Smoothed Aggregation Multigrid for Slowly Mixing Markov Chains

Hans De Sterck, Jamie Pearson

Department of Applied Mathematics, University of Waterloo

Steve McCormick, John Ruge, Tom Manteuffel, Geoff Sanders

Department of Applied Mathematics, University of Colorado at Boulder

University of Waterloo

Copper 2008

1. Simple Markov Chain Example

- start in one state with probability 1: what is the stationary probability vector after ∞ number of steps?

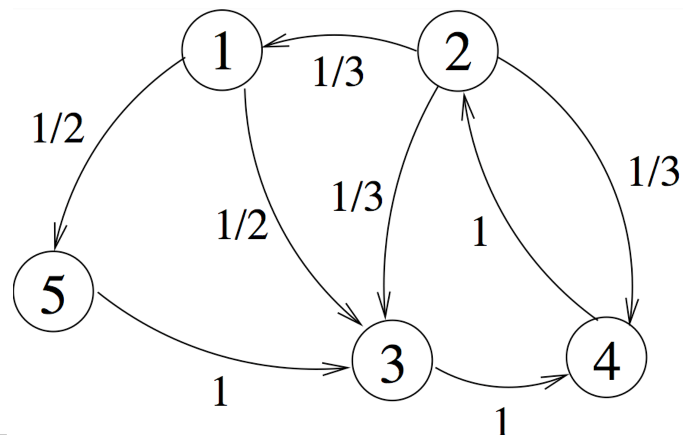
$$\mathbf{x}_{i+1} = B\,\mathbf{x}_i$$

- stationary probability:

$$B\,\mathbf{x} = \mathbf{x} \qquad \|\mathbf{x}\|_1 = 1$$

$$\mathbf{x}^T = [2/19\ 6/19\ 4/19\ 6/19\ 1/19]$$

$$B = \begin{bmatrix} 0 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1/2 & 1/3 & 0 & 0 & 1 \\ 0 & 1/3 & 1 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 \end{bmatrix}$$



Copper 2008
hdesterck@uwaterloo.ca

2. Problem Statement

$$B\,\mathbf{x} = \mathbf{x} \qquad \|\mathbf{x}\|_1 = 1 \qquad x_i \geq 0\,\forall i$$

- *B* is column-stochastic

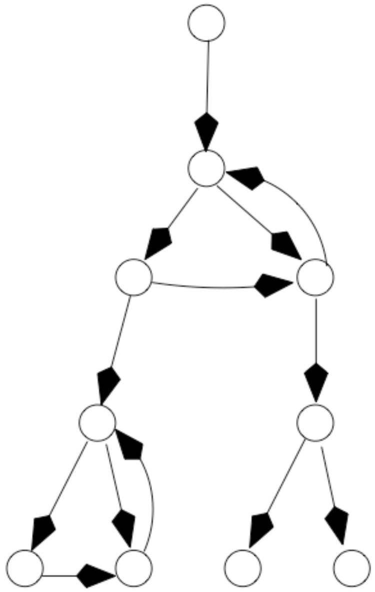
$$0 \leq b_{ij} \leq 1\ \forall i, j \qquad \mathbf{1}^T B = \mathbf{1}^T$$

- *B* is irreducible (every state can be reached from every other state in the directed graph)

$$\Longrightarrow$$

$$\exists!\ \mathbf{x}: \quad B\,\mathbf{x} = \mathbf{x} \qquad \|\mathbf{x}\|_1 = 1 \qquad x_i > 0\ \forall i$$

(no probability sinks!)



University of Waterloo

Copper 2008
hdesterck@uwaterloo.ca

3. Power Method

$$B\,\mathbf{x} = \mathbf{x} \quad \text{or} \quad (I - B)\,\mathbf{x} = 0 \quad \text{or} \quad A\,\mathbf{x} = 0$$

- largest eigenvalue of *B*: $\lambda_1 = 1$
- power method: $\mathbf{x}_{i+1} = B\mathbf{x}_i$
  - convergence factor: $|\lambda_2|$
  - convergence is very slow when

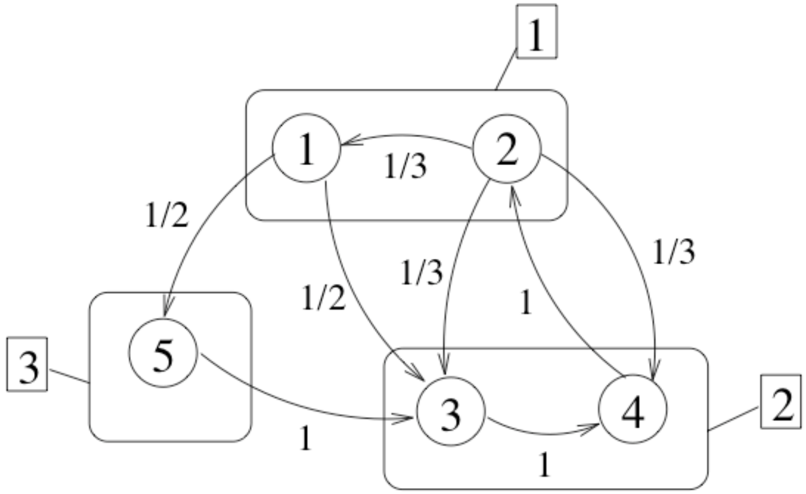
    $$|\lambda_2| \approx 1$$

    (slowly mixing Markov chain) (JAC, GS also slow)

4. Aggregation for Markov Chains

$$B_c \, \mathbf{x}_c = \mathbf{x}_c$$

$$b_{c,IJ} = \frac{\displaystyle\sum_{j \in J} x_j \left(\sum_{i \in I} b_{ij} \right)}{\displaystyle\sum_{j \in J} x_j}$$



$$B_c = Q^T \, B \, \mathrm{diag}(\mathbf{x}) \, Q \, \mathrm{diag}(Q^T \, \mathbf{x})^{-1}$$

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$x_{c,I} = \textstyle\sum_{i \in I} x_i$$

$$\mathbf{x}_c = Q^T \, \mathbf{x}$$

(Krieger, Horton, ... 1990s)

5. Error Equation

- error equation - coarse grid correction:

$$\mathbf{x} = \mathrm{diag}(\mathbf{x}_i)\,\mathbf{e}_i$$
$$A\,\mathrm{diag}(\mathbf{x}_i)\,\mathbf{e}_i = 0$$

$$Q^T A\,\mathrm{diag}(\mathbf{x}_i)\,Q\,\mathbf{e}_c = 0$$
$$A_c\,\mathbf{e}_c = 0$$

$$R = Q^T \qquad P = \mathrm{diag}(\mathbf{x}_i)\,Q$$
$$A_c = R\,A\,P$$

$$\mathbf{x}_{i+1} = P\,\mathbf{e}_c$$

University of Waterloo

Copper 2008
hdesterck@uwaterloo.ca

Error Equation

- important properties of A_c:

$$\mathbf{x} = \mathrm{diag}(\mathbf{x}_i)\,\mathbf{e}_i$$
$$A\,\mathrm{diag}(\mathbf{x}_i)\,\mathbf{e}_i = 0$$

$$Q^T A\,\mathrm{diag}(\mathbf{x}_i)\,Q\,\mathbf{e}_c = 0$$
$$A_c\,\mathbf{e}_c = 0$$

$$R = Q^T \qquad P = \mathrm{diag}(\mathbf{x}_i)\,Q$$
$$A_c = R\,A\,P$$

(1) $\mathbf{1}_c^T A_c = 0 \quad \forall\, \mathbf{x}_i$
(since $\mathbf{1}_c^T R = \mathbf{1}^T$ and $\mathbf{1}^T A = 0$)

