

On the Number of Nonzeros Added when Gaussian Elimination is Performed on Sparse Random Matrices

By I. S. Duff

Abstract. This paper studies the fill-in properties of Gaussian elimination on sparse random matrices. A theoretical study using the concept of random graphs yields formulae from which the fill-in can be evaluated. The predictions are examined for matrices of various orders and densities and are found to be in close agreement with experimental results. A possible consequence of the results of this paper relating to error analysis for sparse systems is given in the concluding section.

1. Introduction. This paper studies the fill-in properties of Gaussian elimination on sparse random matrices. Graph theoretical ideas are introduced and the problem is examined using the concept of random graphs.

This concept of using random graphs is not a new one. They are used as a tool in the development of existence theorems in graph theory [Erdős (1967)] and some work has been done on the structure of a graph where the number of lines is a function of the number of points, the analysis holding in the limit as the number of points tends to infinity [Palásti (1966, 1970)]. Ogilvie (1968) has shown that these asymptotic results hold where the graph has only a small number of vertices (about 10) but none of this work is directly concerned with matrix decomposition or indeed with the adjacency matrix of the graph. Heap (1966) has extended the notion of random graphs to examine properties of their adjacency matrices. However, his results on matrix reducibility are only valid for comparatively dense matrices (probability density ≥ 0.15). A serious attempt to tackle the problem of this paper has been made by Hsieh and Ghausi (1971) where they have tried to make theoretical predictions on the results of Brayton et al. (1970). Their work has, however, several serious defects. Although they make more allowance for interaction than the discussion in Section 2, their formulae only include correlation between elements in the same column ignoring the correlation which exists between elements in the same row. In addition, some of their formulae break down if the density of the matrix increases above 20% and their correlation factor H is chosen empirically to obtain a fit with Brayton's data rather than determined theoretically. The impetus for writing this paper came from remarks by Robert Brayton in the panel at the end of the IBM symposium [Rose and Willoughby (1972)] when he stated that the above problem, although itself not perhaps of devastating importance, has consumed vast

Received November 21, 1972.

AMS (MOS) subject classifications (1970). Primary 05C20, 65F05, 65F99; Secondary 05C99, 60C05, 65G05, 65S05.

Key words and phrases. Sparse random matrix, Gaussian elimination, fill-in, random graphs, legal paths, error analysis.

Copyright © 1974, American Mathematical Society

amounts of man power in its attempted solution. This paper is dedicated to releasing these man hours for more productive work.

Section 2 of this paper makes an 'obvious' initial approach to the problem, indicates some of the difficulties involved in its solution, and gives an upper bound to the quantity it is desired to evaluate. In Section 3, the above-mentioned notion of random graphs is introduced and the elimination problem is interpreted in graph theoretic terms. This interpretation is used in Section 4 to derive formulae relating to the fill-in caused by Gaussian elimination on random graphs. It is seen, in Section 5, that these formulae give results which are in good agreement with experimental data generated by the author and Brayton et al. (1970). The concluding section gives a consequence of the results of this paper and indicates other problems which may be tackled by such techniques.

2. A Statement of the Problem and an Initial Approach. The problem to be considered in this paper is that of Gaussian elimination performed on a matrix which is characterised by having all its diagonal elements nonzero, and with all the off-diagonal elements having the same probability p , say, of being nonzero. The elimination on this random matrix is performed by pivoting down the diagonal in the natural order, and the interest is centred on the probable number of fill-ins that occur. This number will, of course, be some function of n , the order of the matrix, and p only.

Because of the nature of the elimination process, it is immediately seen that a probability P_i can be ascribed to the likelihood of an element being nonzero in the i th reduced matrix. This P_i will be the same for all elements of the reduced matrix.

It is perhaps worthwhile to indicate a natural way to tackle this problem and, by showing its failure, indicate the way in which many 'simple' solutions fail to answer this problem.

