

# Iterative Performance of Monte Carlo Linear Solver Methods

Massimiliano Lupo Pasini (Emory University)

In collaboration with:

Michele Benzi (Emory University)

Thomas Evans (Oak Ridge National Laboratory),

Steven Hamilton (Oak Ridge National Laboratory),

Stuart Slattery (Oak Ridge National Laboratory)

Research supported by DOE (Office of Science)

SIAM Conference on Computational Science and Engineering 2015

March 17, 2015

# Outline

- Motivations
- Standard Monte Carlo vs. Monte Carlo Synthetic Acceleration (MCSA)
- Choice of the preconditioner
- Convergence conditions
- Numerical results

# Outline

- Motivations
- Standard Monte Carlo vs. Monte Carlo Synthetic Acceleration (MCSA)
- Choice of the preconditioner
- Convergence conditions
- Numerical results

# Outline

- Motivations
- Standard Monte Carlo vs. Monte Carlo Synthetic Acceleration (MCSA)
- Choice of the preconditioner
- Convergence conditions
- Numerical results

# Outline

- Motivations
- Standard Monte Carlo vs. Monte Carlo Synthetic Acceleration (MCSA)
- Choice of the preconditioner
- Convergence conditions
- Numerical results

# Outline

- Motivations
- Standard Monte Carlo vs. Monte Carlo Synthetic Acceleration (MCSA)
- Choice of the preconditioner
- Convergence conditions
- Numerical results

# Motivations

## Goal:

- development of reliable algorithms to solve highly dimensional (towards exascale) sparse linear systems in parallel

## Issues:

- occurrence of **faults** (abnormal operating conditions of the computing system which cause a wrong answer)
  - hard faults
  - soft faults

[M. Hoemmen and M.A. Heroux, 2011; G. Bronevetsky and B.R. de Supinski, 2008; P. Prata and J.B. Silva, 1999 ]

# Motivations

**Resilience:** ability to compute a correct output in presence of faults

Different approaches to tackle the problem:

- adaptation of Krylov subspace methods (CG, GMRES, Bi-CGSTab, ...) to fault tolerance by recover-restart strategies
- abandonment of deterministic paradigm in favor of stochastic approaches

[E. Agullo, L. Giraud, A. Guernouche, J. Roman and M. Zounon, 2013]



# Monte Carlo Linear Solvers

## Mathematical setting

Let us consider a sparse linear system

$$Ax = \mathbf{b}, \quad (1)$$

where  $A \in \mathbb{R}^{n \times n}$ ,  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{b} \in \mathbb{R}^n$ .

With left preconditioning (1) becomes

$$P^{-1}Ax = P^{-1}\mathbf{b}, \quad P \in \mathbb{R}^{n \times n}. \quad (2)$$

(2) can be reinterpreted as a fixed point scheme

$$\mathbf{x} = H\mathbf{x} + \mathbf{f}, \quad H = I - P^{-1}A, \quad \mathbf{f} = P^{-1}\mathbf{b}. \quad (3)$$

Assuming  $\rho(H) < 1$ , (3) generates a sequence of approximate solutions  $\{\mathbf{x}_k\}_{k=0}^{\infty}$  which converges to the exact solution of (1).

# Monte Carlo Linear Solvers

## Mathematical setting

The solution to (3) can be written in terms of a power series in  $H$  (Neumann series):

$$\mathbf{x} = \sum_{\ell=0}^{\infty} H^{\ell} \mathbf{f}.$$

By restricting the attention to a single component of  $\mathbf{x}$  we have

$$x_i = \sum_{\ell=0}^{\infty} \sum_{k_1=1}^n \sum_{k_2=1}^n \cdots \sum_{k_{\ell}=1}^n H_{k_0, k_1} H_{k_1, k_2} \cdots H_{k_{\ell-1}, k_{\ell}} f_{k_{\ell}}. \quad (4)$$

(4) can be reinterpreted as the **sampling of an estimator defined on a random walk**.

