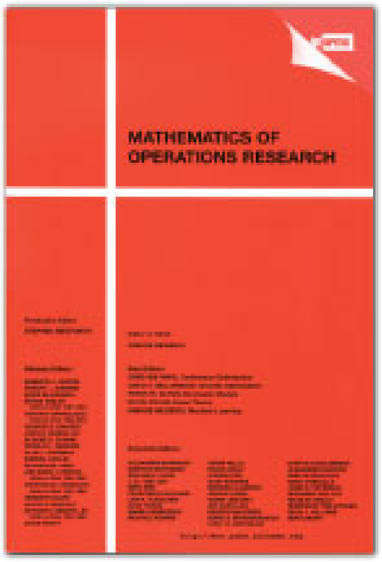

Mathematics of Operations Research

Martingale Inequalities and NP-Complete Problems

Wansoo T. Rhee, Michel Talagrand,

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 12, No. 1, February 1987
Printed in U.S.A.

MARTINGALE INEQUALITIES AND NP-COMPLETE PROBLEMS*

WANSOO T. RHEE† AND MICHEL TALAGRAND‡

We use martingale inequalities in the probabilistic analysis of stochastic models of NP-complete problems. We give examples of applications to the Bin Packing and Traveling-Salesman Problems. In both cases, we obtain sharp bounds on $\Pr(|U_n - EU_n| > t)$, where U_n is the objective function value of the problem with n random data.

1. Introduction. Much of the literature on NP-complete problems concentrates on developing heuristic algorithms (see references on NP-complete [4], [5]). Mostly deterministic problems are studied. However, the stochastic analysis of NP-complete problems has not yet received the same degree of attention and most results available are limited to heuristics under a special probability distribution condition ([1], [6], [7], [10]).

We consider here stochastic analysis of the Bin Packing and Traveling-Salesman Problems (TSP), two important NP-complete problems. J. M. Steele [13] showed that if X_i, $1 \leqslant i < \infty$, are independent uniformly distributed in $[0,1]^2$ and T_n is the length of the shortest closed path connecting $\{X_1, \dots, X_n\}$, then there is a constant $0 < \beta < \infty$ such that for all $\epsilon > 0$,

$$\sum_{n=1}^{\infty} \Pr\left(|T_n/\sqrt{n} - \beta| > \epsilon\right) < \infty.$$

This type of convergence is called complete convergence. B. Weide [15] pointed out that Steele's result is important to the motivation behind Karp's probabilistic analysis of the TSP.

Steele's result is based on an inequality of B. Efron and C. Stein. We shall use here the more powerful technique of martingale inequalities. (In a separate work [12], we actually show that Efron-Stein's inequality can be simply derived using martingale techniques.) This technique is very general and powerful, and it can be reasonably expected to find many other applications in the asymptotic analysis of nonlinear processes. It has been used with great success in the local theory of Banach Spaces (see the forthcoming book of V. Milman and G. Schechman for references [9]).

Consider an increasing sequence $\Sigma_0, \dots, \Sigma_n$ of sub-σ-fields of a basic probability space, where Σ_0 is trivial. A sequence of random variables $(d_i)_{1 \leqslant i \leqslant n}$ is called a martingale difference sequence (m.d.s.) if each d_i is Σ_i-measurable, and if $E(d_i|\Sigma_{i-1}) = 0$ for each i. A trivial but essential observation is that if f is Σ_n-measurable, we can write $f - E(f) = \Sigma_{1 \leqslant i \leqslant n} d_i$, where $d_i = E(f|\Sigma_i) - E(f|\Sigma_{i-1})$, so that (d_i) is a

*Received May 16, 1985; revised October 9, 1985.
AMS 1980 subject classification. Primary: 68C25. Secondary: 90C10, 65K05, 60G48.
IAOR 1973 subject classification. Main: Programming: Integer. Cross References: Probability.
OR/MS Index 1978 subject classification. Primary: 628 Programming/Integer/Algorithms. Secondary: 572 Probability/Stochastic model applications.
Key words. martingale inequalities, stochastic model, tail deviation.
†Ohio State University.
‡University of Paris VI.

0364-765X/87/1104/0177$01.25

martingale difference sequence. Many deep inequalities are known for martingale difference sequences. One of the most remarkable results is the following subgaussian inequality, due to Azuma (see [14, Lemma 4-2-3 and Exercise 4-2-2]).

LEMMA 1. *Let $(d_i)_{1 \le i \le n}$ be a m.d.s. Then for each $t > 0$,*

$$\Pr\left(\left|\sum_{i \le n} d_i\right| > t\right) \le 2\exp\left(-t^2/\left(2\sum_{i \le n} \|d_i\|_\infty^2\right)\right)$$

where $\|d_i\|_\infty$ is the essential sup norm of d_i.

