Smoothed Aggregation Multigrid for Slowly Mixing Markov Chains

Hans De Sterck, Killian Miller, 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

EMG 2008

1. Simple Markov Chain Example

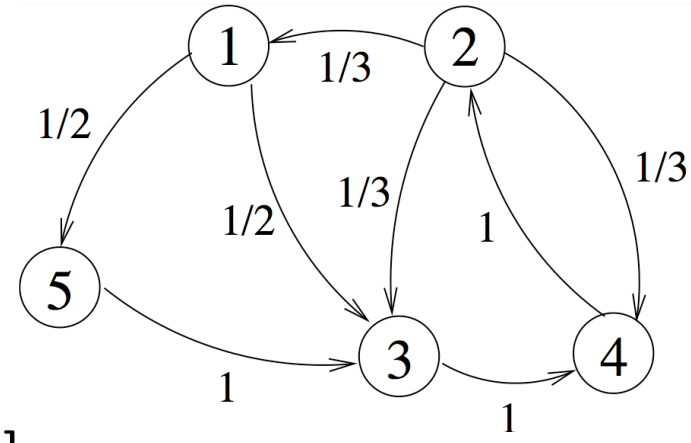
- start in one state with probability 1: what is the stationary probability vector after ∞ number of steps?

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

- stationary probability:

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

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



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

2. Problem Statement

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

- *B* is column-stochastic

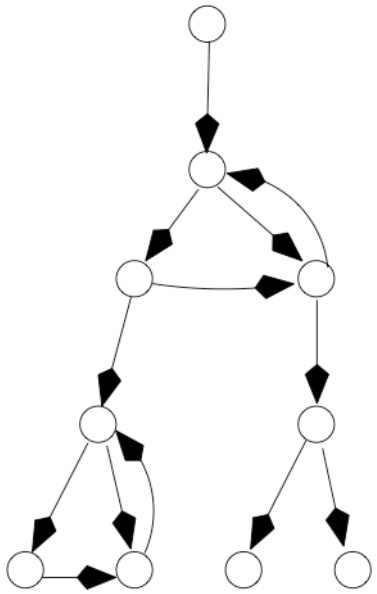
$$0 \leq b_{ij} \leq 1\ \forall i,\ j \qquad \mathbf{1}^T B = \mathbf{1}^T$$

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

$$\Rightarrow$$

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

(no probability sinks!)



3. Power Method

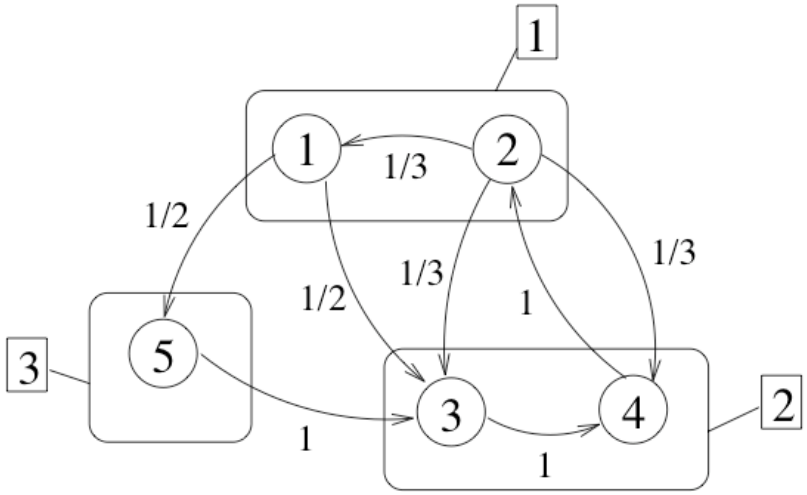
$$B\,\mathbf{x} = \mathbf{x} \quad \text{or} \quad (I - B)\,\mathbf{x} = 0 \quad \text{or} \quad A\,\mathbf{x} = 0$$

- 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

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

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



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

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

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

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

(Krieger, Horton, ... 1990s)

5. Error Equation

- error equation - coarse grid correction:

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

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

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

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

Error Equation

- important properties of A_c:

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

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

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

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

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

6. Multilevel Aggregation Algorithm

Algorithm: Multilevel Adaptive Aggregation method (V-cycle)

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

begin

$\mathbf{x} \leftarrow \text{Relax}(A, \mathbf{x})$ ν_1 times

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

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

$A_c = R\, A\, P$

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

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

$\mathbf{x} \leftarrow \text{Relax}(A, \mathbf{x})$ ν_2 times

end

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

7. Well-posedness: Singular M-matrices

- singular M-matrix:

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

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

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

- 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]).