# Monte Carlo Linear Solvers

Mathematical setting: Forward (Direct) Method

**Goal:** evaluate a functional such as

$$J(\mathbf{x}) = \langle \mathbf{h}, \mathbf{x} \rangle = \sum_{i=1}^n h_i x_i, \quad \mathbf{h} \in \mathbb{R}^n.$$

**Approach:** define random walks to evaluate  $J$ .

Consider a random walk whose state space  $S$  is the set of indices of the forcing term  $\mathbf{f}$ :

$$S = \{1, 2, \dots, n\} \subset \mathbb{N}.$$

[N. Metropolis and S. Ulam, 1949; G. E. Forsythe and A. Leibler, 1950; W. R. Wasow, 1952; J. H. Halton, 1962]

# Monte Carlo Linear Solvers

Mathematical setting: Forward (Direct) Method

Initial probability: pick  $k_0$  as initial step of the random walk:

$$\tilde{p}(k_0 = i) = \tilde{p}_{k_0} = \frac{|h_i|}{\sum_{i=1}^n |h_i|}.$$

Possible choices for the transition probability:

$$p(:, i) : S \rightarrow [0, 1] \quad \forall i \in S$$

$$p(k_i = j | k_{i-1} = i) = P_{i,j} = \frac{|H_{i,j}|}{\sum_{i=1}^n |H_{i,j}|} \quad \text{weighted}$$

$$P_{i,j} = \begin{cases} 0 & \text{if } H_{i,j} = 0 \\ \frac{1}{\#(\text{non-zeros in the row})} & \text{if } H_{i,j} \neq 0 \end{cases} \quad \text{uniform}$$

# Monte Carlo Linear Solvers

Mathematical setting: Forward (Direct) Method

A related sequence of weights is defined by

$$w_{i,j} = \frac{H_{i,j}}{P_{i,j}}$$

in order to build an auxiliary sequence

$$W_0 = \frac{h_{k_0}}{\tilde{p}_{k_0}}, \quad W_i = W_{i-1} w_{k_{i-1}, k_i} \quad k_{i-1}, k_i \in S.$$

This allows us to define a random variable  $X(\cdot) : \Pi \rightarrow \mathbb{R}$ .

$\Pi$  = set of realizations of a random walk  $\gamma$  defined on  $S$ :

$$X(\nu) = \sum_{\ell=0}^{\infty} W_{\ell} f_{k_{\ell}}, \quad \nu = \text{permutation of } \gamma$$

# Monte Carlo Linear Solvers

Mathematical setting: Forward (Direct) Method

Define the expected value of  $X(\cdot)$ :

$$E[X] = \sum_{\nu} P_{\nu} X(\nu) \quad (P_{\nu} = \text{probability of the permutation } \nu).$$

It can be proved that

$$E[W_{\ell} f_{k_{\ell}}] = \langle \mathbf{h}, H^{\ell} \mathbf{f} \rangle$$

and consequently

$$E \left[ \sum_{\ell=0}^{\infty} W_{\ell} f_{k_{\ell}} \right] = \langle \mathbf{h}, \mathbf{x} \rangle .$$

# Monte Carlo Linear Solvers

Mathematical setting: Forward (Direct) Method

If  $\mathbf{h}$  is a vector of the standard basis:  $\mathbf{h} = \mathbf{e}_j$

$$\tilde{p}(K_0 = i) = \delta_{i,j}, \quad i, j = 1, \dots, n$$

and the expected value assumes the following form:

$$E \left[ \sum_{\ell=0}^{\infty} W_{\ell} f_{\mathbf{k}_{\ell}} \right] = x_i = \sum_{\ell=0}^{\infty} \sum_{\mathbf{k}_1=1}^n \sum_{\mathbf{k}_2=1}^n \cdots \sum_{\mathbf{k}_{\ell}=1}^n P_{\mathbf{k}_0, \mathbf{k}_1} P_{\mathbf{k}_1, \mathbf{k}_2} \cdots P_{\mathbf{k}_{\ell-1}, \mathbf{k}_{\ell}} w_{\mathbf{k}_0, \mathbf{k}_1} w_{\mathbf{k}_1, \mathbf{k}_2} \cdots w_{\mathbf{k}_{\ell-1}, \mathbf{k}_{\ell}} f_{\mathbf{k}_{\ell}}. \quad (5)$$

**Drawback:** the estimator is defined entry-wise.

# Monte Carlo Linear Solvers

Mathematical setting: Adjoint Method

Consider the adjoint linear system

$$(P^{-1}A)^T \mathbf{y} = \mathbf{d}.$$

With a fixed point approach, we get

$$\mathbf{y} = H^T \mathbf{y} + \mathbf{d}.$$

By picking  $\mathbf{d} = \mathbf{e}_j$  we have

$$J^*(\mathbf{y}) := \langle \mathbf{y}, \mathbf{f} \rangle = \langle \mathbf{x}, \mathbf{e}_j \rangle = x_j, \quad j = 1, \dots, n.$$