(2) $A_c\,\mathbf{1}_c = 0 \quad$ for $\mathbf{x}_\mathrm{i} = \mathbf{x}$

$$A_c\,(\mathrm{diag}(P^T\,\mathbf{1}))^{-1}$$
$$= R\,(I - B)\,P\,(\mathrm{diag}(P^T\,\mathbf{1}))^{-1}$$
$$= I_c - B_c$$

University of Waterloo

Copper 2008
hdesterck@uwaterloo.ca

6. Multilevel Aggregation Algorithm

Algorithm: Multilevel Adaptive Aggregation method (V-cycle)

$\mathbf{x} = \text{AM\_V}(A, \mathbf{x}, \nu_1, \nu_2)$

begin

$\mathbf{x} \leftarrow \text{Relax}(A, \mathbf{x})$ $\quad \nu_1$ times

build Q based on $\mathbf{x}$ and A $\quad$ (Q is rebuilt every level and cycle)

$R = Q^T$ and $P = \text{diag}(\mathbf{x})\, Q$

$A_c = R\, A\, P$

$\mathbf{x}_c = \text{AM\_V}(A_c\, \text{diag}(P^T\, \mathbf{1})^{-1}, P^T\, \mathbf{1}, \nu_1, \nu_2)$ $\quad$ (coarse-level solve)

$\mathbf{x} = P\, (\text{diag}(P^T\, \mathbf{1}))^{-1} \mathbf{x}_c$ $\quad$ (coarse-level correction)

$\mathbf{x} \leftarrow \text{Relax}(A, \mathbf{x})$ $\quad \nu_2$ times

end

(Krieger, Horton 1994, but no good way to build Q, convergence not good)

Copper 2008
hdesterck@uwaterloo.ca

7. Well-posedness: Singular M-matrices

- singular M-matrix:

$$A = \begin{bmatrix} + & - & - & - & - \\ - & + & - & - & - \\ - & - & + & - & - \\ - & - & - & + & - \\ - & - & - & - & + \end{bmatrix}$$

$A \in \mathbb{R}^{n\times n}$ *is a singular M-matrix* $\Leftrightarrow$

$$\exists\, B \in \mathbb{R}^{n\times n},\ b_{ij} \geq 0\ \forall i,j\ :\ A = \rho(B)\, I - B$$

- our *A=I-B* is a singular M-matrix on all levels

(1) Irreducible singular M-matrices have a unique solution to the problem $A\,\mathbf{x} = 0$*, up to scaling. All components of* $\mathbf{x}$ *have strictly the same sign (i.e., scaling can be chosen s.t.* $x_i > 0\ \forall i$*). (This follows directly from the Perron-Frobenius theorem.)*

(3) Irreducible singular M-matrices have nonpositive off-diagonal elements, and strictly positive diagonal elements ($n > 1$*).*

(4) If A has a strictly positive element in its left or right nullspace and the off-diagonal elements of A are nonpositive, then A is a singular M-matrix (see also [21]).

University of Waterloo

Copper 2008
hdesterck@uwaterloo.ca

Well-posedness: Unsmoothed Method

THEOREM 3.1 (Singular M-matrix property of AM coarse-level operators). *A_c is an irreducible singular M-matrix on all coarse levels, and thus has a unique right kernel vector $\mathbf{e}_c$ with strictly positive components (up to scaling) on all levels.*

(1) $\mathbf{1}_c^T A_c = 0 \quad \forall \mathbf{x}_i$

(since $\mathbf{1}_c^T R = \mathbf{1}^T$ and $\mathbf{1}^T A = 0$)

$$A = \begin{bmatrix} + & - & - & - & - \\ - & + & - & - & - \\ - & - & + & - & - \\ - & - & - & + & - \\ - & - & - & - & + \end{bmatrix}$$

THEOREM 3.2 (Fixed-point property of AM V-cycle). *Exact solution $\mathbf{x}$ is a fixed point of the AM V-cycle.*

(2) $A_c \mathbf{1}_c = 0 \quad$ for $\mathbf{x}_i = \mathbf{x}$

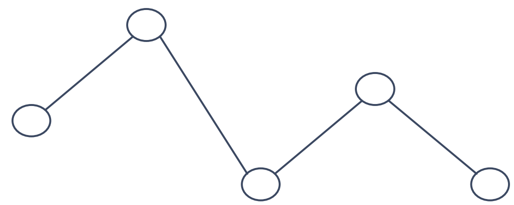
$$A_c \mathbf{e}_c = 0$$
$$\mathbf{x}_{i+1} = P \mathbf{e}_c$$

University of Waterloo

Copper 2008
hdesterck@uwaterloo.ca

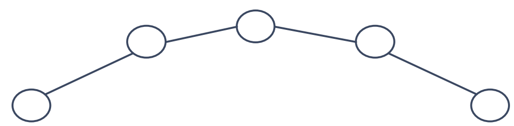
8. We Need ‘Smoothed Aggregation’...

(Vanek, Mandel, and Brezina, Computing, 1996)

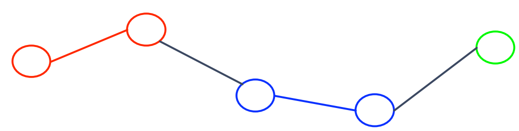


$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

after relaxation:

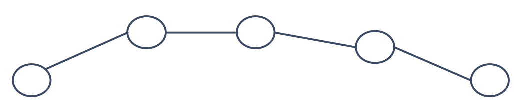


coarse grid correction with Q:



$$Q_s = \begin{bmatrix} \times & 0 & 0 \\ \times & \times & 0 \\ \times & \times & 0 \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}$$

coarse grid correction with Q_s:



University of Waterloo

hdesterck@uwaterloo.ca

Smoothed Aggregation

$$A = D - (L + U)$$

- smooth the columns of *P* with weighted Jacobi:

$$P_s = (1 - w)\,\mathrm{diag}(\mathbf{x}_i)\,Q + w\,D^{-1}\,(L + U)\,\mathrm{diag}(\mathbf{x}_i)\,Q$$

- smooth the rows of *R* with weighted Jacobi:

$$R_s = R\,(1 - w) + R\,w\,(L + U)\,D^{-1}$$

University of Waterloo

Copper 2008
hdesterck@uwaterloo.ca

Smoothed Aggregation

- smoothed coarse level operator:

$$A_{cs} = R_s\,(D - (L+U))\,P_s$$
$$= R_s\,D\,P_s - R_s\,(L+U)\,P_s$$

$$\mathbf{1}_c^T\,A_{cs} = 0 \quad \forall\,\mathbf{x}_i,$$
$$A_{cs}\,\mathbf{1}_c = 0 \quad \text{for } \mathbf{x}_\mathrm{i} = \mathbf{x}$$

- problem: A_{cs} is not a singular M-matrix (signs wrong)

- solution: lumping approach on S in

$$A = \begin{bmatrix} + & - & - & - & - \\ - & + & - & - & - \\ - & - & + & - & - \\ - & - & - & + & - \\ - & - & - & - & + \end{bmatrix}$$

$$A_{cs} = S - G$$

$$\hat{A}_{cs} = \hat{S} - G$$

Copper 2008
hdesterck@uwaterloo.ca

Smoothed Aggregation

$$A_{cs} = S - G \qquad \hat{A}_{cs} = \hat{S} - G$$

$$A = \begin{bmatrix} + & - & - & - & - \\ - & + & - & - & - \\ - & - & + & - & - \\ - & - & - & + & - \\ - & - & - & - & + \end{bmatrix}$$

- we want as little lumping as possible
- only lump 'offending' elements *(i,j)*:

$$s_{ij} \neq 0,\ i \neq j \text{ and } s_{ij} - g_{ij} \geq 0$$

$$\mathbf{1}_c^T \hat{A}_{cs} = 0 \quad \forall \mathbf{x}_i,$$
$$\hat{A}_{cs} \mathbf{1}_c = 0 \quad \text{for } \mathbf{x}_\mathrm{i} = \mathbf{x}$$

(we consider both off-diagonal signs and reducibility here!)