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 $\mathbf{1}^T A = 0$)

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

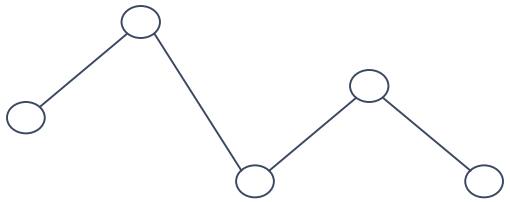
THEOREM 3.2 (Fixed-point property of AM V-cycle). *Exact solution $\mathbf{x}$ is a fixed point of the AM V-cycle.*

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

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

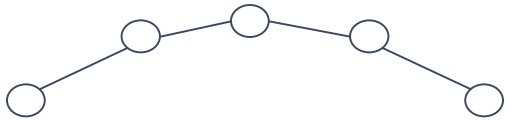
8. We Need 'Smoothed Aggregation'...

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

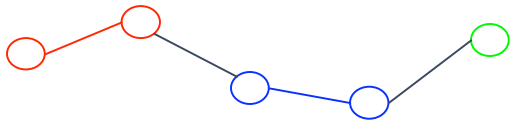


$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

after smoothing:

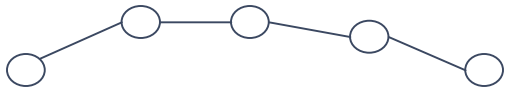


coarse grid correction with Q:



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

coarse grid correction with Q_s:



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:

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

Smoothed Aggregation

- smoothed coarse level operator:

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

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

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

- solution: lumping approach on S in

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

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

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):*

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

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

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

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

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

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

conserves both row and column sums

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.*

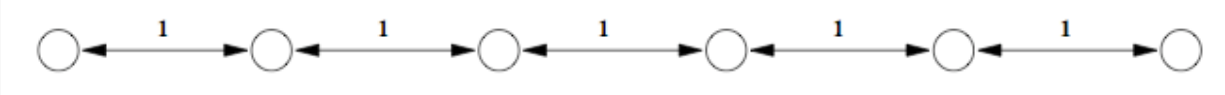
THEOREM 4.2 (Fixed-point property of lumped SAM V-cycle). *Exact solution $\mathbf{x}$ is a fixed point of the SAM V-cycle (with lumping).*

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

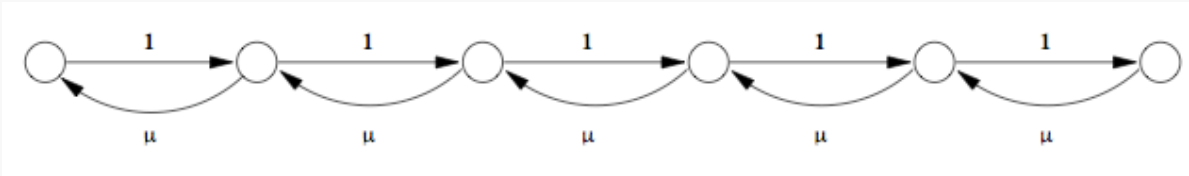
$$A = \begin{bmatrix} + & - & - & - & - \\ - & + & - & - & - \\ - & - & + & - & - \\ - & - & - & + & - \\ - & - & - & - & + \end{bmatrix}$$

