

PARAMETER ESTIMATION IN TOPOLOGICAL ANALYSIS OF BINARY TREE STRUCTURES

■ RONALD W. H. VERWER, JAAP VAN PELT and ANDRÉ J. NOEST
Netherlands Institute for Brain Research, Meibergdreef 33,
1105 AZ Amsterdam, The Netherlands

Different types of random binary topological trees (like neuronal processes and rivers) occur with relative frequencies that can be explained in terms of growth models. It will be shown how the model parameter determining the mode of growth can be estimated with the maximum likelihood procedure from observed data. Monte Carlo simulations were used to study the distributional properties of this estimator which appeared to have a negligible bias. It is shown that the minimum chi-square procedure yields an estimate that is very close to the maximum likelihood estimate. Moreover, the goodness-of-fit of the growth model can be inferred directly from the chi-square statistic. To illustrate the procedures we examined axonal trees from the goldfish tectum. A notion of complete partition randomness is presented as an alternative to our growth hypotheses.

Introduction. The probabilities of occurrence of binary topological trees appear to depend strongly on the mode of branching during the development of the trees. Even simple rules about the growth result in characteristic probabilities for the tree types and small changes in the rules may have a profound effect on the probabilities. As the growth of a tree proceeds by branching events a simple rule would be that branches protrude from any segment in the tree. Van Pelt and Verwer (1985) have elaborated this view by adding an extra assumption, i.e. that intermediate and terminal segments have a different probability to form new branches. The ratio of these two probabilities is the only parameter of the resulting family of growth models. The well-known terminal and segmental growth models (cf. Berry *et al.*, 1980; Van Pelt and Verwer, 1983) are particular cases of these models. In the present report we show how the value of the model parameter, for a random sample of observed binary tree-structures, can be estimated by means of a maximum likelihood or a minimum chi-square procedure. Additionally, we discuss the goodness-of-fit testing of the data against the most likely growth model and the evaluation of the distributional properties of the maximum likelihood estimator of the growth parameter. Application of estimation of the growth parameter to biological material is illustrated with axonal terminal arbors from three different regions of the goldfish optic tectum.

Description of the Growth Models. A tree consists of intermediate and terminal segments and let p_i and p_t denote the probability of protrusion of a new segment from an intermediate or a terminal segment, respectively. The

growth parameter (Q), which is defined as: $Q = p_i / (p_i + p_t)$, is assumed to be constant during the branching process. In Fig. 1 it is shown how a tree of degree 3 might transform into a tree of degree 4. The number of possible topological types increases rapidly (Harding, 1971) with the topological size (i.e. number of terminal segments or degree). The number of partitions (a partition is the division of terminal segments over the two subtrees of a branching point) of the same degree (n) increases only with $n/2$ and it appears to be much more efficient to calculate the probabilities of occurrence of partitions (Van Pelt and Verwer, 1983). For any given value of Q ($0 \leq Q \leq 1$) the relative frequencies of occurrence of the two possible partitions of degree 4 are found to be

$$p(1, 3; Q) = 2p_i + 2p_t = \frac{2}{3-Q} \quad \text{and} \quad p(2, 2; Q) = p_t = \frac{1-Q}{3-Q}.$$

The relation between the probability that n terminal segments are partitioned over the two subtrees of a bifurcation point as $(r, n-r)$ and the parameter Q is given by

$$p(r, n-r; Q) = \prod_{i=1}^{r-1} (i-Q) \cdot \prod_{i=n-r}^{n-1} (i-Q)^{-1} \cdot \left[1 + Q \left\{ \frac{n(n-1)}{2r(n-r)} - 2 \right\} \right] \cdot \binom{n-1}{r-1} \cdot 2^{1-\delta_{r,n-r}} \quad (1)$$

where $n \geq 4$ and $1 \leq r \leq n/2$ (Van Pelt and Verwer, 1985) and $\delta = 1$ if $r \neq n-r$ and $\delta = 0$ if $r = n-r$. By definition the term

$$\prod_{i=1}^{r-1} (i-Q) = 1 \text{ if } r = 1$$

and therefore,

$$p(1, n-1; Q) = \frac{2 + (n-4)Q}{n-1-Q}. \quad (2)$$

Estimation of the Parameter Q. The procedure to find a maximum likelihood estimate (MLE) of the parameter Q may be applied if certain conditions concerning the relations in this section are valid. This will be proved in the Appendix. For general information concerning maximum likelihood estimators the reader is referred to Cox and Hinkley (1974) and Wilks (1962). Under the hypothesis that the tree structures in the observed sample have