- for 'offending' elements *(i,j)*, add $S_{\{i,j\}}$ to S:

$$S_{\{i,j\}} = \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccc} \quad & i & \quad & j & \quad \end{array} \\ \begin{bmatrix} \ddots & \vdots & & \vdots & \\ \cdots & \beta_{\{i,j\}} & \cdots & -\beta_{\{i,j\}} & \cdots \\ & \vdots & & \vdots & \\ \cdots & -\beta_{\{i,j\}} & \cdots & \beta_{\{i,j\}} & \cdots \\ & \vdots & & \vdots & \end{bmatrix} \end{array}$$

$$s_{ij} - g_{ij} - \beta_{\{i,j\}} < 0$$
$$s_{ji} - g_{ji} - \beta_{\{i,j\}} < 0$$

conserves both row and column sums

Copper 2008
hdesterck@uwaterloo.ca

9. Lumped Smoothed Method is Well-posed

THEOREM 4.1 (Singular M-matrix property of lumped SAM coarse-level operators). *$\hat{A}_{cs}$ is an irreducible singular M-matrix on all coarse levels, and thus has a unique right kernel vector $\mathbf{e}_c$ with strictly positive components (up to scaling) on all levels.*

THEOREM 4.2 (Fixed-point property of lumped SAM V-cycle). *Exact solution $\mathbf{x}$ is a fixed point of the SAM V-cycle (with lumping).*

$$\mathbf{1}_c^T \, \hat{A}_{cs} = 0 \quad \forall \, \mathbf{x}_i,$$

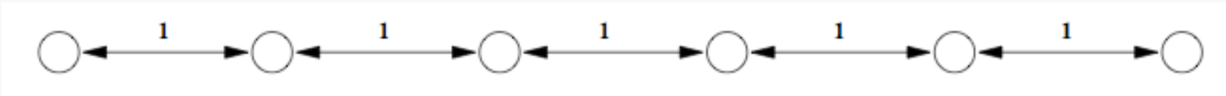
$$\hat{A}_{cs} \, \mathbf{1}_c = 0 \quad \text{for } \mathbf{x}_i = \mathbf{x}$$

$$A = \begin{bmatrix} + & - & - & - & - \\ - & + & - & - & - \\ - & - & + & - & - \\ - & - & - & + & - \\ - & - & - & - & + \end{bmatrix}$$

