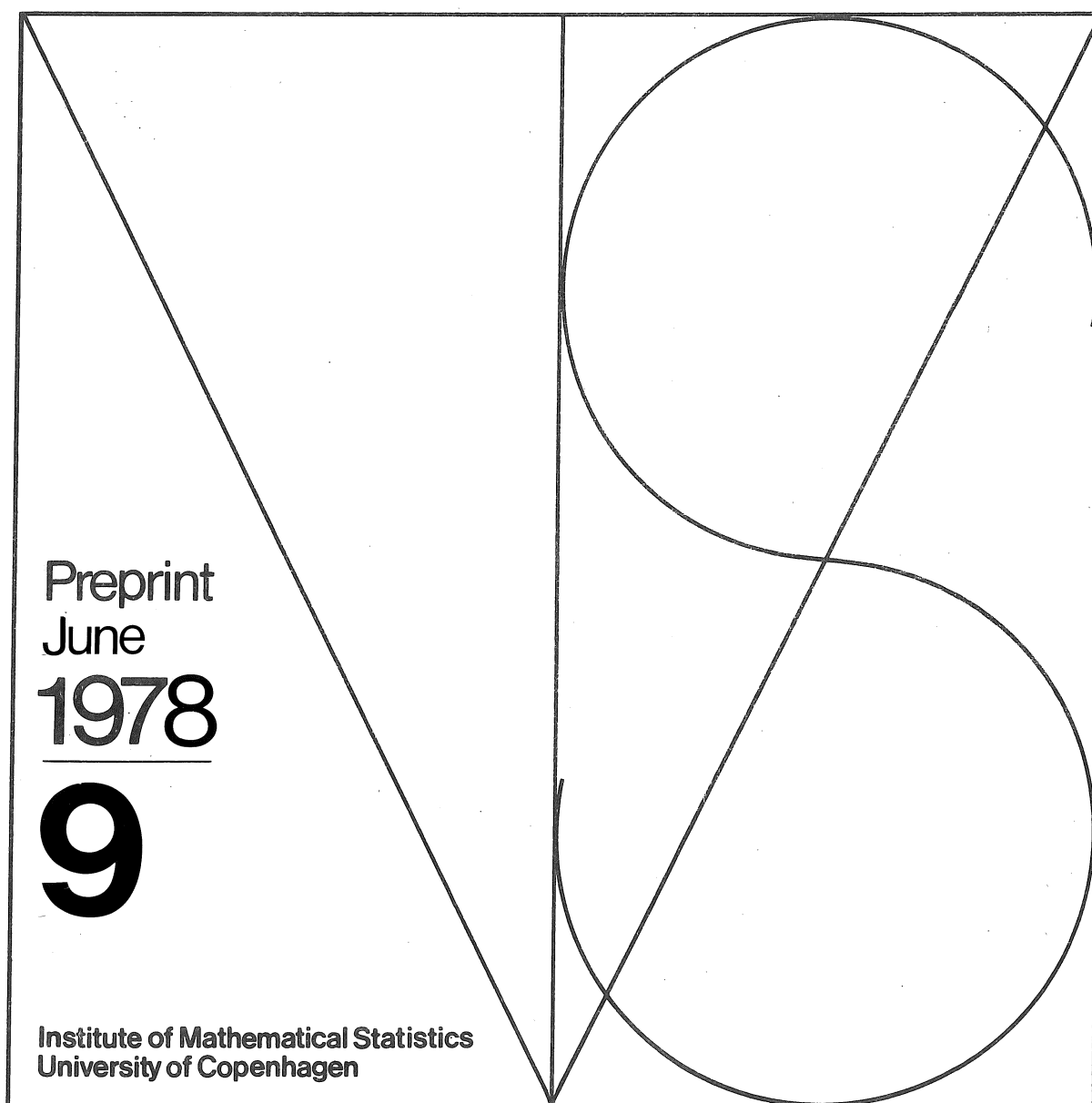


S. L. Lauritzen

T. P. Speed

K. Vijayan

Decomposable Graphs and Hypergraphs



S.L. Lauritzen, T.P. Speed* and K. Vijayan*

DECOMPOSABLE GRAPHS AND HYPERGRAPHS

Preprint 1978 No. 9

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

June 1978

H. C. Ø. - tryk
københavn

Abstract

We define and investigate the notion of a decomposable hypergraph, showing that such a hypergraph always is conformal, i.e. can be viewed as the class of maximal cliques of a graph. We further show that the clique hypergraph of a graph is decomposable if and only if the graph is triangulated and characterise such graphs in terms of a combinatorial identity.

1. INTRODUCTION

There are a number of areas in mathematics in which one needs to consider the combinatorial properties of a non-void class \mathcal{C} of pairwise incomparable subsets of a finite set C . In the theory of games, Vorob'ev [14], the subsets are coalitions; in a measure theoretic problem considered by Kellerer [9] and also Vorob'ev [12], the subsets correspond to prescribed marginals; in the general theory of contingency tables, discussed by Haberman [5], the subsets define the permissible interactions, in graph theory, the subsets are the maximal cliques of a graph with vertex set C ; whilst in the theory of Markov fields over graphs, see Suomela [11] and Vorob'ev [13] the subsets also correspond to maximal cliques of a graph. Certain problems of interest in these fields have led to the definition of a family of such classes \mathcal{C} which, following Haberman, we call decomposable classes, and the main aim of this paper is to unify and extend the combinatorial results known concerning these classes. Further discussion of the particular problems can be found in the references given, or in [10], where the main results of this paper were stated within a broader context.

A pair (C, \mathcal{C}) of the type described above is a hypergraph in the sense of Berge [2], as long as the union of all the members of \mathcal{C} coincides with C , indeed a hypergraph in which no edge is contained in any other edge. Any such hypergraph may be associated with a graph, its 2-section, and we will discuss the relation between the decomposability of \mathcal{C} and properties of this graph. We find, for example, that the family of all decomposable classes \mathcal{C} may be identified with the family of graphs called triangulated

by Berge ([2], p. 368), a family first discussed by Hajnal and Suranyi [6].