Thus we can get bounds on $\Pr(|f - E(f)| > t)$ provided we can control the numbers $\|d_i\|_\infty$.

2. Bin packing. Although we are primarily interested in the Traveling Salesman Problem, we discuss first the much easier case of bin packing. The Bin Packing problem requires finding the minimum number of unit size bins needed to pack a given collection of items with sizes in $[0,1]$. In our model, we have n items to pack, and the size of items are $X_1, \ldots, X_n$, independent identically distributed (i.i.d.) over $[0,1]$. For simplicity of notation, we denote by X_k both item names and item sizes. This will not create any ambiguity. We use a fixed procedure for packing. We assume that this procedure has the following property:

($*$) If the procedure packs $X_1, \ldots, X_n$ in k bins, the number of bins needed to pack $X_1, \ldots, X_l, Y, X_{l+1}, \ldots, X_n$ for any l, and Y in $[0,1]$ is at least k and at most $k+1$.

If the procedure is optimum, that is, it chooses the minimum number of bins, it obviously satisfies ($*$). Another procedure of interest is Next Fit where the bins are filled one at a time and a new bin is started when the current element does not fit in the bin being currently filled. It is easy to see that Next Fit (and hence also Next Fit Decreasing) satisfies a condition similar to ($*$) and that the number of bins needed after addition of one item increases by at most two (so the exponent in Theorem 2 has to be changed to $-t^2/8n$ in that case).

We denote by B the number of bins needed to pack $X_1, \ldots, X_n$. The following improves on a result of W. Rhee [11], where a bound $\exp(-t/\sqrt{n})$ was obtained using Steele's method.

THEOREM 2. *If condition ($*$) holds, then for $t \ge 0$,*

$$\Pr(|B - EB| > t) \le 2\exp(-t^2/(2n)).$$

PROOF. For $0 \le i \le n$, let Σ_i be the σ-field generated by $(X_l)_{l \le i}$. Then B is Σ_n-measurable. Fix i, $1 \le i \le n$, and denote by B^i the number of bins needed to pack $X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n$. By property ($*$), we have $B^i \le B \le B^i + 1$. Thus

$$E(B^i|\Sigma_i) \le E(B|\Sigma_i) \le E(B^i|\Sigma_i) + 1 \quad \text{and}$$

$$E(B^i|\Sigma_{i-1}) \le E(B|\Sigma_{i-1}) \le E(B^i|\Sigma_{i-1}) + 1.$$

However, $E(B^i|\Sigma_i) = E(B^i|\Sigma_{i-1})$. Thus

$$|E(B|\Sigma_i) - E(B|\Sigma_{i-1})| \le 1,$$

that is, $\|d_i\|_\infty \le 1$. The result then follows from Lemma 1. Q.E.D.

In the case where the process is subadditive, in the sense that when the procedure packs $X_1, \ldots, X_n$ in k bins and $Y_1, \ldots, Y_m$ in l bins, then it packs $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ in at most $k + l$ bins, the theory of subadditive processes [8] shows that $\lim_{n \to \infty} B_n/n = c$ exists almost surely (a.s.), where B_n is the number of bins needed to pack the first

n items of an infinite i.i.d. sequence (X_i). The optimal packing and Next Fit are subadditive. A standard 2ϵ-argument shows the following theorem.

THEOREM 3. *There exists $c > 0$ such that for each $\epsilon > 0$, and n large enough,*

$$\Pr\left(|(B_n/n) - c| > \epsilon\right) \leqslant 2\exp\left(-n\epsilon^2/4\right).$$

In the case of Next Fit, this result improves considerably on H. Ong et al., Theorem 3 [10].

3. The Traveling Salesman Problem. In this model of the Traveling-Salesman Problem (TSP), $X_1, \ldots, X_n$ are i.i.d. uniformly distributed over $[0,1]^2$. We denote T (or T_n when the dependence on n needs to be emphasized) the shortest path through $X_1, \ldots, X_n$. We define Σ_i to be the σ-algebra generated by $(X_k)_{k \leqslant i}$. In order to get bounds on $d_i = E(T|\Sigma_i) - E(T|\Sigma_{i-1})$, we use the same method as in the bin packing case. For $1 \leqslant i \leqslant n$, we denote T^i the shortest path through $X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n$ and we note that $E(T^i|\Sigma_i) = E(T^i|\Sigma_{i-1})$. Let $h_i = T - T^i$. Then

$$d_i = E(T|\Sigma_i) - E(T|\Sigma_{i-1}) = E(h_i|\Sigma_i) - E(h_i|\Sigma_{i-1}). \tag{3-1}$$

LEMMA 4. *Let $k \leqslant i \leqslant n$. Then there exist numerical constants C_1 and C_2 such that*

$$E(h_i|\Sigma_k) \leqslant C_1/(n-k-1)^{1/2}, \quad E\left(h_i^2|\Sigma_k\right) \leqslant C_2/(n-k-1) \tag{3-2}$$

for $k < n-1$, and $|h_i| \leqslant C_1, C_2$ for each $i \leqslant n$.