10. Numerical Results: Test Problems

- uniform chain

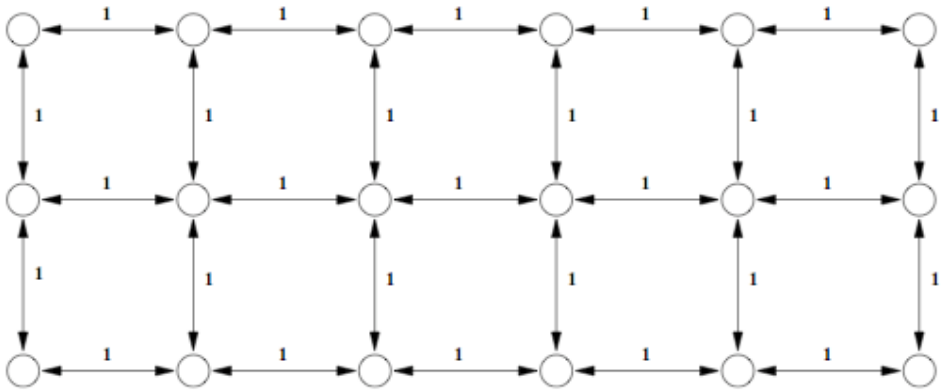


- birth-death chain

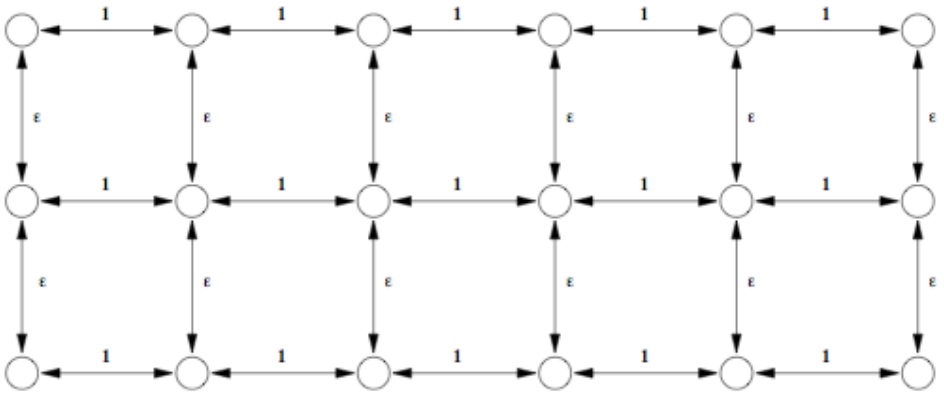


Test Problems

- uniform 2D lattice



- anisotropic 2D lattice



University of Waterloo

EMG 2008
hdesterck@uwaterloo.ca

Test Problems

- tandem queueing network

FIG. 5.6. *Tandem queueing network.*

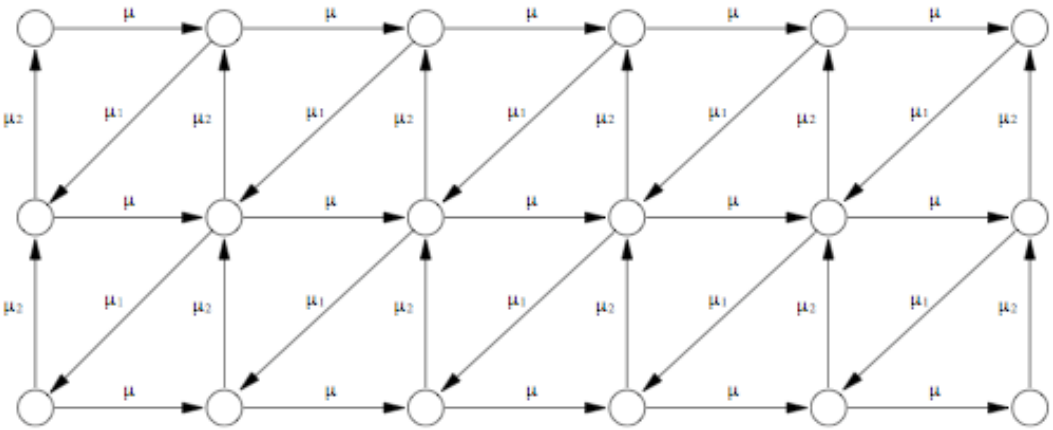
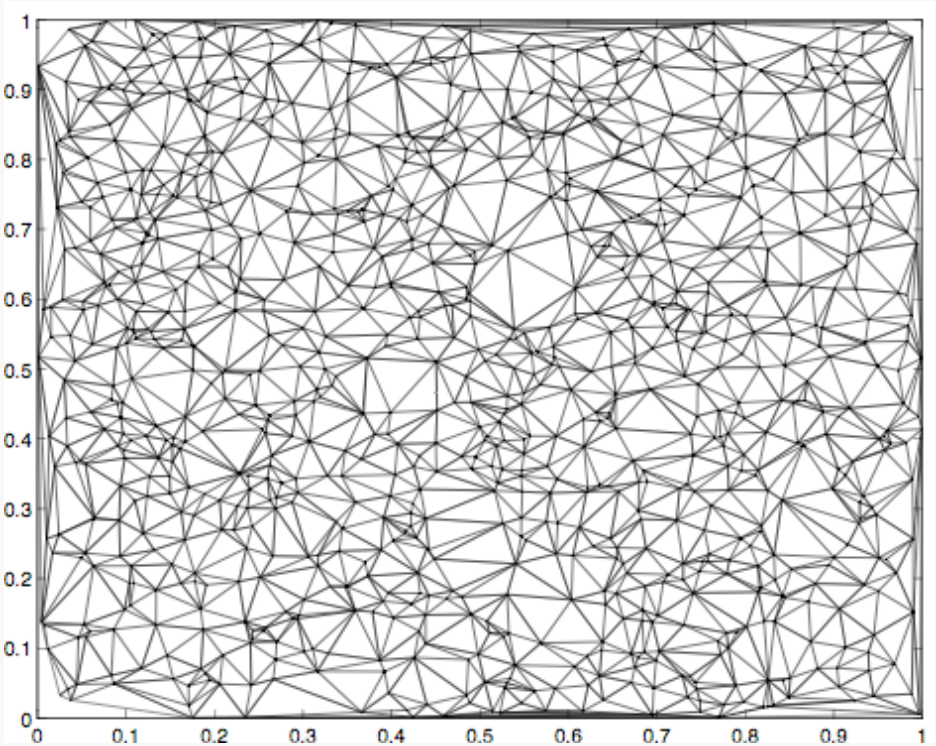


FIG. 5.7. *Graph for tandem queueing network.*

Test Problems

- random walk on triangulation of random points in unit square



11. Numerical Results: Geometric Aggregation (size 3)

n	lev	it	C_{op}	γ	γ_{eff}	it	C_{op}	γ	γ_{eff}
27	2	32	1.32	0.66	0.73	32	1.63	0.66	0.78
81	3	85	1.43	0.87	0.91	52	2.07	0.80	0.90
243	4	>100	1.47	0.95	0.97	73	2.37	0.87	0.94
729	5	>100	1.49	0.98	0.98	>100	2.58	0.92	0.97
2187	6	>100	1.50	0.98	0.98	>100	2.72	0.95	0.98