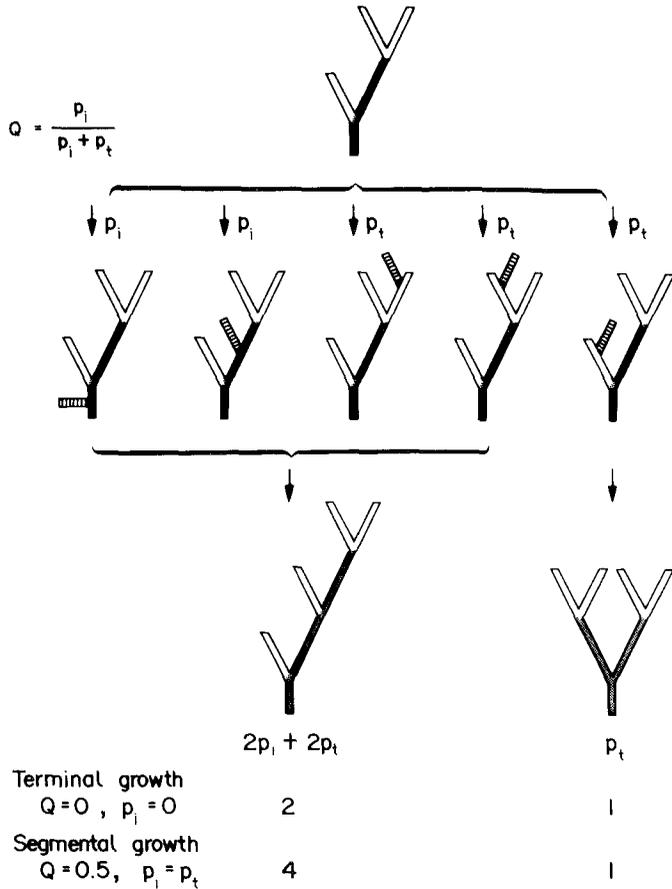


Figure 1. Explanation of the growth hypotheses, with parameter Q , based on the difference in branching probabilities of terminal and intermediate segments. Open bars are terminal segments, the heavily dotted bars are intermediate segments and the striped bar is the new segment that protrudes from an already existing segment due to a branching event. In a binary tree of degree 3 there are five possible branching events, each giving rise to either of two possible tree types of degree 4. Here p_i denotes the probability that a new segment emanates from an intermediate segment and p_t is the probability that the new segment originates from a terminal segment. The relative frequencies with which the two tree types of degree 4 occur as a result of different values of p_i and p_t is indicated.

grown according to one of the present growth models the partition of terminal segments over the two subtrees of any branching point is independent of all the others. Therefore all partitions present in the observed tree structures can be used for the estimation of Q .

Suppose we have a random sample of bifurcation points with partitions $(r_1, n_1 - r_1), (r_2, n_2 - r_2), \dots, (r_m, n_m - r_m)$, each partition being distributed as defined by equation (1), then the likelihood element

$$L_m(Q) = \prod_{j=1}^m p(r_j, n_j - r_j; Q) \tag{3}$$

describes the probability that this particular sample is observed. If certain regularity conditions apply, i.e. if summation and differentiation with respect to Q may be interchanged (see Appendix), we may write

$$\frac{d}{dQ} \ln\{L_m(Q)\} = \sum_{j=1}^m \frac{d}{dQ} \ln p(r_j, n_j - r_j; Q). \tag{4}$$

Then solving

$$\sum_{j=1}^m \frac{d}{dQ} \ln p(r_j, n_j - r_j; Q_0) = 0$$

for Q should yield an efficient estimator (\hat{Q}) of the true value Q_0 . In theory \hat{Q} should have an asymptotically normal distribution with mean Q_0 and variance determined by the expectation of the second derivative of $\ln p(r, n - r; Q)$ (Cox and Hinkley, 1974; Wilks, 1962). However, the restricted interval of Q values and the discrete nature of \hat{Q} for finite sample size implies that the distributional properties of \hat{Q} can only be evaluated by Monte Carlo simulations (see below). If we take the logarithm of equation (1) and differentiate with respect to Q , we find

$$\frac{d}{dQ} \ln p(r_j, n_j - r_j; Q) = - \sum_{i=1}^{r_j-1} (i - Q)^{-1} + \sum_{i=n_j-r_j}^{n_j-1} (i - Q)^{-1} + \frac{C_j}{1 + C_j Q} \tag{5}$$

in which

$$\sum_{i=1}^{r_j-1} (i - Q)^{-1} = 0 \text{ if } r_j = 1 \text{ and } C_j = \left\{ \frac{n_j(n_j - 1)}{2r_j(n_j - r_j)} - 2 \right\}.$$

The parameter Q cannot be made explicit from this expression, therefore we have to rely on an iterative evaluation of the derivative of the log-likelihood function, which at Q_0 equals zero,

$$\frac{d}{dQ} \ln\{L_m(Q_0)\} = \sum_{j=1}^m \left\{ - \sum_{i=1}^{r_j-1} (i - Q_0)^{-1} + \sum_{i=n_j-r_j}^{n_j-1} (i - Q_0)^{-1} + \frac{C_j}{1 + C_j Q_0} \right\} = 0. \tag{6}$$

In Table I the partition frequencies of a sample of retinotectal axon terminal arbors from the central optic tectum of the goldfish are shown (cf. Stuermer, 1984; Verwer *et al.*, 1987). The log-likelihood function of this sample attained its maximum at $\hat{Q}_{MLE} = 0.2436$ (Fig. 2).

