Mathematical Research · Mathematische Forschung

Wissenschaftliche Beiträge
herausgegeben von der
Akademie der Wissenschaften der DDR
Karl-Weierstraß-Institut für Mathematik

Band 38

Parallel Algorithms
and Architectures

Parallel Algorithms and Architectures

Proceedings of the International Workshop on Parallel Algorithms and Architectures held in Suhl (GDR), May 25–30, 1987

edited by
Andreas Albrecht,
Hermann Jung,
and
Kurt Mehlhorn

Akademie-Verlag Berlin 1987

Herausgeber:

Dr. Andreas Albrecht
Doz. Dr. Hermann Jung
Humboldt-Universität zu Berlin
Sektion Mathematik
Prof. Dr. Kurt Mehlhorn
Universität des Saarlandes, Saarbrücken
FB Angewandte Mathematik und Informatik

Die Titel dieser Schriftenreihe werden vom Originalmanuskript der Autoren reproduziert.

ISBN 3-05-500397-7
ISSN 0138-3019

Erschienen im Akademie-Verlag Berlin, DDR-1086 Berlin, Leipziger Str. 3-4

Lizenznummer: 202·100/418/87
Printed in the German Democratic Republic
Gesamtherstellung: VEB Kongreß- und Werbedruck, 9273 Oberlungwitz
LSV 1095
Bestellnummer: 763 760 7 (2182/38)
02800

PREFACE

This volume constitutes of the proceedings of the International Workshop on Parallel Algorithms and Architectures held in Suhl, German Democratic Republic, May 25-29, 1987.
The aim of the conference is to support research activities on parallel architectures and algorithms. The program consists of invited lectures given by acknoledged scientists as well as short communications. It covers both theoretical and practical aspects of parallel computing.
The main topics of the workshop are
- Models of parallel computations;
- New algorithms for individual problems, e.g. from graph theory, logic programming, combinatorics and computational geometry;
- Hardware algorithms, parallel architectures.

We wish to thank all who submitted papers for consideration and all members of the International Program Committee which consisted of A.Albrecht (Berlin), J.von zur Gathen (Toronto), J.Herath (Yokohama), H.Jung (Berlin), R.Kannan (Pittsburgh), A.Konagaya (Tokyo), V.Kotov (Novosibirsk), T.Leighton (Cambridge), K.Mehlhorn (Saarbrücken), J.Miklosko (Bratislava), B.Monien(Paderborn), W.Paul (Saarbrücken), H.Thiele (Berlin), C.D.Thomborson (Duluth), G.Wechsung (Jena), C.K.Yap (New York).

The workshop was organized by the Department of Mathematics of Humboldt-University (Berlin) and chaired by A.Albrecht and H.Jung.
We would like to take this opportunity to express our sincere gratitude to the Organizing Secretary of the workshop T.Zeugmann for his engagement and excellent work.
We are all very grateful for the special support given by the local organizers in Suhl who contributed their generous help in arranging the workshop.
Finally we would like to thank the publishers for their assistance and cooperation in the selection of the present volume.

A. Albrecht, H. Jung, K. Mehlhorn

TABLE OF CONTENTS

Invited Papers

Communications

LATE PAPERS

Invited Papers

Communication

Deterministic Simulation of Idealized Parallel Computers on More Realistic Ones

by

H. Alt* / T. Hagerup** / K. Mehlhorn** / F.P. Preparata***

* Fachbereich Mathematik, Freie Universität Berlin, Arnimallee 2–6, D–1000 Berlin 33, FRG. Part of the research was done while the author was a member of the Mathematical Sciences Research Institute, Berkeley, U.S.A.

** Fachbereich 10, Informatik, Universität des Saarlandes, D–6600 Saarbrücken, FRG.

*** Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801, U.S.A.

Abstract: We describe a non-uniform deterministic simulation of PRAMs on module parallel computers (MPCs) and on processor networks of bounded degree. The simulating machines have the same number n of processors as the simulated PRAM, and if the size of the PRAM's shared memory is polynomial in n, each PRAM step is simulated by $O(\log n)$ MPC steps or by $O((\log n)^2)$ steps of the bounded-degree network. This improves upon a previous result by Upfal and Wigderson. We also prove an $\Omega((\log n)^2/\log\log n)$ lower bound on the number of steps needed to simulate one PRAM step on a bounded-degree network under the assumption that the communication in the network is point-to-point.