# Monte Carlo Linear Solvers

Mathematical setting: Adjoint Method

The initial probability and transition matrix may be defined as

$$\tilde{p}_{k0} = p(k_0 = i) = \frac{|f_i|}{\sum_{i=1}^n |f_i|}, \quad P_{i,j} = \frac{|H_{i,j}^T|}{\sum_{k=1}^n |H_{i,k}^T|} = \frac{|H_{j,i}|}{\sum_{k=1}^n |H_{k,i}|}.$$

Reintroduce weights as before:

$$w_{i,j} = \frac{H_{j,i}}{P_{i,j}} \Rightarrow W_0 = \text{sign}(f_i) \|\mathbf{f}\|_1, \quad W_j = W_{j-1} w_{k_{j-1}, k_j}, \quad k_{j-1}, k_j \in S.$$

Expected value of the estimator for the entire solution vector  $\mathbf{x}$ :

$$E \left[ \sum_{\ell=0}^{\infty} W_{\ell} d_{k_{\ell}} \right] = \sum_{\ell=0}^{\infty} \sum_{k_1}^n \sum_{k_2}^n \cdots \sum_{k_{\ell}}^n f_{k_0} P_{k_0, k_1} P_{k_1, k_2} \cdots P_{k_{\ell-1}, k_{\ell}} W_{k_0, k_1} \cdots W_{k_{\ell-1}, k_{\ell}} \delta_{k_{\ell}, j}. \quad (6)$$

# Monte Carlo Linear Solvers

## Statistical constraints for convergence

Expected value and variance of the estimator must be finite to apply the Central Limit Theorem.

$$\text{Forward Method: } (H^*)_{i,j} = \frac{(H_{i,j})^2}{P_{i,j}}$$

$$E\left[\sum_{\ell=0}^{\infty} W_{\ell} f_{\mathbf{k}_{\ell}}\right] < \infty \Leftrightarrow \rho(H) < 1$$

$$\text{Var}\left[\sum_{\ell=0}^{\infty} W_{\ell} f_{\mathbf{k}_{\ell}}\right] < \infty \Leftrightarrow \rho(H^*) < 1$$

$$\text{Adjoint Method: } (H^*)_{i,j} = \frac{(H_{j,i})^2}{P_{i,j}}$$

$$E\left[\sum_{\ell=0}^{\infty} W_{\ell} d_{\mathbf{k}_{\ell}}\right] < \infty \Leftrightarrow \rho(H) < 1$$

$$\text{Var}\left[\sum_{\ell=0}^{\infty} W_{\ell} d_{\mathbf{k}_{\ell}}\right] < \infty \Leftrightarrow \rho(H^*) < 1$$

[J. H. Halton, 1994; H. Ji, M. Mascagni and Y. Li, 2013]

# Monte Carlo Linear Solvers

The choice of the transition probability

## Theorem (H. Ji, M. Mascagni and Y. Li)

*Let  $H \in \mathbb{R}^{n \times n}$ , where  $\|H\|_\infty < 1$  for the Forward method and  $\|H\|_1 < 1$  for the Adjoint method. Consider  $\nu_k$  as the realization of a random walk  $\gamma$  truncated at the  $k$ -th step. Then, there always exists a transition matrix  $P$  such that  $\text{Var}(X(\nu_k)) \rightarrow 0$  and  $\text{Var}(\sum_\nu X(\nu_k))$  is bounded as  $k \rightarrow \infty$ .*

**Remark** If the hypotheses hold, then  $P_{i,j} = \frac{|H_{i,j}|}{\sum_{k=1}^n |H_{i,k}|}$  (Forward) or  $P_{i,j} = \frac{|H_{i,j}^T|}{\sum_{k=1}^n |H_{i,k}^T|}$  (Adjoint) guarantee  $\|H^*\|_\infty < 1$  or  $\|H^*\|_1 < 1$  respectively.

# Monte Carlo Linear Solvers

## First approaches to a parallelization

First attempts to apply Monte Carlo as a linear solver considered the fixed point formulation of the original linear system.

