

David Aldous Persi Diaconis
Joel Spencer J. Michael Steele
Editors

Discrete Probability and Algorithms



Springer-Verlag

0158
6
David Aldous
Joel Spencer

Persi Diaconis
J. Michael Steele

Editors

Discrete Probability and Algorithms

With 7 Illustrations

江苏工业学院图书馆
藏书章



Springer-Verlag

New York Berlin Heidelberg London Paris
Tokyo Hong Kong Barcelona Budapest

David Aldous
Department of Statistics
University of California
Berkeley, CA 94720 USA

Persi Diaconis
Department of Mathematics
Harvard University
Cambridge, MA 02138 USA

Joel Spencer
Department of Computer Science
Courant Institute
New York University
251 Mercer Street
New York, NY 10012 USA

J. Michael Steele
Department of Statistics
The Wharton School
University of Pennsylvania
3000 Steinberg Hall-Dietrich Hall
Philadelphia, PA 19104-6302 USA

Mathematics Subject Classifications (1991): 60-06, 05-06, 68-06

Library of Congress Cataloging-in-Publication Data

Discrete probability and algorithms / David Aldous . . . [et al.],
editors.

p. cm. — (The IMA volumes in mathematics and its
applications ; v. 72)

Papers from two workshops held at the University of Minnesota in fall
1993.

Includes bibliographical references.

ISBN 0-387-94532-6 (hardcover)

1. Probabilities—Congresses. 2. Algorithms—Congresses.

I. Aldous, D. J. (David J.) II. Series.

QA273.A1D57 1995

519.2—dc20

95-151

Printed on acid-free paper.

© 1995 Springer-Verlag New York, Inc.

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use of general descriptive names, trade names, trademarks, etc., in this publication, even if the former are not especially identified, is not to be taken as a sign that such names, as understood by the Trade Marks and Merchandise Marks Act, may accordingly be used freely by anyone.

Permission to photocopy for internal or personal use, or the internal or personal use of specific clients, is granted by Springer-Verlag, Inc., for libraries registered with the Copyright Clearance Center (CCC), provided that the base fee of \$5.00 per copy, plus \$0.20 per page, is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923, USA. Special requests should be addressed directly to Springer-Verlag New York, 175 Fifth Avenue, New York, NY 10010, USA.

ISBN 0-387-94532-6/1995 \$5.00 + 0.20

Production managed by Hal Henglein; manufacturing supervised by Joe Quatela.

Camera-ready copy prepared by the IMA.

Printed and bound by Braun-Brumfield, Ann Arbor, MI.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-94532-6 Springer-Verlag New York Berlin Heidelberg

The IMA Volumes in Mathematics and its Applications

Current Volumes:

Volume 1: Homogenization and Effective Moduli of Materials and Media
Editors: Jerry Ericksen, David Kinderlehrer, Robert Kohn, and
J.-L. Lions

Volume 2: Oscillation Theory, Computation, and Methods of
Compensated Compactness
Editors: Constantine Dafermos, Jerry Ericksen,
David Kinderlehrer, and Marshall Slemrod

Volume 3: Metastability and Incompletely Posed Problems
Editors: Stuart Antman, Jerry Ericksen, David Kinderlehrer, and
Ingo Muller

Volume 4: Dynamical Problems in Continuum Physics
Editors: Jerry Bona, Constantine Dafermos, Jerry Ericksen, and
David Kinderlehrer

Volume 5: Theory and Applications of Liquid Crystals
Editors: Jerry Ericksen and David Kinderlehrer

Volume 6: Amorphous Polymers and Non-Newtonian Fluids
Editors: Constantine Dafermos, Jerry Ericksen, and
David Kinderlehrer

Volume 7: Random Media
Editor: George Papanicolaou

Volume 8: Percolation Theory and Ergodic Theory of Infinite Particle
Systems
Editor: Harry Kesten

Volume 9: Hydrodynamic Behavior and Interacting Particle Systems
Editor: George Papanicolaou

Volume 10: Stochastic Differential Systems, Stochastic Control Theory,
and Applications
Editors: Wendell Fleming and Pierre-Louis Lions

Volume 11: Numerical Simulation in Oil Recovery
Editor: Mary Fanett Wheeler

Volume 12: Computational Fluid Dynamics and Reacting Gas Flows
Editors: Bjorn Engquist, M. Luskin, and Andrew Majda

Volume 13: Numerical Algorithms for Parallel Computer Architectures
Editor: Martin H. Schultz