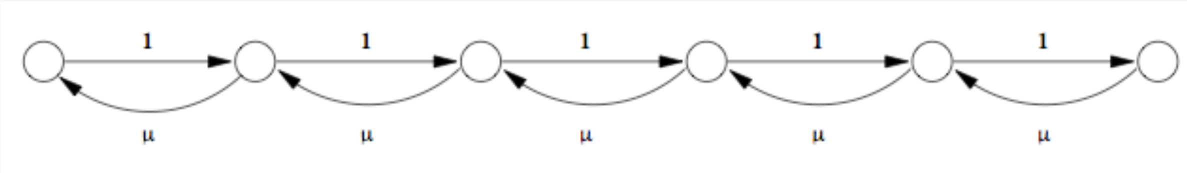
University of Waterloo

10. Numerical Results: Test Problems

- uniform chain



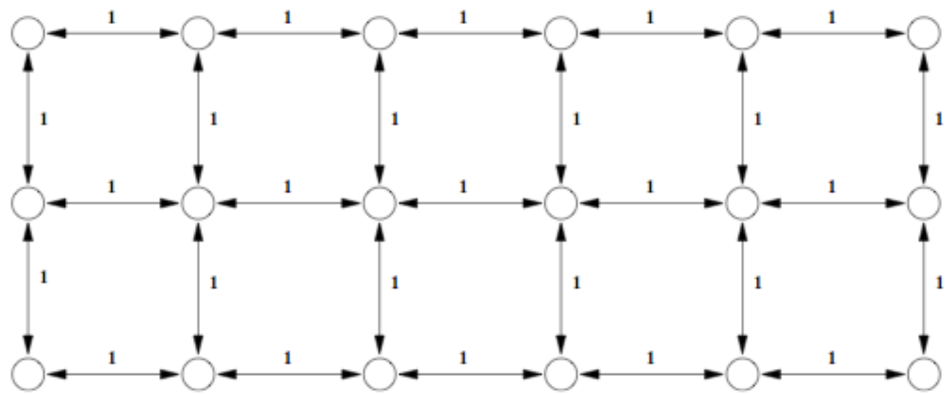
- birth-death chain



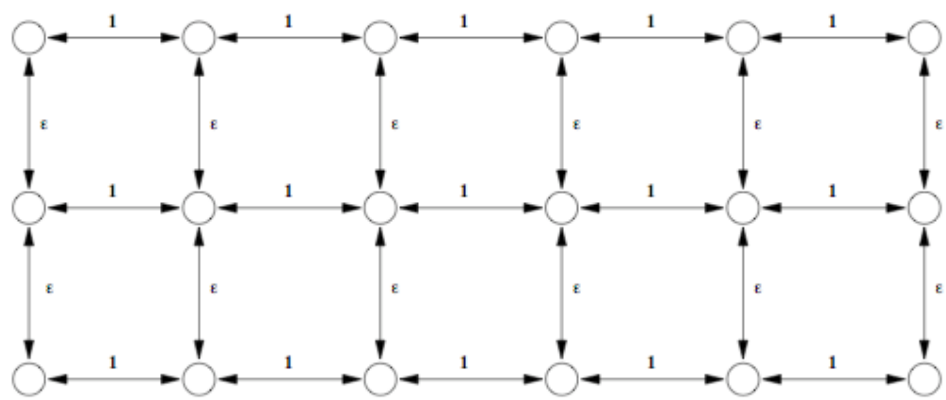
hdesterck@uwaterloo.ca

Test Problems

- uniform 2D lattice



- anisotropic 2D lattice



University of Waterloo

Copper 2008
hdesterck@uwaterloo.ca

Test Problems

- tandem queueing network

FIG. 5.6. *Tandem queueing network.*

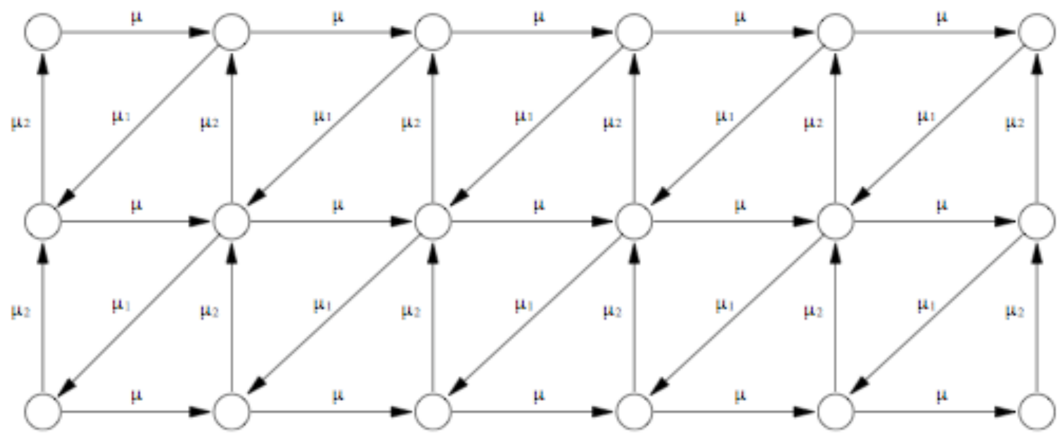


FIG. 5.7. *Graph for tandem queueing network.*

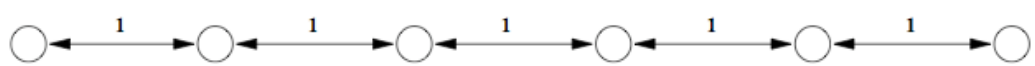
11. Numerical Results: Geometric Aggregation (size 3)

n	γ_{res}	iter	C_{op}	levels
27	0.66	32	1.32	2
81	0.87	85	1.43	3
243	0.95	>100	1.47	4
729	0.98	>100	1.49	5

TABLE 5.1
Uniform chain. G-AM with V-cycles and size-three aggregates. (No smoothing.)

n	γ_{res}	iter	C_{op}	levels	R_{lump}
27	0.27	13	1.32	2	0
81	0.27	13	1.43	3	0
243	0.27	13	1.47	4	0
729	0.27	13	1.49	5	0
2187	0.27	13	1.50	6	0
6561	0.27	13	1.50	7	0

TABLE 5.2
Uniform chain. G-SAM with V-cycles and size-three aggregates. (Smoothing with lumping.)



Numerical Results: Geometric Aggregation (size 3)

n	γ_{res}	iter	C_{op}	levels
27	0.66	33	1.32	2
81	0.88	95	1.43	3
243	0.95	>100	1.47	4
729	0.97	>100	1.49	5

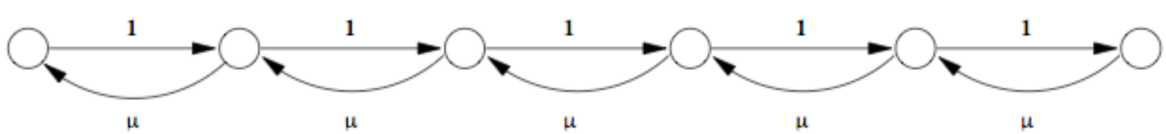
TABLE 5.3

Birth-death chain ($\mu = 0.96$). G-AM with V-cycles and size-three aggregates. (No smoothing.)

n	γ_{res}	iter	C_{op}	levels	R_{lump}
27	0.27	13	1.32	2	0
81	0.27	12	1.43	3	0
243	0.26	13	1.47	4	0
729	0.24	12	1.49	5	0

TABLE 5.4

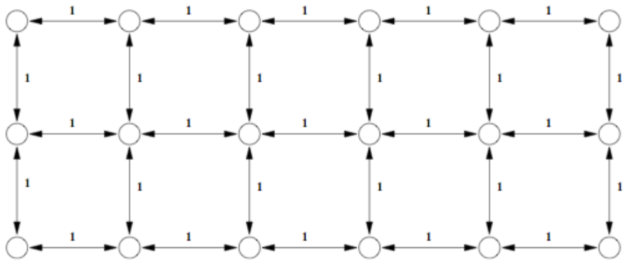
Birth-death chain ($\mu = 0.96$). G-SAM with V-cycles and size-three aggregates. (Smoothing with lumping.)



University of Waterloo

Copper 2008
hdesterck@uwaterloo.ca

Numerical Results: Geometric Aggregation (3x3)



n	γ_{res}	iter	C_{op}	levels
64	0.71	43	1.11	2
100	0.85	72	1.17	3
169	0.86	85	1.15	3
400	0.89	$>$100	1.13	3
900	0.95	$>$100	1.12	4

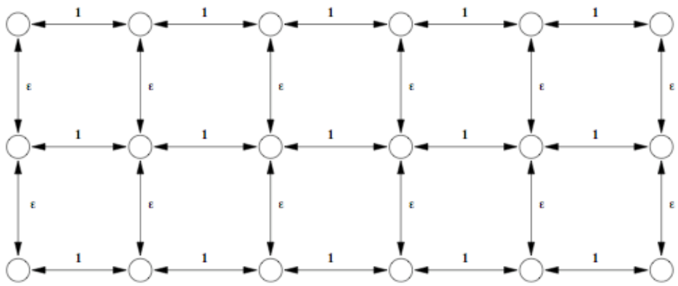
TABLE 5.9
Uniform 2D lattice. G-AM with V-cycles and size-three aggregates. (No smoothing.)

n	γ_{res}	iter	C_{op}	levels	R_{lump}	γ_{res}	iter	C_{op}	levels	R_{lump}
64	0.49	22	1.17	2	0	0.39	17	1.34	2	0
100	0.52	23	1.25	3	0	0.41	18	1.57	3	0
169	0.52	23	1.23	3	0	0.42	18	1.51	3	0
400	0.56	25	1.21	3	1.72e-03	0.44	19	1.48	3	5.64e-03
900	0.56	25	1.21	4	7.58e-04	0.44	18	1.48	4	2.47e-03
1600	0.59	26	1.23	4	0	0.44	19	1.51	4	0
2500	0.59	26	1.22	4	0	0.45	19	1.48	4	0
4900	0.59	26	1.22	4	0	0.44	19	1.50	4	0
6724	0.58	26	1.23	5	9.78e-04	0.44	19	1.53	5	4.08e-03

TABLE 5.10
Uniform 2D lattice. G-SAM with V-cycles (left) and W-cycles (right), using three-by-three aggregates. (Smoothing with lumping.)

hdesterck@uwaterloo.ca

Numerical Results: Geometric Aggregation (3x3)



n	γ_{res}	iter	C_{op}	levels
64	1.00	>100	1.11	2
100	0.92	>100	1.17	3
169	0.95	>100	1.15	3
400	0.98	>100	1.13	3
900	0.99	>100	1.12	4
1600	0.99	>100	1.13	4

TABLE 5.11

Anisotropic 2D lattice ($\epsilon = 1e-6$). G-AM with V-cycles and three-by-three aggregates. (No smoothing.)

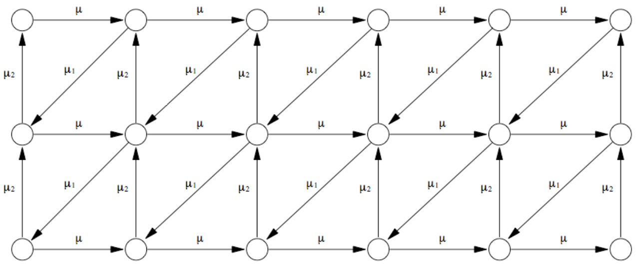
n	γ_{res}	iter	C_{op}	levels	R_{lump}
64	1.00	>100	1.17	2	0
100	1.00	>100	1.25	3	0
169	0.93	>100	1.23	3	0
400	0.97	>100	1.21	3	0
900	0.99	>100	1.21	4	0
1600	0.99	>100	1.23	4	0

TABLE 5.12

Anisotropic 2D lattice ($\epsilon = 1e-6$). G-SAM with V-cycles and three-by-three aggregates. (Smoothing with lumping.)