An alternative procedure to MLE, which is preferred by some authors (Berkson, 1980), is called the minimum chi-square (MCS) method (cf. Berkson, 1980; Conover, 1980, p. 195; Cox and Hinkley, 1974; Cramér, 1946, pp. 424–441). For this procedure the log-likelihood chi-square statistic (T) must be evaluated iteratively and the value of Q for which T attains its absolute minimum is selected as an estimate for Q_0 . It has the advantage that for the minimal T value a level of significance to assess the goodness-of-fit is directly available. However, since we have estimated the parameter from the same data set the test becomes conservative. This conservatism can be corrected by subtracting an extra degree of freedom (Conover, 1980, p. 194; Cramér, 1946, p. 425). A drawback of the chi-square test is that it requires a minimal number of observations in each class, so that usually lumping of classes is necessary, which is arbitrary and may influence the estimation. For this purpose we lumped the observed partitions of the retinotectal axonal trees according to the degree of the smallest (left) subtree (cf. Table I). Additionally the classes for which the degree of the left subtree was ≥ 4 were taken together to obtain a final composite partition distribution. To construct the theoretical composite partition distribution we computed all the necessary individual partition probabilities for each iterated value of Q . Note that the partition probabilities are conditional on the degree of the partitions. This means for instance that, since in the sample of retinotectal axon terminal arbors one partition of degree 32 (i.e. 32(3 29) see Table I) was observed, the probabilities of all possible partitions of degree 32 must be calculated and weighted by $1/w$, where w is the total number of partitions present in the sample. The theoretical composite partition distribution is then obtained by lumping similarly as was done with the observations. In Fig. 3 the behaviour of T with respect to Q is shown. Moreover, it is illustrated how the composite partition distribution of the sample compared with the theoretical composite partition distributions for the parameter values $Q=0, 0.24$ and 0.5 . It appears that for the $\hat{Q}_{\text{MCS}}=0.2428$, T attains its minimal value being 1.24946, which shows a rather good agreement with the Q -value estimated by the maximum likelihood method if we consider the lumping of classes. It was already noted by Berkson (1980) that likelihood minimum chi-square yields the same estimate as the maximum likelihood estimation. The corresponding level of significance for MCS was 0.54. The maximum likelihood procedure does not provide a level of significance indicating the similarity between the theoretical partition distribution with $Q = \hat{Q}_{\text{MLE}}$ and the sample partition distribution.

The Distributional Properties of \hat{Q} . Suppose that we take a finite sample of n th-degree partitions from a population of partitions whose probability distribution is defined by Q_0 . The resulting frequency distribution will deviate somewhat from the original probability distribution and the corresponding

TABLE I
The Partition Scores of the Sample of Retinotectal Axon
Terminals from the Goldfish Central Optic Tectum

		Degree of left subtree (r)									
		1	2	3	4	5	6	7	8	9	10
Degree of right subtree	1										
	2		5								
($n-r$)	3	4	2								
	4	6	1								
	5	3	2	1	1						
	6	3		1							
	7	2	1	1							
	8	1	1				1				
	9	1	3		1						
	10	1									
	11	1	1	1	1						
	12										
	13	1	1								
	14	2									
	15	1	1								1
	16			1							
	17							1			
	18										
	19										1
	20										
	.										
	.										
	29			1							
First lumping	($r, n-r$)	26	18	6	4	—	1	1	—	—	2
Second lumping		26	18	6	8						

maximum likelihood estimate may be found on either side of Q_0 . In order to establish how reliable maximum likelihood estimates of Q_0 are we simulated samples of partitions of a particular degree for particular values of Q_0 . It appeared that for values of Q_0 between 0 and 0.5 the resulting distribution of simulated \hat{Q} values is symmetric, whereas if Q_0 is increasing the simulated distributions become more and more skewed. This occurs because no populations exist for which Q_0 exceeds 1. Even for sample sizes as small as 10 the mean of the simulated \hat{Q} -values was quite close to the Q_0 (Table II). Compared with the standard deviations [$\sigma(\hat{Q})$] the absolute values of the biases $|E(\hat{Q}) - Q_0|$ found in the simulations were negligible [i.e. $|E(\hat{Q}) - Q_0| / \sigma(\hat{Q}) < 0.1$, cf. Cochran (1977, (pp. 12–16))]. Only once a value of the ratio $|bias|/\sigma$ equalled 0.11, which we still considered as negligible. The standard

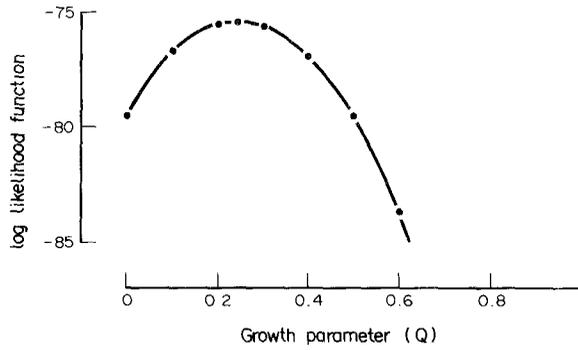


Figure 2. Illustration of the relation of the log-likelihood function and the parameter Q for a sample of six retinotectal axon terminal arborizations (see text and Table I). The maximum likelihood estimate $\hat{Q}_{MLE} = 0.2436$.