We turn now to an outline of the contents of this paper. In §2 we organize the main set-theoretic facts concerning decomposability and in the process prove the equivalence between Haberman's definition and that of Vorob'ev and Kellerer. Also included is a brief discussion of algorithms for checking decomposability. Apart from the definition of decomposability, the material in this section is independent of the rest of the paper. Our main work begins in §3 where we consider the 2-section of any decomposable hypergraph, showing that it is conformal and thereby reducing the discussion to graph theory. A further simplification allows us to consider only connected graphs. In §4 we explore the properties of complete articulation sets, and also give a general graph-theoretic analogue of an index defined by Haberman [5]. After obtaining some properties of this index, we are in a position to draw these ideas together and prove the equivalence of the following properties of a connected graph: (D) the associated clique hypergraph is decomposable; (I) the index satisfies an extremal condition; and (T) the graph is triangulated, i.e. no subset of the vertex set generates a cycle Z_n with $n > 3$.

Our set-theoretic notation is fairly rigorously restricted to the following: elements of base sets are denoted by α, β, γ and δ ; sets of elements i.e. subsets of base sets by a, b, c, e, f and g ; and classes of such sets by A, B, C, E and F . Furthermore the unions of all the sets in the classes, i.e. the base sets, are denoted by A, B, C, E and F , the corresponding upper-case roman letter: i.e. $A = \cup\{a: a \in A\}$. The letter d will be reserved for special use. It

will frequently be necessary to sub- or superscript the foregoing symbols with asterisks, primes etc. Whilst we use the usual notation $A \cup B$, $A \cap B$ and $A \setminus B$ for unions, intersections and differences of classes, we will abbreviate $a \cap b$ and $A \cap B$ by ab and AB when referring to sets. We write $|A|$ for the cardinality of A and denote the empty set by \emptyset . Finally we emphasise that all graphs in this paper are undirected, with no loops and multiple edges; more formally we will speak of a graph $\underline{G} = (V(\underline{G}), E(\underline{G}))$ consisting of a set $V(\underline{G}) = G$ of vertices, and a set $E(\underline{G})$ of unordered pairs of elements of G termed edges. All objects in this paper: sets, graphs, hypergraphs etc. are finite.

2. DECOMPOSABLE HYPERGRAPHS

In this section we give an account of the main set-theoretic properties of decomposable hypergraphs. The results are an integration of those of Haberman [5, Chapter 4], whose terminology we follow, the set-theoretic parts of Kellerer [9], and of Vorob'ev [12].

All of the hypergraphs which we consider in this paper will be a class of pairwise incomparable subsets of a (finite) set, and as such a class is called a generating class in [5], we will call such a hypergraph a generating class hypergraph. More formally

DEFINITION 1 A generating class (abbrev. g.c.) hypergraph is a pair (C, \mathcal{C}) consisting of a finite set C together with a class \mathcal{C} of pairwise incomparable subsets of C whose union coincides with C .

Where no confusion can result we will denote (C, \mathcal{C}) more simply by \mathcal{C} (since $\cup \mathcal{C} = C$ this should cause no problems). For two g.c.

hypergraphs (A, A) and (B, B) we write $(A, A) \leq (B, B)$ if $A \subseteq B$, and if for every $a \in A$ there exists $b \in B$ with $a \subseteq b$. It is easy to see that this relation is a partial order, and if we denote by \mathbb{H} the family of all g.c. hypergraphs, we have

LEMMA 1 The partially ordered set (\mathbb{H}, \leq) is a distributive lattice with zero.

Proof. It is a straightforward calculation to show that the lattice operations are given by $(A, A) \vee (B, B) = (C, C)$ [resp. $(A, A) \wedge (B, B) = (E, E)$] where $C = A \cup B$ [resp. $E = A \cap B$] and C [resp. E] is the class of all maximal elements of the class $A \cup B$ [resp. $\{ab: a \in A, b \in B\}$]. We omit the details. Furthermore, the zero of \mathbb{H} is readily seen to be $(\emptyset, \{\emptyset\})$, where \emptyset denotes the empty set. It only remains for us to check that the inequality $(A \vee B) \wedge F \leq (A \wedge F) \vee (B \wedge F)$ holds, since the reverse inequality is always valid. (Note that we have abbreviated (A, A) , (B, B) and (F, F) by A , B and F respectively.) If $g \in (A \vee B) \wedge F$ then $g = af$, say, for some $a \in A$ and $f \in F$, and our result will follow if we can show that there is no strict inclusion $af \subsetneq bf_1$, $b \in B$, $f_1 \in F$. But if this was the case, such an inclusion would already hold in $(A \vee B) \wedge F$, which is impossible. \square

With this preliminary observation, we return to the basic definition of the paper. It is convenient for a later purpose to formulate it somewhat more generally than in [5].

DEFINITION 2 The g.c. hypergraph C is said to be decomposed into $\{C_i: i \in I\}$ relative to $d \subseteq C$ if $C = \vee\{C_i: i \in I\}$, and if for every pair i, j of distinct elements of I we have $C_i \wedge C_j = \{d\}$.

COROLLARY 1 If C is decomposed into $\{C_i: i \in I\}$ relative to d , then for any ordering i_1, i_2, \dots, i_m of I ($m = |I|$), we have

$$C = (\dots((C_{i_1} \vee C_{i_2}) \vee C_{i_3}) \vee \dots) \vee C_{i_m},$$

each join being a decomposition relative to d .

Proof. This is an immediate consequence of the associativity of the join \vee and of the distributivity (Lemma 1) of \wedge over \vee . \square

Thus we can suppose, where convenient, that our decompositions are sequences of decompositions into two pieces. The simplest kind of g.c. hypergraphs are those of the form $(c, \{c\})$, and following Haberman [5] we give:

DEFINITION 3 A g.c. hypergraph C is said to be decomposable if either $|C| = 1$, or if there exists a decomposition $C = A \vee B$ of C relative to some $d \subseteq C$, with $A \vee B$ both decomposable and $|A| < |C|$, $|B| < |C|$.

It is readily seen that this definition implies that the class of decomposable g.c. hypergraphs exactly is the smallest class of g.c. hypergraphs that contains the simplest ones, i.e. those with $|C| = 1$ and is closed under joins that are decompositions.