P_1 will always be equal to p , the initial probability of the off-diagonal matrix elements being nonzero.

An element in the reduced matrix of order $n - 1$ will be nonzero if:

(i) it is initially nonzero: probability p ;

(ii) it is initially zero and the elements at the head of its row and column in the original matrix are both nonzero thus causing a fill-in to occur: probability $(1 - p) \cdot p^2$, since the events are independent. It thus follows that the relation

$$P_2 = p + (1 - p) \cdot p^2$$

is true.

Now consider the following argument.

An element in the third reduced matrix will be nonzero if it is either nonzero initially or after the first elimination stage (probability P_2), or if it is zero then, but is filled in at the second stage of the elimination (probability $(1 - P_2) \cdot P_2^2$), giving the equation

$$P_3 = P_2 + (1 - P_2) \cdot P_2^2$$

and, by a similar argument, the following recurrence relation holds for P_i ($i = 1, \dots, n - 1$):

$$(1) \quad P_{i+1} = P_i + (1 - P_i) \cdot P_i^2, \quad P_1 = p.$$

The experimental results of Section 5 indicate that the values of P_i obtained from Eq. (1) are clearly overestimates of the true value. Some explanation of this in the case $i = 3$ is now given and then a general theorem relating the P_i to the true value of the probability is stated and proved.

The reason why Eq. (1) fails to give the correct value for P_3 is as follows. Figure 1 should be consulted.

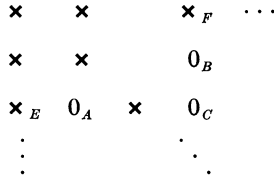


FIGURE 1

It is true to say that the probability of element C being zero after the first stage is $(1 - P_2)$ and that the probability of A and B both being nonzero after the first stage is P_2^2 (since events are independent), but it is certainly not true to claim that these two events are independent as can be seen in Fig. 1, where both the $(1 - P_2)$ and the P_2^2 term depend on the two elements E and F . Thus, if

$\text{Prob}(C \text{ is zero after first stage}) = (1 - P_2)$, and
 $\text{Prob}(A \text{ and } B \text{ nonzero after first stage}) = P_2^2$, then
 $\text{Prob}(C \text{ zero and } A \text{ and } B \text{ nonzero after the first stage}) \neq P_2^2(1 - P_2)$.
 This argument is formalised in the following theorem.

THEOREM 1. *If \bar{P}_i is defined by the relation*

$$\bar{P}_{i+1} = \bar{P}_i + (1 - \bar{P}_i)\bar{P}_i^2, \quad \bar{P}_1 = p,$$

and if P_i is the probability of an element in the i th reduced matrix being nonzero after the decomposition, then the \bar{P}_i form an upper bound for the P_i . In fact,

$$\begin{aligned}
 (2) \quad & \bar{P}_i \geq P_i, \quad (1 \leq i \leq 2), \\
 & \bar{P}_i > P_i, \quad (3 \leq i \leq n - 1).
 \end{aligned}$$

Proof. Certainly, $\bar{P}_i \geq P_i$, $i = 1, 2$, and relation (2) is assumed to hold for all $i \leq r < n - 1$, then, \bar{P}_{r+1} is given by the relation

$$\bar{P}_{r+1} = \bar{P}_r + (1 - \bar{P}_r)\bar{P}_r^2$$

and since $x + (1 - x)x^2$ is a strictly increasing function of x , for $x \in [0, 1]$, the induction hypothesis gives the equation

$$\bar{P}_{r+1} \geq P_r + (1 - P_r)P_r^2.$$

Now,

$$\begin{aligned}
 P_{r+1} &= \text{Prob}\{\text{element in the } (i + 1)\text{th reduced matrix is filled in}\} \\
 &= \text{Prob}\{\text{element in } i\text{th reduced matrix is filled in}\} \\
 &\quad + \text{Prob}\{\text{element is filled in during the } (i + 1)\text{th elimination} \\
 &\quad \quad \quad \text{step and it is zero in the } i\text{th reduced matrix}\}.
 \end{aligned}$$

