Minimal Covers of Maximal Cliques for Interval Graphs

Alain C. Vandal
Department of Mathematics and Statistics, McGill University
Centre for Clinical Epidemiology & Community Studies
SMBD–Jewish General Hospital, Montréal

Marston D.E. Conder
Department of Mathematics, University of Auckland

Robert Gentleman
Fred Hutchison Cancer Research Center

ABSTRACT. We address the problem of determining all sets which form minimal covers of maximal cliques for interval graphs. We produce an algorithm enumerating all minimal covers using the $\subset$-minimal elements of the interval order, as well as an independence Metropolis sampler. We characterize maximal removable sets, which are the complements of minimal covers, and produce a distinct algorithm to enumerate them. We use this last characterization to provide bounds on the maximum number of minimal covers for an interval order with a given number of maximal cliques, and present some simulation results on the number of minimal covers in different settings.

This work was supported in part by the National Sciences and Engineering Research Council of Canada, the Fonds québécois de la recherche sur la nature et les technologies and the New Zealand Marsden Fund.

1 Introduction

An interval order is a partially ordered set, members of which can be identified with intervals on a linear order of the form $[l_i, u_i]$, with the order relation $\prec$ given by $[l_i, u_i] \prec [l_j, u_j]$ whenever $u_i < l_j$. Interval orders and the graph theory associated with their incomparability graphs, also called interval graphs, provide a natural model for the study of scheduling and preference models. They have also recently appeared as a promising abstraction tool in a branch of statistics called nonparametric survival analysis.

In this paper, we characterize the smallest sets of maximal antichains, called *minimal covers*, which cover the elements of an interval graph, and present two main algorithms to enumerate these sets. The maximal antichains of an interval order are the maximal cliques of its interval graph. For simplicity, we select the term *maximal cliques* to indicate both instances; thus such phrases as "maximal cliques of interval orders" carry no ambiguity. As well, we describe an algorithm to generate minimal covers uniformly at random. We also discuss the maximal number of such minimal covers.

Minimal covers are defined in Section 2 along with other necessary concepts associated with interval orders and the description of the substantive problem in statistics which led us to the present investigation. We then consider the enumeration of minimal covers from two points of view. The first, discussed in Section 3, is a backtracking algorithm which constructs all minimal covers of an interval order from the so–called set of $\subset$-minimals of the interval order. We show that this algorithm is a generalization of a classic procedure which generates one maximum chain along with one minimum cover from an interval order. We use the backtracking structure of the algorithm in Section 4 to produce a uniform minimal cover generating algorithm, a necessary extension since, as shown in Section 6, the number of minimal covers is of exponential order with respect to the number of elements in the interval order. The structure of interval orders allows the efficient computation of the minimum and maximum probabilities of generation, thus making it possible to perform perfect simulation from an independence Metropolis sampler. The second perspective, covered in Section 5, is that of a characterization of the complements of minimal covers, termed *maximal removable sets*. We provide properties and algorithmic details concerning maximal removable sets. These, while of perhaps less immediate applicability than the generating algorithm of Section 3, are valuable for the insight they provide on the structure of minimal covers and ultimately provide us with bounds on the maximum number of minimal covers achievable with a given

number of maximal cliques. Section 6 presents the derivation of these bounds as well as some simulation results.

2 Notation and definitions

2.1 Interval orders

Let $\underline{X} = (X, \prec)$ denote a partially ordered set, or *poset*. That is, X is a set together with a binary relation $\prec \subset X \times X$ which is both irreflexive and transitive. An interval order $\underline{X}$ is a poset such that

$$(a \prec x, b \prec y) \Rightarrow (a \prec y \text{ or } b \prec x), \qquad \text{for all } a, b, x, y \in X.$$

We shall use $\sim$ to denote the symmetric complement of $\prec$. That is,

$$x \sim y \text{ if and only if not } (x \prec y) \text{ and not } (y \prec x).$$

The relation $\sim$ is the incomparability relation. The undirected graph $(X, \sim)$ is called the *interval graph* of $\underline{X}$.

A *linear order* is an ordered set $(X, \prec^*)$ such that $x \not\sim y$ for all $x, y \in X$ with $x \neq y$. A *chain* in a poset $(X, \prec)$ is a linear order $\left(C, \prec|_{C \times C}\right)$ with $C \subset X$. A maximum chain is a chain of maximum cardinality in $(X, \prec)$.

Hereinafter, unless otherwise noted, $\underline{X} = (X, \prec)$ will denote an interval order, and we will let $n = |X|$ and $X = \{x_1, \ldots, x_n\}$.

2.2 Maximal cliques and their linear ordering

$M \subset X$ is a *clique* of $\underline{X}$ if $x \sim y$ for all x, y in M. A *maximal clique* is a clique not properly contained in any other clique. We will denote by $\mathcal{M}(\underline{X})$ the set of maximal cliques of an interval order $\underline{X}$.

A crucial characterization of interval orders is that there exists a natural linear ordering on their set of maximal cliques [1]. Specifically, if $M_a, M_b \in \mathcal{M}(\underline{X})$ and we define the relation $\sqsubset$ over $\mathcal{M}(\underline{X})$ by

$$M_a \sqsubset M_b \Leftrightarrow (M_a \setminus M_b) \prec (M_b \setminus M_a),$$

then $(\mathcal{M}(\underline{X}), \sqsubset)$ is a linear order, where the relation $\prec$ is extended to subsets of X in the obvious manner.

Setting $m = |\mathcal{M}(\underline{X})|$, we will assign subscripts $i = 1, \ldots, m$ to the elements of $\mathcal{M}(\underline{X})$ according to their linear ordering, that is with $M_i \sqsubset M_j \Leftrightarrow i < j$. Minima and maxima are thus well-defined elements over subsets of $\mathcal{M}(\underline{X})$.

2.3 Properties of interval order elements

Elements of X have properties which depend on the maximal cliques of $\underline{X}$. If x is contained in only one maximal clique then x is termed a *simplicial element*, since its neighbourhood in the interval graph is complete. A maximal clique containing a simplicial element is said to be essential. It can readily be shown that the first and last maximal cliques in $\mathcal{M}(\underline{X})$, M_1 and M_m, must always be essential (see [14] and [7], Section 2.3).

For $x \in X$, we denote by $x^* = \{M \in \mathcal{M}(\underline{X}) : x \in M\}$ the dual of x with respect to $\mathcal{M}(\underline{X})$. In such a case, the propositions $x \in M$ and $M \in x^*$ equivalently express the fact that x is contained within or covered by maximal clique M. For simplicity, we will write $M \sqsubset x^*$ for $M \sqsubset \min x^*$ and $x^* \sqsubset M$ for $\max x^* \sqsubset M$, though clearly $\sqsubset$ is not a linear order on contiguous subsets of $\mathcal{M}(\underline{X})$.

The dual of every element of an interval order is a contiguous sequence of maximal cliques. This property is referred to as the *consecutive–ones property* of interval orders.

2.4 Clique matrix

The *clique matrix* representation of an interval order is an indicator matrix relating the elements of X to the maximal cliques of $\mathcal{M}(\underline{X})$. Specifically, the clique matrix of $\underline{X}$ is given by $\mathbf{A} \in \{0,1\}^{m \times n}$, where

$$A_{ij} = \begin{cases} 1 & \text{if } x_j \in M_i \\ 0 & \text{if } x_j \notin M_i \end{cases} .$$

The definition implies that the rows of $\mathbf{A}$ are ordered similarly to the elements of $\mathcal{M}(\underline{X})$. Under this ordering, all interval orders will have a unique clique matrix

representation up to the subscript ordering of X or, equivalently, up to ordering of the columns of $\mathbf{A}$.

Our usage of the term "clique matrix" differs from the traditional one ([11], Chapter 3) in that we require the rows of the clique matrix to be ordered according to the maximal clique linear ordering, thus explicitly preserving the consecutive–ones property.

2.5 Covers of maximal cliques

Minimal covers of interval orders form the focus of this paper. We formalize the concept.

Definition 1 *For $\underline{X} = (X, \prec)$ an interval order, we call $\mathcal{C} \subset \mathcal{M}(\underline{X})$ a* cover *of X if $X = \bigcup_{M \in \mathcal{C}} M$. We will call the cover $\mathcal{W}$ a* minimal cover *of X if no proper subset of $\mathcal{W}$ is a cover of X. A minimal cover is a* minimum cover *if it has lowest possible cardinality.*

Clearly all essential maximal cliques are contained in every minimal cover. The fact that minimal covers of various cardinalities exist is not so immediately obvious. Finding them is dealt with in Section 3.

Maximal removable sets (*MRS*s) are closely related to minimal covers.

Definition 2 *A maximal removable set is a set $\mathcal{R}$ of maximal cliques such that $\mathcal{M}(\underline{X}) \setminus \mathcal{R}$ is a minimal cover.*

The properties of MRS's and a direct enumeration method for them are discussed in Section 5.

2.6 Covers and self-consistent estimates

This section briefly describes the substantive statistical problem which motivated the present research. It may be skipped by readers without a break in continuity or understanding.

Survival analysis is a branch of statistics concerned with the analysis of event-time data. Such a datum consists of the time elapsed between an onset and an event, such as between start of chemotherapy and relapse in a cancer patient. Data of this type are often incompletely observed. For example, a study may end before the event of interest is observed, in which case the duration is only known to be longer than a certain value: this situation is referred to as right-censoring. Another common pattern occurs when individuals are monitored periodically for the development of a condition (the event of interest). In such a situation, the event time will only be known to have occurred within a certain time interval, and the exact moment of occurrence remains unknown. Such event-time data are said to be *interval censored*.