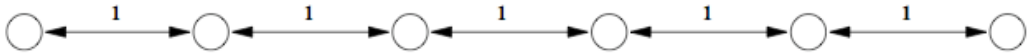
TABLE 5.1

Uniform chain. G-AM with V-cycles (left) and W-cycles (right). (Size-three aggregates, no smoothing.)

n	it	C_{op}	lev	γ	γ_{eff}	R_{lump}
27	13	1.32	2	0.27	0.37	0
81	13	1.43	3	0.27	0.40	0
243	13	1.47	4	0.27	0.41	0
729	13	1.49	5	0.27	0.41	0
2187	13	1.50	6	0.27	0.42	0
6561	13	1.50	7	0.27	0.42	0
19683	13	1.50	8	0.27	0.42	0
59049	13	1.50	9	0.27	0.42	0

TABLE 5.2

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

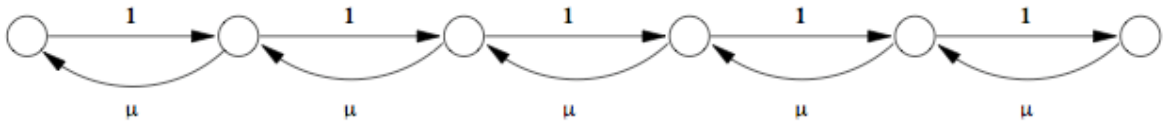


Numerical Results: Geometric Aggregation (size 3)

n	C_{op}	lev	it	γ	γ_{eff}	it	γ	γ_{eff}	R_{lump}
27	1.32	2	33	0.66	0.73	13	0.27	0.37	0
81	1.43	3	95	0.88	0.92	12	0.27	0.40	0
243	1.47	4	>100	0.95	0.97	13	0.26	0.40	0
729	1.49	5	>100	0.97	0.98	12	0.24	0.38	0

TABLE 5.4

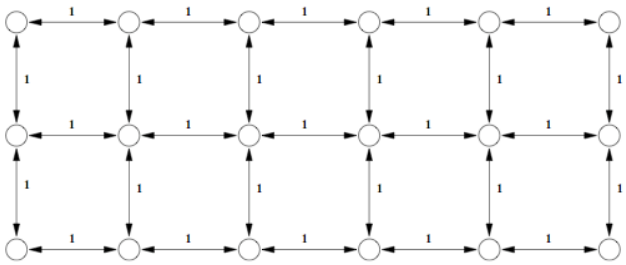
Birth-death chain ($\mu = 0.96$). (left) G-AM (no smoothing) and (right) G-SAM (smoothing with lumping). (V-cycles and size-three aggregates.)



University of Waterloo

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

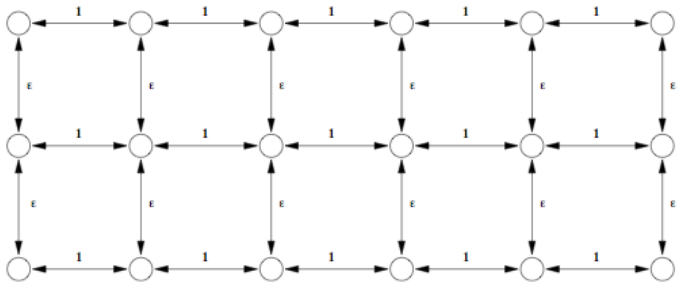
Uniform 2D lattice. G-AM with V-cycles and size-three aggregates. (No smoothing.)

n	lev	it	C_{op}	γ	γ_{eff}	R_{lump}	it	C_{op}	γ	γ_{eff}	R_{lump}
64	2	22	1.17	0.49	0.55	0	17	1.34	0.39	0.49	0
256	3	24	1.22	0.54	0.61	0	17	1.47	0.42	0.55	0
1024	4	25	1.22	0.57	0.63	6.6e-4	18	1.49	0.44	0.58	4.3e-3
4096	4	26	1.23	0.58	0.64	9.7e-4	19	1.52	0.45	0.59	3.6e-3
16384	5	27	1.22	0.61	0.67	4.0e-5	19	1.51	0.45	0.59	1.3e-4
65536	6	27	1.23	0.62	0.67	1.5e-5	19	1.52	0.45	0.59	2.6e-4

TABLE 5.8

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

Numerical Results: Geometric Aggregation (3x3)