By restricting the base set C of a hypergraph (C, C) to a proper subset $E \subset C$, and taking the maximal elements of $\{c^E: c \in C\}$, we obtain the g.c. subhypergraph C^E of C generated by E , cf. Berge [2, p.390]. The family of all decomposable g.c. hypergraphs is closed under this operation, as the next lemma shows.

LEMMA 2 Let (C, C) be a decomposable g.c. hypergraph. Then for any $E \subset C$ the g.c. subhypergraph (E, C^E) is also decomposable.

Proof. The proof is by induction on $|C|$. If $|C| = 1$, the result is certainly true, whilst a decomposable g.c. hypergraph C with $|C| > 1$ is (by definition) decomposable as $C = A \vee B$ relative to some $d \subseteq C$, with A and B both decomposable and $|A| < |C|$, $|B| < |C|$. It is easy to see that C^E is then decomposed into $A^{AE} \vee B^{BE}$ relative to dE , and so the inductive step can be proven, and the result follows. \square

In proving the equivalence of different set-theoretic formulations of decomposability, it is convenient to abstract the following notion, see Vorob'ev [12].

DEFINITION 4 An edge $c^* \in C$ is called extremal in C if C may be decomposed into $\{c^*\} \vee (C \setminus \{c^*\})$ relative to $d^* = c^* \cap U(C \setminus \{c^*\})$; equivalently if there exists $c^{**} \in C \setminus \{c^*\}$ such that $cc^* \subseteq c^{**}c^*$ for every $c \in C \setminus \{c^*\}$.

COROLLARY 2 If c^* is an extremal edge of the decomposable g.c. hypergraph C , then $C \setminus \{c^*\}$ is again decomposable.

Proof. We will see that $C \setminus \{c^*\}$ is just the restriction C^E of C to $E = C \setminus (c^* \setminus d^*) = (C \setminus \{c^*\}) \cup d^*$, and the result will follow from Lemma 2. But this is clear, since none of the edges of $C \setminus \{c^*\}$ intersect $c^* \setminus d^*$ in other than the empty set, and so they all remain pairwise incomparable, whilst $d^* \subset c^{**}$. \square

LEMMA 3 Let C be a decomposable g.c. hypergraph with $|C| \geq 2$. Then there exist at least two extremal edges of C .

REMARK. With a different (but equivalent) form of decomposability, Vorob'ev [12] proved this result as a lemma in § 1.51.

Proof. Again the proof is by induction on $|C|$. All hypergraphs with two incomparable edges are decomposable, and in this case both edges are trivially extremal.

Let C be a decomposable g.c. hypergraph with $|C| > 2$ edges, and suppose the assertion of the lemma is true for all decomposable g.c. hypergraphs with fewer edges. By definition C may be decomposed into $A \vee B$ relative to some $d \subseteq C$, with A and B both decomposable and having fewer edges than C . At least one of them must have two or more edges, say A . Then if we write $d = a^* \vee b^*$, the inductive hypothesis implies that A contains an extremal edge, a' , say, distinct from a^* , and we will see that a' is extremal in C . For if $b \in B$, then

$$a'b = a'(a'b) \subseteq a'd = a'a^*b^* \subseteq a'a^* \subseteq a'a'',$$

where $a'' \in A \setminus \{a'\}$ is such that $a'a \subseteq a'a''$ for all $a \in A \setminus \{a'\}$, (see definition 4). Since the same result is true with b in the above line of inclusions replaced by any $a \in A \setminus \{a'\}$, we have proved that a' is extremal in C . If $|B| \geq 2$, a similar argument proves the existence of an element $b' \in B$ distinct from b^* which is extremal in C , whilst if $|B| = 1$ the edge b^* is itself extremal in C . In either case we have found at least two extremal edges of C and the inductive step is proved. \square

We now have the preliminary results necessary for our first theorem. Part of this theorem is an algorithm which we formulate separately as follows. (i) For a g.c. hypergraph C we choose and fix an edge $\bar{c} \in C$. (ii) If $|C| = n$ we let c_n be any extremal edge of C other than \bar{c} , if such exists; otherwise we put $c_n = \bar{c}$. (iii) If c_n, \dots, c_{m+1} have been determined, $1 < m < n$, we let c_m be any extremal edge of $C \setminus \{c_n, \dots, c_{m+1}\}$ if such exists; otherwise we put $c_m = \bar{c}$. This defines a sequence of edges of C .

THEOREM 1 The following are equivalent for a g.c. hypergraph C with n edges.

- (a) C is decomposable
- (b) The algorithm described above has $c_m \neq \bar{c}$, $1 < m \leq n$, $c_1 = \bar{c}$.
- (c) There exists an ordering of C as $\{c_1, c_2, \dots, c_n\}$ such that for all $m = 1, 2, \dots, n$ there exists $m^* < m$ such that for all $\ell < m$, $c_\ell c_m \subseteq c_{m^*} c_m$.

REMARK The equivalence between (a) and (b) above was essentially proved by Haberman [5], and links his approach with that of Vorob'ev [12], whilst (c) is the form preferred by Kellerer, cf. [9, Satz 3.5].

Proof. (a) implies (b). This implication follows by successively applying Lemma 3 and Corollary 2, each time choosing an extremal element other than \bar{c} , until $m=1$.

(b) implies (c). If c is not chosen until $m=1$, we know that for all m , $1 < m \leq n$, c_m is extremal in $C \setminus \{c_n, \dots, c_{m+1}\} = \{c_1, c_2, \dots, c_m\}$. By definition this means that $\{c_m\} \wedge \{c_1, \dots, c_{m-1}\} = \{d_m\}$, i.e. that $c_\ell c_m \subseteq d_m$ for all $\ell < m$, and also that $d_m = c_{m^*} c_m$ for some $m^* < m$.

(c) implies (a). It is always true that $\{c_1, c_2\}$ is decomposable. Suppose we have proved that for some m between 2 and n in the ordering given by (c), $\{c_1, \dots, c_{m-1}\}$ is decomposable. Then there is a decomposition $\{c_1, \dots, c_m\} = \{c_m\} \vee \{c_1, \dots, c_{m-1}\}$ relative to $c_{m^*} c_m$ into decomposable hypergraphs, and so $\{c_1, \dots, c_m\}$ is decomposable. Continuing until $m=n$ we prove that C is decomposable. \square

We close this section with some remarks concerning algorithms to check decomposability. The procedure given prior to Theorem 1 certainly gives an algorithm which works, but this one is not particularly convenient in practice as it requires searching for an extremal edge, a task which involves repeatedly computing and

comparing many edge intersections. (A hypergraph would normally be stored in a computer as an incidence matrix with rows corresponding to edges and columns corresponding to the elements of the base set.)