PROOF. We follow the argument of J. M. Steele [13]. If there is a path through $X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n$ of length a, the shortest path through $X_1, \ldots, X_n$ is always of length $\leqslant a + 2\sqrt{2}$, so $h_i \leqslant 2\sqrt{2}$. For $k < n-1$, it is of length $\leqslant a + u$, where $u = 2\min\{d(X_i, X_l);\ k < l \leqslant n,\ l \neq i\}$. Since u is independent of Σ_k, we have $E(h_i|\Sigma_k) \leqslant Eu$, $E(h_i^2|\Sigma_k) \leqslant Eu^2$. As observed by J. M. Steele, for some $\alpha > 0$,

$$\Pr(u > t) \leqslant \left(1 - \alpha t^2\right)^{n-k-1} \leqslant \exp\left(-\alpha(n-k-1)t^2\right) \tag{3-3}$$

and this gives $Eu \leqslant C_1(n-k-1)^{-1/2}$, $Eu^2 \leqslant C_2(n-k-1)^{-1}$. Q.E.D.

COROLLARY 5. *There exist numerical constants C_3 and C_4 such that for $k \leqslant i \leqslant n$, we have*

$$\|d_i\|_\infty \leqslant C_3(n-i-1)^{-1/2}, \quad E\left(d_i^2|\Sigma_k\right) \leqslant C_4(n-k-1)^{-1} \quad \text{for } k < i < n-1$$

and also

$$\|d_i\|_\infty \leqslant C_3, \quad E\left(d_i^2|\Sigma_k\right) \leqslant C_4 \quad \text{for each } i, k.$$

PROOF. We have

$$\|d_i\|_\infty \leqslant \left\|E(h_i|\Sigma_i)\right\|_\infty + \left\|E(h_i|\Sigma_{i-1})\right\|_\infty \leqslant 2\left\|E(h_i|\Sigma_i)\right\|_\infty.$$

So $\|d_i\|_\infty \leqslant 2C_1(n-i-1)^{-1/2}$ for $i < n-1$ while $\|d_i\|_\infty \leqslant 2C_1$ for $i = n$ or $i =$

$n-1$. We have for $k<n-1$, $k<i\leqslant n$,

$$E\left(d_i^2|\Sigma_k\right) \leqslant 2\left\{E\left(\left(E(h_i|\Sigma_i)\right)^2|\Sigma_k\right) + E\left(\left(E(h_i|\Sigma_{i-1})\right)^2|\Sigma_k\right)\right\}$$

$$\leqslant 4E\left(h_i^2|\Sigma_k\right) \leqslant 4C_2(n-k-1)^{-1},$$

and also $E(d_n^2|\Sigma_{n-1}) \leqslant 4C_2$. Q.E.D.

Using Corollary 5 and Lemma 1, we get the following:

PROPOSITION 6. *There is a numerical constant C such that for $n \geqslant 2$*

$$\Pr(|T_n - ET_n| > t) \leqslant 2\exp(-t^2/(C\ln n)). \tag{3-4}$$

In particular,

$$\Pr\left(|T_n - ET_n| > \epsilon\sqrt{n}\right) \leqslant 2\exp(-\epsilon^2 n/(C\ln n)). \tag{3-5}$$

J. M. Steele shows that for each p, there is a numerical constant S_p such that $\|T_n - ET_n\|_p \leqslant S_p$ where $\|f\|_p = (E|f|^p)^{1/p}$. Using Chebyshev's inequality, he gives only a bound for $\Pr(|T_n - ET_n| > \epsilon\sqrt{n})$ of order $O(n^{-p})$ for each p. So (3-5) is, in one sense, sharper than Steele's result. On the other hand, for t fixed, the bound in (3-4) becomes useless for $n \to \infty$, so it does not imply even that the sequence $(T_n - ET_n)$ is stochastically bounded. The reason is that, in this case, Lemma 1 is not the best martingale inequality to use.