Thus,

$$P_{r+1} \leq P_r + (1 - P_r) \cdot P_r^2 \leq \bar{P}_r + (1 - \bar{P}_r) \cdot \bar{P}_r^2 = \bar{P}_{r+1}$$

and the result follows by induction. Since $\bar{P}_3 > P_3$, $\bar{P}_i > P_i$, $i = 3, \dots, n - 1$, because of the strict monotonicity of the function $x + (1 - x) \cdot x^2$, $x \in [0, 1]$.

3. A Graph Theoretic Description of the Problem. In this section and the following one, directed graphs as defined by Harary, Norman and Cartwright (1965) are used to analyse the fill-in problem. The terminology used is identical to theirs, the only additional term being that of 'linked'. Point a is *linked* to point b in a directed graph if there exists a directed line of the graph from a to b . A point a is *connected* to a point b if there exists a directed path in the graph from a to b . Thus linked points are connected but not vice versa.

If a global view is taken of the problem, it is natural to ask what possible sequence of events would cause the element $a_{i,i}$ to be filled in. By the discussion of the last section, it is possible, without loss of generality, to consider the case $a_{i,i+1}$. Only the first $i - 1$ stages of the elimination could cause fill-in to this position in the matrix and so only these need be considered in the following discussion.

A random graph on n points is defined to be a directed graph on n points where each line has a probability p ($p \ll 1$, in general) of existing, this probability being the same for all the lines. This directed graph is seen to be none other than the adjacency matrix [Harary (1962b)] for the random matrix described in the first section of this paper.

If the theory of graphs in the elimination process is examined (for example, Tewarson (1967), Harary (1962a), and Parter (1961)), it is seen that one point will be linked to another point in the graphs of subsequent reduced matrices if there exists a path in the initial graph connecting that particular point to the other which does not pass through a remaining point (the term 'legal' path is used to describe this and is attributed to Karp [Willoughby (1971)]). Since the addition of a line in the graphs of the reduced matrices is in 1-1 correspondence with the fill-in of zeros in the original matrix, the following theorem holds.

THEOREM 2. *The probability of the element in the $(i, i + 1)$ th position in the matrix being nonzero after matrix reduction is equal to the probability that there exists a path from the point i to the point $i + 1$ in a random graph on $i + 1$ points if the probability of any line existing in the random graph is the same as the probability of an off-diagonal element of the matrix being nonzero.*

4. The Derivation of the Formula. Let the random graph on n vertices with probability p of any edge existing be denoted by $G(n, p)$. Now, let $P_x(n, p)$ be the probability that a point x is connected to every other point in the graph $G(n, p)$. Thus, for example, $P_x(1, p) = 1$ and $P_x(2, p) = p$. Then the following lemma is true.

LEMMA 1. $P_x(n, p)$ is given by the recurrence relation

$$P_x(n, p) = 1 - \sum_{k=1}^{n-1} \binom{n-1}{k-1} \cdot P_x(k, p) \cdot (1-p)^{k(n-k)}$$

for any n and p .

Proof. Let the points of the graph be partitioned into two subsets, a k -set containing k points and an $(n - k)$ -set containing the remaining $n - k$ points and choose the k points such that x is one of them. Then the probability that x is connected to all the other points in the k -set but to none of the other points is given by the

formula

$$P_x(k, p) \cdot (1 - p)^{k(n-k)}$$

the second part of the formula being present since it is necessary that no lines exist from the points of the k -set to the points of the $(n - k)$ -set. It is possible to select $\binom{n-1}{k-1}$ subsets from the points of $G(n, p)$ such that x is always in the k -set. Furthermore, since the events of x being connected to all points of the k -set but none of the points in the $(n - k)$ -set are mutually exclusive for different selections of k -sets, the probability of x being connected to $k - 1$ other points in the graph $G(n, p)$ is given by the formula

$$\binom{n - 1}{k - 1} P_x(k, p) \cdot (1 - p)^{k(n-k)}.$$

Now, the probability of x being connected to $n - 2$ or less points in the graph is given by the formula

$$\sum_{k=1}^{n-1} \binom{n - 1}{k - 1} P_x(k, p) \cdot (1 - p)^{k(n-k)}$$

since the events of x being connected to $k - 1$ and only $k - 1$ points are mutually exclusive as k varies from 1 to $n - 1$.