As an important part of the simulation of PRAMs on MPCs, we use a new technique for dynamically averaging out a given work load among a set of processors operating in parallel.

A preliminary version of this work was presented at MFCS 86, Bratislava. The full version is available from the authors and will appear in SIAM Journal of Computing.

This work was supported by the DFG, SFB 124, TP B2, VLSI Entwurf und Parallelität, and by NSF Grant ECS-84-10902.

1. Introduction and Models of Computation

Most parallel algorithms in the literature are designed to run on a PRAM (parallel RAM). The PRAM model was introduced by Fortune and Wyllie [FW]. A PRAM consists of some finite number n of sequential processors (RAMs), all of which operate synchronously on a common memory consisting of, say, m storage cells (also called "variables"), cf. fig. 1.

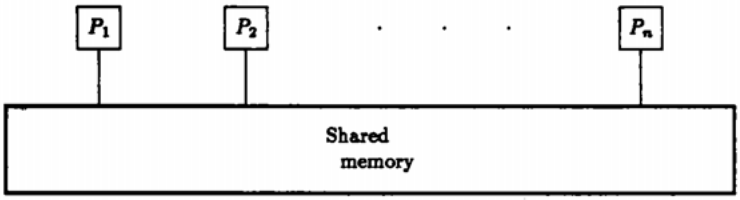

Fig. 1. The PRAM model. $P_1, \ldots, P_n$ are processors.

In every step of the PRAM, each of its processors executes a private RAM instruction. In particular, the processors may all simultaneously access (read from or write into) the common memory. Various types of PRAMs have been defined, differing in the conventions used to deal with read/write conflicts, i.e., attempts by several processors to access the same variable in the same step. In the most restrictive model, exclusive read–exclusive write or EREW PRAMs, no variable may be accessed by more than one processor in a given step. In contrast, CRCW (concurrent read–concurrent write) PRAMs allow simultaneous reading as well as simultaneous writing of each variable, with some rule defining the exact semantics of simultaneous writing.

PRAMs are very convenient for expressing parallel algorithms since one may concentrate on the problem of "parallelizing", i.e., decomposing the problem at hand into simultaneously executable tasks, without having to worry about the communication between these tasks. Indeed, any intermediate result computed by one of the processors will be available to all the others in the next step, due to the shared memory. Unfortunately, the PRAM is not a very realistic model of parallel computation. Present and foreseeable technology does not seem to make it possible for more than a constant number of processors to simultaneously access the same memory module. A model of computation that takes this problem into account is the MPC (*module parallel computer*, [MV]), cf. fig. 2.

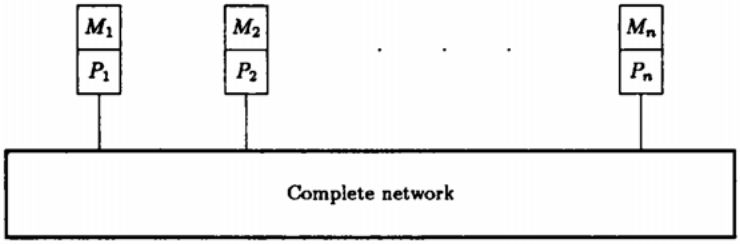

Fig. 2. The MPC model. $P_1, \ldots, P_n$ are processors, $M_1, \ldots, M_n$ memory modules.

An MPC consists of n processors (RAMs), each equipped with a memory module. Every processor may access every memory module via a complete network connecting the processors. However, the memory modules are sequential devices, i.e., able to satisfy only one request at a time. More precisely, a memory module M operates as follows: If several processors try in the same step to access a variable stored in M, exactly one of the processors is allowed to carry out its read or write instruction; the remaining access requests are discarded. All processors are informed of the success or failure of their access attempts. We make no assumptions about how the single successful processor is selected from among the processors competing to access M.

The MPC model is still not realistic for large n because of the postulated complete network connecting the processors. This leads us to consider a third model which we shall call the *network model*. Here the processors are connected via a network of bounded degree, i.e., each processor is linked directly to only a constant number of other processors, cf. fig. 3.