- + embarrassingly parallelizable (ideally we could employ as many processes as the amount of histories)
- huge amount of time to compute a reasonably accurate solution

E.g.  $\|r_0\| = \|\mathbf{b} - A\mathbf{x}_0\| \approx 1$  then  $\frac{\|\mathbf{r}_k\|}{\|\mathbf{b}\|} < 10^{-m} \Rightarrow N_{histories} \approx 10^{2m}$   
because of the Central Limit Theorem.

[S. Branford, C. Sahin, A. Thandavan, C. Weihrauch, V. N. Alexandrov and I. T. Dimov, 2008; P. Jakovits, I. Kromonov and S. R. Srirama, 2011]

# Monte Carlo Linear Solvers

## Monte Carlo Synthetic Acceleration (MCSA)

**Data:**  $A, \mathbf{b}, H, \mathbf{f}, \mathbf{x}_0$

**Result:**  $\mathbf{x}_{num}$

$\mathbf{x}^I = \mathbf{x}_0;$

*while not reached convergence do*

$$\mathbf{x}^{I+\frac{1}{2}} = H\mathbf{x}^I + \mathbf{f};$$

$$\mathbf{r}^{I+\frac{1}{2}} = \mathbf{b} - A\mathbf{x}^{I+\frac{1}{2}};$$

$$\delta\mathbf{x}^{I+\frac{1}{2}} = (I - H)^{-1}\mathbf{r}^{I+\frac{1}{2}};$$

"Solved" with Standard MC

$$\mathbf{x}^{I+1} = \mathbf{x}^{I+\frac{1}{2}} + \delta\mathbf{x}^{I+\frac{1}{2}};$$

*end*

$$\mathbf{x}_{num} = \mathbf{x}^{I+1}$$

[S. Slattery, 2013 (PhD Thesis); T. M. Evans, S. R. Slattery and P. P. H. Wilson, 2013]

# Adaptivity for the number of random walks

Forward Method:

$$\theta_i \in \mathbb{R}$$

$$\theta_i = E \left[ \sum_{l=0}^{\infty} W_l b_{k_l} \right]$$

$$\sigma_i = \sqrt{\text{Var}[\theta_i]}$$

$$\text{Find } \tilde{N}_i \quad \text{s.t.} \quad \frac{\sigma_i^{\tilde{N}_i}}{|\hat{x}_i|} < \varepsilon$$

$$i = 1, \dots, n$$

Adjoint Method:

$$\boldsymbol{\theta} \in \mathbb{R}^n$$

$$\boldsymbol{\theta}_i = E \left[ \sum_{l=0}^{\infty} W_l d_{k_l} \delta_{k_l, i} \right]$$

$$\boldsymbol{\sigma}_i = \sqrt{\text{Var}[\boldsymbol{\theta}]_i}$$

$$\text{Find } \tilde{N} \quad \text{s.t.} \quad \frac{\|\boldsymbol{\sigma}^{\tilde{N}}\|_1}{\|\hat{\mathbf{x}}\|_1} < \varepsilon$$

$$i = 1, \dots, n$$

# Applicable preconditioning techniques

**Remark:** explicit knowledge of  $H_{ij}$  is needed

In fact the entry  $P_{ij}$  of the transition matrix is defined in terms of  $H_{ij}$  (Forward method) or  $H_{ji}$  (Adjoint method).

This limits the viable choices of preconditioners:

- diagonal preconditioners
- block diagonal preconditioners
- sparse approximate inverse preconditioners (AINV)

# Monte Carlo Linear Solvers

## Examples

Strictly diagonally dominant matrices (s.d.d.) with diagonal preconditioning.

$A \in \mathbb{R}^{n \times n}$  s.d.d. by rows:  $P = \text{diag}(A)$

$H = I - P^{-1}A$ ,  $\|H\|_{\infty} < 1$  ( $\Leftrightarrow \|H^*\|_{\infty} < 1$ )

$A \in \mathbb{R}^{n \times n}$  s.d.d. by columns:  $P = \text{diag}(A)$

$H = I - AP^{-1}$ ,  $\|H\|_1 < 1$  ( $\Leftrightarrow \|H^*\|_1 < 1$ )

Strict diag. dominance	Forward Method	Adjoint method
by rows	converges	not guaranteed
by columns	not guaranteed	converges



# Monte Carlo Linear Solvers

## Examples

Other examples:

- A  $M$ -matrix  $\Rightarrow \exists D$  diagonal s.t.  $AD$  is s.d.d.
- with preconditioner  $P = \text{block\_diag}(A)$ :

$$\left\{ \sum_{\substack{j=1 \\ j \neq i}}^p \|A_{ii}^{-1} A_{ij}\|_{\infty} < 1. \quad \forall i = 1, \dots, p \right. \Rightarrow \text{Forward method with block Jacobi converges}$$

$$\left\{ \sum_{\substack{i=1 \\ i \neq j}}^p \|A_{ii}^{-1} A_{ij}\|_1 < 1. \quad \forall j = 1, \dots, p \right. \Rightarrow \text{Adjoint method with block Jacobi converges}$$

## 2D parabolic problems: test case 1

Give an open and bounded set  $\Omega \subset \mathbb{R}^2$  s.t.  $\Omega = (0, 2)^2 \setminus (1, 2)^2$

$$\begin{cases} \frac{\partial u}{\partial t} - \mu \Delta u + \beta(\mathbf{x}) \cdot \nabla u = 0, & \mathbf{x} \in \Omega, \quad t \in (0, T] \\ u(\mathbf{x}, 0) = u_0, & \mathbf{x} \in \Omega \\ u(\mathbf{x}, t) = 0, & \mathbf{x} \in \Gamma_D, \quad t \in (0, T] \\ \frac{\partial u}{\partial n} + \chi u = 0, & \mathbf{x} \in \Gamma_R, \quad t \in (0, T] \end{cases} \quad (7)$$

where  $\mu = \frac{3}{200}$ ,  $\chi = 3$ ,  $\beta(\mathbf{x}) = [x, -y]^T$ .

$\Gamma_R = \{\{x = 1\} \times (1, 2)\} \cup \{(1, 2) \times \{y = 1\}\}$ ,  $\Gamma_D = \partial\Omega \setminus \Gamma_R$ .

Discretization (FreeFem++ employed):

- Triangular FEM
- spatial discretization step  $h_{max} = 0.018$ , 37,177 d.o.f.'s
- time discretization step  $\Delta t = h_{max}^2$

## 2D parabolic problems: test case 1

Preconditioning: sparse factorized AINV with drop tolerance  $\tau = 0.05$ . [M. Benzi and M. Tuma, SISC, 1998]

Numerical setting:

- Adjoint MCSA employed as linear solver
- residual relative tolerance:  $\varepsilon_1 = 10^{-7}$
- weighted transition probability
- maximal # steps per history: 10
- statistical error-based adaptive parameter:  $\varepsilon_2 = 0.5$
- granularity of the adaptive approach:  $n_{histories} = 100$
- simulations run on laptop

## 2D parabolic problems: test case 1

Stiffness matrix:  $A$  ( $n = 37,177$ ).

AINV preconditioner:  $M$ .

Iteration matrix:  $H = I - MA$ .

Solution computed just for a generic time step.

- $\frac{\text{nnz}(H)}{\text{nnz}(A)} = 2.984$
- $\|H\|_1 = 0.597$
- # iterations employed: 9
- relative error =  $3.783 \cdot 10^{-8}$
- # total random walks employed: 10,700

## 2D parabolic problems: test case 2

Give an open and bounded set  $\Omega = (0, 1) \times (0, 1)$

$$\begin{cases} \frac{\partial u}{\partial t} - \mu \Delta u + \beta(\mathbf{x}) \cdot \nabla u = 0, & \mathbf{x} \in \Omega, \quad t \in (0, T] \\ u(\mathbf{x}, 0) = u_0, & \mathbf{x} \in \Omega \\ u(\mathbf{x}, t) = 0, & \mathbf{x} \in \partial\Omega, \quad t \in (0, T] \end{cases} \quad (8)$$

where  $\mu = \frac{3}{200}$ ,  $\beta(\mathbf{x}) = [2, \sin(x)]^T$ .

Discretization (FreeFem++ employed):

- Triangular FEM
- spatial discretization step  $h_{max} = 0.014$ , 40,501 d.o.f.'s
- time discretization step  $\Delta t = h_{max}^2$

## 2D parabolic problems: test case 2

Stiffness matrix:  $A$  ( $n = 40,501$ ).

AINV preconditioner:  $M$ .

