# Some characterization of clique trees and their application to statistical inference

Akimichi Takemura

Graduate School of Information Science and Technology

University of Tokyo

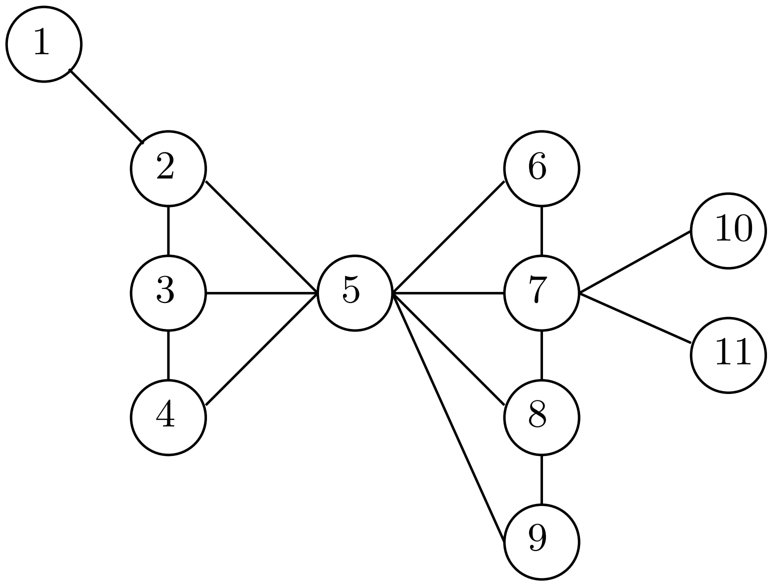
(Joint works with Hisayuki Hara and Satoshi Aoki)

# Contents

# Chordal graphs

- A graph is "chordal" if every circle of length $\geq 4$ has a chord

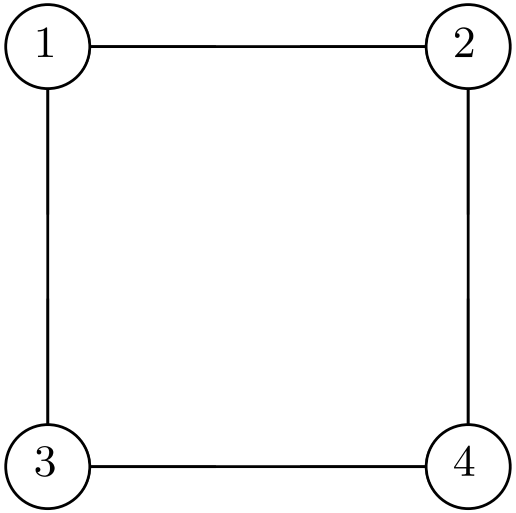
Example of a chordal graph



# Chordal graphs

- A graph is "chordal" if every circle of length $\geq 4$ has a chord

Example of non-chordal graph

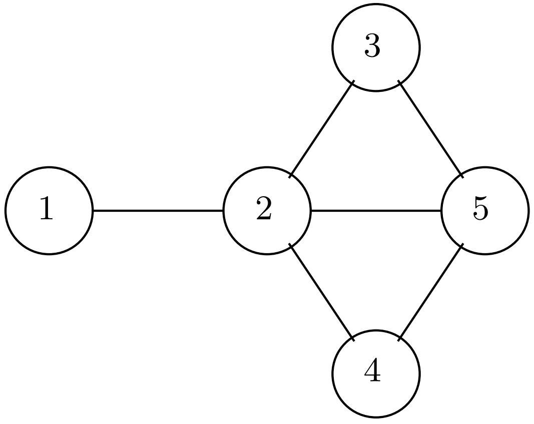


# Terminologies

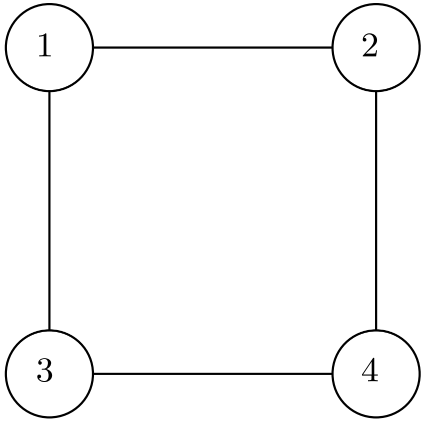
- clique ($\leftrightarrow$ face)
  the set of vertices which induces a complete subgraph
- maximal clique($\leftrightarrow$ facet)

# Minimal vertex separator

- Minimal vertex separator
  The minimal set of vertices which separates some pair of non-adjacent vertices
- A graph is chordal
  $\Leftrightarrow$ any minimal vertex separators induce cliques



