tau' < \infty$$

$$g(\tau) > 0 \text{ for all } \tau \in]0,\tau^*];\ g \text{ decreasing function on }]0,\tau^*]$$

Projection method: Kantorovich

- Approximate $\boxed{T\varphi = z\varphi + f}$ by $\boxed{T_n\varphi_n = z\varphi_n + f}$
  - consider a grid $0 = \tau_{n,0} < \tau_{n,1} < \cdots < \tau_{n,n-1} < \tau_{n,n} = \tau^*$
  - define



$$X_n = span\{e_{n,j},\ j = 1,\ldots,n\},\qquad e_{n,j} \in X$$

- Let π_n be the projection op. $\pi_n x = \sum_{j=1}^{n} \langle x, e^*_{n,j} \rangle e_{n,j}$

$$T_n x = \pi_n T x = \sum_{j=1}^{n} \langle x, \ell_{n,j} \rangle e_{nj},\ \ell_{n,j} = T^* e^*_{n,j}$$

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

Matrix formulation

- 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

$$\boxed{(A_n - zI_n)x_n = b_n}$$

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

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

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

Matrix coefficients: A_n

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

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

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

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

RHS coefficients

for $i \in [1, n]$

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

$$f(\tau) = \begin{cases} -1 & if \quad 0 \le \tau \le \dfrac{\tau^*}{2} \\ 0 & if \quad \dfrac{\tau^*}{2} < \tau \le \tau^* \end{cases}$$

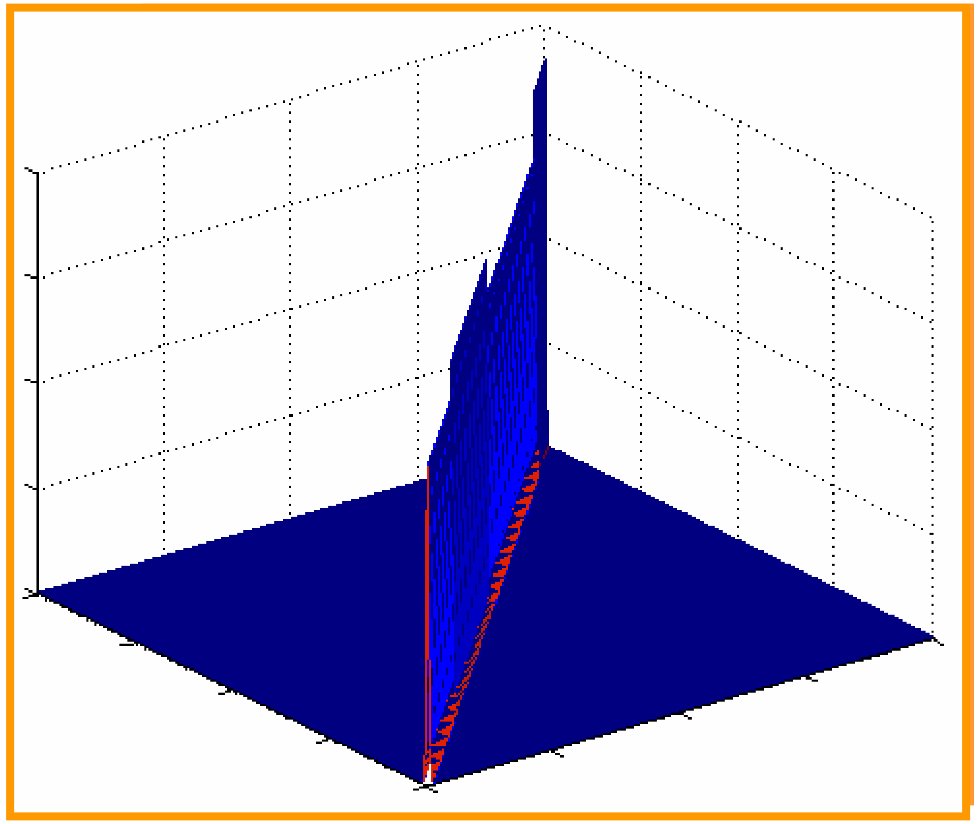
$$= \begin{cases} \dfrac{\varpi}{2h_{n,i}} \left[\begin{array}{c} E_3\left(\dfrac{\tau^*}{2} - \tau_{n,i}\right) - E_3\left(\dfrac{\tau^*}{2} - \tau_{n,i-1}\right) \cdots \\ \cdots + E_3(\tau_{n,i}) - E_3(\tau_{n,i-1}) - 2h_{n,i} \end{array} \right], & \tau_{n,i} \le \dfrac{\tau^*}{2} \\ \dfrac{\varpi}{2h_{n,i}} \left[\begin{array}{c} E_3\left(\tau_{n,i} - \dfrac{\tau^*}{2}\right) - E_3\left(\tau_{n,i-1} - \dfrac{\tau^*}{2}\right) \cdots \\ \cdots - E_3(\tau_{n,i}) + E_3(\tau_{n,i-1}) \end{array} \right], & \tau_{n,i} > \dfrac{\tau^*}{2} \end{cases}$$