University of Waterloo

Numerical Results: Geometric Aggregation (3x3)



n	γ_{res}	iter	C_{op}	levels
121	0.90	>100	1.20	3
256	0.91	>100	1.19	3
676	0.93	>100	1.18	3
1681	0.95	>100	1.19	4

TABLE 5.13
Tandem queueing network. G-AM with V-cycles and three-by-three aggregates. (No smoothing.)

n	γ_{res}	iter	C_{op}	levels	R_{lump}	γ_{res}	iter	C_{op}	levels	R_{lump}
121	0.40	19	1.26	3	5.39e-03	0.29	14	1.60	3	1.42e-02
256	0.44	19	1.28	3	2.43e-02	0.33	14	1.60	3	3.90e-02
676	0.44	19	1.26	3	1.53e-03	0.33	14	1.56	3	2.47e-03
1681	0.46	20	1.27	4	1.20e-03	0.32	14	1.61	4	1.89e-03
3721	0.48	20	1.28	4	6.61e-03	0.33	14	1.63	4	1.14e-02

TABLE 5.14
Tandem queueing network. G-SAM with V-cycles (left) and W-cycles (right), using three-by-three aggregates. (Smoothing with lumping.)

University of Waterloo

Copper 2008
hdesterck@uwaterloo.ca

12. Numerical Results: Algebraic Aggregation

- error equation: $A\,\mathrm{diag}(\mathbf{x}_i)\,\mathbf{e}_i = 0$
- use strength of connection in $A\,\mathrm{diag}(\mathbf{x}_i)$
- define row-based strength (determine all states that strongly influence a row's state, similar to AMG)
- state that has largest value in $\mathbf{x}_i$ is seed point for new aggregate, and all unassigned states influenced by it join its aggregate
- repeat

(our Google SISC paper 2008)

University of Waterloo

Copper 2008
hdesterck@uwaterloo.ca

Numerical Results: Algebraic Aggregation

n	γ_{res}	iter	C_{op}	levels
27	0.75	43	1.71	3
81	0.87	87	1.85	4
243	0.96	>100	1.96	6
729	0.99	>100	1.98	7

TABLE 6.1

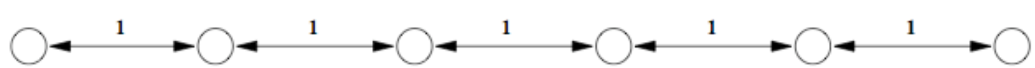
Uniform chain. A-AM with V-cycles and distance-one aggregation. (No smoothing.)

n	γ_{res}	iter	C_{op}	levels	R_{lump}	γ_{res}	iter	C_{op}	levels	R_{lump}
27	0.20	10	2.03	3	2.50e-02	0.32	13	1.30	2	0
81	0.18	10	2.68	4	4.34e-02	0.34	14	1.42	3	0
243	0.19	10	2.78	5	4.75e-02	0.34	13	1.50	4	0
729	0.24	11	3.41	7	5.69e-02	0.30	12	1.51	5	6.08e-04
2187	0.27	11	3.75	8	6.06e-02	0.27	11	1.50	6	2.03e-04
6561	0.31	12	4.03	9	6.39e-02	0.25	11	1.50	7	0

TABLE 6.2

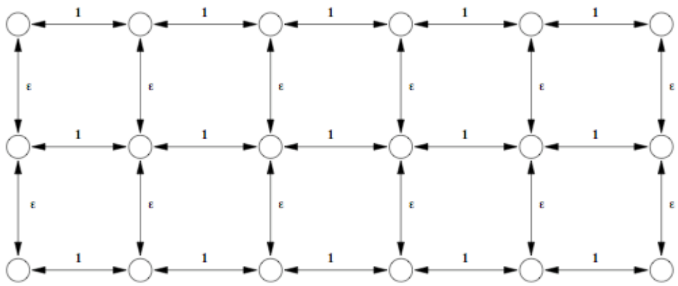
Uniform chain. A-SAM with V-cycles using distance-one aggregation (left) and distance-two aggregation (right). (Smoothing with lumping.)

University of Waterloo



hdesterck@uwaterloo.ca

Numerical Results: Algebraic Aggregation



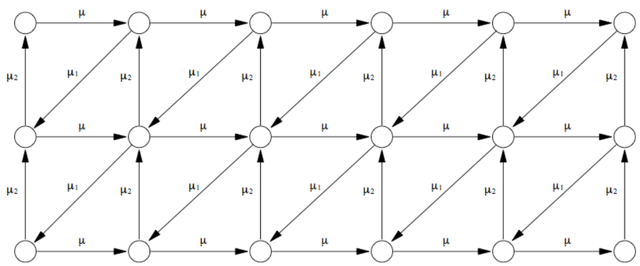
n	γ_{res}	iter	C_{op}	levels
64	0.55	27	1.78	4
100	0.61	31	2.07	5
169	0.67	38	1.98	6
400	0.79	55	2.08	7
900	0.86	79	2.01	8
1600	0.89	>100	2.08	9

TABLE 6.12
Anisotropic 2D lattice ($\epsilon = 1e-6$). A-AM with V-cycles and distance-one aggregation. (No smoothing.)

n	γ_{res}	iter	C_{op}	levels	R_{lump}
64	0.40	18	1.76	3	0
100	0.32	14	1.99	3	2.18e-03
169	0.32	14	2.04	4	2.48e-03
400	0.33	14	2.63	5	2.37e-03
900	0.33	14	2.92	5	1.41e-03
1600	0.33	14	2.99	6	1.20e-03
2500	0.33	14	3.41	7	1.43e-03
4900	0.33	13	3.77	7	6.79e-04
6724	0.33	13	3.90	7	4.93e-04

TABLE 6.13
Anisotropic 2D lattice ($\epsilon = 1e-6$). A-SAM with V-cycles and distance-two aggregation. (Smoothing with lumping.)

Numerical Results: Algebraic Aggregation



n	γ_{res}	iter	C_{op}	levels
121	0.75	39	1.89	4
256	0.88	82	2.35	7
676	0.95	>100	3.28	14
1681	0.93	>100	4.79	20

TABLE 6.14

Tandem queueing network. A-AM with V-cycles and distance-one aggregation. (No smoothing.)

n	γ_{res}	iter	C_{op}	levels	R_{lump}
121	0.38	18	2.04	3	1.44e-01
256	0.39	19	2.27	4	1.13e-01
676	0.48	21	2.47	4	9.09e-02
1681	0.47	21	2.85	5	8.22e-02
3721	0.42	21	3.20	5	7.46e-02

TABLE 6.15

Tandem queueing network. A-SAM with V-cycles and distance-two aggregation. (Smoothing with lumping.)

Conclusions

- SAM: algorithm for stationary vector of slowly mixing Markov chains with near-optimal complexity
- smoothing is essential
- pretty good convergence results
- good theoretical framework (well-posedness)
- are there other ways for choosing R_s, P_s, lumping?
- no theory yet on optimal convergence (non-symmetric matrices)

- Questions?