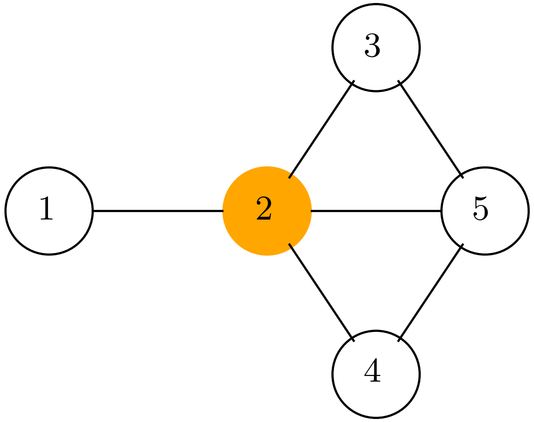

$\mathcal{S} = \{\{2\}, \{2, 5\}\}$



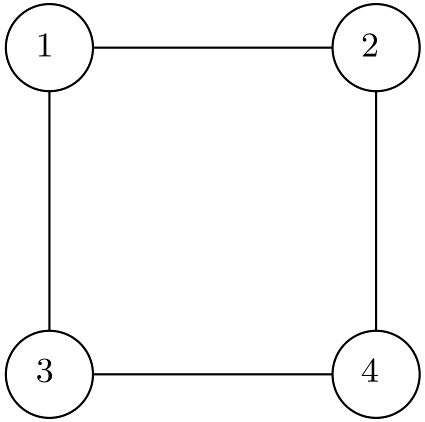

$\mathcal{S} = \{\{1, 4\}, \{2, 3\}\}$

# Minimal vertex separator

- Minimal vertex separator
  The minimal set of vertices which separates some pair of non-adjacent vertices
- A graph is chordal
  $\Leftrightarrow$ any minimal vertex separators induce cliques



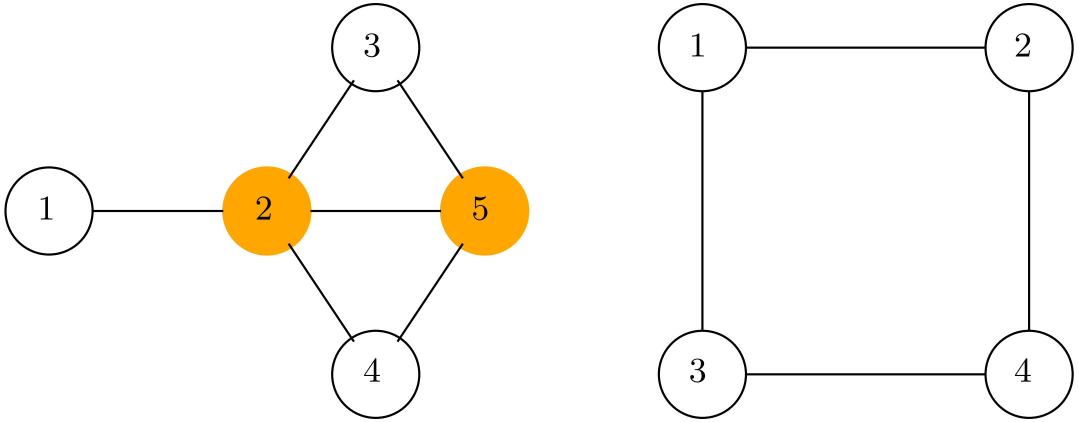

$\mathcal{S} = \{\{2\}, \{2, 5\}\}$




$\mathcal{S} = \{\{1, 4\}, \{2, 3\}\}$

# Minimal vertex separator

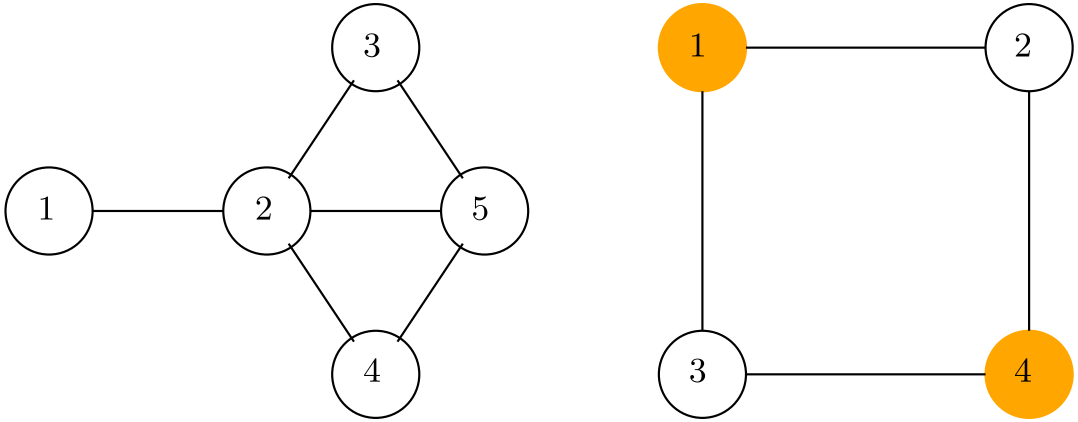
- Minimal vertex separator
  The minimal set of vertices which separates some pair of non-adjacent vertices
- A graph is chordal
  $\Leftrightarrow$ any minimal vertex separators induce cliques