Typical coefficient matrix

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

band and sparse matrix

strong decay in magnitude from the diagonal



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:
  - direct methods,
  - preconditioned nonstationary iterative methods, or
  - iterative refinement methods (Newton-type method):

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

Iterative refinement methods

- Jacobian $(T - zI)^{-1}$ can be approximated by
  - scheme A (Atkinson's algorithm): $(T_n - zI)^{-1}$
  - scheme B (Brakhage's algorithm): $\left(T(T_n - zI)^{-1} - I\right)/z$
  - scheme C (Ahues algorithm): $\left((T_n - zI)^{-1} T - I\right)/z$

- In practice T is not used. The problem is restricted the to $X_m,\ m >> n$, considering a finer projection discretization of T, T_m

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

- T_m restricted to X_n: $\boxed{C = \left(\left\langle e_{m,j}, \ell_{n,i} \right\rangle\right)_{i,j=1}^{n,m}}$ $(n \times m)$

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

Atkinson's scheme

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

repeat until convergence

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

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

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

$$\boxed{\begin{aligned} x_n^{(k+1)} &= x_n^{(0)} + w_n \\ x_m^{(k+1)} &= x_m^{(0)} + w_m \end{aligned}}$$

$$k = k + 1$$

band block LU
or
sparse iterative methods

prolong. w_n

update
x_n and x_m

Solving the problem in the m-D space

- We can solve $T_m\varphi_m = z\varphi_m + f$ for the finer grid approximated matricial problem $A_m - zI_m = b_m$
- Our goal is to experiment with robust and portable algorithm implementations (from the ACTS Collection)
- Direct methods:
  - SuperLU
- Preconditioned nonstationary iterative methods:
  - PETSc
  - Trilinos

Problem specification

- grid τ^* : nonuniform grid (4 zones)
- parameters: $z = 1$, $\varpi = 0.75$ and $\varpi = 0.9$; tol: $\varepsilon \leq 10^{-12}$
- machines: located at LBNL/NERSC
  - SGI Altix 350: 32 64-bit 1.4 GHz Intel Itanium-2 processors, with 192 GBytes of shared memory
  - AMD Opteron Cluster: 356 dual-processor nodes, 2.2 GHz/node, 6 GB/node, interconnected with a high-speed InfiniBand network
  - IBM SP: 380 compute nodes with 16 Power 3+ processors/node, 16 GB memory/node.
- software:
  - MPI, F77 & F95, PETSc, SuperLU

Normalized times for the generation phase and system solution with SuperLU, for various matrix sizes (m), on the SGI Altix

$\varpi = 0.75$	generation	solution	
m		factor	solve
1000	3,26E+03	6,95E+01	1,00E+00
2000	2,12E+04	1,65E+02	3,00E+00
4000	9,71E+04	3,59E+01	6,00E+00
8000	4,26E+05	7,51E+02	1,80E+01
16000	1,80E+06	1,54E+03	3,00E+01
32000	7,36E+06	3,12E+03	5,35E+01

Normalized times and nb. it. for various matrix sizes (m) on up to 32 processors (p) on the Opteron cluster

a constant memory use per node allows efficiency to be maintained

$\varpi = 0.75$ m	p	generation	GMRES 22 iterations	BiCGStab 14 iterations
10000	1	5,40E+03	7,54E+00	7,95E+00
	2	2,67E+03	4,02E+00	4,58E+00
	4	1,39E+03	2,32E+00	2,56E+00
	8	6,90E+02	1,80E+00	1,97E+00
	16	3,51E+02	1,15E+00	1,25E+00
	32	1,79E+02	1,15E+00	1,36E+00
25000	4	8,41E+03	5,42E+00	5,61E+00
	8	4,28E+03	3,02E+00	3,15E+00
	16	2,16E+03	2,05E+00	1,83E+00
	32	1,07E+03	1,00E+00	1,15E+00
50000	16	8,57E+03	3,14E+00	3,20E+00
	32	4,24E+03	1,53E+00	1,86E+00

Normalized times and nb. it. for various matrix sizes (m) on up to 32 processors (p) on the Opteron cluster