An alternative algorithm, originally introduced by Goodman in the context of contingency tables, is much better suited to computer implementation. See [4], and it is also described in [3] and [8, pp 49-50]. To motivate this algorithm we note that for any extremal edge c^* of a g.c. hypergraph \mathcal{C} , the elements of $c^* \setminus d^*$, where $d^* = c^* \cap \bigcup (\mathcal{C} \setminus \{c^*\}) [= c^* c^{**} \text{ for some } c^{**} \in \mathcal{C} \setminus \{c^*\}]$ belong to precisely one edge of \mathcal{C} , namely the extremal edge c^* . The converse to this observation: "an edge containing elements belonging to no other edge is extremal" is false in general, but it is near enough to true for a simple algorithm checking decomposability to exist. An example which rules out the possible converse is $\mathcal{C} = \{\{1,2\}, \{2,3,4\}, \{4,5\}\}$, in which 3 belongs only to the non-extremal edge $\{2,3,4\}$. However, such elements are always associated with a decomposition which may, in turn be associated with a restriction (cf. the proof of Corollary 1). We formulate the idea as follows:

PROPOSITION 1 Let h be a subset of an edge c^* of a g.c. hypergraph \mathcal{C} consisting of elements belonging to precisely one edge of \mathcal{C} . Put $d = c^* \setminus h$, $\bar{\mathcal{C}} = \{d\} \vee (\mathcal{C} \setminus \{c^*\})$ and $\bar{c} = \bigcup \bar{\mathcal{C}}$. Then \mathcal{C} may be decomposed into $\{c^*\} \vee \bar{\mathcal{C}}$ relative to d , and $c^{\bar{\mathcal{C}}} = \bar{\mathcal{C}}$.

Proof. It is easy to see that $\{c^*\} \vee \bar{\mathcal{C}} = \mathcal{C}$, and if $c \in \mathcal{C} \setminus \{c^*\}$, $cc^* \subseteq d$, whilst $c^*d = d$. As in the proof of Corollary 1, distinct elements of $\mathcal{C} \setminus \{c^*\}$ remain incomparable when restricted to \bar{c} , because they do not intersect $h = c \setminus \bar{c}$. \square

COROLLARY 3 If C is decomposable, then so also is \bar{C} . \square

Thus we may check C for decomposability by searching for one [or more] element[s] belonging to exactly one edge of C - a very easy task computationally - and suppressing that element [those elements], in the sense that we form \bar{C} as above. We then repeat the procedure. If C is decomposable, this will continue until no elements are left, whilst it cannot do so if C is not decomposable.

This includes our general set-theoretic discussion of decomposability.

3 CONFORMAL HYPERGRAPHS

The aim of this section is to reduce the study of decomposable g.c. hypergraphs to the study of certain connected graphs. We do this by discussing the graph known as the 2-section of a g.c. hypergraph (C, C) , here denoted by \mathcal{C}_C , which (following [2, p.396]) is defined to be the graph which has vertex set C , and as edges the set of all unordered pairs $\{\alpha, \beta\}$ for which there exists an element $c \in C$ with $\{\alpha, \beta\} \subseteq c$.

For a general subclass $A \subseteq C$ we need to consider the 2-section \mathcal{A}_A of the hypergraph (A, A) , and ask about its relation to the subgraph of \mathcal{C}_C generated by [2, p.7], equivalently, induced by [7, p. 11] $A \subseteq C$, here denoted by $\langle A \rangle$. In general $\langle A \rangle$ need not coincide with \mathcal{A}_A , but there is an important special case in which it does so.

LEMMA 4 If we have a decomposition $C = \bigvee_{i \in I} C_i$ relative to $d \subseteq C$, then the subgraphs $\langle C_i \rangle$ of the 2-section \mathcal{C}_C generated by the subsets $C_i = \bigcup C_i$ coincide with the 2-sections \mathcal{C}_i .

Proof. By Corollary 1 we only have to prove the result for pairwise decompositions $C = A \vee B$. It is clear that the vertices of $\langle A \rangle$ and \underline{A}_A coincide, and it is equally clear that if α and α' are adjacent in \underline{A}_A , i.e. if $\{\alpha, \alpha'\} \subseteq a$ for some $a \in A$, then α and α' are adjacent in \underline{C}_C and hence in $\langle A \rangle$.

On the other hand, if $\{\alpha, \alpha'\}$ is an edge in $\langle A \rangle$, then $\{\alpha, \alpha'\} \subseteq c$ for some $c \in C$. If $c \in A$, then we have shown that $\{\alpha, \alpha'\}$ is an edge of \underline{A}_A , whilst $c \in B$, then $\{\alpha, \alpha'\} \subseteq AB \subseteq d$, and so $\{\alpha, \alpha'\}$ is still an edge of \underline{A}_A . \square

With this lemma proved we can turn to the main result of this section. Recall that a clique in a simple graph is a maximal complete subgraph [7, p.20], although some writers including [2, p.7] do not require maximality, and hence speak of maximal cliques. Further, a g.c. hypergraph C is called conformal [2, p.396] if the class of all (maximal) cliques of the 2-section \underline{C}_C of C coincides with C .

PROPOSITION 2 Let the g.c. hypergraph C be decomposed into $\{C_i : i \in I\}$ relative to $d \subseteq C$. Then C is conformal if and only if for all $i \in I$, C_i is conformal.

Proof As before it is enough to consider pairwise decompositions $C = A \vee B$. Suppose that A and B are conformal and let c be a (maximal) clique in \underline{C}_C . We first note that we must have $c \subseteq A$ or $c \subseteq B$ for if this was not the case and $\alpha \in c \setminus B$, $\beta \in c \setminus A$ there must be a $c' \in C$ with $\{\alpha, \beta\} \subseteq c'$. But $c' \in C$ implies $c' \in A \cup B$ and $c' \in A$ implies $\beta \in A$, $c' \in B$ implies $\alpha \in B$, in both cases a contradiction. But if $c \subseteq A$, c is a clique in A by Lemma 4 and thus $c \in A$ by assumption. The maximality of c implies $c \in C$. Similarly, if $c \subseteq B$ we get

$c \in C$, which was to be proved.