Under interval censoring, *self-consistent estimates* (SCEs) form an important class of event-time distribution estimates[12,16,25]. SCEs are crucially related to the maximal clique covers of the interval order underlying the data. SCEs will only place probability mass on the real representation of the maximal cliques of the interval graph of the data, that is, on the maximal intersections $H_i = \bigcap_{X_j \in M_i} X_j$, where the X_j, $j = 1, \ldots, n$ are taken to be the real intersections forming the data [17,22]. If $\mathbf{A}$ is the clique matrix of the data, p_i the probability mass placed on H_i and $\mathbf{p} = [p_1, \ldots, p_m]^\top$, then an SCE of $\mathbf{p}$ is any vector $\hat{\mathbf{p}} \geq 0$ (elementwise) satisfying $\sum_{i=1}^{m} \hat{p}_i = 1$ and

$$n\hat{\mathbf{p}} = \mathbf{D}_{\hat{\mathbf{p}}}\mathbf{A}(\mathbf{A}^\top \hat{\mathbf{p}})^{-\mathbf{I}}, \tag{2.1}$$

where $\mathbf{D}_{\mathbf{x}}$ is the diagonal matrix with diagonal $\mathbf{x}$ and $\mathbf{x}^{-\mathbf{I}} = [1/x_1, \ldots, 1/x_n]^\top$ [10]. Existence of an SCE follows from the existence of a unique nonparametric maximum likelihood estimate (NPMLE), which is itself an SCE [9].

Clearly an incidental requirement on $\hat{\mathbf{p}}$ from (2.1) is that $\mathbf{A}^\top \hat{\mathbf{p}} > 0$ elementwise, that is, $\mathcal{C}_{\hat{\mathbf{p}}} \doteq \{M_i \in \mathcal{M}(\underline{X}) : \hat{p}_i > 0\}$ must form a cover for the data. We can also establish the converse as follows. A basic algorithm to obtain an SCE is the EM algorithm [2], which in this instance consists of iterating Equation (2.1) with $\hat{\mathbf{p}}$ replaced by r^{th} iterate $\mathbf{p}^{(r)}$ on the right-hand side and by $(r+1)^{\text{th}}$ iterate $\mathbf{p}^{(r+1)}$ on the left-hand side, starting from some initial value $\mathbf{p}^{(0)}$. Iterations are pursued until convergence to a fixed point of (2.1). If $\mathbf{p}^{(0)} > 0$ elementwise, the incomplete multinomial nature of the data [23] causes $\mathbf{p}^{(r)}$ to converge to the NPMLE of $\mathbf{p}$ through this iterative procedure [24].

If some of the entries of $\mathbf{p}^{(0)}$ are identically zero but $\mathcal{C} = \mathcal{C}_{\mathbf{p}^{(0)}}$ remains a cover for the data, then $\mathbf{p}^{(r)}$ will converge to a unique SCE $\mathbf{p}_{\mathcal{C}}$, regardless of the values of the non-zero entries of $\mathbf{p}^{(0)}$. This can be seen by positing artificial interval censored data

with clique matrix $\mathbf{A}^*$ consisting of the rows of $\mathbf{A}$ corresponding to the maximal cliques in $\mathcal{C}$ (see Lemma 3.4). The limit $\mathbf{p}_{\mathcal{C}}$ is then isomorphic to the unique NPMLE given these artificial interval censored data.

Thus every cover of $\underline{X}$ corresponds to a unique self-consistent estimate, although several "starting value covers" $\mathcal{C}_{\mathbf{p}^{(0)}}$ may yield the same SCE. If $\mathcal{C}_{\mathbf{p}^{(0)}}$ is a minimal cover, the EM algorithm will yield an SCE $\hat{\mathbf{p}}$ with $\mathcal{C}_{\hat{\mathbf{p}}} = \mathcal{C}_{\mathbf{p}^{(0)}}$, since in the artificial data posited above, every maximal clique is essential. Any empirical investigation of SCEs on the semi-lattice of maximal clique covers must start with an investigation of the minimal covers, which form the base of this semi-lattice. This is what we propose in the present article.

3 Enumerating minimal covers

In this section, we first present a simple algorithm in Construction 3.2 which enables us to find a single *minimum* cover. This procedure is similar to that found in [8]; its expression is couched in the language of comparability rather than incomparability, and serves to introduce its generalization to Construction 3.3.

The following result is the well-known dual expression of Dilworth's Decomposition Theorem [3] and concerns the cardinality of minimum covers.

Theorem 3.1 *Let $(X, \prec)$ be a partially ordered set. The length of the longest chain of $(X, \prec)$ equals the minimum number of cliques required to cover the elements of X.*

Since the minimum number of cliques can be no smaller than the minimum number of maximal cliques which cover a poset, the theorem tells us that a minimum cover has cardinality equal to the length of the longest chain in the poset.

This result can be shown constructively for interval orders, with the added bonus that we produce a minimum cover and a maximum chain in the process. In essence, the following Construction and Theorem state that there exists a minimum cover consisting of a set of maximal cliques, each of which uniquely contains a certain element. The set of these elements forms a maximum chain in the interval order.

Construction 3.2 *Let $\{M_1, \ldots, M_m\}$ be the linearly ordered set of maximal cliques of interval order $\underline{X} = (X, \prec)$. Let y_1 be a simplicial element of X belonging to*

the first maximal clique M_1, and let $\mu_1 = M_1$. Form the sets $Y_1' = \{y_1\}$ and $\mathcal{W}_1' = \{\mu_1\}$. Then form the sets Y_i' and $\mathcal{W}_i'$ for $i > 1$ as follows:

While i is such that $\{y \in X : \mu_{i-1} \sqsubset y^*\} \neq \emptyset$, let $Y_i' = Y_{i-1}' \bigcup \{y_i\}$, where

$$y_i \in \left\{ y \in X : \operatorname*{argmin}_{y} \max_{\mu_{i-1} \sqsubset y^*} y^* \right\},$$

is a particular selection of y_i, and let $\mathcal{W}_i' = \mathcal{W}_{i-1}' \bigcup \{\mu_i\}$, where

$$\mu_i = \max y_i^*.$$

Theorem 3.2 *Under Construction 3.2, there exists I such that $\{y : \mu_I \sqsubset y^*\} = \emptyset$ and $\mathcal{W}' = \mathcal{W}_I'$ is a minimum cover for X.*

Proof. The proofs of coverage and minimality for $\mathcal{W}'$ are specializations of the proofs of Claims 1 and 2 in Theorem 3.3, possible since Construction 3.2 is a special case of Construction 3.3. Proof that $\mathcal{W}'$ is of minimal cardinality is simple viewed from the context of Theorem 3.3. □

The initial inclusion of $\{M_1\}$ in the sequence of $\mathcal{W}_i'$'s in Construction 3.2 is legitimate and necessary since M_1 is always essential, and therefore required in any cover. We note without proof that Y' is a maximum chain in $\underline{X}$.

Example 3.1 (In the following example as in subsequent ones, elements of X [columns] are identified by regular subscripts and maximal cliques of $\underline{X}$ [rows] by bolded subscripts.) *Consider the clique matrix*

$$\mathbf{A} = \begin{array}{c} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \\ \mathbf{6} \\ \mathbf{7} \\ \mathbf{8} \end{array} \begin{array}{c} \begin{array}{cccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \end{array} \\ \left[\begin{array}{cccccccccccccccc} \underline{\mathbf{1}} & \underline{1} & \underline{1} & \underline{1} & \underline{1} & \underline{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & \underline{\mathbf{1}} & \underline{1} & \underline{1} & \underline{1} & \underline{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & \underline{\mathbf{1}} & \underline{\mathbf{1}} & \underline{1} & \underline{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & \underline{\mathbf{1}} \end{array} \right] \end{array}$$

Construction 3.2 yields $\mathcal{W}' = \{\mathbf{1}, \mathbf{4}, \mathbf{7}, \mathbf{8}\}$, with $I = 4$. Members of these maximal cliques which do not *belong to maximal cliques previously chosen in the construction*

have their clique matrix entries underlined above. Y' could be either of the sets $\{1, 7, 12, 16\}$ or $\{1, 7, 13, 16\}$. The occurrences of the y's in the maximal cliques for which they cause inclusion in the minimum cover are bolded in the matrix.

Before proceeding with minimal covers, it is necessary to perform a reduction of the interval order. We call $y \in X$ a $\subset$-minimal element, or a $\subset$-minimal for short, if its dual, y^*, is minimal in the subset ordering, or equivalently if y^* has no other dual as a proper subset. Let $ME(\underline{X})$ denote the set of $\subset$-minimals of $\underline{X}$; in particular, $ME(\underline{X})$ will include all simplicial elements. As an illustration, $ME(\underline{X}) = \{1, 7, 10, 13, 16\}$ for the interval order $\underline{X}$ of Example 3.1.