But this is the probability that x is not connected to all the other points of the graph and so the lemma has been proved.

The idea of connectedness can now be extended in the following manner. If $P_{xy}(n, p)$ is the probability that there exists a directed path from x to y in $G(n, p)$ then the following lemma holds.

LEMMA 2. For all n and p , $P_{xy}(n, p)$, as defined above, is given by

$$P_{xy}(n, p) = 1 - \sum_{k=1}^{n-1} \binom{n - 2}{k - 1} P_x(k, p) (1 - p)^{k(n-k)}$$

where the $P_x(k, p)$ have the same meaning as in Lemma 1.

Proof. If, as in the proof of Lemma 1, the points of the graph are thought of as being partitioned into two subsets, a k -set containing x , and an $(n - k)$ -set containing y , then the probability that x is connected to every point in the k -set but to none in the $(n - k)$ -set (including y) is $P_x(k, p) \cdot (1 - p)^{k(n-k)}$ as before.

But this time the $k - 1$ other points of the k -set must only be selected from the $(n - 2)$ points of the graph omitting x and y , and so the probability that x is connected to any $k - 1$ points of the graph and only $k - 1$ points and is, in addition, not connected to y is given by the formula

$$\binom{n - 2}{k - 1} P_x(k, p) (1 - p)^{k(n-k)}$$

and so the probability that x is connected to $n - 2$ other points or less but is not connected to y is given by the formula

$$\sum_{k=1}^{n-1} \binom{n - 2}{k - 1} P_x(k, p) (1 - p)^{k(n-k)}$$

and, since this is just the probability that x is not connected to y in $G(n, p)$, the lemma has now been proved.

The following theorem follows immediately from Theorem 2 and Lemmas 1 and 2.

THEOREM 3. *If P_i is the probability that an element in the i th reduced matrix is nonzero during the LU decomposition of a random matrix, as defined in this paper, then the following relation holds*

$$(3) \quad P_i = 1 - \sum_{k=1}^i \binom{i-1}{k-1} H_k (1-p)^{k(i+1-k)}$$

where the H_k are given by

$$(4) \quad H_k = 1 - \sum_{j=1}^{k-1} \binom{k-1}{j-1} H_j (1-p)^{j(k-j)}.$$

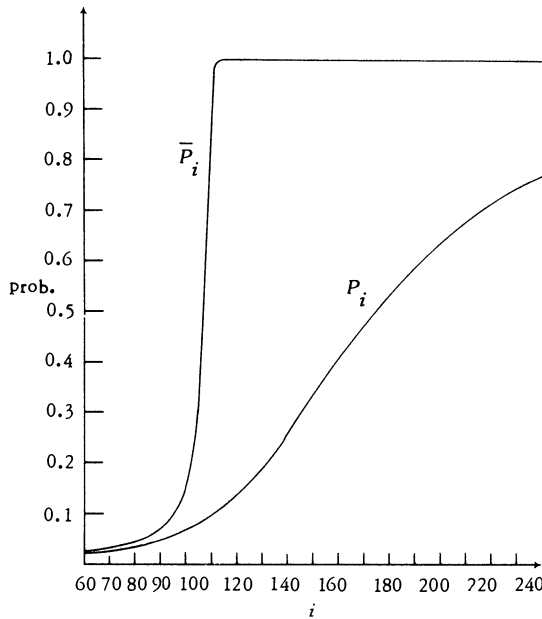
The next theorem then follows immediately.