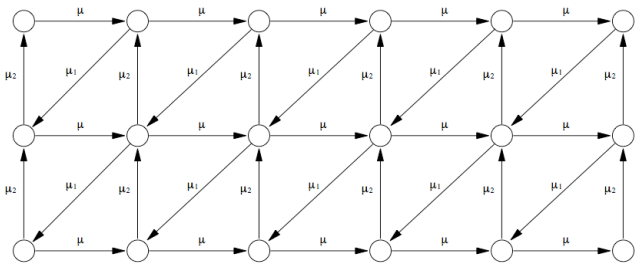


n	lev	it	C_{op}	γ	γ_{eff}	it	C_{op}	γ	γ_{eff}	R_{lump}
64	2	>100	1.11	1.00	1.00	>100	1.17	1.00	1.00	0
256	3	>100	1.14	0.97	0.97	>100	1.22	0.95	0.96	0
1024	4	>100	1.13	0.99	0.99	>100	1.22	0.99	0.99	0
4096	4	>100	1.13	0.99	0.99	>100	1.23	0.99	0.99	0

TABLE 5.9

Anisotropic 2D lattice ($\epsilon = 1e{-}6$). (left) G-AM (no smoothing) and (right) G-SAM (smoothing with lumping) with V-cycles and three-by-three aggregates.

Numerical Results: Geometric Aggregation (3x3)



n	it	C_{op}	lev	γ	γ_{eff}
256	>100	1.19	3	0.91	0.93
1024	>100	1.19	4	0.96	0.97
4096	>100	1.19	4	0.96	0.97

TABLE 5.10

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

n	lev	it	C_{op}	γ	γ_{eff}	R_{lump}	it	C_{op}	γ	γ_{eff}	R_{lump}
256	3	19	1.28	0.44	0.52	2.4e-2	14	1.60	0.33	0.50	3.9e-2
1024	4	19	1.27	0.46	0.55	2.0e-3	14	1.62	0.32	0.50	5.9e-3
4096	4	20	1.29	0.49	0.57	7.4e-3	14	1.65	0.33	0.51	1.4e-2
16384	5	21	1.28	0.56	0.63	8.1e-4	14	1.63	0.33	0.51	2.0e-3
65536	6	21	1.28	0.51	0.59	1.6e-3	14	1.64	0.33	0.51	3.8e-3

TABLE 5.11

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

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 paper on PageRank in SISC, 2008)

Numerical Results: Algebraic Aggregation

n	it	C_{op}	lev	γ	γ_{eff}
27	39	1.71	3	0.74	0.84
81	83	1.85	4	0.87	0.93
243	>100	1.96	6	0.96	0.98
729	>100	1.98	7	1.00	1.00

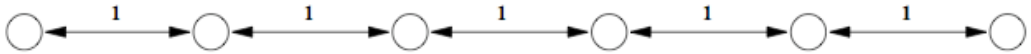
TABLE 6.1

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

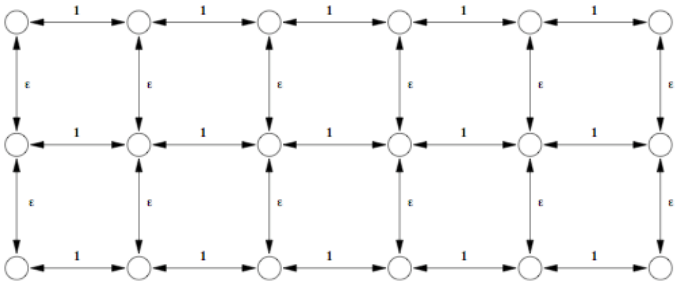
n	it	C_{op}	lev	γ	γ_{eff}	R_{lump}	it	C_{op}	lev	γ	γ_{eff}	R_{lump}
27	10	2.09	3	0.20	0.46	2.4e-2	13	1.33	2	0.32	0.43	0
81	10	2.09	4	0.18	0.44	0.0e-0	13	1.44	3	0.32	0.45	0
243	10	2.20	5	0.21	0.49	1.3e-3	12	1.46	4	0.32	0.46	0
729	11	2.23	6	0.21	0.50	3.3e-3	12	1.49	5	0.31	0.45	0
2187	11	2.24	6	0.23	0.52	2.5e-3	12	1.49	6	0.31	0.46	0
6561	11	2.24	8	0.24	0.53	1.4e-3	12	1.49	7	0.31	0.46	0
19683	11	2.25	8	0.26	0.55	1.4e-3	12	1.49	8	0.32	0.47	0
59049	12	2.25	9	0.30	0.59	1.0e-3	12	1.50	9	0.32	0.47	0

TABLE 6.2

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



Numerical Results: Algebraic Aggregation

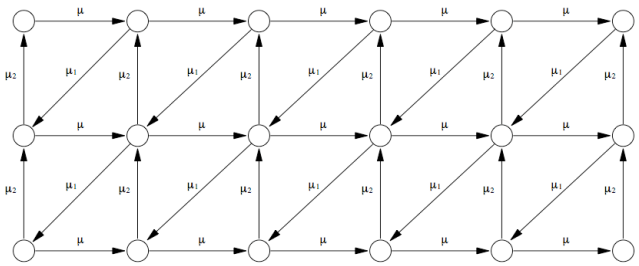


n	it	C_{op}	lev	γ	γ_{eff}	it	C_{op}	lev	γ	γ_{eff}	R_{lump}
64	27	1.78	4	0.55	0.72	17	1.76	3	0.40	0.59	0.0e-0
256	45	1.93	6	0.74	0.86	15	2.23	4	0.33	0.61	7.4e-4
1024	80	2.01	8	0.86	0.93	14	2.81	5	0.33	0.68	1.6e-3
4096	>100	2.04	10	0.93	0.96	14	3.43	7	0.33	0.73	4.9e-4
16384	>100	2.06	11	0.96	0.98	13	4.17	7	0.33	0.77	2.5e-4
65536	>100	2.16	13	0.97	0.99	13	4.80	9	0.32	0.79	7.6e-5