Iteration matrix:  $H = I - MA$ .

Solution computed just for a generic time step.

- $\frac{\text{nnz}(H)}{\text{nnz}(A)} = 3.106$
- $\|H\|_1 = 0.185$
- # iterations employed: 9
- relative error =  $6.904 \cdot 10^{-8}$
- # total random walks employed: 4,500

## 2D parabolic problems: test case 3

Give an open and bounded set  $\Omega = (0, 1) \times (0, 1)$

$$\begin{cases} \frac{\partial u}{\partial t} - \mu \Delta u + \beta(\mathbf{x}) \cdot \nabla u = 0, & \mathbf{x} \in \Omega, \quad t \in (0, T] \\ u(\mathbf{x}, 0) = u_0, & \mathbf{x} \in \Omega \\ u(\mathbf{x}, t) = u_D(\mathbf{x}), & \mathbf{x} \in \partial\Omega, \quad t \in (0, T] \end{cases} \quad (9)$$

where  $\mu = \frac{3}{200}$ ,  $\beta(\mathbf{x}) = [2y(1 - x^2), -2x(1 - y^2)]^T$ ,  
 $u_D = 0$  on  $\{\{x = 0\} \times (0, 1)\}, \{(0, 1) \times \{y = 0\}\}, \{(0, 1) \times \{y = 1\}\}$ .

Discretization (IFISS toolbox employed):

- Quadrilateral FEM
- spatial discretization step  $h = 2^{-8} \Rightarrow 66,049$  d.o.f.'s
- time discretization step  $\Delta t = h^2$

## 2D parabolic problems: test case 3

Stiffness matrix:  $A$  ( $n = 66,049$ ).

AINV preconditioner:  $M$ .

Iteration matrix:  $H = I - MA$ .

Solution computed just for a generic time step.

- $\frac{\text{nnz}(H)}{\text{nnz}(A)} = 4.736$
- $\|H\|_1 = 0.212$
- # iterations employed: 12
- relative error =  $9.193 \cdot 10^{-8}$
- # total random walks employed: 9,300



# Conclusions and future developments

## Conclusions

- MC solvers motivated by resilience
- hypotheses for convergence difficult to satisfy in general
- currently best results obtained using AINV

## Future developments

- extension of the set of matrices for which MC solvers is guaranteed to converge a priori
- analysis of how  $\rho(H)$  and  $\rho(H^*)$  affect convergence
- refinement of the adaptive selection of histories

# Bibliography I



I. Dimov, V. Alexandrov, and A. Karaivanova.

Parallel resolvent Monte Carlo algorithms for linear algebra problems.

*Mathematics and Computers in Simulation*, 55(55):25–35, 2001.



T.M Evans, S.W. Mosher, S.R. Slattery, and S.P. Hamilton.

A Monte Carlo synthetic-acceleration method for solving the thermal radiation diffusion equation.

*Journal of Computational Physics*, 258(November 2013):338–358, 2014.

# Bibliography II



T.M. Evans, S.R. Slattery, and P.P.H Wilson.

Mixed Monte Carlo parallel algorithms for matrix computation.

*International Conference on Computational Science 2002*, 2002.



T.M. Evans, S.R. Slattery, and P.P.H Wilson.

A spectral analysis of the domain decomposed Monte Carlo method for linear systems.

*International Conference on Mathematics and Computational Methods Applied to Nuclear Science & Engineering*, 2013.



J.H. Halton.

Sequential Monte Carlo.

*Mathematical Proceedings of the Cambridge Philosophical Society*, 58(1):57–58, 1962.

# Bibliography III



J.H. Halton.

Sequential Monte Carlo techniques for the solution of linear systems.

*Journal of Scientific Computing*, 9(2):213–257, 1994.



J. Hao, M. Mascagni, and L. Yaohang.

Convergence analysis of Markov chain Monte Carlo linear solvers using Ulam-von Neumann algorithm.

*SIAM Journal on Numerical Analysis*, 51(4):2107–2122, 2013.



N. Metropolis and S. Ulam.

The MOnTe Carlo ethod.

*Journal of th American Statistical Association*,  
44(247):335–341, 1949.

# Bibliography IV



S. Slattery.

*Parallel Monte Carlo Synthetic Acceleration Methods For Discrete Transport Problems.*

PhD thesis, University of Wisconsin-Madison, 2013.