We note that $ME(\underline{X})$ is an embedded semi-order in $\underline{X}$ [7, Chapter 6]. We will order $\subset$-minimals according to the relation $<^-$, where for $y_1, y_2 \in ME(\underline{X})$,

$$y_1 <^- y_2 \Leftrightarrow \min y_1^* \sqsubset \min y_2^* \Leftrightarrow \max y_1^* \sqsubset \max y_2^*.$$

We assume henceforth that no two $\subset$-minimals share the same dual, an assumption that is easy to enforce in practice. The ordering $<^-$ corresponds to both the left– and the right–endpoint orderings $\prec^-$ and $\prec^+$ described in [7, §2.2]; it is a linear ordering on $ME(\underline{X})$. It is obvious that the minimal covers of $ME(\underline{X})$ are exactly the minimal covers of $\underline{X}$.

Construction 3.2 will yield a single *minimum* cover. The natural generalization from this setting is to broaden the choice for μ_i without changing the property that no two elements of Y are covered by the same maximal clique. This generalization leads to Construction 3.3.

Construction 3.3 *Let y_1 be a simplicial element of $ME(\underline{X})$ belonging to the first maximal clique M_1, and let $\mu_1 = M_1$. Form the sets $Y_1 = \{y_1\}$ and $\mathcal{W}_1 = \{\mu_1\}$. Then form the sets Y_i and $\mathcal{W}_i$ for $i > 1$ as follows:*

While i is such that $\{y \in ME(\underline{X}) : \mu_{i-1} \sqsubset y^*\} \neq \emptyset$, let $Y_i = Y_{i-1} \bigcup \{y_i\}$, where

$$y_i \in \left\{ y \in ME(\underline{X}) : \operatorname*{argmin}_y \max_{\mu_{i-1} \sqsubset y^*} y^* \right\},$$

is a particular selection of y_i, and let $\mathcal{W}_i = \mathcal{W}_{i-1} \bigcup \{\mu_i\}$, where

$$\mu_i \in y_i^* \setminus y_{i-1}^*$$

is a particular selection of μ_i.

The requirement that the y's be limited to the set of $\subset$-minimals can be illustrated using the interval order of Example 3.1. Suppose that Construction 3.3 is being applied and has so far generated $Y_2' = \{1, 7\}$ and $\mathcal{W}_2' = \{\mathbf{1}, \mathbf{4}\}$. If we do not restrict y_3 to belong to $Y(\underline{X})$, then we could use $y_3 = 12$. If the selection of maximal clique yields $\mu_3 = \mathbf{5}$, the next step of the Construction will force $\mathbf{7} \prec y_4$, bypassing element 13. It should be clear that maximal clique $\mathbf{5}$ belongs to no minimal cover, in this example.

Theorem 3.3 *For every sequence of pairs $(\mathcal{W}_i, Y_i), i = 1, \ldots$ formed in Construction 3.3, there exists $I \leq m$ such that $\{y : \mu_I \sqsubset y^*\} = \emptyset$. Defining $(\mathcal{W}, Y) = (\mathcal{W}_I, Y_I)$, the class of minimal covers of X is exactly the class of sets $\mathcal{W}$ which can be produced by Construction 3.3.*

A simple Lemma is necessary for the proof of Theorem 3.3.

Lemma 3.4 *Let $\mathcal{M}' \subset \mathcal{M}(\underline{X})$. Then there exists an interval ordering $\prec'$ such that $\mathcal{M}' = \mathcal{M}(X, \prec')$.*

Proof of Lemma 3.4. Since $(\mathcal{M}(\underline{X}), \sqsubset)$ is linearly ordered, $\left(\mathcal{M}', \sqsubset|_{\mathcal{M}' \times \mathcal{M}'}\right)$ is a linearly ordered set of maximal cliques. Therefore $\mathcal{M}'$ is the set of maximal cliques for some interval order $\prec' \subset X \times X$ [1]. □

Proof of Theorem 3.3. First we must show that I is well-defined for any selection of $y_i, \mu_i, i = 1, \ldots, I$ compatible with Construction 3.3. Since $\mu_{i-1} \sqsubset y_i^*$ and since $\mu_i \in y_i^*$, it must be that $\mu_{i-1} \sqsubset \mu_i$. Thus in at most m steps we will reach $\mu_I = M_m$. Since $M_m = \max \mathcal{M}(\underline{X})$ we obtain $\{y : \mu_I \sqsubset y^*\} = \emptyset$, and the construction terminates.

We now show the theorem's statement in three steps:

1. Every $\mathcal{W}$ produced by Construction 3.3 is a cover for X.
2. Every cover $\mathcal{W}$ produced by Construction 3.3 is minimal for X.
3. Every minimal cover for X is a set $\mathcal{W}$ compatible with the selection method of Construction 3.3.

- **Claim 1**: *Every $\mathcal{W}$ is a cover for X.*

Let $Y = Y_I$ and $\mathcal{W} = \mathcal{W}_I$ be particular realizations of Construction 3.3, and let $x \in X$ be arbitrary. We must show that $\mu \in x^*$ for some $\mu \in \mathcal{W}$. If $x \in Y$, x is covered by construction, so assume $x \in X \setminus Y$. Since $\mu_1 = M_1 \sqsubset x^* \sqsubset M_m = \mu_I$, there must be some i such that $\mu_i \sqsubset x^*$ but $\mu_{i+1} \not\sqsubset x^*$. Also, since we know y_{i+1} minimizes $\max\limits_{\mu_i \sqsubset y^*} y^*$ and since $\mu_i \sqsubset x^*$, we deduce that $\max y^*_{i+1} \leq \max x^*$.

We can now show that $\mu_{i+1} \in x^*$. Assume not; then we need $\max x^* \sqsubset \mu_{i+1}$. But then $\max x^* \sqsubset \mu_{i+1} \leq \max y^*_{i+1} \leq \max x^*$, a contradiction. Therefore x is contained in μ_{i+1}, so $\mathcal{W}$ is a cover.

- **Claim 2**: *Every cover $\mathcal{W}$ is minimal for X.*

Let Y and $\mathcal{W}$ be particular realizations of Construction 3.3 as before, and let $\mathcal{W}^{(i)} = \mathcal{W} \setminus \{\mu_i\}$ for some $i \in \{2, \ldots, I-1\}$, since neither $\mu_1 = M_1$ nor $\mu_I = M_m$ can be removed. We show that $\mathcal{W}^{(i)}$ is not a cover.

Since $\mu_{i-1} \sqsubset y^*_i$, we know that $\mu_{i-1} \notin y^*_i$, which implies that $\{\mu_1, \ldots, \mu_{i-1}\} \bigcap y^*_i = \emptyset$, by the linear ordering of the maximal cliques. By construction, $\mu_{i+1} \notin y^*_i$, which again implies that $\{\mu_{i+1}, \ldots, \mu_t\} \bigcap y^*_i = \emptyset$. Hence $\mathcal{W}^{(i)}$ does not cover y_i, and thus $\mathcal{W}$ is a minimal cover.

- **Claim 3**: *Every minimal cover for X is produced by the algorithm of the theorem.*

Let $V = \{\nu_1 = M_1, \nu_2, \ldots, \nu_{t-1}, \nu_t = M_m\}$ be a minimal cover. We need to supply a set $Y = \{y_1, \ldots, y_t\}$ such that V and Y satisfy Construction 3.3. For this it is enough to show that

$$\text{if } y_i \in \left\{ y : \operatorname*{argmin}_{y} \max_{\nu_{i-1} \sqsubset y^*} y^* \right\} \text{ for } i = 1, \ldots, I, \text{ then } \nu_i \in y^*_i \setminus y^*_{i-1}. \qquad (3.2)$$

Apply Construction 3.2 to V, which by Lemma 3.4 is the maximal clique set of some interval order. Since V is minimal, it is its own minimum cover. The set Y' produced in the construction satisfies requirement (3.2). $\square$

Though Construction 3.3 involves an arbitrary selection from a set, the available set from which to select μ_i does not depend on the particular choices of y_{i-1} and y_i, since only $\max y^*_{i-1}$ and $\max y^*_i$ play a role in the determination of that set. Algorithm 3.5 **ListMinCovers** below explicitly recognizes this fact by retaining only this information from the elements of Y. On the other hand, the choice of μ_{i-1} does affect the set of available y_i's at every step.

We can translate Theorem 3.3 to the following algorithm, which returns the set of all minimal covers of X when called with arguments $(X_{\min}, \emptyset)$, with $X_{\min}$ the set of $\subset$-minimals of X. The second argument is $\emptyset$ only on the first call, with the convention that $\emptyset \sqsubset M$ for all $M \in \mathcal{M}(\underline{X})$; otherwise it is a maximal clique, μ_0, which corresponds to $\max y^*_{i-1}$ and contributes to defining the set $y^*_i \setminus y^*_{i-1}$ of Construction 3.3.

Algorithm 3.5

ListMinCovers(Y, μ_0)
Arguments: Y, a set of $\subset$-minimals ordered according to $<^-$;
μ_0, a subset of Y.

```
    begin
    if Y = ∅
        return {∅}
    else
        M ← ∅
L1:     y_0 ← argmin_y max_{y∈Y} y*
L2:     for each μ ∈ {M : M ∈ y*_0, μ_0 ⊏ M}
R1:         M_0 ← ListMinCovers({y : μ ⊏ y*}, max y*_0)
L3:         for each W ∈ M_0
                M ← M ⋃ {{μ_0} ⋃ W}
    return M
    end
```

The call **ListMinCovers**$(ME(\underline{X}), \emptyset)$ will return the list of minimal covers of $\underline{X}$.