Conversely suppose that C is conformal and let a be a (maximal) clique in $\underline{A}_A = \langle A \rangle$. We just have to show that there is an $a' \in A$ such that $a \subseteq a'$ (the maximality of a will then imply $a = a'$). By Lemma 4, a is a complete subset of \underline{C} and by the conformality of C , there is a $c \in C$ such that $a \subseteq c$. If $c \in B$, $a = ac \subseteq AB = d$ and there is thus an $a' \in A$ such that $a' \supseteq a$. If $c \in A$ already, we can use c as a' . \square

The following result is essentially part (4⁰) of Theorem 2.2 of Vorob'ev [12] and is Theorem 5 of Andersen [1].

COROLLARY 4 Every decomposable hypergraph is conformal.

Proof. Since the 2-section of a hypergraph C with $|C| = 1$ is a complete graph, any such hypergraph is conformal. The corollary then follows directly from the definition and Proposition 2. \square

As a consequence of this proposition we need only discuss those decomposable hypergraphs C which consist of the class of all (maximal) cliques of a graph \underline{C} . We will write $C_{\underline{C}}$ for the hypergraph of all maximal cliques of the graph \underline{C} . The following discussion shows how we can, without loss of generality, restrict ourselves even further to consider only connected graphs. It may be contrasted with [2, p.391].

For any pair a, b of edges of a hypergraph (C, C) with 2-section $C_{\underline{C}}$ write $a \equiv b$ if there exists a sequence $a = c_1, c_2, \dots, c_m = b$ of edges such that $c_{k-1}c_k \neq \emptyset$, $1 < k \leq m$. This is easily seen to be an equivalence relation on C and we denote by $\{C_t : t \in T\}$ the equivalence classes of C under \equiv . Put $C_t = \cup C_t$ and let \underline{C}_t denote the

2-section of the hypergraph (C_t, C_t) , $t \in T$. In these terms we have:

LEMMA 5 The connected components of \mathcal{C}_C are precisely the graphs $\{C_t : t \in T\}$.

Proof: We begin by noting that each graph C_t is connected. If $\alpha, \beta \in C_t$ with $\alpha \in a$ and $\beta \in b$ say, then there should exist a chain $a = c_1, c_2, \dots, c_m = b$ with $c_{k-1}c_k \neq \phi$, $1 < k \leq m$. Choosing $\lambda_k \in c_{k-1}c_k$, $1 < k \leq m$ we see that $\alpha = \lambda_1, \lambda_2, \dots, \lambda_m, \lambda_{m+1} = \beta$ is a chain in C_t , thus proving that C_t is connected.

If α and β are connected in \mathcal{C}_C there exists a chain $\alpha = \lambda_1, \lambda_2, \dots, \lambda_n = \beta$ such that $\{\lambda_{k-1}, \lambda_k\} \subseteq c_k \in C$, $1 < k \leq n$. But this means that $c_1 = c_n$ and so there exists $t \in T$ with $\{\alpha, \beta\} \subseteq C_t$. Thus the C_t are connected components of \mathcal{C}_C and since $\bigcup_{t \in T} C_t = C$, $\bigcup_{t \in T} C_t = C$ (union of classes), we have described all of the connected components and the proof is complete. \square

Our next lemma shows that non-trivial decompositions of clique hypergraphs are associated with complete articulation sets, where a subset $d \subseteq C$ of a connected graph $\mathcal{C} = (C, E(\mathcal{C}))$ is called an articulation set if $\langle C \setminus d \rangle$ is disconnected, [2, p.8].

LEMMA 6 Let $C = A \vee B$ be a decomposition of the clique hypergraph C of the connected graph \mathcal{C} relative to $d \subseteq C$, then every path from $A \setminus B$ to B must contain an element of d .

Now suppose that $\alpha \in A \setminus B$ and $\beta \in B$. If $\alpha = \gamma_0, \gamma_1, \dots, \gamma_m = \beta$ is a connecting path, then for each i , $1 \leq i \leq m$, there exists $c_i \in C$ with $\{\gamma_{i-1}, \gamma_i\} \subseteq c_i$. Since $\alpha \in A \setminus B$ we must have $c_1 \in A$, and so $k = \min(i : c_i \in B)$ must satisfy $1 < k \leq m$. But then $c_{k-1} \in A$ and so $\gamma_{k-1} \in c_{k-1}c_k \subseteq d$. \square

COROLLARY 5 Any cycle in C which intersects both $B \setminus A$ and $A \setminus B$ must contain two non-consecutive elements of d .

Proof. Let the cycle contain $\alpha \in A \setminus B$ and $\beta \in B \setminus A$. By arguing as above in (say) the clockwise direction, we get an element $\delta_1 \in d$, and by arguing in the counter clockwise direction we find an element $\delta_2 \in d$. These elements cannot be consecutive in the cycle for α and β separate them. \square

COROLLARY 6 The graphs $\mathcal{A}_A = \langle A \rangle$ and $\mathcal{B}_B = \langle B \rangle$ are connected subgraphs of \mathcal{C} with clique hypergraphs A and B , prospectively.

Proof. Let α, α' be distinct elements of A . Since \mathcal{C} is connected there is a path in \mathcal{C} between α and α' , and we will see that any such path of shortest length must lie entirely within A . For if this was not the case, it would have to meet $B \setminus A$ and so pass through d twice. But then the two elements of d could be joined (within A) thereby shortening the path. Thus A is a connected subset. The remaining assertions are consequences of the foregoing Lemma 4 and Proposition 2. \square

We close this section with some remarks on the relation between our notion of decomposition applied to the clique hypergraph of a connected graph, and to the separation into pieces of such a graph relative to a complete articulation set [2, p.329]. Let d be a complete articulation set of a connected graph \mathcal{C} , i.e. d is complete and $\langle \mathcal{C} \setminus d \rangle$ is disconnected, and suppose that $\langle \mathcal{C} \setminus d \rangle$ has connected components $\{E_i : i \in I\}$. Then the pieces of \mathcal{C} relative to d are the subgraphs $\{\mathcal{C}_i : i \in I\}$ where $\mathcal{C}_i = \langle E_i \cup d \rangle$, $i \in I$. Finally, let C and $\{C_i : i \in I\}$ be the clique hypergraphs corresponding to \mathcal{C} and $\{\mathcal{C}_i : i \in I\}$ respectively.