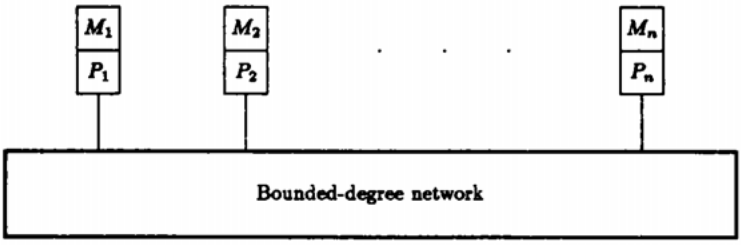

Fig. 3. The network model. $P_1, \ldots, P_n$ are processors, $M_1, \ldots, M_n$ memory modules.

Since each step of a completely interconnected processor network may be simulated by $O(\log n)$ steps of a bounded-degree network ([AKS], [L]), efficient algorithms for the MPC model translate into asymptotically efficient algorithms for the network model.

The simulation of the idealized parallel machine, the PRAM, on the more realistic one, the MPC, has been considered in several previous papers. A naive approach represents each variable x of the PRAM by one variable $\psi(x)$ of the MPC. Now if a PRAM step accesses the variables $x_1, \ldots, x_l$, collisions may occur in the simulating machine because $\psi(x_1), \ldots, \psi(x_l)$ are not necessarily located in distinct memory modules. If $m \leq n$, the m variables may be allocated to m different memory modules, and a trivial $O(1)$-time simulation is possible. However, we are concerned with the case in which m is considerably larger than n. Here a major problem is to find a memory correspondence between the PRAM and the MPC such that, for all possible access patterns of the PRAM, the maximum number of variables requested from a single MPC memory module is kept low. Note that, for specific PRAM algorithms such as matrix multiplication, there may be very efficient ways of assigning variables to modules; we refer the reader to Section 4 of the survey paper by Kuck [K]. Here we are interested in universal simulations which work efficiently no matter which algorithm is executed by the PRAM.

Some results have been obtained previously using *probabilistic* methods: Mehlhorn and Vishkin [MV] used universal hashing to define the memory correspondence. They obtained several upper bounds, for example an average of $O(\log n)$ MPC steps to simulate one PRAM step, with the total amount of memory used by the MPC larger than the PRAM memory by a factor of $O(\log n)$. Upfal [U] found a probabilistic simulation of $O((\log n)^2)$ average time for one PRAM step on a bounded-degree network; this was recently improved to $O(\log n)$ by Karlin and Upfal [KU].

This paper is concerned with *deterministic* simulations. We define the *slow-down* of a simulation as the number of steps needed by the simulating machine in the worst case to simulate one step of the simulated machine. Note that if $m \geq n^2$, the simple scheme outlined above (x is represented by $\psi(x)$) performs poorly: An adversary could make the PRAM step access n variables $x_1, \ldots, x_n$ with $\psi(x_1), \ldots, \psi(x_n)$ all in the same module. Hence the slow-down is $\Omega(n)$. This reasoning shows that each PRAM variable must be represented by several "copies" stored in different modules. Mehlhorn and Vishkin [MV] showed that read instructions can be handled very efficiently using this idea. However, they did not know how to deal with write instructions. In a beautiful paper Upfal and Wigderson [UW] resolved this problem and exhibited a simulation which uses $\Theta(\log n)$ copies of each PRAM variable. If m is polynomial in n, the slow-down is $O(\log n(\log\log n)^2)$. They also showed an $\Omega(\log n/\log\log n)$ lower bound on the slow-down for a large class of simulations.

Using similar techniques, this paper improves the upper bound to $O(\log m)$. If m is polynomial in n, this is $O(\log n)$. Consequently, a PRAM step may be simulated in $O(\log n \log m)$ time on a bounded-degree network. On the other hand, we show that $\Omega(\log n \log m/\log\log m)$ time is necessary under certain assumptions on any bounded-degree network whose communication is restricted to be point-to-point. A similar result was also obtained by Karlin and Upfal [KU]. The assumption of point-to-point communication is not satisfied by our simulation algorithm which uses more general communication patterns.

The PRAM simulations which we consider will be based on *emulations* of the PRAM's shared memory. We conceptually retain the n PRAM processors while replacing (or, equivalently, implementing) the PRAM's (physically infeasible) shared memory by a (more feasible) suitably programmed MPC or bounded-degree network with n processors, called the emulating processors. Each PRAM processor, which was formerly connected to the shared memory, is now instead connected to one of the emulating processors called its *associated* processor, each emulating processor being associated with exactly one PRAM processor, cf. fig. 4. We require the change to be completely transparent, i.e., all PRAM programs must run without change (though possibly slower) on the modified machine.