$\varpi = 0.90$ $\varpi = 0.75$ 22 14

m	p	generation	GMRES 24 iterations	BiCGStab 37 iterations
10000	1	2,83E+03	6,14E+00	6,99E+00
	2	1,36E+03	3,62E+00	4,12E+00
	4	7,20E+02	2,20E+00	2,31E+00
	8	3,59E+02	[illegible]	1,76E+00
	16	1,80E+02	1,11E+00	1,30E+00
	32	[illegible]	1,02E+00	1,00E+00
25000	4	[illegible]	4,96E+00	5,37E+00
	8	2,22E+03	2,96E+00	3,39E+00
	16	1,11E+03	1,78E+00	2,11E+00
	32	5,55E+02	1,33E+00	1,46E+00
50000	16	4,36E+03	2,82E+00	3,14E+00
	32	2,21E+03	2,15E+00	2,05E+00

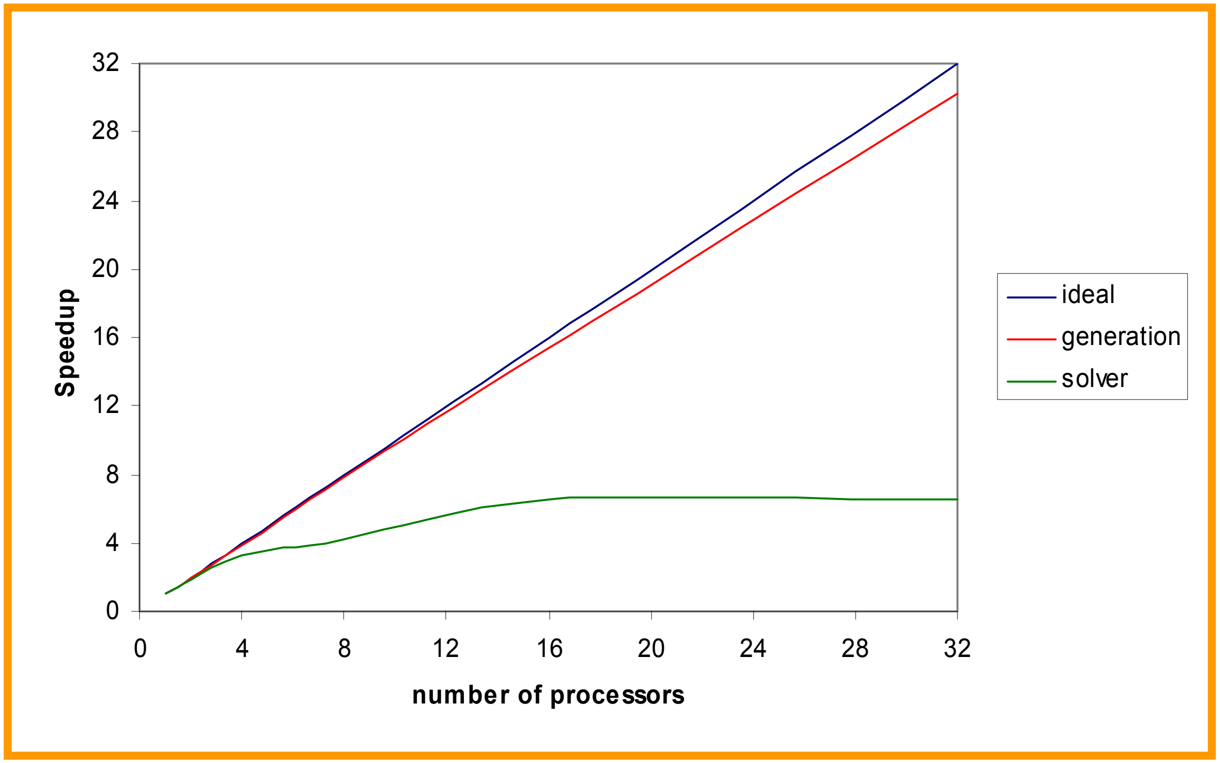
1.5x-2.5x smaller

Speedup up to 32 processors on the Opteron cluster

$$t_p = \text{elapsed time using } p \text{ processors}$$

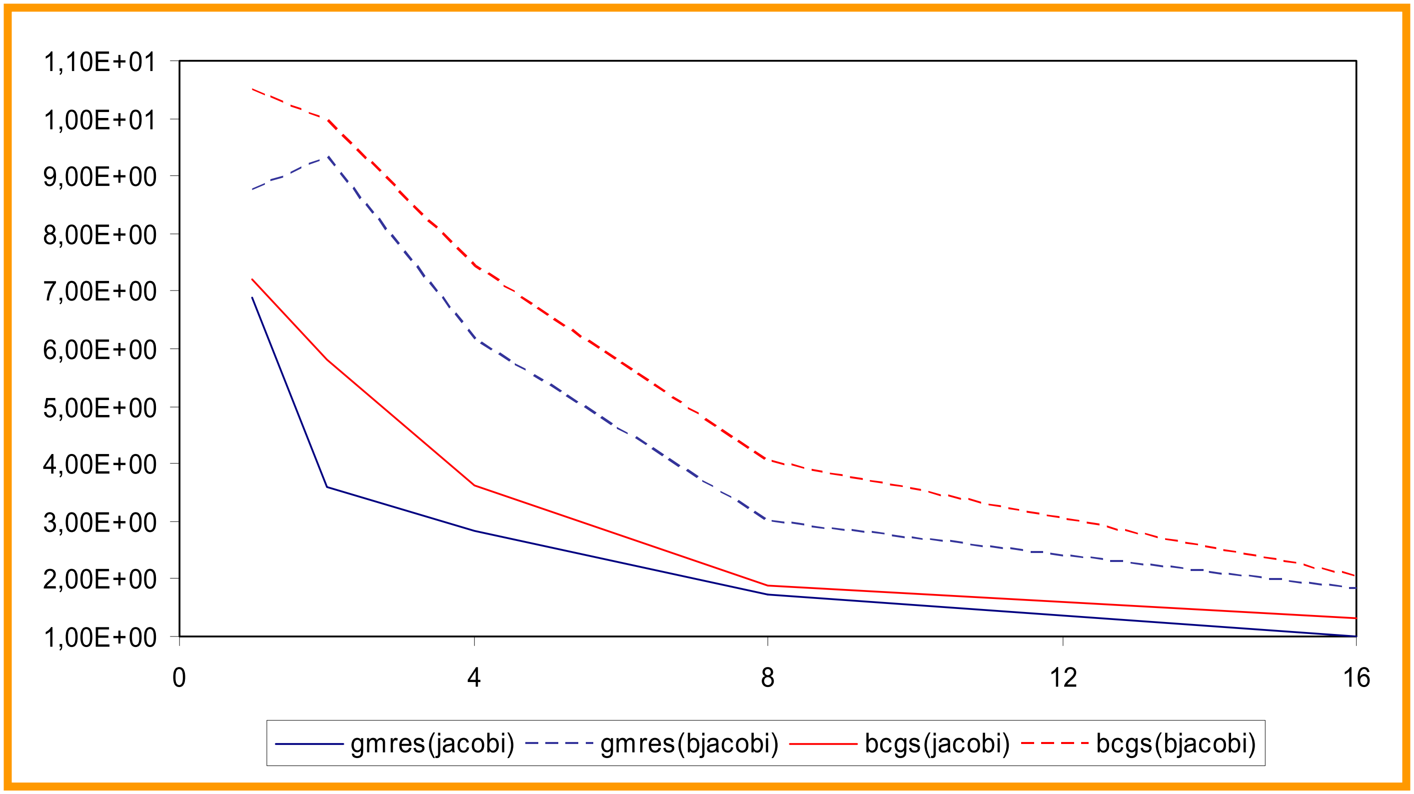
$$S_p = \frac{t_1}{t_p}$$

$$m = 10000$$



Normalized times for Jacobi and block Jacobi preconditioners on the Opteron cluster

$m = 50000$



Conclusions

- We discussed the numerical solution of a radiative transfer equation for modelling the emission of photons in stellar atmospheres.
- The parallelization of the generation phase greatly reduces the time to solution and enables the solution of large systems.
- The selection of appropriate linear solvers is important for delivering performance and portability.
- Compared to iterative refinement techniques, the present approach
  - leads to 40% savings in time in the generation phase (for m=50000 and np=5)
  - reduces the number of communications required for mapping the coarse problem into the fine one (up to 5x for Atkinson and 4x for Brakhage and Ahues' schemes for m=50000 and np=5)

Main references

- 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
- B. Rutily, Multiple scattering theoretical and integral equations, *Integral methods in Science and Engineering: Analytic and Numerical, Techniques*, Birkhaüser, 211-231, 2004
- L.A. Drummond and O. Marques, An Overview of the Advanced CompuTational Software (ACTS) Collection, *ACM TOMS*, 31:282–301, 2005
- P.B. Vasconcelos and F. D. d'Almeida, Performance evaluation of a parallel algorithm for a radiative transfer problem, *LNCS*, 3732: 864–871, 2006

Motivation

- In this work we consider the numerical solution of a radiative transfer equation for modeling the emission of photons in stellar atmospheres.
- Mathematically, the problem is formulated in terms of a weakly singular Fredholm integral equation defined on a Banach space.
- Computational approaches to solve the problem are discussed, using direct and iterative strategies that are implemented in open source packages.

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

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

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

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

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

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

$k = k + 1$

---

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 $\left(w_n - x_n^{(k)}\right)$ from P_0

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

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

$k = k + 1$

send

broadcast