PROPOSITION 3. $C = \vee\{C_i : i \in I\}$ is a decomposition relative to d .

Proof. This is a straightforward checking of definitions and so is omitted. \square

4 THE INDEX

We have seen that any decomposable hypergraph C gives rise to a graph \underline{C}_C , its 2-section, whose class of (maximal) cliques is C . Further, we have seen how we may restrict ourselves to those hypergraphs which derive in this way from connected graphs. Thus we may begin afresh by supposing given a connected graph $\underline{C} = (C, E(\underline{C}))$ and denoting its class of (maximal) cliques by $C = \underline{C}_C$. We also use the notation $\underline{C} \setminus s$ for the subgraph $\langle C \setminus s \rangle$ generated by $C \setminus s$ where $s \subseteq C$.

The main purpose of this section is to define an index associated with the complete subsets of C and derive some of its basic properties. Such an index was defined in quite a different way by Haberman [5, p.174], where it was called the adjusted replication number.

DEFINITION 5. For any complete subset d , let $\beta'_0(\underline{C} \setminus d)$ denote the number of pieces of \underline{C} relative to d in which d is not a (maximal) clique. Let

$$\nu(d) = 1 - \beta'_0(\underline{C} \setminus d).$$

The notation β'_0 is intended to suggest a modification of the number of connected components β_0 (= 0'th Betti number) of a graph (= 1-complex), see [7].

LEMMA 7

 $v(d) = 1$ if d is a clique $v(d) = 0$ if d is not an articulation setand not a clique $v(d) < 0$ implies that d is an articulation set.

Proof. If d is a clique, it will be a clique in all the pieces of C relative to d and

$$v(d) = 1 - \beta'_0(C \setminus d) = 1 - 0 = 1.$$

If d is not a clique, nor a articulation set $\beta'_0(C \setminus d) = 1$ and thus $v(d) = 0$.

If $v(d) < 0$, $\beta'_0(C \setminus d) \geq 2$, and d must be an articulation set. \square

Our major result in this section relates our index across decompositions. More precisely, let the clique hypergraph C of a connected graph be decomposed into $C_i, i \in I$ relative to $d^* \subseteq C$ and let v_i denote the indices associated with $\langle C_i \rangle$. Let also $v_i(d) = 0$ if $d \not\subseteq C_i$.

LEMMA 8 For any complete subset $d \subseteq C$ we have, with the notation above:

$$v(d) = \begin{cases} \sum_{i \in I} v_i(d) & \text{if } d \neq d^* \\ \sum_{i \in I} v_i(d^*) - |I| + 1 & \text{if } d = d^*. \end{cases}$$

Proof. We readily see, that we as usual can restrict ourselves to the case with $|I| = 2$, i.e. $C = A \vee B$, $A \wedge B = \{d^*\}$. We then consider four cases: i) $d = d^*$, ii) $d \subsetneq d^*$, iii) $d \supsetneq d^*$, iv) $d \not\subseteq d^*$ and $d \not\supseteq d^*$.

i) $d = d^*$. If d^* is removed from A we get pieces A_1, \dots, A_k ,

$A_{\sim k+1}, \dots, A_{\sim k+m}$, with $A_{\sim 1}, \dots, A_{\sim k}$ containing d^* as a clique and $A_{\sim k+1}, \dots, A_{\sim k+m}$ not as a clique. Similarly we get pieces $B_{\sim 1}, \dots, B_{\sim p}, B_{\sim p+1}, \dots, B_{\sim p+q}$.

But the pieces of \mathcal{C} obtained by removing d must be the same since B_i is always separated from A_j by d^* according to lemma 6.

Thus we have

$$v(d^*) = 1 - k - p$$

$$v_A(d^*) = 1 - k, \quad v_B(d^*) = 1 - p,$$

whereby we see that our formula holds.

ii) $d \subsetneq d^*$. Let the pieces of \mathcal{A} and \mathcal{B} relative to d be $A_{\sim 1}, \dots, A_{\sim k}, A_{\sim k+1}, \dots, A_{\sim k+m}, B_{\sim 1}, \dots, B_{\sim p}, B_{\sim p+1}, \dots, B_{\sim p+q}$, as before.

Exactly one A-piece and one B-piece contain points of $d^* \setminus d$. Because if there were more, these pieces would be connected via $d^* \setminus d$ when d was removed, thus contradicting the notion of a piece, d^* is not a clique in such a piece since $d \subsetneq d^*$, with d^* complete. So, let those pieces be $A_{\sim k+1}$ and $B_{\sim p+1}$. The pieces of \mathcal{C} relative to d are then

$$A_{\sim k+1} \cup B_{\sim p+1}, A_{\sim k}, \dots, A_{\sim k+2}, \dots, A_{\sim k+m}, B_{\sim 1}, \dots, B_{\sim p}, B_{\sim p+2}, \dots, B_{\sim p+q}.$$

Since any two of these cannot be connected via $d^* \setminus d$ and therefore only are via d , again by lemma 6. Thus

$$v_A(d) = 1 - m, \quad v_B(d) = 1 - q,$$

$$v(d) = 1 - [(m-1) + (q-1) + 1] = v_A(d) + v_B(d).$$

iii) $d \supsetneq d^*$. Then we must either have $d \subseteq A$ or $d \subseteq B$. Suppose $d \subseteq A$. Let $A_{\sim 1}, \dots, A_{\sim k}, A_{\sim k+1}, \dots, A_{\sim k+m}$ be the pieces of \mathcal{A} relative to d .

Let $B_1^*, \dots, B_p^*, B_{p+1}^*, \dots, B_{p+q}^*$ be the pieces of B relative to d^* .