We now establish the time complexity of Algorithm 3.5. Let $n_0 \doteq |ME(\underline{X})|$ and $m_0 \doteq \max_{y \in Y} |y^*|$. Define $T(i)$ to be the time complexity of Algorithm 3.5 when called with a set $Y_i = \{y_{i+1}, y_{i+1}, \ldots, y_{n_0}\} \subset ME(\underline{X})$ of $\subset$-minimals, $i =$

$0, \ldots, n_0 - 1$. Let $N_{\max}(m)$ be the maximum number of minimal covers for an interval order with m maximal cliques. We show in §6.1 that $N_{\max}(m) \leq O(1.84^{m-1})$.

We use $k_0, k_1, k_2, k_3 \ldots$ to denote unknown constants. With **ListMinCovers** called with argument Y_i, $i < n_0$, statement L1 is executed in constant time k_1, since the ordering properties of $<^-$ on Y_i imply that $\operatorname{argmin}_y \max_{y \in Y} y^* = y_i$. The number of iterations of L2 is bounded above by $m - i - 1$, $i = 1, \ldots, m-1$. At the j^{th} iteration of L2, statement R1 will require time bounded above by $T(i+j) + k_2$, returning a set containing at most $N_{\max}(m - i - j)$ maximal cliques, $j = 1, \ldots, n_0 - i$. The total execution time of L3 within iteration j of L2 is thus bounded above by $k_3 N_{\max}(m - i - j)$. Thus $T(i) \leq k_1 + \sum_{j=1}^{n_0 - i} [T(i+j) + k_2 + k_3 N_{\max}(m - i - j)]$ for $i = 0, \ldots, n_0 - 1$. Positing $T(n_0) = k_0$, and given the bound on $N_{\max} m$, standard manipulations shows that the overall time complexity $T(0)$ is bounded above by $O(2^{\max(n_0, m-2)})$. We note that the $\subset$-minimality of elements of $ME(\underline{X})$ forces $n_0 \leq m = |\mathcal{M}(\underline{X})|$. Part 1 of Observation 3.6, below, shows that we can remove all simplicial elements from the set of $\subset$-minimals to run the algorithm, yielding an effective bound of $n_0 \leq m - 2$. Hence the time complexity of Algorithm 4.1 is bounded by $O(2^{m-2})$.

In practice, this exponential time complexity shows that tractability can be maintained as $n \to +\infty$ if m increases slowly enough with respect to n. In general, however, enumeration is impractical and we will need a random generation alternative to enumeration. In § 4, we see how Algorithm 3.5 can be adapted to a uniform Metropolis independence sampler, thereby converting issues of tractability to issues of efficiency.

In some settings involving a moderate amount of data, enumeration may remain a viable option. Apart from the reduction of the original X to its $\subset$-minimals, several other simplifications can be applied to the minimal cover enumeration procedure of Construction 3.3. These simplifications may provide substantial gain in practice, particularly in counting, rather than enumerating, minimal covers.

Observation 3.6

1. Every essential maximal clique in the minimal cover $\mathcal{W}$ may be included *a priori* in a listed cover, thereby restricting the $\subset$-minimal set $ME(\underline{X})$ to non-simplicial elements, as these elements can only generate their own essential maximal clique in Construction 3.3.

2. Assuming still that $\subset$-minimals are sorted according to the order of the initial maximal clique of their dual, we need only find minimal covers for $\subset$-minimals belonging to connected components of the interval graph, and combine them at the end of the procedure.

3. Maximal cliques which overlap exactly the same $\subset$-minimals will be mutually exclusive and interchangeable in any minimal cover produced by Construction 3.3. The classes of such elements can be kept track of and all but one element from each class deleted from the problem, to be reinstated after the covers are listed.

The above considerations lead to straightforward modifications to Algorithm 3.5 In the time complexity analysis, parts 1 and 2 of the above Observation effectively decrease n_0, while part 3 effectively decreases m_0, though the worst-case time-complexity remains unchanged.

4 Uniform sampling of minimal covers

4.1 Random generation

Since the number of minimal covers often precludes enumeration, a procedure for pseudo-random generation must be devised. A random minimal cover generation algorithm can be obtained from Algorithm 3.5. To generate one minimal cover, a maximal clique is selected uniformly at random from those available at each iteration. Algorithm 4.1 returns both a minimal cover and its probability of generation by the algorithm.

The call **RandomMinCover**$(ME\,(\underline{X}))$ described in Algorithm 4.1 will return a random minimal cover $\mathcal{W}$, the corresponding random chain of $\subset$-minimals $\mathcal{Y}$ and the probability of generation p.

Algorithm 4.1 does not generate minimal covers with uniform probability, but can be adapted to this task. The simplest way to do so is to base an independence Metropolis sampler [13] on the algorithm. A useful characteristic of the independence Metropolis sampler in our context is that it allows perfect sampling, unlike most instances of MCMC where only bounds on the total variation or some other measure of convergence are available.

Algorithm 4.1

RandomMinCover$(ME\,(\underline{X}))$
Argument: $ME\,(\underline{X})$, the set of $\subset$-minimals of interval order $\underline{X}$;
begin
$Y \leftarrow ME\,(\underline{X})$
$\mathcal{W} \leftarrow \emptyset$
$\mathcal{Y} \leftarrow \emptyset$
$\mu_0 \leftarrow \emptyset$
$p \leftarrow 1$
while $Y \neq \emptyset$
    begin
    $y_0 \leftarrow \underset{y}{\operatorname{argmin}} \max_{y^* \in Y} y^*$
    $T \leftarrow \{M \in y_0^*; \mu_0 \sqsubset M\}$
    $\mu_0 \leftarrow$ **ChooseRandom** (T)
    $p \leftarrow \dfrac{p}{|T|}$
    $\mathcal{W} \leftarrow \mathcal{W} \cup \{\mu_0\}$
    $\mathcal{Y} \leftarrow \mathcal{Y} \cup \{y_0\}$
    $Y \leftarrow \{y \in Y; \max y_0^* \sqsubset y^*\}$
    end
return $(\mathcal{W}, \mathcal{Y}, p)$
end

We briefly describe the Metropolis independence sampler for our application. We call *trial probability* of a minimal cover the probability that it is generated by Algorithm 4.1, and denote by $p(\mathcal{W})$ the trial probability of $\mathcal{W}$. The Markov chain of minimal covers is denoted $\mathcal{W}_k$, $k = 0, 1, , \ldots$. The independence Metropolis sampler starts with a minimal cover $\mathcal{W}_0$ taken from the $p(\cdot)$ distribution. Thereafter, given current chain state $\mathcal{W}_k$, Algorithm 4.1 generates a proposal $\mathcal{W}$ and a random uniform variate $U_k \in (0, 1)$. The next state in the chain is determined by

$$\mathcal{W}_{k+1} = \begin{cases} \mathcal{W}_k & \text{if } U_k \geq p(\mathcal{W}_k)/p(\mathcal{W}) \\ \mathcal{W} & \text{otherwise.} \end{cases} \tag{4.3}$$

The stationary distribution of the chain is uniform.

The ability to perform perfect simulation is a property of all Metropolis independence samplers for which the minimum and maximum trial probabilities as well as

their corresponding states are known. If we assign a weak ordering of the state space corresponding to the natural ordering of the trial probabilities, it is easy to see that the Markov chain given by transition rule (4.3) is monotone in the sense of Propp & Wilson [18]. Their method of coupling from the past can be used to achieve stationarity of the chain. Given a partial order on the minimal covers, coupling from the past will start two chains from a maximal and a minimal state at some fixed time in the past, and run them using the same proposals and the same sequence of U_k's. If the chains are in the same state at time zero, the distribution of the resulting minimal cover is exactly uniform. An overview of this form of coupling for independence Metropolis samplers is provided in [15].

4.2 Minimum and maximum trial probabilities

Though the implementation just described is perhaps the simplest possible form of coupling from the past, it requires knowledge of the minimum and maximum trial probabilities. In order to find an expression for these, we first reduce the problem along the lines of Observation 3.6, parts 1 and 2. Specifically, we remove all simplicial elements from the list of $\subset$-minimals, and we consider the problem only for a list of $\subset$-minimals which forms a connected incomparability graph. Otherwise, we can consider each component of the incomparability graph in turn, as the maximal cliques and $\subset$-minimals returned by Algorithm 4.1 from one component do not affect those returned from another component.

Algorithm 4.1 can be thought of as returning a list of pairs of maximal cliques and $\subset$-minimals, specifically a set $\{(\mu_{i_k}, y_{j_k}) : \mu_{i_k} \in \mathcal{W}, y_{j_k} \in \mathcal{Y}\}$ where $i_k < i_l$ and $j_k < j_l$ whenever $k < l$. We call such a set a minimal cover list. Since minimal cover lists and minimal covers are in one-to-one relation, we can analyze trial probabilities using minimal cover lists.