THEOREM 4. *The expected value for the total number of nonzeros in the LU decomposition of a random matrix, as in Theorem 3, is given by the formula*

$$(5) \quad 2 \cdot \sum_{i=1}^{n-1} (n-i) P_i + n$$

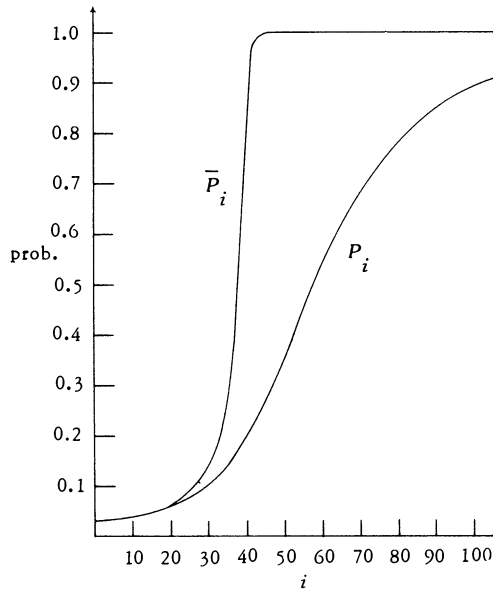
where the P_i are given by Eqs. (3) and (4) and the unit nonzeros in the diagonal of U are excluded since they are not stored.

5. Experimental Results. In order to examine the defects in using the upper bound approximation of Eq. (1) as an estimate for the probability density of the LU decomposition of a random matrix, the results of computing the recurrence relation for various starting values of p were compared with the probabilities obtained from Eqs. (3) and (4). These results are shown in Graphs 1 to 5 for values of p equal



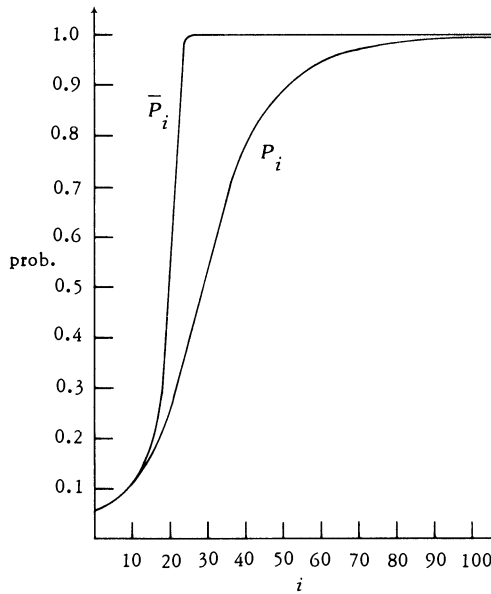
GRAPH 1

A comparison of the true probability density of the reduced form (P_i) with that derived from the recurrence relation (1) (\bar{P}_i) with starting value of $p = 0.01$. $\bar{P}_{60} = 0.023418$ while $P_{60} = 0.022709$.



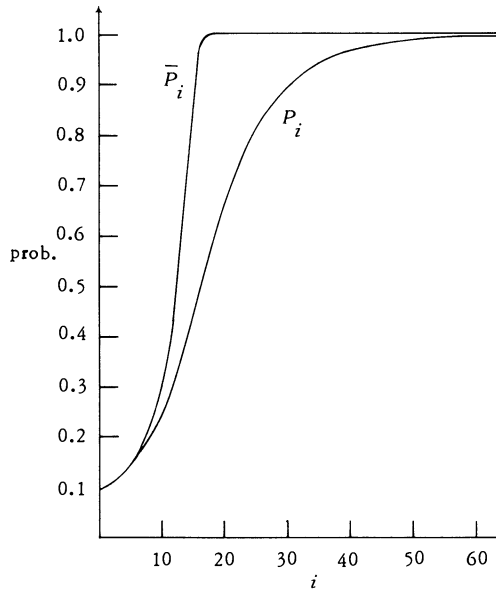
GRAPH 2

A comparison of the true probability density of the reduced form (P_i) with that derived from the recurrence relation (1) (\bar{P}_i) with starting value of $p = 0.03$.