# Monte Carlo Linear Solvers

The choice of the transition probability

## Theorem (Hao J. et al.)

*Let  $H \in \mathbb{R}^{n \times n}$  with spectral radius  $\rho(H) < 1$ . Let  $H^+ \in \mathbb{R}^{n \times n}$  where  $H_{i,j}^+ = |H_{i,j}|$ . Consider  $\nu_k$  as the realization of a random walk  $\gamma$  truncated at the  $k$ -th step. If  $\rho(H^+) > 1$ , there does not exist a transition matrix  $P$  such that  $\text{Var}\left(X(\nu_k)\right)$  converges to zero as  $k \rightarrow \infty$ .*

# Numerical experiments

## Simplified $P_N$ equations

Steady-state, multigroup, one-dimensional, eigenvalue-form of Boltzmann transport equation:

$$\begin{aligned} \mu \frac{\partial \psi^g(x, \mu)}{\partial x} + \sigma^g(x) \psi^g(x, \mu) = & \sum_{g'=1}^{N_g} \int_{4\pi} \sigma_s^{gg'}(x, \hat{\Omega} \cdot \hat{\Omega}') \psi^{g'}(x, \Omega') d\Omega' + \\ & + \frac{1}{k} \sum_{g'=1}^{N_g} \frac{\chi^g}{4\pi} \int_{4\pi} \nu \sigma_f^{g'}(x) \psi^{g'}(x, \Omega') d\Omega'. \end{aligned} \quad (10)$$

$\psi^g(x, \mu)$  = angular flux for group  $g$

$\sigma^g$  = total interaction cross section

$\sigma_s^{gg'}(x, \hat{\Omega} \cdot \hat{\Omega}')$  = scattering cross section from group  $g' \rightarrow g$

# Numerical experiments

## Simplified $P_N$ equations

- Legendre polynomial spectral discretization ( $P_N$  equations)
- consider just odd sets of  $P_N$  equations
- removing lower order gradient terms from each equation ( $SP_N$  equations)

$$-\nabla \cdot \mathbb{D}_n \nabla \mathbf{U}_n + \sum_{m=0}^{\frac{N+1}{2}} \mathbb{A}_{nm} \mathbf{U}_m = \frac{1}{k} \sum_{m=1}^{\frac{N+1}{2}} \mathbb{F}_{nm} \mathbf{U}_m, \quad n = 1, \dots, \frac{N+1}{2} \quad (11)$$

$\mathbf{U}_m$  is a linear combination of moments.

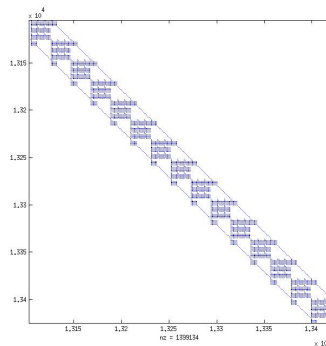
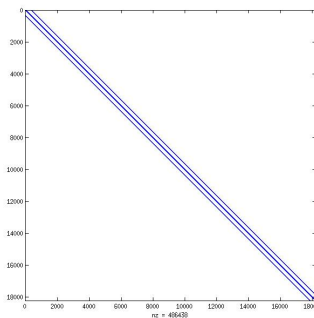
- Discretized with Finite Volumes



# Numerical experiments

## Simplified $P_N$ equations

### Sparsity pattern and zoom on the structure of a $SP_1$ matrix



# Numerical experiments

## Simplified $P_N$ equations

Properties of matrices:

- discretization of  $SP_1$  equations
- all nonzero diagonal entries

Numerical treatment:

- block Jacobi preconditioning

Kind of matrix	size	nnz	Prec. block size	nnz after prec
$SP_1$ (a)	18,207	486,438	63	2,644,523
$SP_1$ (b)	19,941	998,631	69	1,774,786

[Matrices provided by courtesy by Thomas Evans, Steven Hamilton, Stuart Slattery, ORNL]

# Numerical experiments

## Simplified $P_N$ equations

MCSA parameter setting:

- Adjoint method
- residual relative tolerance:  $\varepsilon_1 = 10^{-7}$
- almost optimal transition probability
- maximal # steps per history: 10
- statistical error-based adaptive parameter:  $\varepsilon_2 = 0.5$
- granularity of the adaptive approach:  $n_{histories} = 1,000$
- simulations run in serial mode