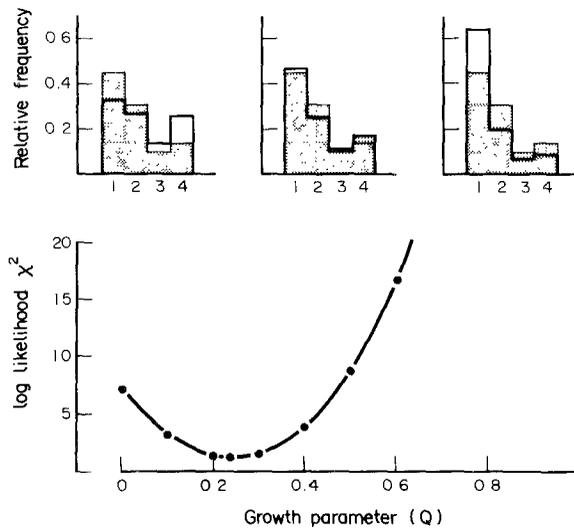


Figure 3. Minimum chi-square estimation of retinotectal axon terminal arbors. The observed partitions were lumped as indicated in Table I and all classes over 4 were added to class 4. The resulting distribution is presented as the dotted area in the upper three figures. The curve shows the relation between the log-likelihood chi-square statistic (T) and the value of Q . The upper three figures illustrate the comparison between the observed and the theoretical distributions at the values of $Q = 0, 0.24$ and 0.5 , respectively. The theoretical distribution is indicated by the thick line.

deviations $[\sigma(\hat{Q})]$ appeared to be related to both the degree and the number of partitions in each sample (Table III), that is standard deviations decreased when either the number of observations or the degree of the partitions was increased. In both Tables II and III the corresponding Monte Carlo confidence

TABLE II
 Monte Carlo Simulations of 10 Partitions of Degree 10
 for Different Values of Q_0

(A)	Q_0	Percentage of trimming	Nominal CI	Critical values	95% interval for CI
	0	8.5	0.82	-0.2495; 0.2473	0.79-0.84
	0.1	3	0.93	-0.2342; 0.4700	0.91-0.94
	0.2	1	0.95	-0.1933; 0.6218	0.94-0.96
	0.3	2.7	0.95	-0.0761; 0.6986	0.93-0.96
	0.4	—	0.95	0.0170; 0.7705	0.94-0.96
	0.5	—	0.95	0.1177; 0.8781	0.93-0.96
	0.6	—	0.87	0.3048; 0.8912	0.85-0.89
	0.7	—	0.90	0.4206; 1.0000	0.88-0.92
	0.8	—	0.70	0.6570; 1.0000	0.67-0.73
	0.9	—	—	—	—

(B)	Q_0	Bias mean(\hat{Q}) - Q_0	s.d. $\sigma(\hat{Q})$	bias /s.d.
	0	-0.0146	0.13	0.11
	0.1	0.0049	0.17	0.03
	0.2	-0.0045	0.19	0.02
	0.3	-0.0003	0.20	0.002
	0.4	0.0079	0.20	0.04
	0.5	-0.0077	0.19	0.04
	0.6	0.0045	0.19	0.02
	0.7	-0.0014	0.16	0.01
	0.8	0.0002	0.14	0.001
	0.9	0.0003	0.11	0.003

For each Q_0 1000 simulations were performed. Note that the nominal confidence interval (CI) is determined by the amount of trimming and the discrete character of \hat{Q} . For $Q_0=0.9$ the nominal confidence interval is too small to be relevant.

intervals (Buckland, 1984) are shown as well. Indicated for each set of simulations is the range in which the actual confidence level will lie with approximately 95% confidence. In the Appendix it is shown that although Q is only defined on the interval $[0, 1]$ negative values of Q correspond to regular partition probability distributions if $Q > -2/(n-4)$. This means that if the estimated \hat{Q} is larger than $-2/(n-4)$ the frequency distribution of observed partitions agrees with a regular, though not necessarily interpretable, probability distribution. On the other hand there are also perfectly legitimate samples of partitions for which no reliable \hat{Q} -value can be estimated. If such samples occurred we trimmed the distribution equally on both sides such that the distribution did not include any values of $\hat{Q} < -2/(n-4)$ and then calculated the so-called trimmed mean (cf. Lehmann, 1983). Thus, if

TABLE III

Monte Carlo Simulations of Samples of Hypothetical Partitions. The Influence of the Degree and the Number of Partitions on the Dispersion (Standard Deviation and 95% Confidence Interval) of the Maximum Likelihood Estimate (\hat{Q}) is Shown

Degree	Number of observations	Percentage of trimming	Critical values of 95% nominal confidence interval	Bias (mean(\hat{Q}) - Q_0)	s.d. $\sigma(\hat{Q})$	bias /s.d.
10	10	1	-0.2495; 0.6102	-0.0142	0.19	0.07
10	25	—	-0.0493; 0.4532	0.0028	0.13	0.02
10	50	—	0.0252; 0.3793	0.0032	0.09	0.04
10	100	—	0.0766; 0.3220	-0.0014	0.06	0.02
5	100	—	-0.0541; 0.4286	-0.0051	0.12	0.04
15	100	—	0.1041; 0.3068	0.0031	0.05	0.06
20	100	—	0.1190; 0.2937	0.0027	0.05	0.05
40	100	—	0.1351; 0.2757	0.0033	0.04	0.08

For each combination of degree and number of observations 1000 simulations were performed with $Q_0=0.2$.