Smoothed Aggregation

$$A_{cs} = S - G$$

$$\hat{A}_{cs} = \hat{S} - G$$

$$A = \begin{bmatrix} + & - & - & - & - \\ - & + & - & - & - \\ - & - & + & - & - \\ - & - & - & + & - \\ - & - & - & - & + \end{bmatrix}$$

$$s_{ij} - g_{ij} - \beta_{\{i,j\}} < 0$$
$$s_{ji} - g_{ji} - \beta_{\{i,j\}} < 0$$

- for 'offending' elements *(i,j),* choose $\eta \in (0,1]$ s.t.

$$s_{ij} - g_{ij} - \beta^{(1)}_{\{i,j\}} = -\eta\, g_{ij}$$
$$s_{ji} - g_{ji} - \beta^{(2)}_{\{i,j\}} = -\eta\, g_{ji}$$

with $\beta_{\{i,j\}} = \max(\beta^{(1)}_{\{i,j\}}, \beta^{(2)}_{\{i,j\}})$

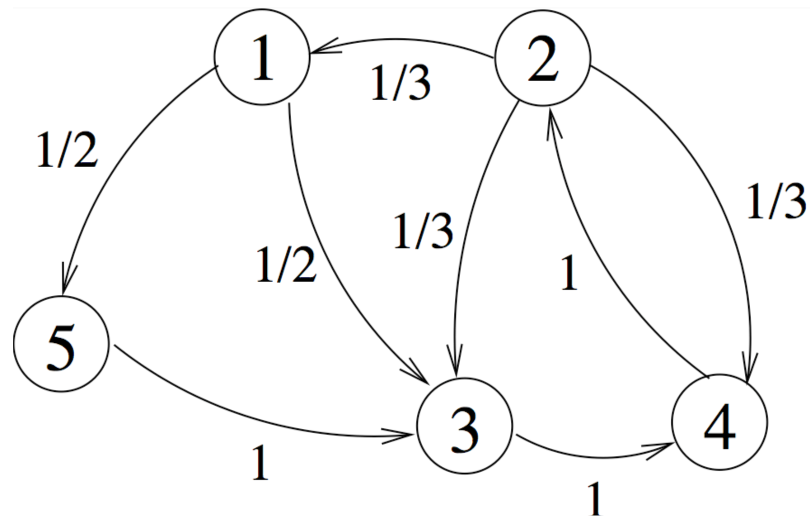
- η=1 means lump full value of offending elements of *S* ($\hat{s}_{ij}$ = 0)

University of Waterloo

hdesterck@uwaterloo.ca

1. Simple Markov Chain Example

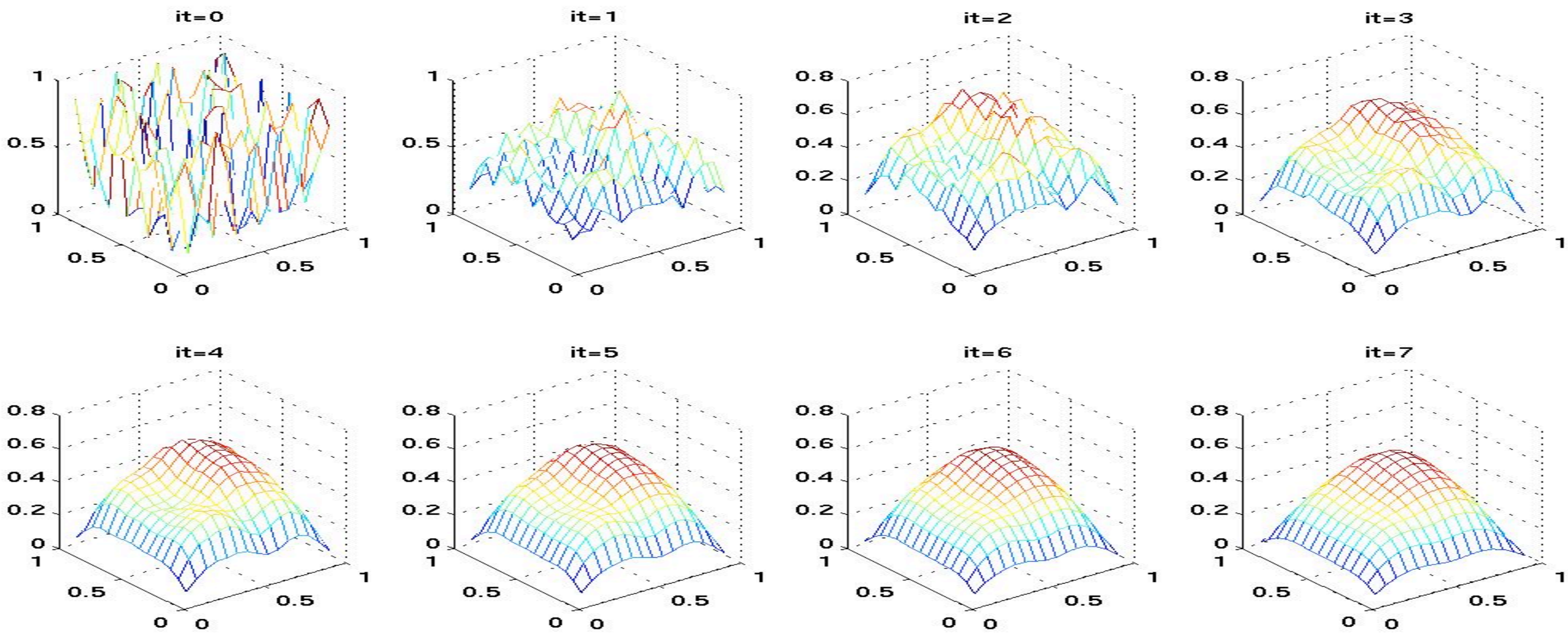
- 5 states
- each outgoing edge same probability (random walk on directed graph)



Copper 2008
hdesterck@uwaterloo.ca

Principle of Multigrid (for PDEs)

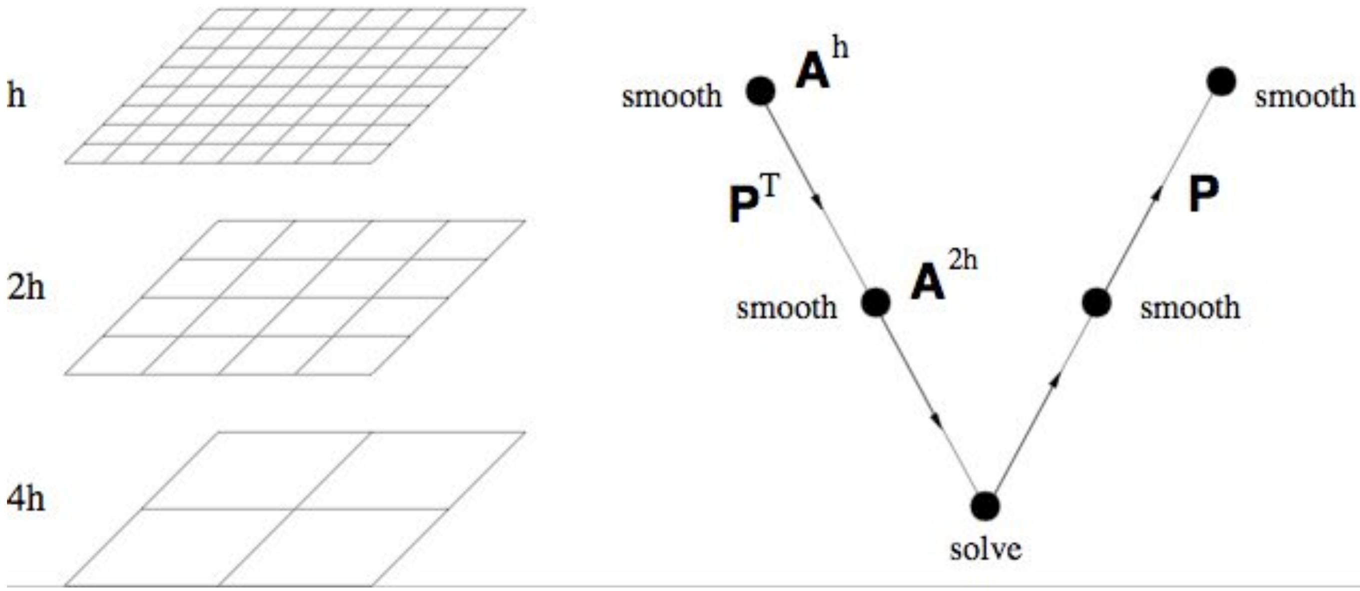
$$-u_{xx} - u_{yy} = f(x,y) \qquad Ax = b$$