# Numerical experiments

Simplified  $P_N$  equations

matrix	$\rho(H)$	$\rho(H^*)$	relative err.	# iterations
$SP_1$ (a)	0.9779	0.9758	$9.97 \cdot 10^{-6}$	340
$SP_1$ (b)	0.9798	0.9439	$3.89 \cdot 10^{-5}$	209

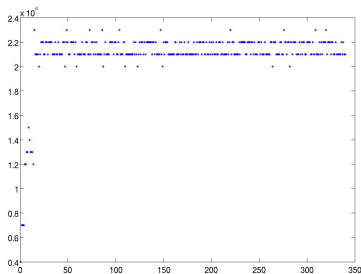


Figure 1 : # histories for (a) at each iteration.

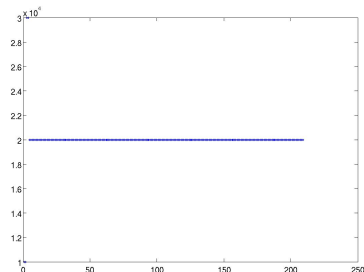


Figure 2 : # histories for (b) at each iteration.

# Numerical experiments

## Simplified $SP_N$ equations

Issues raised for a general  $SP_N$  matrix:

- hard to find a block Jacobi preconditioner s.t.  $\rho(H^*) < 1$  for every  $SP_N$  problem
- attempt to use Approximate Inverse preconditioners was not effective
  - for sparse preconditioners  $\rho(H^*) < 1$  is not respected
- attempt to use ILU preconditioners was not effective
  - for ILU(0) we got  $\rho(H^*) > 1$
  - massive fill-in with ILUT
- reordering and scaling did not facilitate the convergence requirements

# Numerical experiments

## Diagonally dominant matrices

Set of matrices obtained by a diagonal shift of matrices associated with  $SP_1$ ,  $SP_3$  and  $SP_5$  equations.

$$(A + sD), \quad s \in \mathbb{R}^+, \quad D = \text{diag}(A).$$

All the matrices have been turned into s.d.d. by columns.

Initial matrix	$s$	$\rho(H)$	$Forward - \rho(H^*)$	$Adjoint - \rho(H^*)$
$SP_1$ (a)	0.3	0.7597	0.7441	0.6983
$SP_1$ (b)	0.4	0.7046	1.1448	0.5680
$SP_3$	0.9	0.5869	0.4426	0.3727
$SP_5$	1.6	0.5477	0.3790	0.3431

# Numerical experiments

## Diagonally dominant matrices

$$(A + sD), \quad s \in \mathbb{R}^+, \quad D = \text{diag}(A).$$

matrix	nnz	s	relative err.	# iterations
$SP_1$ (a)	486,438	0.3	$6.277 \cdot 10^{-7}$	36
$SP_1$ (b)	998,631	0.4	$9.885 \cdot 10^{-7}$	21
$SP_3$	846,549	0.9	$5.919 \cdot 10^{-7}$	18
$SP_5$	1,399,134	1.6	$4.031 \cdot 10^{-7}$	19

# Numerical experiments

## Diagonal shift

Strict diagonally dominance is a sufficient condition for convergence but not necessary.

$$(A + sD), \quad s \in \mathbb{R}^+, \quad D = \text{diag}(A)$$

such that  $\rho(H) < 1$  and  $\rho(H^*) < 1$ .

Initial matrix	$s$	$\rho(H)$	$Forward - \rho(H^*)$	$Adjoint - \rho(H^*)$
$SP_1$ (a)	0.2	0.8230	0.8733	0.8195
$SP_1$ (b)	0.2	0.8220	1.5582	0.7731
$SP_3$	0.3	0.8126	0.9459	0.7961
$SP_5$	0.7	0.8376	0.8865	0.8026



# Numerical experiments

## Diagonal shift

$$(A + sD), \quad s \in \mathbb{R}^+, \quad D = \text{diag}(A).$$

matrix	$s$	relative err.	# iterations
$SP_1$ (a)	0.2	$6.394 \cdot 10^{-7}$	48
$SP_1$ (b)	0.2	$2.59 \cdot 10^{-6}$	36
$SP_3$	0.3	$5.35 \cdot 10^{-7}$	45
$SP_5$	0.7	$4.21 \cdot 10^{-7}$	70