$\hat{Q} < -2/(n-4)$ for x out of y simulated \hat{Q} -values, then from both the left- and the right-side of the distribution of \hat{Q} the x most extreme values were deleted. The range of the remaining \hat{Q} -values provides a $100(y-2x)/y$ per cent confidence interval, whose own approximate 95% interval can be calculated according to Buckland (1984). It must be kept in mind that \hat{Q} can only assume a finite number of values for finite sample size. Therefore, if Q_0 is larger than about 0.5 and the sample size is small to moderate (< 50) the discrete character of the value domain of \hat{Q} becomes more pronounced at the right tail of the distribution of simulated \hat{Q} values, which must be accounted for when establishing a confidence interval.

The sample of central retinotectal axon terminals from which the partitions of Table I were observed consisted of six trees. The degrees of these trees were 12, 15, 18, 24, 25 and 32. Using the Monte Carlo procedure we have simulated 100 samples of partitions for these trees under the hypothesis that they were formed by a growth model for which $Q_0=0.2436$ (cf. Verwer *et al.*, 1987). The resulting mean value of \hat{Q} was 0.2422 (Fig. 4). The critical values of the approximate 95% confidence interval are 0.0882 and 0.4052, whose own 95% confidence interval is between 93.65 and 96.35% (cf. Buckland, 1984).

Test for Complete Partition Randomness. Along the line of reasoning in Diggle (1983) we introduce here a concept of complete randomness of the distribution of partitions of terminal segments over the subtrees of a bifurcation point. For our purpose let complete partition randomness be represented by

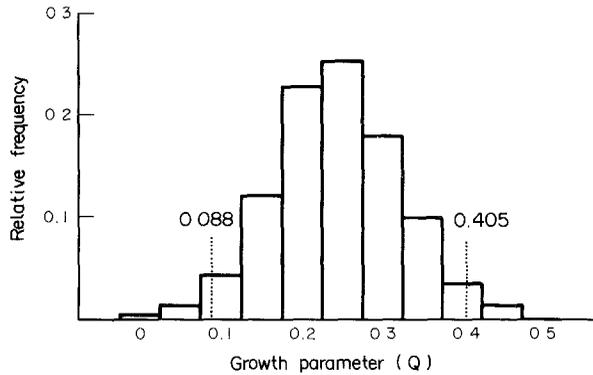


Figure 4. The distribution of \hat{Q} values of 1000 simulated samples of six trees, with 12, 15, 18, 24, 25 and 32 terminal segments, respectively, for which $Q_0=0.2436$. The dotted lines indicate the critical values of the nominal 95% confidence interval.

the binomial distribution. There seems to be no reason to proceed with topological analysis of binary trees in terms of our present growth models if the distribution of the observed partitions corresponds to a truncated binomial distribution with $P=1/2$. The binomial distribution is truncated because the partition $(0, n)$ is non-existent. Also notice that any value of P other than $1/2$ would be meaningless, since we cannot label the two subtrees of a bifurcation point. In terms of the branching process complete partition randomness can best be conceived of as that the addition of a new segment occurs with equal probability at either subtree of a particular bifurcation point, irrespective of how many terminal segments each of the two subtrees possess at a certain moment. By contrast, we would expect a growth model based on Q to favour the subtree with the largest degree (Van Pelt and Verwer, 1985). Other conceivable growth models might for instance favour the smallest subtree (cf. Van Pelt and Verwer, 1986) or one of the subtrees because of its particular shape. Note, however, that shape is not a label because it is subject to changes due to the branching process. The probability that a partition $(r, n-r)$ occurs in the case of complete partition randomness is

$$p(r, n-r; B) = \binom{n}{r} \frac{2^{1-\delta_{r,n-r}}}{(2^n-2)} = \binom{n}{r} \frac{2^{-\delta_{r,n-r}}}{(2^{n-1}-1)} \tag{7}$$

since

$$\sum_{r=1}^{n-1} \binom{n}{r} = 2^n - 2.$$

Here $\delta = 1$ if $r \neq n-r$ and $\delta = 0$ if $r = n-r$. In practice we first test the partition data with the chi-square test for goodness-of-fit against complete randomness,

before we continue to estimate the parameter of a growth model. The corresponding chi-square statistic for the sample of goldfish central tectum axon terminals was $T = 38.8$, which for 3-degrees-of-freedom leads to rejection of complete partition randomness.