TABLE 6.7

Anisotropic 2D lattice ($\epsilon = 1e-6$). (left) A-AM (no smoothing, distance-one aggregation) and (right) A-SAM (smoothing with lumping, distance-two aggregation) with V-cycles.

Numerical Results: Algebraic Aggregation



n	it	C_{op}	lev	γ	γ_{eff}
256	>100	1.86	5	0.92	0.96
1024	>100	1.96	6	0.91	0.95
4096	>100	2.03	7	0.98	0.99
16384	>100	2.09	9	0.98	0.99

TABLE 6.8

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

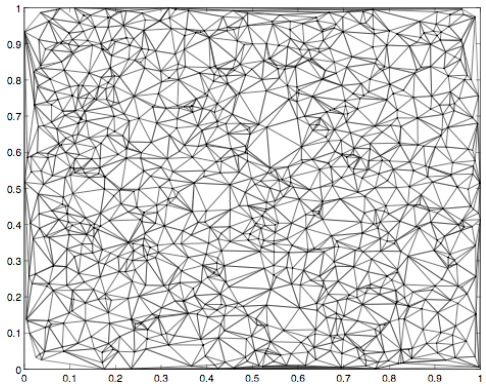
n	it	C_{op}	lev	γ	γ_{eff}	R_{lump}
256	18	1.94	4	0.39	0.61	1.1e-1
1024	20	2.04	4	0.41	0.64	7.6e-2
4096	24	2.12	5	0.45	0.69	5.5e-2
16384	30	2.18	6	0.56	0.77	5.3e-2
65536	37	2.37	6	0.71	0.86	1.3e-1

TABLE 6.9

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

University of Waterloo

Numerical Results: Algebraic Aggregation



$n = 32768$	it	C_{op}	lev	γ	γ_{eff}
multilevel A-AM	>100	1.26	8	0.9827	0.9862
two-level A-AM	>100	1.20	2	0.9833	0.9861
one-level method	>100	1.00	1	0.9846	0.9846

TABLE 6.11

Unstructured planar graph with $n = 32768$ nodes. Comparison of multilevel A-AM, two-level A-AM (aggregation-disaggregation), and one-level relaxation. (V-cycles and distance-one aggregation, no smoothing. The V(1,1)-cycles have two relaxations per level per cycle.))

n	it	C_{op}	lev	γ	γ_{eff}	R_{lump}
1024	20	1.69	5	0.5265	0.6848	2.6e-02
2048	19	1.68	5	0.5200	0.6779	2.1e-02
4096	21	1.80	5	0.6069	0.7578	2.4e-02
8192	22	1.92	7	0.6360	0.7904	2.5e-02
16384	30	2.03	7	0.7610	0.8744	2.4e-02
32768	28	2.08	7	0.7399	0.8649	2.4e-02

TABLE 6.10

Unstructured planar graph. A-SAM with V-cycles and distance-one aggregation. (Smoothing with lumping.))

13. AMG version of algorithm also works

$$\mathbf{v} \leftarrow V(A, \mathbf{x})$$

1. Relax a_1 times on $A\mathbf{u} = \mathbf{0}$ with initial guess $\mathbf{x}$.
2. If you have reached the coarsest level, then go to step 4.
   Else

   Set $\bar{A} \leftarrow A\,\mathrm{diag}(\mathbf{x})$.
   Compute the set of coarse grid points C.
   Construct the interpolation operator P.
   Construct the coarse grid operator $\bar{A}_c \leftarrow P^T \bar{A} P$.
   Obtain lumped coarse grid operator $\hat{A}_c \leftarrow Lump(\bar{A}_c, \eta)$.
   Recursively call the algorithm $\mathbf{e}_c \leftarrow V(\hat{A}_c, \mathbf{1}_c)$.
3. Correct $\mathbf{x} \leftarrow \mathrm{diag}(\mathbf{x}) P \mathbf{e}_c$.
4. Relax a_2 times on $A\mathbf{u} = \mathbf{0}$ with initial guess $\mathbf{x}$.

14. Conclusions

- SAM (Smoothed Aggregation for Markov Chains): algorithm for stationary vector of slowly mixing Markov chains with near-optimal complexity
- smoothing is essential for aggregation for many problems
- pretty good convergence results, operator complexity may grow
- AMG works too
- appropriate theoretical framework (well-posedness)
- are there other ways for choosing R, P, lumping?
- no theory yet on optimal convergence (non-symmetric matrices)

- Questions?