Then the pieces of C relative to d must be

$$A_1, \dots, A_{k+m}, (B_1^* \cup d), \dots, (B_{p+q}^* \cup d),$$

since $d \cap B = d^*$. But d must be a clique of all the B -pieces, because no vertices in $B \setminus d^*$ are adjacent to those in $d \setminus d^*$ by lemma 6.

$$v(d) = 1 - m = v_A(d) \text{ and } d \notin B$$

implies $v_B(d) = 0$, i.e. that the formula is correct.

iv). $d \not\subseteq d^*$ and $d^* \not\subseteq d$. Again, let us assume $d \subset A$, i.e. $d \notin B$.

Let A_0 be the A -piece relative to d containing $d^* \setminus d \neq \emptyset$. Then the pieces of C relative to d are

$$A_0 \cup B, A_1, \dots, A_k,$$

where A_0, \dots, A_k are the A -pieces relative to d . Note that d is a clique in $A_0 \cup B$ iff it is in A_0 , since no vertices in B are adjacent to vertices in $d \setminus d^* \neq \emptyset$. Thus $v(d) = v_A(d)$ and since $v_B(d) = 0$, the proof is complete. \square

COROLLARY 7 For any connected graph C with the class of (maximal) cliques C ,

$$\sum_{\substack{\text{all complete} \\ \text{subsets } d}} v(d) \geq 1$$

Proof. By induction on $|C|$. If $|C| = 1$ the result is clearly true. Suppose that C is a connected graph with $|C| > 1$ and that the assertion is true for all connected graphs with fewer than $|C|$ cliques. Then either $v(d) \geq 0$ for all d , in which case the result is true because $v(c) = 1$ for all $c \in C$ by Lemma 7, or there is a

d^* with $v(d^*) < 0$. But then d^* is an articulation set by lemma 7 and C can be decomposed into C_i , $i \in I$ relative to d^* , where C_i are the clique hypergraphs of the pieces (Proposition 3).

Clearly, $|C_i| < |C|$ so the inductive hypothesis and the preceding lemma gives us

$$\begin{aligned} \sum_{d \neq d^*} v(d) &= \sum_{d \neq d^*} \left(\sum_{i \in I} v_i(d) \right) + \sum_{i \in I} v_i(d^*) - |I| + 1 \\ &= \sum_{i \in I} \left(\sum_d v_i(d) \right) - |I| + 1 \geq |I| - |I| + 1 = 1. \quad \square \end{aligned}$$

5 DECOMPOSABLE GRAPHS

In this section we draw together the notions introduced in the previous two and show that graphs \tilde{C} whose clique hypergraphs $C_{\tilde{C}}$ are decomposable have other interesting properties. We use the notation Z_n for the graph known as the n -cycle [7,p.13].

THEOREM 2 The following properties of a connected graph \tilde{C} are equivalent

- (D) The clique hypergraph $C_{\tilde{C}}$ is decomposable.
- (I) $\sum_d v(d) = 1$
- (T) No subset $s \subseteq C$ generates a cyclic subgraph $\langle s \rangle \approx Z_n$ with $n > 3$.

REMARKS Vorob'ev derived condition (T) in his discussion of this topic, see [12, Theorem 2,2], see also Kellerer [9, Satz 3,2], and we note that such graphs are called triangulated by Berge [2,p.368]. An easy reformulation of (T) is (T') : every polygon in \tilde{C} of length $k \geq 4$ has a chord. Graphs with these properties were apparently first studied by Hajnal and Suranyi [6].

Proof. (D) implies (I). This is an easy induction on $|C|$ using the (Index) Lemma 8. The conclusion is clearly true for $|C| = 1$, and so we take a decomposable clique hypergraph $C = C_C$ with $|C| > 1$, supposing that the conclusion is true for all decomposable clique hypergraphs with fewer than $|C|$ edges. Then there must exist a decomposition $C = A \vee B$ relative to a subset $d^* \subseteq C$, of C into decomposable hypergraphs. By Lemma and Corollary 6, A and B are both the clique hypergraphs of connected graphs with fewer elements than C . If (I) is true for A and for B , as it must be by the inductive hypothesis, it remains true for C by using Lemma 8, since

$$\begin{aligned} \sum_d v(d) &= \sum_{d \neq d^*} (v_A(d) + v_B(d)) + v_A(d^*) + v_B(d^*) - 1 \\ &= 1 + 1 - 1 = 1. \end{aligned}$$

(I) implies (T). Again the proof is by induction on $|C|$. If $|C| = 1$, the corresponding graph is complete and (T) always holds. Suppose now $|C| > 1$ and that the assertion is true for all connected graphs with fewer than $|C|$ cliques. If (I) holds for \underline{C} and $|C| > 1$ there must be a d^* with $v(d^*) < 0$ since $v(c) = 1$ for all $c \in C$ by lemma 7. As in Corollary 7 we deduce that there is a decomposition of C into C_i , $i \in I$ relative to d^* and with $|C_i| < |C|$. Using the inductive hypothesis, Lemma 8 and Corollary 7, we deduce that \underline{C}_i satisfy (T). That \underline{C} satisfies (T) now follows from Corollaries 1 and 6.

(T) implies (D). The final implication is also proved by induction on $|C|$. As before it is easy to see that the conclusion desired is true when $|C| = 1$ and so we make the now familiar inductive hypothesis. Then with $|C| > 1$ there is either (i) a decomposition $C = A \vee B$ of C relative to some $d \subseteq C$, or (ii) there is no such decomposition. Since property (T) is preserved upon passing to

generated subgraphs, we note that in case (i) \underline{A} and \underline{B} must satisfy (T). But then the inductive hypothesis implies that A and B are both decomposable, and so we conclude that C is decomposable.

Our proof will be complete when we show that case (ii) cannot arise. To prove this, let \mathcal{D} be the set of intersections of distinct cliques and let d be an element in \mathcal{D} which is maximal in \mathcal{D} under set inclusion. We shall show that d is an articulation set and hence by Proposition 2 defines a decomposition. Suppose $\underline{C} \setminus d$ is connected. Since $d \in \mathcal{D}$ there are $a, b \in C$ such that $ab = d$, $a \setminus d \neq \emptyset$, $b \setminus d \neq \emptyset$ and $a \setminus d$ is connected to $b \setminus d$ outside d . Amongst the pairs $\alpha \in a \setminus d$ and $\beta \in b \setminus d$, select a pair, α^* , β^* , say, for which the shortest connecting path is of shortest length. Then

$$\alpha^* = \gamma_0, \gamma_1, \dots, \gamma_m = \beta^*$$

is of length $m \geq 2$, and because it is a shortest path, $\gamma_i \notin a$ for $i > 0$.