Results and Discussion. For this study we used the *camera lucida* drawings of retinotectal axon terminal arbors from the three different regions of the tectum of the adult goldfish (*Carassius auratus*) displayed in Figs 13, 14 and 15 in Stuermer (1984). We refer to this article for details concerning the maintenance of the animals and the preparation for light microscopical visualization of the axon terminal arborizations. The goldfish tectum is a flat brain region that is curved into a dome-like shape, whose top is called the center. A relatively large area around the top is called the central tectum. A narrow annular strip at the base of the dome is called the peripheral region, and a more or less equally wide annular strip located between the central and the peripheral tectum is the intermediate tectum. Ganglion cells in the retina send their axons to the tectum, where they terminate in a way that is homotopic with the location of the ganglion cell bodies. Thus each tiny region in the retina is connected in an orderly way to a corresponding tiny region in the tectum. We have analyzed nine observed retinotectal arbors from the peripheral, six from the intermediate and six from the central tectum under the hypothesis that they were formed according to our growth models (cf. Van Pelt and Verwer, 1985). First, we checked whether the trees could have been formed by a completely random process. All three groups strongly rejected complete partition randomness, as can be seen from the chi-square statistic that was 38.8 for central axons, 14.3 for intermediate axons and 143.2 for peripheral axons (3-degrees-of-freedom in each case). The results of the analysis concerning the growth models are summarized in Table IV. It appeared that the maximum likelihood estimate of the growth parameter was 0.4170 for the peripheral arbors, 0.0888 for the intermediate arbors and 0.2436 for the central arbors. The estimations obtained with the minimum chi-square method (cf. Table IV) are in close agreement considering the necessary grouping of the partition data. The minimal value of the chi-square statistic was used to assess the goodness-of-fit between the observed and the expected partition distributions. One extra degree-of-freedom was subtracted, to correct for the estimation of Q (Conover, 1980, p. 194), before determining the corresponding level of significance. Direct comparison of the partition distributions showed that the peripheral arbors were significantly different from the intermediate arbors, whereas neither of them differed significantly from the central arbors (Verwer *et al.*, 1987). Without further knowledge our interpretation would be that the peripheral and intermediate axons grow in different ways and that the central axonal arbors have some kind of intermediate branching mode. However, evidence

TABLE IV

Distributions of Partitions from Observed Axonal Trees from Three Areas of the Goldfish Optic Tectum and the Estimated Values of the Growth Parameter

(A) Partition Distributions after Lumping:

Degree of left subtree	Peripheral tectum		Intermediate tectum		Central tectum	
	Number of obs.	Exp. freq.	Number of obs.	Exp. freq.	Number of obs.	Exp. freq.
1	89	86.9	15	16.4	26	27.4
2	27	31.9	14	12.3	18	14.4
3	12	12.2	9	6.0	6	6.4
≥ 4	23	20.1	5	8.3	8	9.8

(B) Estimated Values of Growth Parameter:

	Peripheral tectum	Intermediate tectum	Central tectum
\hat{Q}_{MLE}	0.4170	0.0888	0.2436
\hat{Q}_{MCS}	0.4150	0.0728	0.2428
Level of significance	0.54	0.20	0.54

\hat{Q}_{MLE} : the maximum likelihood estimate of Q , \hat{Q}_{MCS} : the minimum chi-square estimate of Q . The expected frequencies of the partition distributions for each group are based on their respective minimum chi-square estimates. The corresponding chi-square statistics yield the presented levels of significance, indicating the goodness-of-fit, for 2-degrees-of-freedom.

exists that throughout a large part of the lifetime of the goldfish both the retina and the tectum continue to grow in size, though each in a different fashion. In order to maintain the homotopy between the retinal ganglion cells and their axon terminal arbors in the tectum, the axonal arbors appear to shift to adjust their position [see Easter and Stuermer (1984) and refs therein]. Essential features of this phenomenon are that new axons enter the tectum peripherally and the already existing peripheral axons move to more intermediate positions. The axons in the intermediate and central tectum also move to maintain the existing relative positions. The values we obtain for the growth parameters for the three groups are, therefore, interpreted as follows: Growing retinotectal axons form new branches with more or less equal probability at both intermediate and terminal segments (i.e. $p_i \approx p_t$, see definition of Q) and this applies to the peripheral axonal terminal trees. The process of shifting necessarily involves withdrawal of branches and our models do not account for such developmental actions. However, the growth parameter appears to be sensitive to such changes in the topological properties. Thus the withdrawal of branches is more extensive than the outgrowth of new segments for the axonal

arbors found in the intermediate tectum resulting in a lower value of \hat{Q} . The reason why withdrawal is accompanied by a decrease in the value of \hat{Q} is at present unknown. For instance, a random withdrawal would have a negligible effect on the value of Q (see Van Pelt *et al.*, 1986). The more centrally located axonal trees are in a kind of equilibrium between outgrowth and withdrawal as indicated by the Q value which is located halfway between those of the intermediate and the peripheral axons. This interpretation fits well with the existing ideas about shifting axons (Easter and Stuermer, 1984; Stuermer, 1984) and is further supported by direct comparisons between the topological properties of the trees in the three tectal areas (cf. Verwer and Van Pelt, 1986; Verwer *et al.*, 1987).

We have shown that the parameter of the family of growth models (cf. Van Pelt and Verwer, 1985) can be estimated for a sample of observed binary trees by either a maximum likelihood or a minimum chi-square procedure. The maximum likelihood estimator has a negligible bias as was determined with Monte Carlo simulations. It is expected that this also holds for the minimum chi-square estimator, because it is essentially the same as the maximum likelihood estimator (cf. Berkson, 1980; Cramér, 1946). The maximum likelihood estimate uses more information from the data than the minimum chi-square estimate which practically always involves some kind of grouping of the data. On the other hand the goodness-of-fit of the model can be determined directly with the minimum chi-square method. The resulting estimate can either be used to explain the mode of growth of a group of trees (i.e. the axonal trees in the peripheral tectum) or describe changes in topological properties (i.e. for intermediate and central axonal trees) when effects are involved that cannot be explained by our growth models. Further we have defined a notion of complete partition randomness that may be used to establish whether topological analysis in terms of growth models is useful.