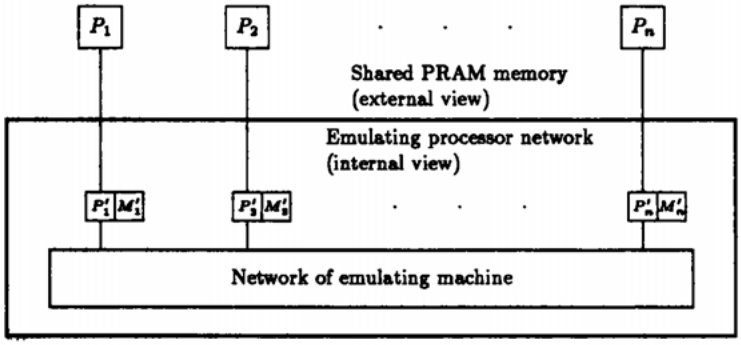

Fig. 4. Emulation of the shared memory of a PRAM.
For $i = 1, \ldots, n$, P_i is a PRAM processor and P'_i its associated emulating processor.

Note that although the most direct PRAM simulation implied by a memory emulation as above uses a total of $2n$ processors, it is a trivial matter to reduce the number of processors to n by coalescing each pair of associated processors into a single processor. For expository reasons we prefer to keep the clean separation between PRAM processors and (emulated) shared memory.

Our simulation algorithms are non-uniform. This means that they are not given explicitly. Instead we merely prove that algorithms with the desired properties exist. For fixed values of n and m, such algorithms may be found by exhaustive search in a large but finite set. We return to this question in the concluding section.

It has been known since Adleman's work [A] that probabilistic algorithms may be converted into non-uniform deterministic ones. Hence the result by Karlin and Upfal [KU] automatically translates into a non-uniform deterministic simulation of PRAMs on a bounded-degree network. However, if the translation is based on Karlin and Upfal's analysis of their algorithm and uses known techniques, it introduces an $\Omega(n)$-increase in the product of time and number of processors [R, Theorem 6]. Since it is not difficult to devise a uniform deterministic simulation which uses $O(n^2/(\log n)^2)$ processors and has a slow-down of $O(\log n)$ (the construction is similar to one presented in the remark ending Section 3), deterministic algorithms derived from Karlin and Upfal's probabilistic simulation are of little interest. The same is true of all other known probabilistic solutions.

The remaining part of the paper is structured as follows: In Section 2 we describe our simulation of PRAMs on MPCs and show that its slow-down is $O(\log m)$. As part of the development of the algorithm, we define and solve a so-called "redistribution problem". Section 3 considers the simulation of PRAMs on bounded-degree networks and establishes upper and lower bounds of $O(\log n \log m)$ and $\Omega(\log n \log m / \log\log m)$, respectively. In Section 4 we return to the redistribution problem and prove a stronger result than what was needed in Section 2. Finally, Section 5 addresses some interesting and important open issues. Sections 2 to 5 can be found in the full paper.

References

[A]: L. Adleman: "Two Theorems on Random Polynomial Time". Proc. 19'th Symp. Found. of Comp. Sci. (1978), 75–83.

[AKS]: M. Ajtai, J. Komlós, E. Szemerédi: "An $O(n \log n)$ Sorting Network". Proc. 15'th ACM Symp. Theory of Comp. (1983), 1–9.

[B]: K.E. Batcher: "Sorting networks and their applications". Proc. AFIPS Spring Joint Comp. Conf. **32** (1968), 307–314.

[BH]: A. Borodin, J.E. Hopcroft: "Routing, Merging, and Sorting on Parallel Models of Computation". Proc. 14'th ACM Symp. Theory of Comp. (1982), 338–344.

[FW]: S. Fortune, J. Wyllie: "Parallelism in Random Access Machines". Proc. 10'th ACM Symp. Theory of Comp. (1978), 114–118.

[GG]: O. Gabber, Z. Galil: "Explicit Constructions of Linear Size Concentrators and Superconcentrators". Proc. 20'th Symp. Found. of Comp. Sci. (1979), 364–370.

[K]: D.J. Kuck: "A Survey of Parallel Machine Organization and Programming". *Computing Surveys* **9**:1 (1977), 29–59.