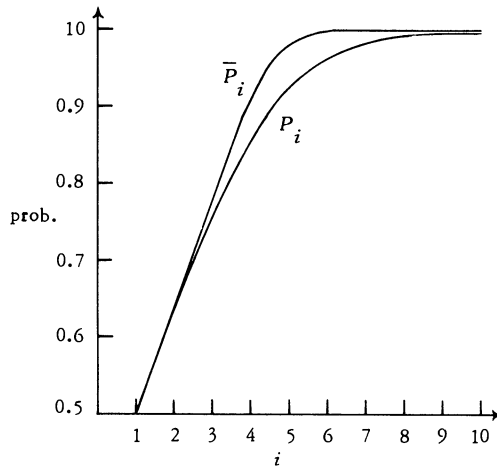
GRAPH 3

A comparison of the true probability of the reduced form (Eq. (3)) with that derived from the recurrence relation (1) with starting value of $p = 0.06$.



GRAPH 4

A comparison of the true probability density of the reduced form (Eq. (3)) with that derived from the recurrence relation (1) with starting value of $p = 0.10$.



GRAPH 5

A comparison of the true probability density of the reduced form with that derived from the recurrence relation (1) with starting value of $p = 0.50$.

to 0.01, 0.03, 0.06, 0.10, and 0.50, resp. It is seen immediately from these results that the \bar{P}_i generated from Eq. (1) give an upper bound to the true value for the probability density. The strict inequality indicated by Theorem 1 is not always evident because of the scale of the graph but the computed results from which the graphs were obtained do show such an inequality. For the lowest value of p in Graph 1, it is seen that the difference only becomes evident as i increases in size. For the middle

values of p in Graphs 2 to 4, the difference between the upper bound for P_i and its true value becomes evident at lower values of i while by the time p is as large as 0.5, the effect is nearly dwarfed by the increase in size of P_i and \bar{P}_i for increasing i .

The effect of using these two sets of values for the probability on the estimate of fill-in to a random matrix is shown in Table 1. In both cases, Eq. (5) was used

TABLE 1

A comparison of the average fill-in caused by Gaussian elimination on several different random matrices (average of 20 runs at each order and density) with the theoretical prediction given in Theorems 2 and 3 and the upper bound given by Eq. (1)

n	P_1	NUMBER OF OFF-DIAGONAL NON-ZEROS IN FINAL FORM			STANDARD DEVIATION OF EXPERIMENTAL RESULTS
		\bar{P}	P	EXPERIMENTAL	
100	0.01	176	166	160	14.8
50	0.01	30	29	29	6.25
30	0.01	10	10	10	2.54
25	0.01	6	6	6	0.89
20	0.01	4	4	4	1.80
100	0.03	4259	2253	2309	239
50	0.03	290	168	162	33.4
30	0.03	39	38	37	6.18
25	0.03	24	24	22	3.45
20	0.03	14	14	13	1.06
100	0.06	6763	5204	5211	334
50	0.06	1062	656	641	71.3
30	0.06	182	119	123	33.8
25	0.06	86	68	69	17.1
20	0.06	39	36	33	5.33
100	0.10	7966	7022	6925	246
50	0.10	1553	1212	1193	77.5
30	0.10	388	274	264	34.9
25	0.10	222	157	161	22.4
20	0.10	105	79	82	13.0
100	0.50	9662	9623	9637	47.5
50	0.50	2334	2315	2317	28.3
30	0.50	802	792	788	14.6
25	0.50	544	536	537	15.5
20	0.50	336	330	333	10.9