The authors are grateful to H. Stoffels and G. Van der Meulen for their help with the figures.

APPENDIX

Proof of Regularity of the Growth Probability Distribution. In order that the maximum likelihood estimation may be applied the growth probability distribution should be regular in its first and second derivative with respect to Q (Wilks, 1962). In the present case regularity in the first derivative is defined as

$$E\left\{\frac{d}{dQ} \ln p(r, n-r; Q)\right\} = \frac{d}{dQ} \sum_{r=1}^{\lfloor n/2 \rfloor} p(r, n-r; Q) = 0 \quad (\text{A1})$$

and regularity in the second derivative is defined as

$$\begin{aligned}
 & E\left\{\frac{d^2}{dQ^2} \ln p(r, n-r; Q)\right\} + E\left\{\left[\frac{d}{dQ} \ln p(r, n-r; Q)\right]^2\right\} \\
 &= \frac{d^2}{dQ^2} \sum_{r=1}^{[n/2]} p(r, n-r; Q) = 0.
 \end{aligned} \tag{A2}$$

Here $[n/2]$ is the largest integer in $n/2$. And further it is clear that

$$E\left\{\frac{d}{dQ} \ln p(r, n-r; Q)\right\} = \sum_{r=1}^{[n/2]} \frac{d}{dQ} \ln\{p(r, n-r; Q)\} \cdot p(r, n-r; Q)$$

and

$$\begin{aligned}
 & E\left\{\frac{d^2}{dQ^2} \ln p(r, n-r; Q)\right\} + E\left\{\left[\frac{d}{dQ} \ln p(r, n-r; Q)\right]^2\right\} \\
 &= \sum_{r=1}^{[n/2]} \left\{\frac{d^2}{dQ^2} \ln p(r, n-r; Q) + \left[\frac{d}{dQ} \ln p(r, n-r; Q)\right]^2\right\} p(r, n-r; Q).
 \end{aligned}$$

First we note that for finite n the partition probability $p(r, n-r; Q)$ consists of a finite product of functions of Q each of which is infinitely many times continuously differentiable with respect to Q . Therefore, this also holds for $p(r, n-r; Q)$. Further it is easily verified that $p(r, n-r; Q)$ is always positive if

$$\frac{-2}{n-4} < Q < 1, \quad 1 \leq r \leq \frac{1}{2}n \quad \text{and} \quad n \geq 4.$$

If $p(r, n-r; Q)$ is positive and once continuously differentiable, then

$$\frac{d}{dQ} p(r, n-r; Q) = p(r, n-r; Q) \cdot \left\{\frac{d}{dQ} \ln p(r, n-r; Q)\right\} \tag{A3}$$

and therefore,

$$E\left\{\frac{d}{dQ} \ln p(r, n-r; Q)\right\} = \sum_{r=1}^{[n/2]} \frac{d}{dQ} p(r, n-r; Q) = \frac{d}{dQ} \sum_{r=1}^{[n/2]} p(r, n-r; Q) = 0$$

which proves assertion (A1). If moreover $p(r, n-r; Q)$ is twice continuously differentiable with respect to Q the proof of assertion (A2) follows immediately by differentiating both sides of equation (A3) with respect to Q :

$$\begin{aligned}
 \frac{d^2}{dQ^2} p(r, n-r; Q) &= \left\{\frac{d}{dQ} \left[\frac{d}{dQ} \ln p(r, n-r; Q)\right]\right\} p(r, n-r; Q) \\
 &\quad + \left\{\frac{d}{dQ} \ln p(r, n-r; Q)\right\} \left\{\frac{d}{dQ} p(r, n-r; Q)\right\} \\
 &= \left\{\frac{d^2}{dQ^2} \ln p(r, n-r; Q) + \left[\frac{d}{dQ} \ln p(r, n-r; Q)\right]^2\right\} p(r, n-r; Q),
 \end{aligned}$$

which results in

$$\begin{aligned} E\left\{\frac{d^2}{dQ^2} \ln p(r, n-r; Q)\right\} + E\left\{\left[\frac{d}{dQ} \ln p(r, n-r; Q)\right]^2\right\} \\ = \sum_{r=1}^{\lfloor n/2 \rfloor} \frac{d^2}{dQ^2} p(r, n-r; Q) \\ = \frac{d^2}{dQ^2} \sum_{r=1}^{\lfloor n/2 \rfloor} p(r, n-r; Q) = 0 \end{aligned}$$

as required for assertion (A2). Since $E\{d/dQ \ln p(r, n-r; Q)\}$ equals zero for each finite n and the maximum likelihood estimation can be applied without restriction to conditional probabilities (cf. Cox and Hinkley, 1974) we obtain an efficient estimate (\hat{Q}) of the true value Q_0 for any combination of trees or subtrees by solving equation (6).