Let $S(s,t)$ be the set of all sequences of the form

$$[(\mu_{i_1}, y_{j_1} = y_1), (\mu_{i_2}, y_{j_2}), ..., (\mu_{i_r} = \mu_s, y_{j_r} = y_t)]$$

where $\mu_{i_k} \in y^*_{j_k}$, $k = 1, \ldots, r$, $\mu_{i_k} \in \mathcal{M}(\underline{X})$ and $y_{j_k} \in ME(\underline{X})$ for $k = 1, \ldots, r$, and $\mu_{i_1} \sqsubset \mu_{i_2} \sqsubset \ldots \sqsubset \mu_{i_r} = \mu_s$ and $y_1 = y_{j_1} <^- y_{j_2} <^- \cdots <^- y_{j_r} = y_t$. Note that $S(s,t)$ is larger than the set of minimal cover lists, in that it contains minimal cover sublists as well as non-minimal and non-covering lists.

We set a probability measure P on $S(s,t)$ based on the trial probabilities. Specifically, for $\sigma \in S(s,t)$, $P[\sigma]$ is the probability that the Algorithm 4.1 produces the list or sublist σ at any step in its execution.

Define $g_{s,t}$ by

$$g_{s,t} = \begin{cases} 0 & \text{if } P[\sigma] = 0 \text{ for all } \sigma \in S(s,t) \\ \min\{P[\sigma] > 0 : \sigma \in S(s,t)\} & \text{otherwise.} \end{cases}$$

It must then be true that

$$g = \min\{g_{s,t} > 0 : 1 \leq s \leq m, 1 \leq t \leq q\}$$

corresponds to the minimum trial probability of a minimal cover. To see why, suppose that $\sigma(s,t)$ has minimum non-zero trial probability. If the corresponding maximal clique set $\{\mu_{i_1}, \ldots, \mu_s\}$ is not a minimal cover, there exist a superset of it which is, otherwise $P[\sigma(s,t)] = 0$. Since the probability of generating a minimal cover list is never greater than that of generating a sublist included in it, there is therefore a minimal cover with minimal non-zero trial probability.

We adopt the convention that $[\mu_s]$, $[y_t]$ and $[(\mu_s, y_t)]$ denote the events that μ_s, y_t and (μ_s, y_t), respectively, are generated by Algorithm 4.1 as part of a returned minimal cover list. Consider now the following decomposition of $P[\sigma]$ for $\sigma \in S(s,t)$:

$$\begin{aligned} P[\sigma] &= P\left[(\mu_s, y_t), (\mu_{i_{r-1}}, y_{j_{r-1}}), \ldots, (\mu_{i_1}, y_1)\right] \\ &= P\left[\mu_s |\, y_t, y_{j_{r-1}}\right] P\left[y_t |\, (\mu_{i_{r-1}}, y_{j_{r-1}})\right] P\left[(\mu_{i_{r-1}}, y_{j_{r-1}}), \ldots, (\mu_{i_1}, y_1)\right] \end{aligned}$$

In the above expression,

$$\begin{aligned} P\left[\mu_s |\, y_t, y_{j_{r-1}}\right] &= \mathbb{1}\left[\mu_s \in y_t^* \setminus y_{j_{r-1}}^*\right] \left| y_t^* \setminus y_{j_{r-1}}^* \right|^{-1} \\ P\left[y_t |\, (\mu_{i_{r-1}}, y_{j_{r-1}})\right] &= \mathbb{1}\left[y_t = \operatorname*{argmin}_y \left\{y : \mu_{i_{r-1}} \sqsubset y^*\right\}\right]. \end{aligned}$$

where the latter indicator is 0 if $\left\{y : \mu_{i_{r-1}} \sqsubset y^*\right\} = \emptyset$.

Now set

$$z_{s,t} = \max_{\substack{s_0 < s \\ t_0 < t}} \left\{ \mathbb{1}\left[\mu_s \in y_t^* \setminus y_{t_0}^*\right] \mathbb{1}\left[y_t = \operatorname*{argmin}_y \{y : \mu_{s_0} \sqsubset y^*\}\right] g_{s_0,t_0} \right\},$$

and we can write

$$g_{s,t} = \begin{cases} 0 & \text{if } z_{s,t} = 0 \\ \min\limits_{\substack{s_0<s \\ t_0<t}} \dfrac{g_{s_0,t_0}}{|y_t^* \setminus y_{t_0}^*|} & \text{otherwise.} \end{cases} \tag{4.4}$$

Thus, given all g_{s_0,t_0}, it is possible to scan all pairs (μ_{s_0}, y_{t_0}) for $s_0 < s$ and $t_0 < t$ to obtain the minimum trial probability of generating a sequence ending with (μ_s, y_t).

The maximum trial probability can be determined in much the same way. Following an argument similar to the one leading to (4.4), we define

$$h_{s,t} = \max_{\sigma(s,t) \in S(s,t)} P[\sigma(s,t)].$$

The following recurrence relation then holds:

$$h_{s,t} = \max_{\substack{s_0<s \\ t_0<t}} \frac{h_{s_0,t_0}}{|y_t^* \setminus y_{t_0}^*|}.$$

Determining the covers corresponding to the minimum and maximum trial probabilities requires a small amount of bookkeeping associating minimal covers with the relevant index pair (s,t). Because the procedure used follows the $<^-$ ordering of the $\subset$-minimals, the relevant index pair will be such that $\mu_s \in y_q^*$, where $q = |ME(\underline{X})|$. The minimum trial probability minimal cover can thus be obtained from the index pair

$$(s_0, t_0) = \operatorname*{argmin}_{(s,t)} \left\{ g_{s,t} > 0 : \mu_s \in y_q^* \right\}$$

while the index pair of the maximal trial probability minimal cover will be

$$(s_1, t_1) = \operatorname*{argmax}_{(s,t)} \left\{ h(s,t) : \mu_s \in y_q^* \right\}.$$

Aside from the possibility of performing perfect sampling for minimal covers, the question of the efficiency of the independence Metropolis sampler has been addressed by [15] and [20]. In the present case, putting $p^{(k)}(\mathcal{W})$ to be the probability that $\mathcal{W}_k = \mathcal{W}$ in a chain started from some distribution $p(\cdot)$ and letting N be the

number of minimal covers, we obtain the geometric convergence rate

$$\left| p^{(n)}(\mathcal{W}) - 1/N \right| \leq (1 - N p_{\min})^k$$

where $p_{\min}$ is the minimum trial probability. An expression of this bound wholly in terms of q or m and of total variation is still missing.

5 Maximal removable sets

We now turn our attention to maximal removable sets (MRS). In Definition 2, an MRS was defined to be a set of maximal cliques such that its complement, with respect to $\mathcal{M}(\underline{X})$, is a minimal cover. An equivalent characterization of minimal covers can be obtained in terms of removable sets. A set of maximal cliques is removable if no sequence of maximal cliques in the set is equal to x^* for any $x \in X$. A removable set is maximal if adding any maximal clique creates such a sequence. Hence, an MRS is maximal among removable sets under inclusion ordering.

Algorithm 5.2 is discussed in this section for expository purposes. It stands in stark contrast to the backtracking approach of Algorithm 3.5, and may eventually suggest novel methods for random generation.

5.1 Simplifying Assumptions

For any minimal cover or maximal removable set problem, we can assume that the only essential maximal cliques are M_1 and M_m. If any other maximal clique, say M_r, is essential, the problem can be split into two subproblems: one dealing with M_1 through M_r and the other with maximal cliques M_r through M_m. The solution to the large problem is simply the union of the solutions to the two subproblems.

For a minimal cover this statement would be negated if there existed a maximal clique which belonged to the minimal cover of one subproblem but which could be removed once the union were formed. This would imply that the element x covered by this maximal clique had now been covered by a maximal clique from the other subproblem. However, by the contiguity (consecutive-one's property) of the dual, any such maximal clique would be covered by M_r, and hence a contradiction would arise. Since every MRS is the complement of a minimal cover, the problems are equivalent and the same simplification is obtained.

We also assume that any x such that x^* contains an essential maximal clique has been removed from the problem, since it is covered by the essential maximal clique. In other words, since the essential maximal clique must have been retained, the element x imposes no other restrictions on the MRS.

5.2 MRS generation algorithm

Once the simplifying assumptions of Section 5.1 have been applied we may assume that we are dealing with maximal cliques M_2 through M_{m-1} and that none of these maximal cliques are essential.

The following simple Lemma can be proven under these assumptions.

Lemma 5.1 *For any i such that $2 \leqslant i \leqslant m-3$ both M_i and M_{i+2} can be removed together.*

Proof. Suppose not. Then there must exist an x such that x^* includes M_i and M_{i+2} but not M_{i+1}. For all interval orders x^* must be contiguous hence no such x can exist. □

The basis of our MRS generating algorithm lies in the following two properties of interval orders.

Property I *If M_i is the largest maximal clique in a removable set, then M_{i+2} can be removed for $i < m-2$.*

Property II *There can exist MRSs containing both M_i and M_{i+3} and neither M_{i+1} nor M_{i+2}.*

Property I is simply a special case of Lemma 5.1 and hence holds. Property II can hold when $x_1^* = \{M_a, \ldots, M_i, M_{i+1}\}$, $a \leq i$, and $x_2^* = \{M_{i+2}, M_{i+3}, \ldots, M_b\}$, $i+3 \leq b$. In this case, if $\{M_a, \ldots, M_i\} \bigcup \{M_{i+3}, \ldots, M_b\}$ is removable and removed then neither M_{i+1} nor M_{i+2} can be removed.

In principle, our algorithm consists of enumerating all removable sets and then deleting from the enumeration those that are proper subsets of any other removable set.

The efficiency of the algorithm can be optimized by performing the deletions while enumeration is taking place.