to calculate the values given in the appropriate columns. In order that the theoretical predictions of this chapter might be compared with experiment, some runs were done performing Gaussian elimination on random matrices, twenty runs being made at each value of order and density. The average number of nonzeros in the decomposed form and the standard deviation from this mean were calculated and the results are shown in the appropriate columns of the table. It is seen that, in all cases, the number of nonzeros predicted by the theory is in very close agreement with that obtained from the runs on random matrices, and the predicted value, in every case, falls well within the distribution peak defined by the experimental standard deviation. As expected, the use of the upper bound for P_i gives an overestimate of the fill-in which is particularly marked when n is high, except for the case $p = 0.5$ when the fill-in, in both cases, is almost total.

An examination of the results in the paper by Brayton et al. (1970) indicates that most of his experiments were performed with random matrices of a higher order and lower density than the ones used in this paper. In addition to the core storage restrictions for large systems (there is also a limitation if the matrix is held in packed form because of the amount of fill-in during the elimination process), another difficulty was encountered in the calculation of fill-in to matrices of low initial density. This difficulty was caused by rounding errors in the computation of P_i from Eqs. (3) and (4). These errors are introduced by the subtraction of the sum in Eq. (4) from 1. This particular error is relatively a very large one since the value of the summation is very close to 1. For example, if it is assumed that the computer can store up to 10 decimal digits, then a summation of value $1 - 10^{-7}$ will give an error in H_k in the third significant figure; and summation values of $1 - 10^{-20}$ are obtained if $p = 0.01$!. The use of double precision merely delays the catastrophe.

In an attempt to avoid this calamity, the author has developed a set of procedures to deal with the calculation of P_i from Eqs. (3) and (4) in rational arithmetic. These procedures are the subject of a future paper by the author and work to infinite precision (to within the core size of the machine). J. H. Griesmer of IBM has also computed values for H_k on a machine at Yorktown Heights using a rational arithmetic package in a system called SCRATCHPAD [Griesmer and Jenks (1971)]. Unfortunately, both the IBM system and the procedures of the author take a long time to evaluate H_k for large k (because at stage k , $k - 1$ infinite precision terms have to be calculated for the summation in Eq. (4)), and so it is impractical to try to test the results in Brayton et al. (1970) for low density and high n . For $p = 0.01$, the author was fortunate in being able to use a multiple precision arithmetic package developed by Seraphin and Milazzo (Stony Brook) from which the results in Graph 1 were obtained. These results can be seen to be consistent with the experimental calculations in Brayton et al. (1970) for their higher density matrices. If any readers of this paper have a good multiple precision package and unlimited computing time, they would be welcome to tackle the recurrence relation for lower values of p .

6. Conclusions and a Possible Consequence. This paper has indicated some of the power of graph theory in the analysis of large sparse random systems. It is quite probable that other problems concerning singularity, reducibility and other elimination schemes can be tackled in a similar fashion. One thing which the results of this paper do indicate is the need for sparsity pivoting in Gaussian elimination.

The fill-in of a matrix from an average of 3 off-diagonal nonzeros per row to more than 20 is rather drastic and would be reaching the limit of advisability of packed storage were it much larger.

A possible consequence of the results of this paper might be to obtain some estimate of the possible error introduced in the elimination process. The error analysis of Wilkinson deals in terms of upper bounds rather than estimates of errors and his factor of n^{2-2t} [Wilkinson (1965), p. 248] is not very helpful for large sparse systems. It is evident from these calculations of Wilkinson that the determining factor for these bounds is the number of operations being performed on the various positions of the matrix.

If it is possible to calculate a probabilistic estimate of this number of operations, then it is feasible that a more useful estimate of the actual error involved in the solution of sparse matrix equations could be obtained than the crude upper bound given above.

The following theorem is a first step in this direction.

THEOREM 5. *The expected number of operations on an element in row i , column j , $j > i$, is given by the formula*

$$\sum_{k=2}^i (P_k - P_{k-1}) / (1 - P_{k-1})$$

where the P_k are defined in Eq. (3).