Smoothed Aggregation

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

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

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

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

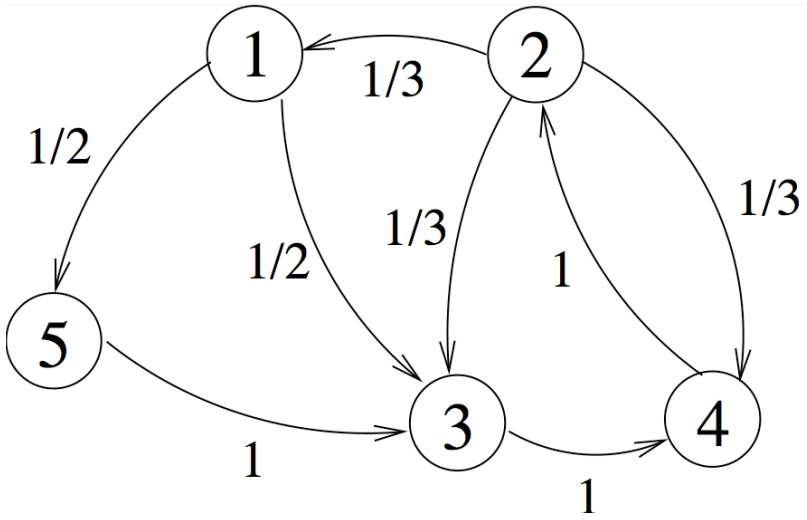
$$s_{ij} - g_{ij} - \beta^{(1)}_{\{i,j\}} = -\eta\, g_{ij}$$
$$s_{ji} - g_{ji} - \beta^{(2)}_{\{i,j\}} = -\eta\, g_{ji}$$

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)

1. Simple Markov Chain Example

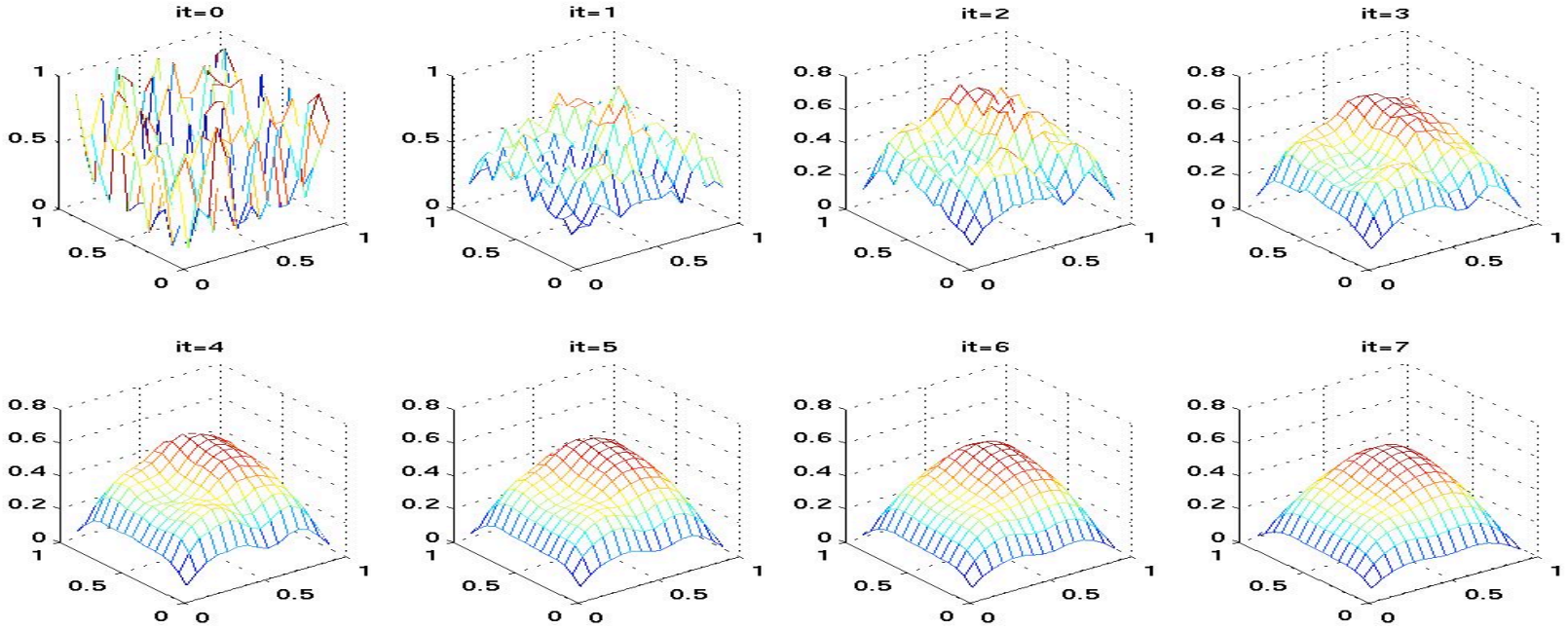
- 5 states
- each outgoing edge same probability (random walk on directed graph)