We first note the form of removable sets. We maintain the subscript ordering described in Section 2.2. Consider the subscripts of any two adjacent elements of an MRS. They must follow one of three patterns: (M_i, M_{i+1}), (M_i, M_{i+2}), or (M_i, M_{i+3}). It is not possible for (M_i, M_{i+4}) to occur in sequence in an MRS because, by a variant of Property I, it is possible to remove M_{i+2} for any removable set which includes (M_i, M_{i+4}), and hence the removable set is not maximal.

This points to a further important property of MRSs. Consider constructing an MRS starting with either M_2 or M_3 and subsequently adding maximal cliques with larger indices. The partially constructed list will be termed a candidate set. Define the *head* of a candidate set as the set of maximal cliques with subscripts contiguous with the largest maximal clique in that candidate set. Any two candidate sets with the same head will evolve in exactly the same manner. A gap in the sequence of maximal cliques making up a removable set provides a kind of independence: the future evolution of a candidate set does not depend on those maximal cliques preceding the gap.

For example, suppose two candidate sets are $\{M_2, M_3, M_5\}$ and $\{M_3, M_5\}$. There is then no point in pursuing $\{M_3, M_5\}$ since its evolution can be no different from that of $\{M_2, M_3, M_5\}$ and therefore it cannot be maximal. On the other hand if one candidate is $\{M_2, M_3, M_4\}$ and a second is $\{M_3, M_4\}$ then such elimination is not possible. While the smaller is a proper subset of the larger, they do not share the same head.

From the preceding argument we can deduce that there are at most 3^{m-2} sequences to be examined. However, the number is actually much smaller. One needs to consider (M_j, M_{j+3}) only if (M_j, M_{j+1}) is not possible. Hence the true upper bound is 2^{m-2}. A further substantial saving comes from identifying and deleting candidates that will become proper subsets of other candidates.

Algorithm 5.2

ListMaxRemovableSets(S)
Arguments: S a list of candidates.

```
begin
    if S = ∅ then
        return ListMaxRemovableSets({{M2, M\M2}}, {M3, M \ {M2, M3}})
    else
        begin
        S* ← ∅
        for each w = {s, M} in S
            begin
                if isCovered (s, S*, S) break
                S* = Append ( S*, {(s ∪ M2), M\ {M1, M2}})
                if isRemovable (s ∪ M1)
                    S* = Append ( S*, {(s, M1), M\M1})
                else if isRemovable (s ∪ M3)
                    S* = Append (S*, {(s, M3), M\ {M1, M2, M3}})
            end
        if Finished (S*)
            return (S*)
        else
            return ListMaxRemovableSets(S*)
        end
end
```

Our algorithm *grows* candidate lists starting at the left end. By Property I either M_2 or M_3 must be the first element of the MRS. From this point elements are added sequentially to the candidate lists according to Properties I and II and the observations made above.

Any candidate which is a proper subset of another candidate can be deleted if both have the same head ; this is detected by the function **isCovered**. The function **isRemovable** determines whether a given candidate can be removed. Finally the function **Finished** determines whether all candidates have been pursued to M_m.

Since the workings of Algorithm 5.2 are slightly more complex than those of Algorithms 3.5, we provide an example in Appendix A.

6 The number of minimal covers of an interval order

How many minimal covers does $\underline{X}$ have? The question is not only of theoretical interest. In applications, if this number grows too large for the size of $\underline{X}$, we might opt for random generation of minimal covers instead of enumeration.

The number of minimal covers will depend on four characteristics of the interval order: the number of maximal cliques, the number of $\subset$-minimals, the cardinality of each $\subset$-minimal and the amount of overlap between the $\subset$-minimals. The problem of determining the number of minimal covers for a general interval order is thus fairly complex.

We aim at bounding the maximal number of minimal covers $N_{\max}(m)$ for an interval order with m maximal cliques. The next section provides lower and upper bounds on $N_{\max}$, while the section following provides some simulation results designed to supply an empirical approximation for this number and to determine the applicability of Algorithm 3.5 in a realistic situation.

6.1 Bounds on $N_{\max}(m)$

Theorem 6.1 *For $m \geq 2$ and under the simplifying assumptions of § 5.1,*

$$\left\lfloor (0.69)1.44^{m-1} \right\rfloor \leq N_{\max}(m) \leq \left\lceil (0.62)1.84^{m-1} + 0.5 \right\rceil \tag{6.5}$$

Proof. An immediate lower bound on the maximum number of minimal covers $N_{\max}(m)$ for a given m can be determined if we assume no overlap between $\subset$-minimals. The number of minimal covers ${}^{\text{no}}N_{\max}$ in the case of no overlap is then simply the product of the cardinalities of the $\subset$-minimals. Since M_1 and M_m are essential and both contain a $\subset$-minimal of cardinality 1, we need to partition $m-2$ elements into consecutive sequences such that the product of the lengths of these sequences is maximized. The problem can be reformulated as that of maximizing $\prod_{k=1}^{K} a_k$ subject to $\sum_{k=1}^{K} a_k = m-2$ for K and $a_1, \ldots, a_k$, for which a solution is exposed in [19, Theorem 4-6]. For all values of $m \geq 2$, ${}^{\text{no}}N_{\max} \geq 3^{\frac{m-2}{3}} = (3^{-1/3})(3^{1/3})^{m-1}$.

For the upper bound, we bound the number of MRSs compatible with m maximal cliques by using the properties of MRSs described in Section 5.2. Specifically, we

determine the number B_m of ways in which a sequence of $m-1$ (ordered) items can be partitioned exactly in consecutive groups of 1, 2 or 3 items, such that a group of 3 items occurs neither at the beginning nor at the end of the sequence. If we take the items to be the linearly ordered maximal cliques of an interval order beginning at M_2, then a set formed by the final maximal clique of each group, bar the last group, will form a set of maximal cliques of the requisite structure for an MRS. The number of such partitions thus forms an upper bound on the number of MRSs achievable with m maximal cliques.

The reason for starting the sequence at M_2 and omitting to include the final maximal clique from the last group is that M_1 and M_m are essential, and thus cannot be removed. The requirement that a group of 3 neither start nor end a sequence of groups formed in this manner ensures that at least one of maximal cliques M_2 and M_3, and at least one of M_{m-2} and M_{m-1}, will be included in the candidate set, the necessity of which is stated in Section 5.2.

To determine the number of groupings of $m-1$ items in sequence which satisfy the above conditions, we start with the determination of the number G_m of general groupings in 1, 2 or 3 of $m-1$ items which do not necessarily satisfy the condition on the first and last group. Since each such sequence must start with a group of length $i = 1, 2$ or 3, and since the remainder of the sequence of $m-i-1$ elements must satisfy the grouping requirements as well, it is clear that $G_m = G_{m-1} + G_{m-2} + G_{m-3}$ for $m > 4$. To determine the number B_m of group sequences of $m-1$ items which satisfy the requirement of not starting or ending with a group of length 3, we simply subtract the number of unconstrained group sequences which start or end with a grouping of 3. There are $2G_{m-3} - G_{m-6}$ such groupings, so that $B_m = G_m - 2G_{m-3} + G_{m-6}$ for $m > 4$. A simple algebraic verification shows that $B_m = B_{m-1} + B_{m-2} + B_{m-3}$ for $m > 4$. Values for $2 \leq m \leq 4$ are depicted in the following diagrams. Only the endpoint of each group is identified by a circle ($\circ$).

M_2	$\circ$

$m = 2$, $B_2 = 1$

M_2	$\circ$	
M_3	$\circ$	$\circ$

$m = 3$, $B_3 = 2$

M_2	$\circ$	$\circ$	
M_3	$\circ$		$\circ$
M_4	$\circ$	$\circ$	$\circ$

$m = 4$, $B_4 = 3$

Since $B_2 = 1$, $B_3 = 2$, $B_4 = 3$, and $B_m = B_{m-1} + B_{m-2} + B_{m-3}$ for $m > 4$, $B_m = T_{m-1}$, where T_m is the m^{th} Tribonacci number [4]. It was shown in [21] that

$T_m = \lceil \alpha\rho^m + 0.5 \rceil$, where α and ρ are given by

$$\begin{aligned}
\alpha &= \frac{1}{9\sqrt{33}}\left(\kappa_1^2 + \kappa_2^2\right) + \frac{5}{2}\left(\kappa_1 + \kappa_2\right) + \frac{1}{3} \approx 0.6184, \\
\rho &= \frac{1}{3}\left(\kappa_1 + \kappa_2 + 1\right) \approx 1.8393 \\
\text{for } \kappa_1 &= \sqrt[3]{19 + 3\sqrt{33}} \\
\text{and } \kappa_2 &= \sqrt[3]{19 - 3\sqrt{33}},
\end{aligned}$$

which completes the proof. □

6.2 Simulations and simulation results

The above theorem shows that the maximal number of minimal covers grows exponentially large quite rapidly, lying somewhere between the curves $y = (0.48)1.44^m$ and $y = (0.33)1.84^m$ for $m > 1$. Having in mind practical applications for minimal covers, we wish to determine whether, in practice, the number of minimal covers tends to reach these large values. We adopted a simulation approach with this purpose in mind, with the objectives of assessing the bounds mentioned above and of determining the behaviour of the number of minimal covers of an interval order likely to occur as a real data set. The simulation results we present are based on two pseudo-random interval order generation mechanisms designed to provide answers to both of these objectives.