- high-frequency error is removed by relaxation (weighted Jacobi, Gauss-Seidel, ...)
- low-frequency-error needs to be removed by coarse-grid correction

hdesterck@uwaterloo.ca

Multigrid Hierarchy: V-cycle



- multigrid V-cycle:
  - relax (=smooth) on successively coarser grids
  - transfer error using restriction ($R=P^T$) and interpolation (P)
- W=O(n)

Smoothed Aggregation

$$\begin{aligned} A_{cs} &= R_s \left(D - (L+U)\right) P_s \\ &= R_s \, D \, P_s - R_s \, (L+U) \, P_s \end{aligned}$$

$$A = \begin{bmatrix} + & - & - & - & - \\ - & + & - & - & - \\ - & - & + & - & - \\ - & - & - & + & - \\ - & - & - & - & + \end{bmatrix}$$

$$A_{cs} = S - G$$

$$\hat{A}_{cs} = \hat{S} - G$$

- we want to retain crucial properties

$$\begin{aligned} \mathbf{1}_c^T \, \hat{A}_{cs} &= 0 \quad \forall \, \mathbf{x}_i, \\ \hat{A}_{cs} \, \mathbf{1}_c &= 0 \quad \text{for } \mathbf{x}_\mathrm{i} = \mathbf{x} \end{aligned}$$

- we can lump to diagonal in symmetric way, conserving both row and column sums

University of Waterloo

hdesterck@uwaterloo.ca

Numerical Results: Geometric Aggregation (size 3)

n	γ_{res}	iter	C_{op}	levels
54	0.86	75	1.43	3
162	0.95	>100	1.47	4
486	0.97	>100	1.49	5
1458	0.98	>100	1.50	6

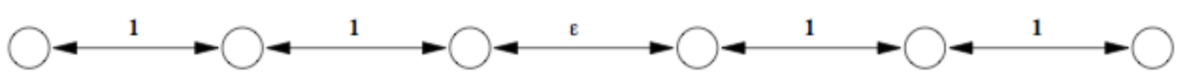
TABLE 5.5

Uniform chain with two weak links ($\epsilon = 0.001$). G-AM with V-cycles and size-three aggregates. The two weak links occur between aggregates at all levels. (No smoothing.)

n	γ_{res}	iter	C_{op}	levels	R_{lump}
54	0.26	12	1.43	3	0
162	0.27	13	1.47	4	0
486	0.27	13	1.49	5	0
1458	0.27	13	1.50	6	0
4374	0.27	13	1.50	7	0

TABLE 5.6

Uniform chain with two weak links ($\epsilon = 0.001$). G-SAM with V-cycles and size-three aggregates. The two weak links occur between aggregates at all levels. (Smoothing with lumping.)



Numerical Results: Geometric Aggregation (size 3)

n	γ_{res}	iter	C_{op}	levels
27	1.00	>100	1.32	2
81	1.00	>100	1.43	3
243	0.98	>100	1.47	4
729	0.98	>100	1.49	5

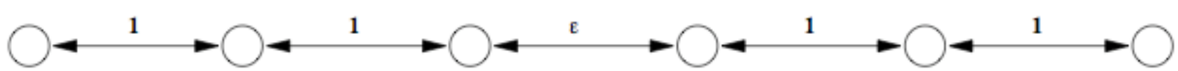
TABLE 5.7

Uniform chain with two weak links ($\epsilon = 0.001$). G-AM with V-cycles and size-three aggregates. The two weak links occur inside an aggregate on the finest level. (No smoothing.)

n	γ_{res}	iter	C_{op}	levels	R_{lump}
27	1.00	>100	1.32	2	0
81	1.00	>100	1.43	3	0
243	1.00	>100	1.47	4	0
729	1.00	>100	1.49	5	0

TABLE 5.8

Uniform chain with two weak links ($\epsilon = 0.001$). G-SAM with V-cycles and size-three aggregates. The two weak links occur inside an aggregate on the finest level. (Smoothing with lumping.)



University of Waterloo

Numerical Results: effect of η

n	γ_{res}	iter	C_{op}	levels
27	0.75	43	1.71	3
81	0.87	87	1.85	4
243	0.96	>100	1.96	6
729	0.99	>100	1.98	7

TABLE 6.1
Uniform chain. A-AM with V-cycles and distance-one aggregation. (No smoothing.)

n	γ_{res}	iter	C_{op}	levels	γ_{res}	iter	C_{op}	levels
27	0.20	10	2.03	3	0.19	10	2.03	3
81	0.27	11	2.69	4	0.25	11	2.59	4
243	0.32	12	3.33	6	0.32	12	3.33	6
729	0.51	14	3.73	8	0.64	17	3.75	8
2187	0.75	21	3.95	9	0.77	24	3.94	9

TABLE 6.3
Uniform chain. A-SAM with V-cycles using distance-one aggregation, lumping only the off-diagonal elements of $R_s\,D\,P_s$ that cause nonnegative off-diagonal elements of A_{cs}. Lumping their full value ($\eta = 1$, left), and part of their value ($\eta = 0.75$, right).

Copper 2008
hdesterck@uwaterloo.ca

Numerical Results: effect of η

n	γ_{res}	iter	C_{op}	levels	γ_{res}	iter	C_{op}	levels
27	0.19	10	2.03	3	0.19	10	1.96	3
81	0.19	10	2.51	4	0.20	10	2.61	4
243	0.24	11	2.96	5	0.22	11	3.20	5
729	0.37	12	3.63	7	0.29	12	3.29	6
2187	0.44	14	3.84	8	0.28	12	3.35	7

TABLE 6.4

Uniform chain. A-SAM with V-cycles using distance-one aggregation, lumping part of the value of the off-diagonal elements of $R_s\, D\, P_s$ that cause nonnegative off-diagonal elements of A_{cs}: $\eta = 0.25$, left, and, $\eta = 0.1$, right.

n	γ_{res}	iter	C_{op}	levels	γ_{res}	iter	C_{op}	levels
27	0.20	10	2.03	3	0.20	10	2.03	3
81	0.18	10	2.68	4	0.18	10	2.71	4
243	0.19	10	2.78	5	0.20	10	3.03	5
729	0.24	11	3.41	7	0.24	11	3.50	7
2187	0.27	11	3.75	8	0.26	11	3.81	8

TABLE 6.5

Uniform chain. A-SAM with V-cycles using distance-one aggregation, lumping part of the value of the off-diagonal elements of $R_s\, D\, P_s$ that cause nonnegative off-diagonal elements of A_{cs}: $\eta = 0.01$, left, and, $\eta = 1e-6$, right.