[KU]: A.R. Karlin, E. Upfal: "Parallel Hashing – An Efficient Implementation of Shared Memory". Proc. 18'th ACM Symp. Theory of Comp. (1986), 160–168.

[L]: T. Leighton: "Tight Bounds on the Complexity of Parallel Sorting". Proc. 16'th ACM Symp. Theory of Comp. (1984), 71–80.

[MV]: K. Mehlhorn, U. Vishkin: "Randomized and Deterministic Simulations of PRAMs by Parallel Machines with Restricted Granularity of Parallel Memories". *Acta Informatica* **21** (1984), 339–374.

[O]: O. Ore: "Theory of Graphs". American Mathematical Society, Providence, Rhode Island (1962).

[PV]: F.P. Preparata, J. Vuillemin: "The Cube-Connected Cycles: A Versatile Network for Parallel Computation". *Communications of the ACM* **24**:5 (1981), 300–309.

[R]: J.H. Reif: "On the Power of Probabilistic Choice in Synchronous Parallel Computations". Proc. 9'th Int. Coll. Automata, Languages and Programming (1982), 442–450.

[U]: E. Upfal: "A Probabilistic Relation Between Desirable and Feasible Models of Parallel Computation". Proc. 16'th ACM Symp. Theory of Comp. (1984), 258–265.

[UW]: E. Upfal, A. Wigderson: How to Share Memory in a Distributed System". Proc. 25'th Symp. Found. of Comp. Sci. (1984), 171–180.

[VW]: U. Vishkin, A. Wigderson: "Dynamic Parallel Memories". *Information and Control* **56** (1983), 174–182.

CONVEX HULL OF RANDOMLY CHOSEN POINTS FROM A POLYTOPE
(Preliminary Version)

Rex Dwyer * and Ravi Kannan **

* Computer Science dept., Carnegie-Mellon University.
Supported by NSF Grant ECS 8418392
** Computer Science Department, Carnegie-Mellon University and Institut für Operations Research, Universität Bonn. Supported by NSF Grant ECS 8418392 and the Alexander von Humboldt Stiftung, Bonn

1. Abstract

Suppose we pick independently n random points $X_1, X_2, \ldots, X_n$ from a d dimensional polytope P. (i.e., X_i are independent identically distributed random variables each with density $= 1/volume(P)$ in P and 0 outside.) Let E_n be the convex hull of $X_1, X_2, \ldots X_n$. The following questions arise naturally :

1) What is the value of V_n the expected ratio of the volume of $P \setminus E_n$ to the volume of P ?

2) What is the expected number of extreme points of the polytope E_n ?

We show an upper bounds of $\frac{C(P)}{n}(\log n)^{d+1}$ on V_n and $C(P)(\log n)^{d+1}$ on M_n where $C(P)$ is a constant that depends only on P (not on n). In both cases elementary arguments will only give a bound that replaces the power of $\log n$ by a power (less than one) of n. Previously, similar results were known only for the case of $d = 2$. (Buchta (1984) and Rényi and Solanke (1963, 1964)). There has been substantial amount of work on the problem for spheres as well as for other quantities depending on E_n in two dimensions. (see for example W.M. Schmidt (1968), G.Buchta, J.Müller and R.F.Tichy (1985), P.M.Gruber (1983) and I.Bárány and Z.Füredi (1986)) In case the polytope P has at least one vertex with exactly d adjacent vertices, we prove lower bounds of

$$d(P)(\log n)^{d-1} \ /n \text{ on } V_n \text{ and } d(P)(\log n)^{d-1} \text{ on } M_n.$$

Using the bounds,we are able to show that certain simple divide and conquer algorithms for finding the set of all extreme points have good sequential (linear time) and parallel (polylog time) complexitites in the expected case when the points are chosen at random independently from a polytope in a fixed number of dimensions.

The results are based on a natural notion of centrality which we introduce for convex sets.

2. The volume of the central region

Definiton : For any positive real number ϵ and a convex set P, a point p in P is ϵ-central for P, if for any hyperplane H through p, the volume of P in each of the half spaces determined by H is at least ϵ times the total volume of P.

Clearly, no point is more than $\frac{1}{2}$ central and the center of gravity is the only $\frac{1}{2}$ central point. It is also clear that if p is ϵ central then it is δ central for any $\delta \leq \epsilon$. We show below that the volume of the subset of points that are not ϵ central cannot be too high. (It is obvious that the set of non ϵ central points is Lesbeg measurable; we use volume to denote the Lesbeg measure.)