Proof.

The probability that an element is nonzero before stage k
 = The probability that it is nonzero before stage $k - 1$
 + The probability that it is zero before stage $k - 1$ and is filled in at stage $k - 1$.

Hence

$$P_k = P_{k-1} + (1 - P_{k-1})X$$

where X is the probability that the element is filled in (or operated on) at stage $k - 1$.

Hence the result.

Acknowledgements. This is part of the work done for the author's D. Phil. Thesis [Duff (1972)]. The author is indebted to the Carnegie Trust for the Universities of Scotland for finance to pursue this research and to his supervisor Dr. J. K. Reid for suggesting improvements to the original manuscript. He also thanks the Instruction and Research Support Group of the Computing Center of the State University of New York at Stony Brook for their assistance in computing P_i for $p = 0.01$.

Computing Laboratory
 19 Parks Road
 Oxford, England

R. K. BRAYTON, F. G. GUSTAVSON & R. A. WILLOUGHBY (1970), "Some results on sparse matrices," *Math. Comp.*, v. 24, pp. 937-954. MR 43 #1396.

I. S. DUFF (1972), *Analysis of Sparse Systems*, D. Phil. Thesis, Oxford University.

P. ERDÖS (1967), *Applications of Probabilistic Methods to Graph Theory*, Seminar on Graph Theory (edited by F. Harary), Holt, Rinehart and Winston, New York, pp. 60-64. MR 36 #68.

- J. H. GRIESMER & R. D. JENKS (1971), *A Set of SCRATCHPAD Examples*, Publication of Thomas J. Watson Research Center, Yorktown Heights, N.Y.
- F. HARARY (1962a), "A graph theoretic approach to matrix inversion by partitioning," *Numer. Math.*, v. 4, pp. 128-135. MR 25 #2977.
- F. HARARY (1962b), "The determinant of the adjacency matrix of a graph," *SIAM Rev.*, v. 4, pp. 202-210. MR 26 #1876.
- F. HARARY, R. Z. NORMAN & D. CARTWRIGHT (1965), *Structural Models: An Introduction to the Theory of Directed Graphs*, Wiley, New York. MR 32 #2345.
- B. R. HEAP (1966), "Random matrices and graphs," *Numer. Math.*, v. 8, pp. 114-122. MR 33 #2565.
- H. Y. HSIEH & M. S. GHAUSI (1971), *Probabilistic Approach to Optimal Pivoting and Prediction of Fill-In for Random Sparse Matrices*, IBM Tech. Report TR 22.1248.
- J. C. OGLIVIE (1968), *The Distribution of Number and Size of Connected Components in Random Graphs of Medium Size*, Proc. IFIP Conf., Edinburgh. Applications 3, Booklet H, pp. 89-92.
- I. PALÁSTI (1966), "On the strong connectedness of directed random graphs," *Studia Sci. Math. Hungar.*, v. 1, pp. 205-214. MR 34 #7403.
- I. PALÁSTI (1970), "On some structured properties of given types of random graphs," *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.*, v. 19, pp. 33-72. (Hungarian) MR 41 #5236.
- S. V. PARTER (1961), "The use of linear graphs in Gauss elimination," *SIAM Rev.*, v. 3, pp. 119-130. MR 26 #908.
- D. J. ROSE & R. A. WILLOUGHBY (Editors) (1972), *Sparse Matrices and their Applications*, Proc. Conf. IBM Research Center, Plenum Press, New York.
- R. P. TEWARSON (1967), "On the product form of inverses of sparse matrices and graph theory," *SIAM Rev.*, v. 9, pp. 91-99. MR 36 #1092.
- J. H. WILKINSON (1965), *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford. MR 32 #1894.
- R. A. WILLOUGHBY (1971), Talk at Computer Sci. Dept., Univ. of Waterloo. (Unpublished.)