The generation mechanisms we present warrant a preliminary explanation. It was shown in [6] that all interval orders $\underline{X}$ can be represented as sets of intervals on the real line

$$\left\{[f(x), f(x) + \rho(x)] : x \in X, f : X \mapsto \mathbb{R}, \rho : X \mapsto \mathbb{R}^+\right\},$$

characterized by left endpoint function f and non-negative length function ρ, and such that for $x, y \in X$, $x \prec y$ if and only if $f(x) + \rho(x) < f(y)$. Thus it is enough, in order to generate random finite interval orders, to generate left and right endpoints defining real intervals.

In simulation series A, we systematically varied the number n of elements in X between 10 and 115, and the ratio of expected left-endpoint placement to inter-

val length in the real representation. We produced real representations of interval orders by generating left endpoints according to an Exponential distribution with mean 1, then generating lengths according to an Exponential distribution with mean $\mu = 1/\lambda$, where λ took on the values $0.4, 0.6, \ldots, 3.0$. A simple calculation shows that the probability of overlap between any two intervals in such a setup is $1/(1+\lambda)$, and so varied between 0.25 and 0.72 in the course of our simulation. By way of comparison, the proportion of pairs of overlapping intervals in the breast cosmesis data presented in [5], was approximately 0.45. For each pair (n, λ), 50 interval orders were generated, thus yielding 14,300 interval orders in total. Simulation series A was designed to produce a large variety of overlapping patterns which would help assess the bounds of Equation (6.5) (see Figure 1).

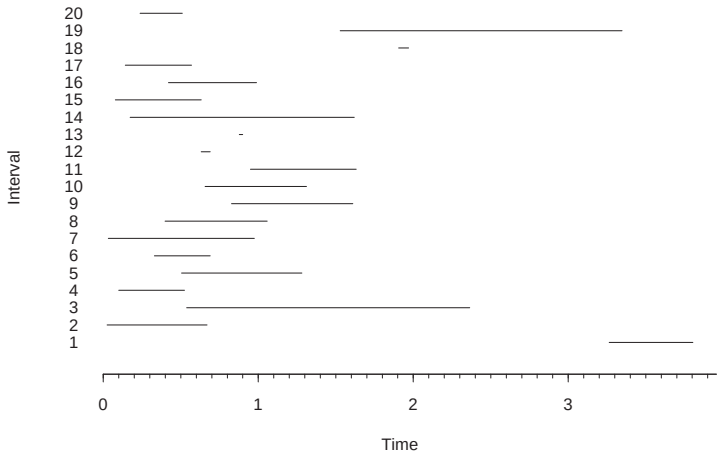

Figure 1: Series A sample interval set; $n = 20, \lambda = 0.5$.

For simulation series B, we generated intervals by setting potential inspection times at $t = 0, 1, 2, \ldots, 30$. For each interval , the inspection times were retained with a probability of 1 for $t = 0$, a probability of 0.4 for $t = 2, \ldots, 6$, and a probability of 0.1 for $t = 7, \ldots, 30$. A number n of event times were generated according to an Exponential distribution with mean μ, $\mu = 2, 2 \times (1.25) = 3.5, 2 \times (1.25)^2 = 3.125, \ldots, 2 \times (1.25)^{15} = 56.8$, for values of $n = 10, 20, \ldots, 150$. Intervals were formed by using the largest inspection time smaller than the event time as the left endpoint, and the smallest inspection time larger than the event time as the right endpoint. This setup mimics a long-term prospective study in which a condition is monitored at fixed inspection times which may be missed; event time corresponds to the

moment of change in condition. This simulation setup can produce intervals without a finite right endpoint, i.e. right-censored data, with a probability of $\exp(-30/\mu)$. This probability ranged over $0.0667, 0.0833, \ldots, 0.53$. Sixty simulations were run for each pair (n, μ), thus yielding a total of 13,500 simulated interval orders. Simulation series B was designed to mimic typical data from a long term prospective study where a condition is periodically monitored for change (see Figure 2).

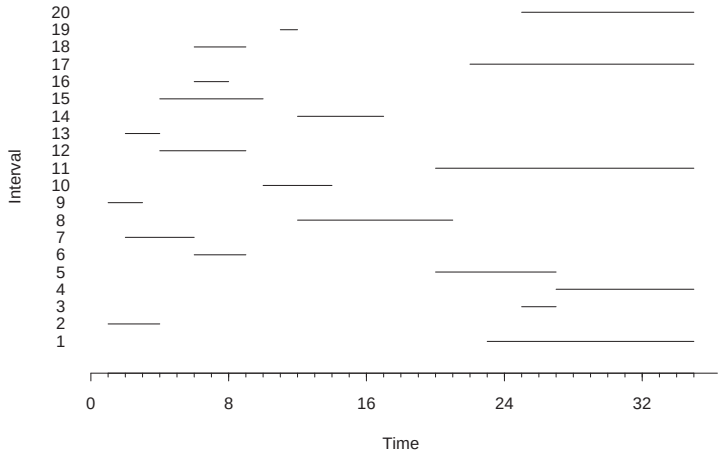

Figure 2: Series B sample interval set; $n = 20, \mu = 15$. (Intervals 1, 4, 11, 17 and 20 are right censored)

Simulation results are illustrated graphically using boxplots. We adhered to standard conventions in the use of boxplots. The range represented by the inner shaded box corresponds to values of N lying between the first and third quartile of the data, with the line within this box indicating the median of the values. The difference between third and first quartiles is called the *interquartile distance*. Whiskers are drawn below to the nearest value which does not fall short of the first quartile minus 1.5 times the interquartile distance, and above to the nearest value which does not exceed the third quartile plus 1.5 times the interquartile distance. Values exceeding the whiskers above and below are drawn individually as small line segments. Details on boxplots, their use and interpretation can be found in any good elementary applied statistics textbook.

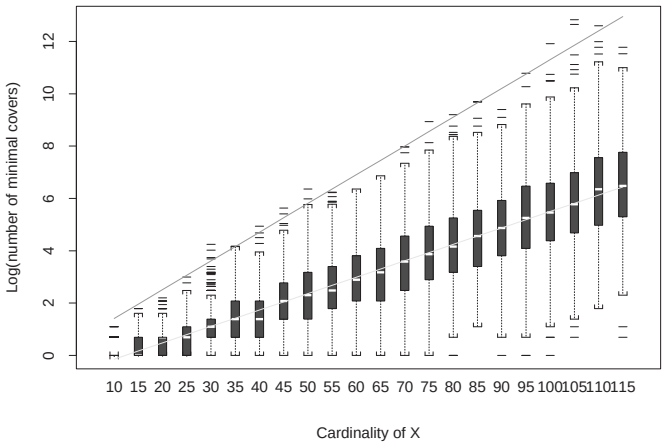

Figure 3: N as a function of n for simulation series A. The lower solid line is a least-squares regression line of N on n (intercept $\approx \log(0.461)$, slope $\approx \log(1.065)$); the upper solid line is a least-squares regression line of $\max_n N$ on n [intercept $\approx \log(1.37)$, slope $\approx \log(1.12)$].

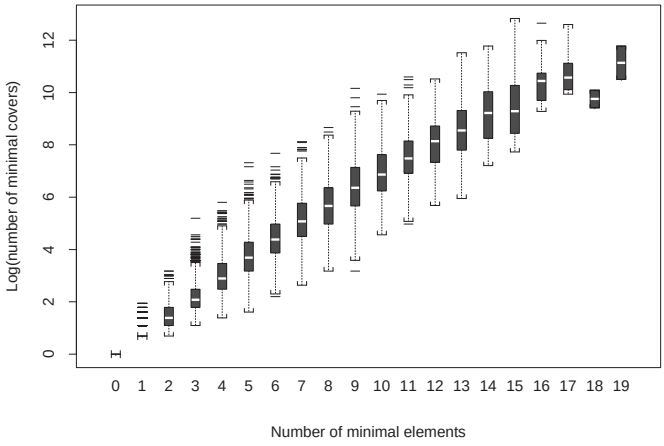

Figure 4: N as a function of k for simulation series A.

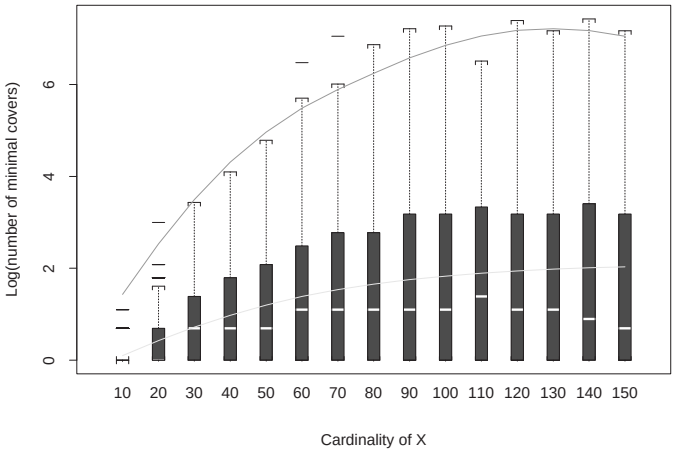

Figure 5: N as a function of n for simulation series B. Smooth curves were computed using local regression (*loess*).