PROPOSITION 7 ([3, Theorem III-1.1]). *There exists a numerical constant C_5 such that for each m.d.s. $(d_i)_{1\leqslant i\leqslant n}$, and each $t>0$,*

$$\Pr(|\Sigma d_i| > t) \leqslant 2\exp(-t/C_5 A)$$

where $A = \mathrm{Max}_{1\leqslant k\leqslant n}\|E(\Sigma_{k\leqslant i\leqslant n} d_i^2|\Sigma_k)\|_\infty^{1/2}$.

THEOREM 8. *There exists a numerical constant γ such that for each n,*

$$\Pr(|T_n - ET_n| > t) \leqslant 2\exp(-\gamma t). \tag{3-6}$$

PROOF. We have

$$E\left(\sum_{k\leqslant i\leqslant n} d_i^2|\Sigma_k\right) \leqslant E\left(d_k^2|\Sigma_k\right) + \sum_{k<i\leqslant n} E\left(d_i^2|\Sigma_k\right).$$

So, from Corollary 5, we have

$$\left\|E\left(\sum_{k\leqslant i\leqslant n} d_i^2|\Sigma_k\right)\right\|_\infty \leqslant C_4 + (n-k-1)C_4(n-k-1)^{-1} \leqslant 2C_4$$

and the result follows from Proposition 7.

REMARKS. (a) It seems reasonable to conjecture that (3-6) is essentially the best possible.

(b) From (3-6) follows that for each p, $\|T_n - ET_n\|_p \leqslant C_7 p$, where C_7 is numerical constant. On the other hand, the constant S_p obtained by J. M. Steele are very fast growing functions of p.

(c) When t is significantly larger than $\ln n$, the bound in (3-4) is better than the bound in (3-6).

References

[1] Coffman, E. G., So, K., Hofri, M. and Yao, A. C. (1980). A Stochastic Model of Bin Packing. *Inform. and Control* **44** 105–115.

[2] Efron, B. and Stein, C. (1981). The Jackknife Estimate of Variance. *Ann. Statist.* **9** 586–596.

[3] Garsia, A. M. (1973). Martingale inequalities. *Seminar Notes on Recent Progresses*. W. A. Benjamin, New York.

[4] Garey, M. R. and Johnson, D. S. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Co., San Francisco.

[5] Karp, R. M. (1972). Reducibility among Combinatorial Problems. *Complexity of Computer Computations*. R. E. Miller and J. W. Thatcher (Eds.), Plenum Press, New York, 85–103.

[6] ______. (1976). The Probabilistic Analysis of Some Combinatorial Search Algorithms and Complexity. *New Directions and Recent Results*. J. F. Taub (Ed.), Academic Press, New York, 1–19.

[7] ______. (1977). Probabilistic Analysis of Partitioning Algorithms for the Traveling Salesman Problem in the Plane. *Math. Oper. Res.* **2** 209–224.

[8] Kingman, J. F. C. (1976). Subadditive Processes. Lecture Notes in Math., **539**, Springer-Verlag, Berlin and New York, 168–222.

[9] Milman, V. and Schechman, G. (1986). Asymptotical Theory of Finite Dimensional Norm Space, Springer-Verlag, *Lecture Notes in Math*. 1200.

[10] Ong, H. L., Magazine, M. J. and Wee, T. S. (1984). Probabilistic Analysis of Bin Packing Heuristics. *Oper. Res.* **32**, **5** 983–998.

[11] Rhee, W. (1985). Convergence of Optimal Stochastic Bin Packing. *Oper. Res. Lett.* **4**, **3** 121–123.

[12] ______ and Talagrand, M. (1986). Martingale Inequalities and the Jackknife Estimate of Variance. *Statist. Probab. Lett.* **4** 5–6.

[13] Steele, J. M. (1981). Complete Convergence of Short Paths and Karp's Algorithm for the TSP. *Math. Oper. Res.* **6**, **3** 374–378.

[14] Stout, W. F. (1974). *Almost Sure Convergence*. Academic Press, New York.

[15] Weide, B. W. (1978). Statistical Methods in Algorithm Design and Analysis. Thesis, Department of Computer Science, Carnegie-Mellon University.

RHEE: FACULTY OF MANAGEMENT SCIENCES, THE OHIO STATE UNIVERSITY, 1775 COLLEGE ROAD, COLUMBUS, OHIO 43210

TALAGRAND: EQUIPE D'ANALYSE–TOUR 46, UNIVERSITY OF PARIS VI, 4 PLACE JUSSIEU, 75230 PARIS CEDEX 05, FRANCE