Theorem 1 *For any polytope P in R^d, there is a constant $C(P)$ depending only upon P and d such that for every positive ϵ, the volume of the set of non ϵ central points is at most*

$$C(P)\,\epsilon\,(log\,\tfrac{1}{\epsilon})^d \cdot \text{ volume of } P$$

Remark It is possible to see that the volume of noncentral points is at most $C(P)(\epsilon)^{1/d}$ times the volume of P by elementary arguments - any point at distance at least $c.(\epsilon)^{1/d}$ from the boundary of P has a sphere around it of radius $c.(\epsilon)^{1/d}$, any hyperplane through the point leaves a half of this sphere (which is of sufficient volume) in either half space. Note that the bound in the theorem is stronger - there ϵ is multiplied by a power of $log\,\frac{1}{\epsilon}$ whereas the elementary argument gives a bound where ϵ is multiplied by a power of $\frac{1}{\epsilon}$.

Remark The theorem is not true when P is a general convex set. For example it is easy to show that when P is a sphere, the ϵ central region is a concentric sphere of radius $(1 - c_d\epsilon^{1/d})$ times the radius of the original sphere whence the volume of the smaller sphere is $1 - O(\epsilon^{1/d})$ times that of the whole sphere.

Proof : With each unit vector c in R^d, we can associate a function $f_c : R^d \to R$ as follows : $f_c(x) =$ the volume of the set $\{y : c \cdot y \geq c \cdot x\}$, i.e., $f_c(x)$ is the volume of the half space above x determined by the hyperplane orthogonal to c through x. Further let c_o be the minimum of the linear functional $c \cdot x$ over P. It will be useful to define also $g_c : R \to R_+$ as $g_c(\lambda) =$ the volume of the set $\{x : c \cdot x \leq c_o + \lambda\}$. The following is a direct consequence of the Brunn-Minkowski theorem (Bonnesen and Fenchel (1934)).

Proposition 1 *Suppose λ, δ are real numbers such that $0 \leq \lambda \leq \delta$. Then, with the notation of the last paragraph,*

$$g_c(\lambda) \geq \left(\frac{\lambda}{\delta}\right)^d g_c(\delta)$$

Proof : Without loss of generality, we may translate P so that $c_o = 0$. For any positive real number α let $A(\alpha)$ be the $d-1$ dimensional volume of the intersection

of $\{x : c \cdot x = \alpha\}$ and P. The Brunn-Minkowski theorem asserts that the $d-1$ st root of $A(\cdot)$ is a concave function. From this it follows that for any α in the interval $[\lambda \quad \delta]$,

$$A(\lambda) \geq \left(\frac{\lambda}{\alpha}\right)^{d-1} A(\alpha)$$

Integrating $A(\alpha)$ from λ to δ and using the above we get that

$$g_c(\delta) - g_c(\lambda) \leq \frac{\lambda A(\lambda)}{d}\left(\left(\frac{\delta}{\lambda}\right)^d - 1\right)$$

Further, since there is some point q in P with $c \cdot q = 0$, the convex hull of q and $P \cap \{x : c \cdot x = \lambda\}$ contributes at least $A(\lambda)\lambda/d$ to $g_c(\lambda)$, so we have

$$g_c(\lambda) \geq A(\lambda)\lambda/d$$

These two inequalities together give us the proposition.

□

Let $F_i = \{x : c_i \cdot x \leq d_i\} \quad i = 1, 2, \ldots k$ be the defining inequalities of P where the c_i are unit vectors. We call a hyperplane of the form $\{x : c_i \cdot x = d_i - 2^{k/d}\epsilon\}$ where k is a natural number, a "copy" of F_i provided it intersects P. If any subset of d of the hyperplanes among the facets or their copies intersect at a point, we call the point a "grid point". It is clear that the number of grid points is at most $C(P)(log\ \frac{1}{\epsilon})^d$ for some constant $C(P)$ independent of ϵ. The facets and their copies subdivide the polytope P into what we may call "regions" - each region is a set of the form $\{x : d_i - 2^{k_i/d} \geq c_i \cdot x \geq d_i - 2^{k_i+1/d}$ for $i = 1, 2, \ldots k\}$ where the k_i are some natural numbers. It is clear that each region is the convex hull of some grid points. Finally, define for each x in P the unit vector $c(x)$ to be such that $f_{c(x)}(x) = \min\{f_c(x) : c \text{ a unit vector }\}$. The proof of the following proposition follows closely the lines of a proof of Lovász and Scarf (1986).