The Range of Values of the Parameter Q. Since the parameter is defined as $Q = p_i/(p_i + p_i)$, with $0 \leq p_i \leq 1/n$ and $0 \leq p_i \leq 1/(n-1)$, the formal interval for Q is $[0, 1]$. However, the probabilities $p(r, n-r; Q)$ still constitute a regular probability distribution for $-2/(n-4) < Q < 0$, although the interpretation in terms of p_i and p_i is impossible. It is obvious that as n increases $-2/(n-4)$ approaches 0. For negative values of Q the corresponding probability distributions describe tree populations in which symmetrical partitions occur with higher frequency than we get for terminal growth ($Q=0$). Thus if n_{\max} is the degree of the largest partition present in a sample then we will only use estimated \hat{Q} -values from the maximum likelihood or minimum chi-square procedure and from Monte Carlo simulations if $-2/\{n_{\max}-4\} < \hat{Q} < 1$. It may be remarked that for $Q=1$, $p(1, n-1; Q=1)=1$ and for each $r > 1$, $p(r, n-r; Q=1)=0$ and that the maximum likelihood procedure is not allowed, but it is also actually unnecessary. At the other extreme, that is if $Q = -2/(n-4)$, $p(1, n-1; Q = -2/(n-4))=0$ while for each $r > 1$, $p(r, n-r; Q = -2/(n-4))$ being positive and $\sum_{r=2}^{\lfloor n/2 \rfloor} p(r, n-r; Q = -2/(n-4))=1$. For $Q < -2/(n-4)$ negative probabilities occur. If a sample appears to have more symmetrical partitions than can be accounted for by distributions with $Q > -2/\{n_{\max}-4\}$ one might use growth models described by Van Pelt and Verwer (1986) in which an extra parameter, incorporating the order of the segments in the branching probabilities, is included. In these growth models also populations containing only highly symmetrical trees can be described. But parameter estimation becomes more complicated than in the situation reported here, because the branching probabilities can only be expressed in terms of recurrent relations. Based on the results in the present report it can be shown that it is also permissible to apply maximum likelihood estimation to the two parameter growth models.

LITERATURE

- Berkson, J. 1980. "Minimum Chi-Square, Not Maximum Likelihood!" *Ann. Statist.* **12**, 457-487.
- Berry, M., P. McConnell and J. Sievers. 1980. "Dendritic Growth and the Control of Neuronal Form." In *Current Topics in Developmental Biology*, R. K. Hunt (Ed.), Vol. 15, pp. 67-101. New York: Academic Press.
- Buckland, S. T. 1984. "Monte Carlo Confidence Intervals." *Biometrics* **40**, 811-817.
- Cochran, W. G. 1977. *Sampling Techniques*, 3rd Edn. New York: J. Wiley.
- Conover, W. J. 1980. *Practical Nonparametric Statistics*, 2nd Edn. New York: J. Wiley.
- Cox, D. R. and D. V. Hinkley. 1974. *Theoretical Statistics*. London: Chapman and Hall.
- Cramér, H. 1946. *Mathematical Methods of Statistics*. Princeton: Princeton University Press.
- Diggle, P. J. 1983. *Statistical Analysis of Spatial Point Patterns*. London: Academic Press.

- Easter, S. S., Jr. and C. A. O. Stuermer. 1984. "An Evaluation of the Hypothesis of Shifting Terminals in Goldfish Optic Tectum". *J. Neurosci.* **4**, 1052–1063.
- Harding, E. F. 1971. "The Probabilities of Rooted Tree-shapes Generated by Random Bifurcation." *Adv. appl. Prob.* **3**, 44–77.
- Lehmann, E. L. 1983. *Theory of Point Estimation*. New York: J. Wiley.
- Stuermer, C. A. O. 1984. "Rules for Retinotectal Terminal Arborizations in the Goldfish Optic Tectum." *J. comp. Neurol.* **229**, 214–232.
- Van Pelt, J. and R. W. H. Verwer. 1983. "The Exact Probabilities of Branching Patterns Under Terminal and Segmental Growth Hypotheses." *Bull. math. Biol.* **45**, 269–285.
- and ——— 1985. "Growth Models (Including Terminal and Segmental Branching) for Topological Binary Trees." *Bull. math. Biol.* **47**, 323–336.
- and ——— 1986. "Topological Properties of Binary Trees Grown with Order-dependent Branching Probabilities." *Bull. math. Biol.* **48**, 197–211.
- , ——— and H. B. M. Uylings. 1986. "Application of Growth Models to the Topology of Neuronal Branching Patterns." *J. Neurosci. Meth.* **18**, 153–165.
- Verwer, R. W. H., J. Van Pelt and H. B.M. Uylings. 1985. "A Simple Statistical Test for the Vertex Ratio Using Monte Carlo Simulation." *J. Neurosci. Meth.* **14**, 137–142.
- and ———. 1986. "Descriptive and Comparative Analysis of Geometrical Properties of Neuronal Tree Structures." *J. Neurosci. Meth.* **18**, 179–206.
- , ——— and C. A. O. Stuermer. 1987. "Topological Analysis of Goldfish Retinotectal Axon Terminals." (submitted).
- Wilks, S. S. 1962. *Mathematical Statistics*. New York: J. Wiley.

RECEIVED 13-8-86

REVISED 29-12-86