Volume 14: Mathematical Aspects of Scientific Software
Editor: J.R. Rice

Volume 15: Mathematical Frontiers in Computational Chemical Physics
Editor: D. Truhlar

Volume 16: Mathematics in Industrial Problems
by Avner Friedman

Volume 17: Applications of Combinatorics and Graph Theory to the
Biological and Social Sciences
Editor: Fred Roberts

Volume 18: q -Series and Partitions
Editor: Dennis Stanton

Volume 19: Invariant Theory and Tableaux
Editor: Dennis Stanton

Volume 20: Coding Theory and Design Theory Part I: Coding Theory
Editor: Dijen Ray-Chaudhuri

Volume 21: Coding Theory and Design Theory Part II: Design Theory
Editor: Dijen Ray-Chaudhuri

Volume 22: Signal Processing: Part I - Signal Processing Theory
Editors: L. Auslander, F.A. Grünbaum, J.W. Helton, T. Kailath,
P. Khargonekar, and S. Mitter

Volume 23: Signal Processing: Part II - Control Theory and Applications
of Signal Processing
Editors: L. Auslander, F.A. Grünbaum, J.W. Helton, T. Kailath,
P. Khargonekar, and S. Mitter

Volume 24: Mathematics in Industrial Problems, Part 2
by Avner Friedman

Volume 25: Solitons in Physics, Mathematics, and Nonlinear Optics
Editors: Peter J. Olver and David H. Sattinger

- Volume 26: Two Phase Flows and Waves
Editors: Daniel D. Joseph and David G. Schaeffer
- Volume 27: Nonlinear Evolution Equations that Change Type
Editors: Barbara Lee Keyfitz and Michael Shearer
- Volume 28: Computer Aided Proofs in Analysis
Editors: Kenneth Meyer and Dieter Schmidt
- Volume 29: Multidimensional Hyperbolic Problems and Computations
Editors: Andrew Majda and Jim Glimm
- Volume 30: Microlocal Analysis and Nonlinear Waves
Editors: Michael Beals, R. Melrose, and J. Rauch
- Volume 31: Mathematics in Industrial Problems, Part 3
by Avner Friedman
- Volume 32: Radar and Sonar, Part I
by Richard Blahut, Willard Miller, Jr., and Calvin Wilcox
- Volume 33: Directions in Robust Statistics and Diagnostics: Part I
Editors: Werner A. Stahel and Sanford Weisberg
- Volume 34: Directions in Robust Statistics and Diagnostics: Part II
Editors: Werner A. Stahel and Sanford Weisberg
- Volume 35: Dynamical Issues in Combustion Theory
Editors: P. Fife, A. Liñán, and F.A. Williams
- Volume 36: Computing and Graphics in Statistics
Editors: Andreas Buja and Paul Tukey
- Volume 37: Patterns and Dynamics in Reactive Media
Editors: Harry Swinney, Gus Aris, and Don Aronson
- Volume 38: Mathematics in Industrial Problems, Part 4
by Avner Friedman
- Volume 39: Radar and Sonar, Part II
Editors: F. Alberto Grünbaum, Marvin Bernfeld, and
Richard E. Blahut
- Volume 40: Nonlinear Phenomena in Atmospheric and Oceanic Sciences
Editors: George F. Carnevale and Raymond T. Pierrehumbert

Volume 41: Chaotic Processes in the Geological Sciences

Editor: David A. Yuen

Volume 42: Partial Differential Equations with Minimal Smoothness and Applications

Editors: B. Dahlberg, E. Fabes, R. Fefferman, D. Jerison, C. Kenig, and J. Pipher

Volume 43: On the Evolution of Phase Boundaries

Editors: Morton E. Gurtin and Geoffrey B. McFadden

Volume 44: Twist Mappings and Their Applications

Editors: Richard McGehee and Kenneth R. Meyer

Volume 45: New Directions in Time Series Analysis, Part I

Editors: David Brillinger, Peter Caines, John Geweke, Emanuel Parzen, Murray Rosenblatt, and Murad S. Taqqu

Volume 46: New Directions in Time Series Analysis, Part II

Editors: David Brillinger, Peter Caines, John Geweke, Emanuel Parzen, Murray Rosenblatt, and Murad S. Taqqu