Proposition 2 *If a point x of P lies at distance at least $2^{1/d}\epsilon$ from every facet, then there is a grid point y such that*

$$f_{c(x)}(x) \geq f_{c(y)}(x)/2$$

Proof Under the hypothesis, it is clear that x is in an interior region R. Suppose R is the convex hull of grid points $y_1, y_2, \ldots y_k$. One of the y_i - call it y for short satisfies

$$f_{c(x)}(x) \geq f_{c(x)}(y)$$

Further,

$$f_{c(z)}(y) \geq f_{c(y)}(y)$$

Suppose the straight line *from x to y* intersects the boundary of P at z. By the definition of "copies", it is clear that $|x - z| \leq 2^{1/d}|y - z|$. So by proposition 1, we have

$$f_{c(y)}(y) \geq f_{c(y)}(x)/2$$

The three inequalities together establish the proposition.

Now we go back to the proof of the theorem. Let $\{f_{c(y)} : y \text{ a grid point }\}$ be $\{f^{(1)}, f^{(2)} \ldots f^{(s)}\}$ where $s \leq C(P)(log\ \frac{1}{\epsilon})^d$.

$$T^{(i)} = \{x : f^{(i)}(x) \leq 2\epsilon\} \text{ for } i = 1, 2, \ldots s$$

If $x \in P$ is at least $2^{1/d}\epsilon$ away from the boundary of P and $x \notin T^{(i)}$ for $i = 1, 2, \ldots s$, we wish to assert that x is ϵ central for P. This is because, for such x, there exists some grid point y such that $f_{c(z)}(x) \geq f_{c(y)}(x)/2 \geq \epsilon$, whence of course $f_c(x) \geq \epsilon$ for every unit vector c. Obviously, $\mu(T^{(i)}) = 2\epsilon$ (here μ denotes the volume) for each i, further, the the volume of the set of points at distance at most ϵ from the boundary of P is at most $C(P)'\epsilon$ for some constant $C(P)'$. Thus the volume of the set of non ϵ-central points in P is at most $2\epsilon s + \epsilon C(P)' \leq C(P)''\epsilon(log\ \frac{1}{\epsilon})^d$. So we have proved the theorem.

3. Probablilstic results

The following simple lemma together with the theorem of the last section gives us the probabilistic results.

Lemma 1 *Suppose n points are picked at random independently from a polytope P in R^d and suppose E_n denotes their convex hull. If x is an ϵ -central point of P, then the probability that x does not belong to E_n is at most $\binom{n}{d-1}(1-\epsilon)^{n-d+1}$.*

Proof: If x is not in E_n, then it lies on a facet F of the convex hull of $E_n \cup \{x\}$. F contains at least $d-1$ of n randomly picked points. Since x is ϵ central for P, the volume of P on either side of F is at most $1 - \epsilon$. Thus the lemma folows.

We use the lemma as follows : Let $t = \frac{n(d+1)}{n-d+1}$ and choose $\epsilon \sim \frac{t\ log\ n}{n}$. Then a calculation shows that if x is ϵ central for P, then the probability that $x \notin E_n$ is at most $1/n^2$.

Thus the expected volume of $P \setminus E_n$ is at most $\frac{1}{n^2}$ plus the volume of the non ϵ central region. The latter is at most $C(P)(d+1)(\log n)^{d+1}/(n-d+1)$ as a simple calculation shows. Summing, we get that the expected volume of $P \setminus E_n$ is at most $C(P)(\log n)^{d+1}/n$ for $n \geq 2d$ for a suitable $C(P)$.

To get the expected number of extreme points of E_n, note that the probability that one of the chosen n points is extreme equal to the probability that it does not belong to E_{n-1}. Thus the central region contributes at most $n/(n-1)^2 < 1$ to the expected number of extreme points , while the non central region contributes at most n times its expected volume overall at most $O((\log n)^{d+1})$. So we have proved the theorem promised in the introduction.