Copper 2008
hdesterck@uwaterloo.ca

Numerical Results: Algebraic Aggregation

n	γ_{res}	iter	C_{op}	levels
27	0.75	50	1.71	3
81	0.87	92	1.85	4
243	0.96	>100	1.96	6
729	0.97	>100	1.99	8

TABLE 6.6

Birth-death chain ($\mu = 0.96$). A-AM with V-cycles and distance-one aggregation. (No smoothing.)

n	γ_{res}	iter	C_{op}	levels	R_{lump}
27	0.27	12	1.32	2	0
81	0.35	15	1.43	3	0
243	0.35	15	1.47	4	0
729	0.35	15	1.49	5	0

TABLE 6.7

Birth-death chain ($\mu = 0.96$). A-SAM with V-cycles and distance-two aggregation. (Smoothing with lumping.)

Copper 2008
hdesterck@uwaterloo.ca

Numerical Results: Algebraic Aggregation

n	γ_{res}	iter	C_{op}	levels
64	0.73	41	1.73	3
100	0.80	56	1.83	4
169	0.85	77	1.85	4
400	0.89	>100	1.96	6
900	0.96	>100	1.96	6

TABLE 6.10

Uniform 2D lattice. A-AM with V-cycles and distance-one aggregation. (No smoothing.)

n	γ_{res}	iter	C_{op}	levels	R_{lump}
64	0.42	18	1.30	3	0
100	0.45	19	1.26	3	3.46e-03
169	0.44	18	1.33	3	9.47e-03
400	0.47	20	1.46	4	9.27e-03
900	0.46	18	1.59	4	1.72e-02
1600	0.48	19	1.60	4	1.16e-02
2500	0.48	19	1.67	5	1.44e-02
4900	0.48	18	1.75	5	1.21e-02
6724	0.49	18	1.76	5	1.36e-02

TABLE 6.11

Uniform 2D lattice. A-SAM with V-cycles and distance-two aggregation. (Smoothing with lumping.)

4. Aggregation for Markov Chains

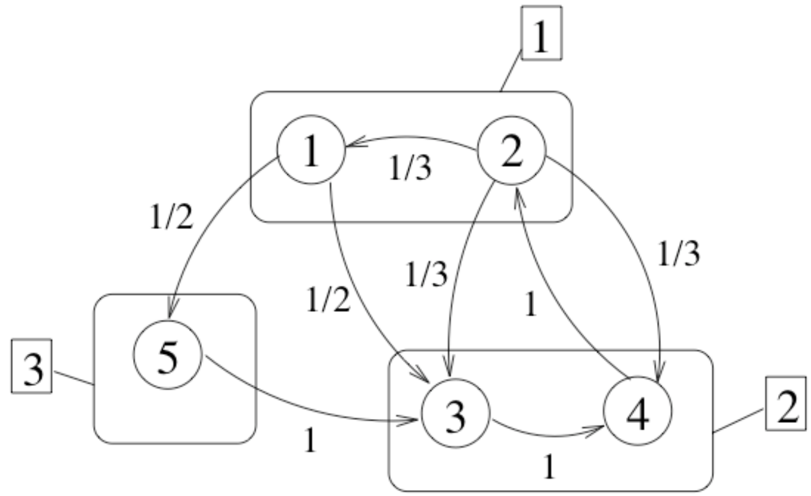
- form three coarse, aggregated states

$$x_{c,I} = \sum_{i \in I} x_i$$

$$\mathbf{x}_c^T = [8/19 \;\; 10/19 \;\; 1/19]$$

$$B_c \mathbf{x}_c = \mathbf{x}_c$$

$$b_{c,IJ} = \frac{\sum\limits_{j \in J} x_j \left(\sum\limits_{i \in I} b_{ij} \right)}{\sum\limits_{j \in J} x_j}$$



$$B_c = \begin{bmatrix} 1/4 & 3/5 & 0 \\ 5/8 & 2/5 & 1 \\ 1/8 & 0 & 0 \end{bmatrix}$$

(Simon and Ando, 1961)

University of Waterloo

Copper 2008
hdesterck@uwaterloo.ca

5. Error Equation

- multiplicative correction: error equation, coarse level error equation, and coarse grid correction

$$\mathbf{x} = \mathrm{diag}(\mathbf{x}_i)\,\mathbf{e}_i$$
$$A\,\mathrm{diag}(\mathbf{x}_i)\,\mathbf{e}_i = 0$$

$$Q^T A\,\mathrm{diag}(\mathbf{x}_i)\,Q\,\mathbf{e}_c = 0$$
$$A_c\,\mathbf{e}_c = 0$$

$$R = Q^T \qquad P = \mathrm{diag}(\mathbf{x}_i)\,Q$$
$$A_c = R\,A\,P$$

University of Waterloo

Copper 2008
hdesterck@uwaterloo.ca

5. Error Equation

- error equation - coarse grid correction:

$$\mathbf{x} = \mathrm{diag}(\mathbf{x}_i)\,\mathbf{e}_i$$
$$A\,\mathrm{diag}(\mathbf{x}_i)\,\mathbf{e}_i = 0$$

$$Q^T A\,\mathrm{diag}(\mathbf{x}_i)\,Q\,\mathbf{e}_c = 0$$
$$A_c\,\mathbf{e}_c = 0$$

$$R = Q^T \qquad P = \mathrm{diag}(\mathbf{x}_i)\,Q$$
$$A_c = R\,A\,P$$

$$\mathbf{x}_{i+1} = P\,\mathbf{e}_c$$

$$\mathbf{x}_c = \mathrm{diag}(P^T\,\mathbf{1})\,\mathbf{e}_c$$

$$A_c\,(\mathrm{diag}(P^T\,\mathbf{1}))^{-1}\,\mathbf{x}_c = 0$$

$$\mathbf{x}_{i+1} = P\,(\mathrm{diag}(P^T\,\mathbf{1}))^{-1}\mathbf{x}_c$$

Copper 2008
hdesterck@uwaterloo.ca

Numerical Results: Algebraic Aggregation

n	γ_{res}	iter	C_{op}	levels
54	0.83	67	1.86	4
162	0.93	>100	1.91	5
486	0.96	>100	1.98	7
1458	0.97	>100	1.99	9

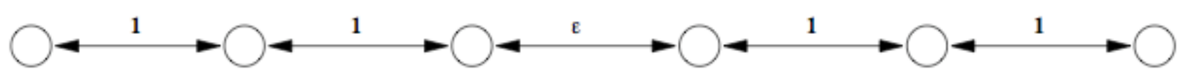
TABLE 6.8

Uniform chain with two weak links ($\epsilon = 0.001$). A-AM with V-cycles and distance-one aggregation. (No smoothing.)

n	γ_{res}	iter	C_{op}	levels	R_{lump}
54	0.33	14	1.51	3	0
162	0.33	13	1.50	4	5.51e-03
486	0.34	13	1.50	5	0
1458	0.29	12	1.49	6	3.07e-04
4374	0.27	11	1.50	7	3.05e-04

TABLE 6.9

Uniform chain with two weak links ($\epsilon = 0.001$). A-SAM with V-cycles and distance-two aggregation. (Smoothing with lumping.)



University of Waterloo