Volume 47: Degenerate Diffusions

Editors: Wei-Ming Ni, L.A. Peletier, and J.-L. Vazquez

Volume 48: Linear Algebra, Markov Chains, and Queueing Models

Editors: Carl D. Meyer and Robert J. Plemmons

Volume 49: Mathematics in Industrial Problems, Part 5

by Avner Friedman

Volume 50: Combinatorial and Graph-Theoretic Problems in Linear Algebra

Editors: Richard A. Brualdi, Shmuel Friedland, and Victor Klee

Volume 51: Statistical Thermodynamics and Differential Geometry of Microstructured Materials

Editors: H. Ted Davis and Johannes C.C. Nitsche

Volume 52: Shock Induced Transitions and Phase Structures in General Media

Editors: J.E. Dunn, Roger Fosdick, and Marshall Slemrod

Volume 53: Variational and Free Boundary Problems

Editors: Avner Friedman and Joel Spruck

- Volume 54: Microstructure and Phase Transitions
Editors: David Kinderlehrer, Richard James, Mitchell Luskin, and Jerry L. Ericksen
- Volume 55: Turbulence in Fluid Flows: A Dynamical Systems Approach
Editors: George R. Sell, Ciprian Foias, and Roger Temam
- Volume 56: Graph Theory and Sparse Matrix Computation
Editors: Alan George, John R. Gilbert, and Joseph W.H. Liu
- Volume 57: Mathematics in Industrial Problems, Part 6
by Avner Friedman
- Volume 58: Semiconductors, Part I
Editors: W.M. Coughran, Jr., Julian Cole, Peter Lloyd, and Jacob White
- Volume 59: Semiconductors, Part II
Editors: W.M. Coughran, Jr., Julian Cole, Peter Lloyd, and Jacob White
- Volume 60: Recent Advances in Iterative Methods
Editors: Gene Golub, Anne Greenbaum, and Mitchell Luskin
- Volume 61: Free Boundaries in Viscous Flows
Editors: Robert A. Brown and Stephen H. Davis
- Volume 62: Linear Algebra for Control Theory
Editors: Paul Van Dooren and Bostwick Wyman
- Volume 63: Hamiltonian Dynamical Systems: History, Theory, and Applications
Editors: H.S. Dumas, K.R. Meyer, and D.S. Schmidt
- Volume 64: Systems and Control Theory for Power Systems
Editors: Joe H. Chow, Petar V. Kokotovic, and Robert J. Thomas
- Volume 65: Mathematical Finance
Editors: Mark H.A. Davis, Darrell Duffie, Wendell H. Fleming, and Steven E. Shreve
- Volume 66: Robust Control Theory
Editors: Bruce A. Francis and Pramod P. Khargonekar

Volume 67: Mathematics in Industrial Problems, Part 7

by Avner Friedman

Volume 68: Flow Control

Editor: Max D. Gunzburger

Volume 69: Linear Algebra for Signal Processing

Editors: Adam Bojanczyk and George Cybenko

Volume 70: Control and Optimal Design of Distributed Parameter Systems

Editors: John E. Lagnese, David L. Russell, and Luther W. White

Volume 71: Stochastic Networks

Editors: Frank P. Kelly and Ruth J. Williams

Volume 72: Discrete Probability and Algorithms

Editors: David Aldous, Persi Diaconis, Joel Spencer, and J. Michael Steele

Forthcoming Volumes:

1992 Summer Program: *Environmental Studies: Mathematical, Computational, and Statistical Analysis*

1992–1993: *Control Theory*

Robotics

Nonsmooth Analysis & Geometric Methods in Deterministic Optimal Control

1993 Summer Program: *Modeling, Mesh Generation, and Adaptive Numerical Methods for Partial Differential Equations*

1993–1994: *Emerging Applications of Probability*

Random Discrete Structures

Mathematical Population Genetics

Stochastic Problems for Nonlinear Partial Differential Equations

Image Models (and their Speech Model Cousins)

Stochastic Models in Geosystems

Classical and Modern Branching Processes

1994 Summer Program: *Molecular Biology*

1994-1995: *Waves and Scattering*

Computational Wave Propagation

Wavelets, Multigrid and Other Fast Algorithms (Multipole, FFT)
and Their Use In Wave Propagation

Waves in Random and Other Complex Media

is based on the proceedings of the workshop "Probability and Algorithms" and "The Fast Fourier Transform" which were an integral part of the 1993-94 Summer Program. The workshop "Probability and Algorithms" was organized by David Aldous and J. Michael Steele for organizing the workshop. We also take this opportunity to thank the National Science Foundation, the Air Force Office of Scientific Research, the Office of Naval Research, and the National Security Agency for their support in making the workshop possible.

Avner Friedman
Willard Miller, Jr.