EMG 2008
hdesterck@uwaterloo.ca

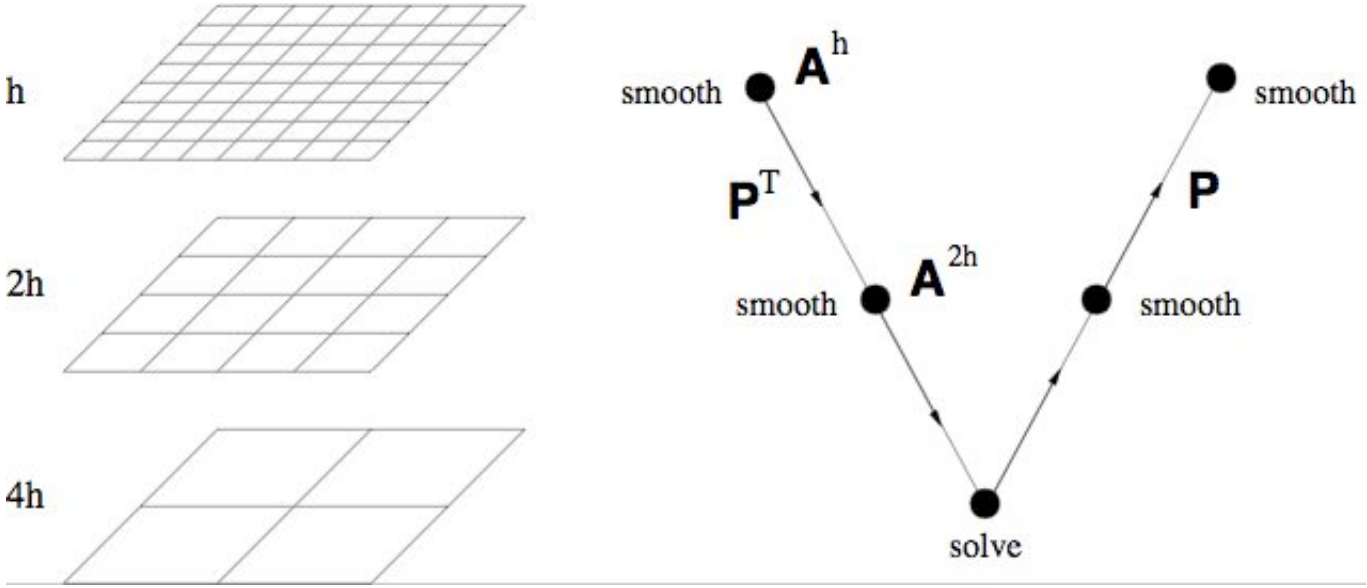
Principle of Multigrid (for PDEs)

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



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

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

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

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

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

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

- we want to retain crucial properties

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

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

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

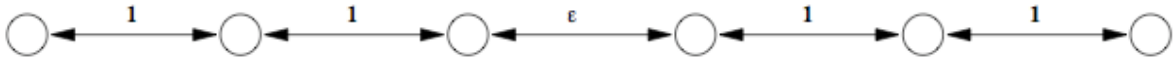
TABLE 5.5

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

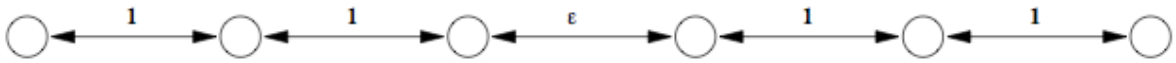
TABLE 5.7

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.)



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).

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.

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.)

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

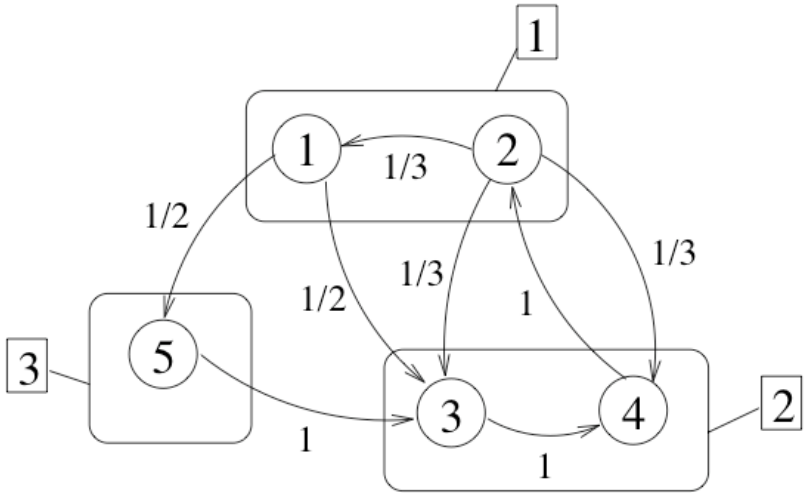
- form three coarse, aggregated states

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

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

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

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



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

(Simon and Ando, 1961)

5. Error Equation

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

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

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

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

5. Error Equation

- error equation - coarse grid correction:

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

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

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

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

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

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

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

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.)