Let us note that γ_1 cannot be adjacent in \underline{C} to every $\delta \in d = ab$. For if this was the case, $\{\gamma_1, d^*\} \cup d$ would be a complete subset of \underline{C} and so contained in a clique $c \in C \setminus \{a, b\}$ then we would have $\mathcal{D} \ni ac \supseteq d \cup \{\alpha^*\}$, contradicting the maximality of d . Thus there exists $\delta^* \in d = ab$ with $\{\gamma_1, \delta^*\} \notin E(\underline{C})$.

Now let $k = \min \{j : \{\gamma_j, \delta^*\} \in E(\underline{C})\}$. By the foregoing, $j \geq 2$ and since $\beta^* = \gamma_m$ is adjacent to δ^* , $j \leq m$. Then $\alpha^* = \gamma_0, \gamma_1, \dots, \gamma_{j-1}, \gamma_j, \delta^*, \alpha^*$ is a cycle of length $j + 2 \geq 4$ in \underline{C} . It has no chords, since the path from α^* to β^* has shortest length. But this contradicts (T) and so $\underline{C} \setminus d$ must be connected. The proof is now complete. \square

Acknowledgements.

The authors would like to thank Søren Tolver Jensen for reading the manuscript and making helpful criticisms, and the Danish Natural Science Research Council for financial support.

REFERENCES

- 1 A.H. ANDERSEN Multidimensional contingency tables. Scand. J. Statist. 1 (1974) 115-127.
- 2 C. BERGE "Graphs and Hypergraphs". Translated from French by Edward Minieka. North-Holland Publishing Company, Amsterdam, London. American Elsevier Publishing Company Inc. New York, 1973.
- 3 Y.M.M. BISHOP, S.E. FIENBERG and P.W. HOLLAND. "Discrete Multivariate Analysis". MIT Press, Cambridge, Mass. 1975.
- 4 L.A. GOODMAN Partitioning of Chi-square, analysis of marginal contingency tables, and estimation of expected frequencies in multidimensional contingency tables. Journal of the American Statistical Association, 66 (1971) 339-344.
- 5 S.J. HABERMAN "The Analysis of Frequency Data". The University of Chicago Press, Chicago, London, 1974.
- 6 A. HAJNAL and J. SURANYI Über die Auflösung von Graphen in vollständige Teilgraphen. Ann. Univ. Sci. Budapest 1 (1958) 113-121.
- 7 F. HARARY "Graph Theory". Addison-Wesley Publishing Company, Reading, Massachusetts; Menlo Park, California; London; Don Mills, Ontario, 1969.
- 8 S.T. JENSEN "Flersidede kontingenstabeller". Institute of Mathematical Statistics, University of Copenhagen, 1978.
- 9 H.G. KELLERER Verteilungsfunktionen mit gegebenen Marginalverteilungen. Z. Wahrscheinlichkeitstheorie 3 (1964) 247-270.

- 10 T.P. SPEED Decompositions of graphs and hypergraphs.
To appear in the Proceedings of an International Conference on the Theory of Graphs, Canberra, Australia, August-September 1977.
- 11 P. SUOMELA "Construction of Nearest-Neighbour Systems".
Ann. Acad. Sci. Fenn. A, Mathematica Dissertationes 10
(1976).
- 12 N.N. VOROB'EV Consistent families of measures and their
extensions. Theor. Prob. Appl. 7 (1962) 147-163.
- 13 N.N. VOROB'EV Markov measures and Markov extensions.
Theor. Prob. Appl. 8 (1963) 420-429.
- 14 N.N. VOROB'EV Coalition games.
Theor. Prob. Appl. 12 (1967) 250-266.

PREPRINTS 1977

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE
INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5,
2100 COPENHAGEN Ø, DENMARK.

- No. 1 Asmussen, Søren & Keiding, Niels: Martingale Central Limit Theorems and Asymptotic Estimation Theory for Multitype Branching Processes.
- No. 2 Jacobsen, Martin: Stochastic Processes with Stationary Increments in Time and Space.
- No. 3 Johansen, Søren: Product Integrals and Markov Processes.
- No. 4 Keiding, Niels & Lauritzen, Steffen L. : Maximum likelihood estimation of the offspring mean in a simple branching process.
- No. 5 Hering, Heinrich: Multitype Branching Diffusions.
- No. 6 Aalen, Odd & Johansen, Søren: An Empirical Transition Matrix for Non-Homogeneous Markov Chains Based on Censored Observations.
- No. 7 Johansen, Søren: The Product Limit Estimator as Maximum Likelihood Estimator.
- No. 8 Aalen, Odd & Keiding, Niels & Thormann, Jens: Interaction Between Life History Events.
- No. 9 Asmussen, Søren & Kurtz, Thomas G.: Necessary and Sufficient Conditions for Complete Convergence in the Law of Large Numbers.
- No. 10 Dion, Jean-Pierre & Keiding, Niels: Statistical Inference in Branching Processes.

PREPRINTS 1978

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE
INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5,
2100 COPENHAGEN Ø, DENMARK.

- No. 1 Tjur, Tue: Statistical Inference under the Likelihood Principle.
- No. 2 Hering, Heinrich: The Non-Degenerate Limit for Supercritical Branching Diffusions.
- No. 3 Henningsen, Inge: Estimation in M/G/1-Queues.
- No. 4 Braun, Henry: Stochastic Stable Population Theory in Continuous Time.
- No. 5 Asmussen, Søren: On some two-sex population models.
- No. 6 Andersen, Per Kragh: Filtered Renewal Processes with a Two-Sided Impact Function.
- No. 7 Johansen, Søren & Ramsey, Fred L.: A Bang-Bang Representation for 3x3 Embeddable Stochastic Matrix.
- No. 8 Braun, Henry: A Simple Method for Testing Goodness of Fit in the Presence of Nuisance Parameters.
- No. 9 Lauritzen, S.L. & Speed, T.P. & Vijayan, K. : Decomposable Graphs and Hypergraphs.