FOREWORD

This IMA Volume in Mathematics and its Applications

DISCRETE PROBABILITY AND ALGORITHMS

is based on the proceedings of two workshops, "Probability and Algorithms" and "The Finite Markov Chain Renaissance" that were an integral part of the 1993-94 IMA program on "Emerging Applications of Probability." We thank David Aldous, Persi Diaconis, Joel Spencer, and J. Michael Steele for organizing these workshops and for editing the proceedings. We also take this opportunity to thank the National Science Foundation, the Air Force Office of Scientific Research, the Army Research Office, and the National Security Agency, whose financial support made the workshop possible.

Avner Friedman

Willard Miller, Jr.

PREFACE

Discrete probability theory and the theory of algorithms have become close partners over the last ten years, though the roots of the partnership go back much longer. There are many reasons that underlie the coordination of these two fields, but some sense of the driving principles can be evoked by considerations like the following:

- When the use of a rule in an algorithm might lead to locking conflicts, randomization often provides a way to avoid stalemate.
- When a combinatorial object cannot be easily constructed, one can still often show the existence of the object by showing that under a suitable probability model such an object (or one close enough for appropriate modification) will exist with positive probability.
- When one needs to make a random uniform selection from an intractably large set, one can sometimes succeed by making clever use of a random walk (or other Markov chain) that has for its stationary measure the desired distribution.
- Finally, in many large systems that are driven by elements of chance, one often finds a certain steadiness that can be expressed by limit laws of probability theory and that can be exploited in the design of algorithms.

All of the chapters in this volume touch on one or more of these themes. The method of probabilistic construction is at the heart of the paper by Spencer and Tetali on Sidon sets as well as that of Godbole, Skipper, and Sunley, which traces its roots back to one of the first great successes of the “probabilistic method” — Erdős’s pioneering analysis of the central Ramsey numbers.

The theme of steadiness in large random structures is evident in almost all of the volume’s chapters, but it is made explicit in the paper by Fill and Dobrow on the move-to-front rule for self-organizing lists, the chapter by Yukich on Euclidean functionals (like the TSP), and in the paper by Steele that explores the limit theory that has evolved from the Erdős-Szekeres theorem on monotone subsequences. The chapter by Alon also shows how to find algorithmically useful “order in chaos” by developing a basic criterion of network connectivity in random graph models with unequal probabilities for edges.

The theme of “uniform selection by walking around” is perhaps most explicitly illustrated in the two chapters by Diaconis and Gangolli and Diaconis and Holmes. The first of these shows how one can use the ideas of the “Markov Chain Renaissance” to make progress on the difficult problem of the enumeration of integer tables with specified row sums and column sums. The second paper shows how new developments emerging from the Markov chain renaissance can be brought to bear on problems of concern

in statistics, computer science, and statistical mechanics. Aldous also makes a contribution in the thick of the new theory of finite Markov chains by showing in his chapter that one can simulate an observation from a chain's stationary distribution (quickly, though approximately)—all the while not knowing the transition probabilities of the chain except through the action of a “take a step from state x ” oracle.

The two further chapters in this collection are surveys that call on all of the basic themes recalled above. They are also tightly tied to the central concerns of the theory of probabilistic complexity. The first of these is the survey of A. Karlin and P. Raghavan on random walks in undirected graphs—a notion that is present in many of the collection's chapters. The second is the survey by D. Welsh on randomized approximation schemes for Tutte-Gröthendieck invariants, which are remarkable polynomials whose values at special points give precise information about such basic graph theoretic problems as the number of connected subgraphs, the number of forest subgraphs, the number of acyclic orientations, and much more.

All the papers in this volume come from two Workshops, “Probability and Algorithms” and “The Finite Markov Chain Renaissance,” that were held during the Special Year in Emerging Applications of Probability at the Institute for Mathematics and Its Applications at the University of Minnesota during the fall of 1993. The IMA provided a singularly congenial environment for productive scientific exchange, and with any luck the chapters of this volume will convey a sense of the excitement that could be felt in the progress that was reported in these IMA Workshops.

It is a pleasure to thank Avner Friedman, Willard Miller, Jr., and the IMA staff for their efficient organization of the workshops and the entire program, and to thank Patricia V. Brick for administering the preparation of this volume.