Theorem 2 *Suppose P is a polytope in R^d and E_n is the convex hull of n randomly and independently picked points in P . Then the expected volume of $P \setminus E_n$ is $O((\log n)^{d+1}/n)$ and the expected number of extreme points of E_n is $O((\log n)^{d+1})$ where the hidden constants depend only upon P and d.*

4. Lower bounds

We prove lower bounds on the volume of the non ϵ central region. These, we will see, also imply lower bounds on V_n and M_n.

Theorem 3 *If P is a polytope of nonzero volume in R^d, such that P has at least one vertex with precisly d neighbouring vertices, then there exists a constant $e(P)$ depending only upon P such that for any ϵ, the volume of the set of points that are not ϵ central to P is at least $e(P)\epsilon(\log \frac{1}{\epsilon})^{d-1}$ times the volume of P.*

Proof. : First, we give the argument for P = the cube $\{x : 0 \leq x_i \leq 1 \text{ for } i = 1,2,\ldots d\}$. For any point $p = (p_1, p_2, \ldots, p_d)$ in the cube, consider the hyperplane $\{x : \sum x_i/p_i = d\}$. This hyperplane, call it H, passes through p. Further the region in the cube "below" H is contained in the simplex with vertices $(0,0,\ldots,0), (dp_1,0,0,\ldots,0), (0,dp_2,0,\ldots,0), \ldots, (0,0,\ldots,0,dp_d)$, whose volume is $d^d \prod p_i/d!$. Thus if the product of the p_i is less than $d!\epsilon/d^d$, then p is not ϵ central to P. The following claim finishes the proof of the lemma in the case of cubes.

Claim : The volume of the set $S_d(\delta) = \{p \in R^d : 0 \leq p_i \leq 1, \prod p_i \leq \delta\}$ is at least $\delta(\log \frac{1}{\delta})^{d-1}/(d-1)!$.

Proof. : The proof is by induction on d. For $d = 1$, it is clear. In general, let $v(d, \delta)$ be the volume of $S_d(\delta)$. Then, we have

$$v(d, \delta) = \delta + \int_{x_1 = delta}^{1} v(d-1, \delta/x_1) dx_1$$

Now the claim follows by induction.

Suppose now P is any polytope with a vertex p of degree d. By trnaslating P, assume that $p = 0$. Let $p^1, p^2, \ldots p^d$ be the vertices adjacent to 0. Then the parallelopiped $Q = \{x : x = \sum \lambda_i p^i; 0 \leq \lambda_i \leq 1/d\}$, is contained in P and furthermore, P is contained in the cone $\{x : x = \lambda_i p^i, \lambda_i \geq 0\}$. So if a point x in Q has a hyperplane passing through it so that the volume of the cone below the hyperplane is less than ϵ times the volume of P, then clearly, x is not ϵ central to P. Let τ be a linear transformation that sends Q to the unit cube. τ preserves ratios of volumes, so the theorem follows from the argument for cubes plus the fact that the ratio of the volume of Q to the volume of P is a constant that depends only upon P.

To derive lower bounds on M_n, V_n, we prove the simple converse to Lemma 1.

Proposition 3 *If x is not ϵ central to P, then the probability that x does not belong to E_n is at least $(1-\epsilon)^n$.*

Proof. : The proof is straight forward and is ommited.

Now we choose $\epsilon = t/n$ for an as yet unspecified constant t. Then the non ϵ central points comprise a set of volume at least $e(P)(\log n)^{d-1}/n)$ for some constant $e(P)$ depending only upon P (and t which is but a constant) . For each x such that x is not ϵ central to P, the probability that x is not in E_n is at least $e^{-t} = O(1)$ for t independent of n. So the expected volume of $P \setminus E_n$ is at least $O((\log n)^{d-1}/n)$. The lower bound on M_n is proved as follows. By Chernoff bounds, the probability that the number of non-ϵ central points out of the n randomly picked points is less than $s = e(P)(\log n)^{d-1}(1-e^{-1})/(2n)$ is at most $1/2$ (this is a very crude estimate). Thus with probability at least a half s or more points are picked from the non central region. Each such point is an extreme point of E_n with probability a constant greater than 0. So we have shown the following :

Theorem 4 *Let E_n be the convex hull of n indpendently and randomly picked points from a d dimensional polytope P in R^d, that has at least one vertex with precisely d adjacent vertices. The expected value of the ratio of the volume of $P \setminus E_n$ to the volume of P is at least $\Omega((\log n)^{d-1}/n)$; the expected number of extreme*