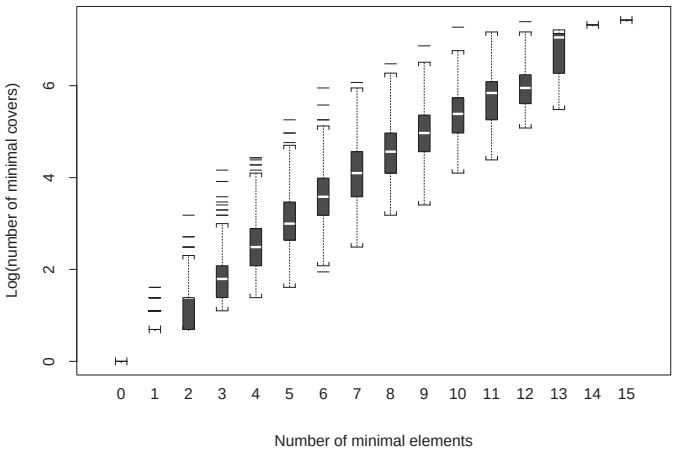

Figure 6: N as a function of k for simulation series B.

The results of simulation series A are shown in Figures 3 and 4. The exponential growth of the number of minimal covers N with the cardinality n of $\underline{X}$ is manifest in both the average case and the maximal case, placing an exponential lower bound on this growth for the general interval order. An approximate value of ${}^{\mathrm{A}}N_{\max}(m)$, the maximum number of minimal covers in terms of m for simulation series A, can be determined by a Poisson regression of m on n with identity link. Such a regression yields $m \approx 0.366n + 0.805$, from which we can derive the approximation ${}^{\mathrm{A}}N_{\max}(m) \approx (0.485)1.023^m$ (see also the results quoted in the caption of Figure 3). However, Figure 4 shows, as was expected, that this growth is more strongly associated with the growth of the number k of $\subset$-minimals in X than with the growth of n, though both quantities are positively correlated. Because simulations were not run an equal number of times for each value of k, the range of values of N as a function of k should not be interpreted as meaningful on the boxplot.

By contrast, simulation series B shows that the rate of increase of N with respect to n diminishes with increasing n in the average case, while the maximum value of N itself remains more or less constant for $n \geq 80$ (Figure 5). These results are explained by the fact that the simulation series B setup creates an expected proportion of right censored values which increases with μ; this increase causes the number of maximal cliques to converge in probability to 1 as μ grows larger. In the limit, all intervals overlap, forming a single maximal clique and a single minimal cover. Thus N tends to 1 in probability. Figure 6, by contrast, shows that the relationship between the number of $\subset$-minimals and the number of minimal covers remains roughly exponential, which indicates that right-censoring curbs the number of minimal covers by preventing the creation of large numbers of $\subset$-minimals.

6.3 Perspectives

The results of simulation series A confirm a rapid exponential growth in the number of minimal covers, which will preclude their enumeration even for modestly sized interval orders. However, simulation series B provides an indication that some realistic censoring mechanisms, at least, will maintain the number of minimal covers at manageable levels. In this case, the censoring mechanism is based on fixed inspection times, forces the number of maximal cliques, and thus the number of $\subset$-minimals, to be at most the number of inspection times. We can expect this phenomenon to curb the value of N to manageable values in some applications.

Refinining the bound on $N_{\max}(m)$ remains an open problem of mostly theoretical interest.

References

[1] G. BEHRENDT (1988). Maximal antichains in partially ordered sets. *Ars Combin.* **25C**, 149–151.

[2] A. P. DEMPSTER, N. M. LAIRD & D. B. RUBIN (1977). Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion). *J. Roy. Statist. Soc. Ser. B* **39**, 1–38.

[3] R. P. DILWORTH (1950). A decomposition theorem for partially ordered sets. *Ann. of Math.* **51**, 161–166.

[4] M. FEINBERG (1963). Fibonacci-tribonacci. *Fibonacci Quart.* **1**, 71–74.

[5] D. M. FINKELSTEIN (1986). A proportional hazards model for interval-censored failure time data. *Biometrics* **42**, 845–854.

[6] P. C. FISHBURN (1973). Interval representations for interval orders and semiorders. *J. Math. Psych.* **10**, 91–105.

[7] P. C. FISHBURN (1985). *Interval Orders and Interval Graphs*. Wiley, New York.

[8] F. GAVRIL (1972). Algorithms for minimum coloring, maximum clique, minimum covering by cliques, and maximum independent set of a chordal graph. *SIAM J. Comput.* **1**, 180–187.

[9] R. GENTLEMAN & C. J. GEYER (1994). Maximum likelihood for interval censored data: Consistency and computation. *Biometrika* **81**, 618–623.

[10] R. GENTLEMAN & A. C. VANDAL (2001). Computational algorithms for censored data using intersection graphs. *J. Comput. Graph. Statist.* **10**, 403–421.

[11] M. C. GOLUMBIC (1980). *Algorithmic Graph Theory and Perfect Graphs*. Academic Press, New York.

[12] M. G. GU & C.-H. ZHANG (1993). Asymptotic properties of self-consistent estimators based on doubly censored data. *Ann. Statist.* **21**, 611–624.

[13] W. K. HASTINGS (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika* **57**, 97–109.

[14] C. G. LEKKERKERKER & J. C. BOLAND (1962). Representations of a finite graph by a set of intervals on the real line. *Fundam. Math.* **51**, 45–64.

[15] J. S. LIU (1996). Metropolized independent sampling. *Statistics and computing* pages 113–119.

[16] P. A. MYKLAND & J.-J. REN (1996). Algorithms for coomputing self-consistent and maximum likelihood estimators with doubly censored data. *Ann. Statist.* **24**, 1740–1764.

[17] R. PETO (1973). Experimental survival curves for interval censored data. *Appl. Statist.* **22**, 86–91.

[18] J. G. PROPP & D. B. WILSON (1996). Exact sampling with coupled Markov

chains and applications to statistical mechanics. In *Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995)*, volume 9, pages 223–252.

[19] T. L. SAATY (1970). *Optimization in Integers and Related Extremal Problems*. McGraw-Hill, New York.

[20] R. L. SMITH & L. TIERNEY (1996). Exact transition probabilities for the independence metropolis sampler. From the MCMC preprint server.

[21] W. R. SPICKERMAN (1982). Binet's formula for the tribonacci sequence. *Fibonacci Quart.* **20**, 118–120.

[22] B. W. TURNBULL (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *J. R. Statist. Soc. B* **38**, 290–295.

[23] A. C. VANDAL, R. GENTLEMAN & X. LIU (2005). Constrained estimation and likelihood intervals for censored data. *Can. J. Statist.* **33**, 71–83.

[24] C. F. J.WU (1983). On the convergence properties of the EM algorithm. *Ann. Statist.* **11**, 95–103.

[25] Q. YU, L. LI & G. Y. C. WONG (2000). Consistency of the self-consistent estimator of survival functions with interval-censored data. *Scand. J. Statist.* **27**, 35–44.

A Example of the MRS algorithm

Consider the output of Algorithm 5.2 applied to the data of Example 3.1. In that example only M_1 and M_8 are essential.

Step 1 $S_0 = \{M_1, M_2\}$

$T_1 = \{M_2, M_3\}$	removable
$T_2 = \{M_2, M_4\}$	removable
$T_3 = \{M_3, M_4\}$	removable
$T_4 = \{M_3, M_5\}$	removable

Step 2 $S_1 = \{\{M_2, M_3\}, \{M_2, M_4\}, \{M_3, M_4\}, \{M_3, M_5\}\}$

$T_1 = \{M_2, M_3, M_4\}$	**not removable** $x_7^* \subset T_1$
$T_1' = \{M_2, M_3, M_6\}$	removable
$T_2 = \{M_2, M_3, M_5\}$	removable
$T_3 = \{M_2, M_4, M_5\}$	removable
$T_4 = \{M_2, M_4, M_6\}$	removable
$T_5 = \{M_3, M_4, M_5\}$	**not removable** $x_{10}* = T_5$
$T_5' = \{M_3, M_4, M_7\}$	removable

$T_6 = \{M_3, M_4, M_6\}$ removable

$\{M_3, M_5\}$ is not considered because it has the same head as T_2.

Step 3 $S_2 = \{\{M_2, M_3, M_6\}, \{M_2, , M_3, M_5\}, \{M_2, M_4, M_5\},$
$\{M_2, M_4, M_6\}, \{M_3, M_4, M_7\}, \{M_3, M_4, M_6\}\}$

$T_1 = \{M_2, M_3, M_6, M_7\}$ **not removable** $x^*_{13} \subset T_1$

$T'_1 = \{M_2, M_3, M_5, M_6\}$ removable

$T_2 = \{M_2, M_3, M_5, M_7\}$ removable

$T_3 = \{M_2, M_4, M_5, M_6\}$ removable

$T_4 = \{M_2, M_4, M_5, M_7\}$ removable

$T_5 = \{M_2, M_4, M_6, M_7\}$ **not removable** $x^*_{13} \subset T_5$

$T'_5 = \{M_3, M_4, M_6\}$ removable

$T_6 = \{M_3, M_4, M_7\}$ removable

Done $S = \{\{M_2, M_3, M_5, M_6\}, \{M_2, M_3, M_5, M_7\}, \{M_2, M_4, M_5, M_6\},$
$\{M_2, M_4, M_5, M_7\}, \{M_3, M_4, M_6\}, \{M_3, M_4, M_7\}\}$