David Aldous
Persi Diaconis
Joel Spencer
J. Michael Steele

CONTENTS

Foreword	xiii
Preface	xv
On simulating a Markov chain stationary distribution when transition probabilities are unknown.....	1
<i>David Aldous</i>	
A note on network reliability	11
<i>Noga Alon</i>	
Rectangular arrays with fixed margins.....	15
<i>Persi Diaconis and Anil Gangolli</i>	
Three examples of Monte-Carlo Markov chains: at the interface between statistical computing, computer science, and statistical mechanics	43
<i>Persi Diaconis and Susan Holmes</i>	
The move-to-front rule for self-organizing lists with Markov dependent requests	57
<i>Robert P. Dobrow and James Allen Fill</i>	
The asymptotic lower bound on the diagonal Ramsey numbers: A closer look	81
<i>Anant P. Godbole, Daphne E. Skipper, and Rachel A. Sunley</i>	
Random walks and undirected graph connectivity: A survey	95
<i>Anna R. Karlin and Prabhakar Raghavan</i>	
Sidon sets with small gaps	103
<i>Joel Spencer and Prasad Tetali</i>	
Variations on the monotone subsequence theme of Erdős and Szekeres	111
<i>J. Michael Steele</i>	
Randomised approximation schemes for Tutte-Gröthendieck invariants	133
<i>Dominic Welsh</i>	
Quasi-additive Euclidean functionals	149
<i>J.E. Yukich</i>	

ON SIMULATING A MARKOV CHAIN STATIONARY DISTRIBUTION WHEN TRANSITION PROBABILITIES ARE UNKNOWN*

DAVID ALDOUS†

Abstract. We present an algorithm which, given a n -state Markov chain whose steps can be simulated, outputs a random state whose distribution is within ε of the stationary distribution, using $O(n)$ space and $O(\varepsilon^{-2}\tau)$ time, where τ is a certain “average hitting time” parameter of the chain.

1. Introduction. Our topic is a small corner of the region where algorithms and Markov chains meet. While perhaps not relevant to the main theoretical or practical issues in that region, it seems interesting enough to be worth recording. My motivation came from a remarkable result of Asmussen, Glynn and Thorisson [2], restated as Theorem 1 below.

Consider a Markov chain on states $\{1, 2, \dots, n\}$ with irreducible transition matrix $\mathbf{P} = (p(i, j))$ and hence with a unique stationary distribution $\pi_{\mathbf{P}} = (\pi_{\mathbf{P}}(i))$. Suppose we have a subroutine that simulates steps from \mathbf{P} , i.e. given any state i as input it outputs a random state J_i with $P(J(i) = j) = p(i, j) \forall j$, independent of previous output. The problem is to devise an algorithm which terminates in some random state ξ such that, regardless of \mathbf{P} ,

$$(1.1) \quad \frac{1}{2} \sum_j |P(\xi = j) - \pi_{\mathbf{P}}(j)| \leq \varepsilon.$$

The point is that the algorithm is not allowed to know \mathbf{P} . That is, given the first s steps $(i_1, j_1), (i_2, j_2), \dots, (i_s, j_s)$, we must specify a rule by which we either terminate and output j_s , or else specify a state i_{s+1} to be the next input to the subroutine, and this rule can use only $(i_1, j_1, i_2, j_2, \dots, j_s)$ and external randomization. Write $\mathcal{A}(\varepsilon)$ for the class of algorithms A which satisfy (1.1) for all irreducible \mathbf{P} (we suppress dependence on n here). For $A \in \mathcal{A}(\varepsilon)$ let $c(A, \mathbf{P})$ be the mean number of steps simulated by the algorithm.

It is obvious that (even for $n = 2$) there is no algorithm $A \in \mathcal{A}(\varepsilon)$ such that $\sup_{\mathbf{P}} c(A, \mathbf{P}) < \infty$, by considering “almost reducible” chains.

Two different methods for attempting to construct algorithms in $\mathcal{A}(\varepsilon)$ suggest themselves.

Matrix perturbation theory gives bounds for $\pi_{\mathbf{Q}} - \pi_{\mathbf{P}}$ in terms of bounds for $\mathbf{Q} - \mathbf{P}$. Fix some integer m , simulate m steps from each state i to get an empirical estimate

* Research supported by N.S.F. Grant DMS92-24857.

† Department of Statistics, University of California, Berkeley, CA 94720.

$Q_m(i, j)$ of the $p(i, j)$ together with an estimate of the error $Q_m - P$. Then calculate numerically the stationary distribution π_m corresponding to Q_m and a confidence interval for the error $\pi_m - \pi$. If this interval is too large, increment m and repeat.

Let's call this a *matrix-based* method, in contrast to a *pure simulation* method below.

Markov chain theory says that if we simulate the chain from an arbitrary initial state then at a sufficiently large time t (randomized to avoid periodicity) the current state will have approximately distribution π . So fix t , simulate the chain for t steps, perform some test on the observed path to check if t is sufficiently large, and if not then increment t and repeat.

It is of course not entirely clear how to turn these vague ideas into algorithms which are provably in $\mathcal{A}(\varepsilon)$. Using a rather different idea, Asmussen et al ([2] Theorem 3.1) showed that in fact one can simulate π_P exactly with no knowledge of P .

THEOREM 1. *There exists an $A_0 \in \mathcal{A}(0)$ with $c(A_0, P) < \infty \forall P$.*

Briefly, given a way of simulating exactly the distribution π restricted to $B_j = \{1, 2, \dots, j\}$, then they describe a procedure to simulate exactly the distribution π restricted to B_{j+1} .

Unfortunately it seems difficult to give an informative upper bound for $c(A_0, P)$ in terms of P . In section 2 we present and analyze a slightly simpler "pure simulation" algorithm which does permit a natural upper bound, at the cost of producing approximate rather than exact stationarity. Theorem 2 states the precise result. Section 3 gives a lower bound for the performance of any algorithm, and section 4 contains further discussion.

2. The algorithm. We start by outlining the idea of the algorithm. Write $E_i T_j$ for the mean hitting time on state j , starting from state i . Define the averaged hitting time $\tau = \tau_P$ by

$$(2.1) \quad \tau = \sum_i \sum_j \pi(i) E_i T_j \pi(j).$$

Because τ is essentially an upper bound on the time taken to approach stationarity, it is enough to be able to estimate the value of τ by simulation in $O(\tau)$ steps, for then we can run another simulation for $O(\tau)$ steps and output the final state. The estimation of τ is done via a "coalescing paths" routine: run the chain from an arbitrary start until a specified state j is hit, keeping track of states visited; start again from some unvisited state and run until visiting some state hit on a previous run; and so on until every state has been visited. Then the number of steps used in this procedure is $\Theta(\tau)$, for a typical initial target j .

Here is the precise algorithm. We are given $\varepsilon > 0$, states $\{1, \dots, n\}$,

and the ability to simulate a step of the Markov chain from any specified state.

Algorithm A_ε

1. Let $t_0 \leftarrow n$.
2. Pick U random, uniform on $\{1, \dots, t_0\}$.
3. Simulate the chain, starting at state 1, for U steps. Let j be the final state.
4. Start a counter at 0 and count steps as they are simulated in the stages below. If the count exceeds εt_0 before this algorithm starts stage 8, let $t_0 \leftarrow 2t_0$ and go to stage 2.
5. Let $B \leftarrow \{j\}$.
6. Simulate the chain, starting at state $\min\{i : i \notin B\}$, until the chain hits B , keeping track of the set B' of states visited.
7. Let $B \leftarrow B \cup B'$. If $B \neq \{1, \dots, n\}$ go to stage 6.
8. Pick U random, uniform on $\{1, \dots, \lceil t_0/\varepsilon \rceil\}$. Simulate the chain, starting at state 1, for U steps. Output the final state ξ .

Stages 5,6,7 are the “coalescing walks” routine. It is clear that the algorithm requires only $O(n)$ space, to track which states are hit during this routine.

THEOREM 2. Fix $0 < \varepsilon < 1/4$. Then $A_\varepsilon \in \mathcal{A}(4\varepsilon)$ and

$$c(A_\varepsilon, \mathbf{P}) \leq \frac{81\tau\mathbf{P}}{\varepsilon^2} \forall \mathbf{P}.$$

One might guess that some variation of the construction would lead to an algorithm where the bound is polynomial in $\log 1/\varepsilon$, but I have not pursued that possibility.

The rest of the section contains the proof of Theorem 2. To start with some notation, write $\|\cdot\|$ for variation distance between distributions

$$\|\theta - \mu\| = \frac{1}{2} \sum_i |\theta(i) - \mu(i)|.$$

Write

$$s(j) = \max_i E_i T_j$$

$$s_* = \min_j s(j).$$

Write C_j for the “coalescing paths” time, that is the number of steps required to complete stages 5,6,7 of the algorithm, ignoring the cut-off rule in 4. Finally, recall two standard facts: the *right-averaging principle* ([1] Chapter 2)

$$(2.2) \quad \sum_j E_i T_j \pi(j) = \tau \quad \forall i$$