$\mathcal{S} = \{\{2\}, \{2, 5\}\}$ $\mathcal{S} = \{\{1, 4\}, \{2, 3\}\}$

# Minimal vertex separator

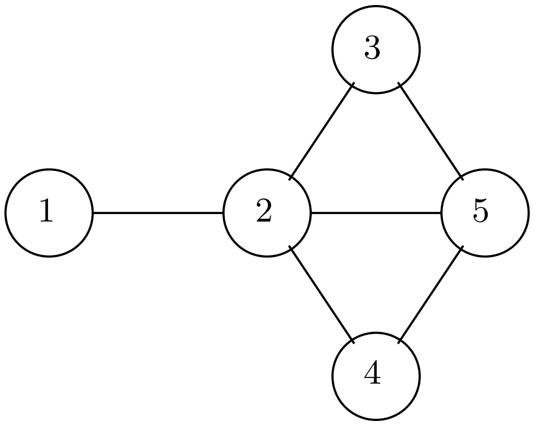
- Minimal vertex separator
  The minimal set of vertices which separates some pair of non-adjacent vertices
- A graph is chordal
  $\Leftrightarrow$ any minimal vertex separators induce cliques



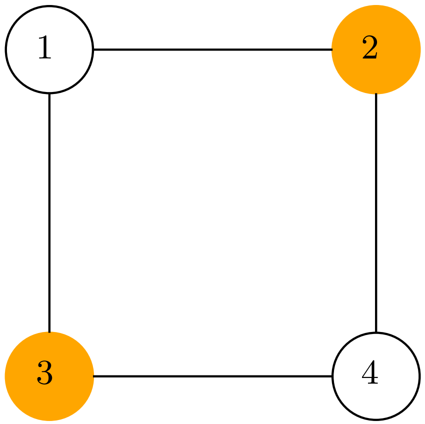

$\mathcal{S} = \{\{2\}, \{2, 5\}\}$ $\mathcal{S} = \{\{1, 4\}, \{2, 3\}\}$

# Minimal vertex separator

- Minimal vertex separator
  The minimal set of vertices which separates some pair of non-adjacent vertices
- A graph is chordal
  $\Leftrightarrow$ any minimal vertex separators induce cliques




$\mathcal{S} = \{\{2\}, \{2, 5\}\}$

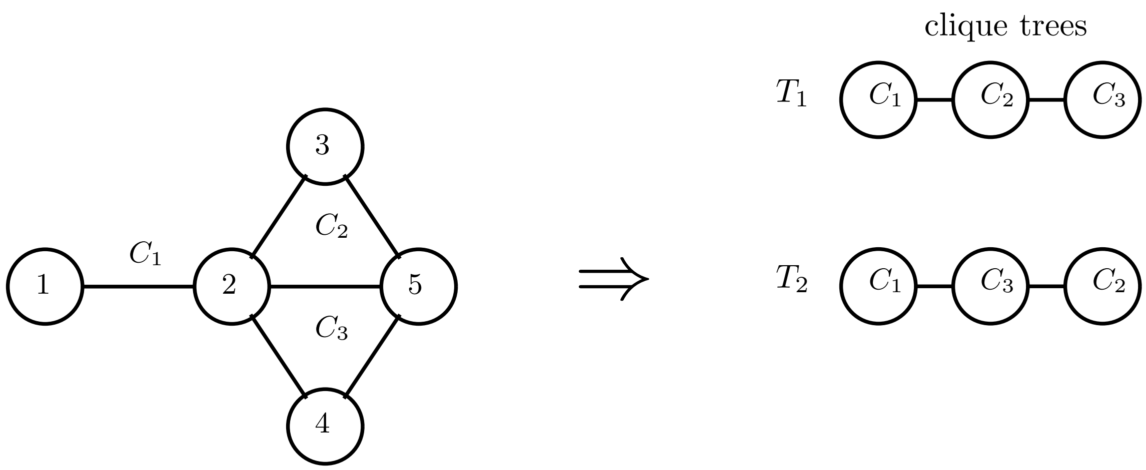



$\mathcal{S} = \{\{1, 4\}, \{2, 3\}\}$

# Clique tree

- A clique tree $T = (\mathcal{C}, \mathcal{E})$
  $\forall C, \forall C' \in \mathcal{C}$ and any $C'' \in \mathcal{C}$ on the unique path in $T$ between $C$ and $C'$

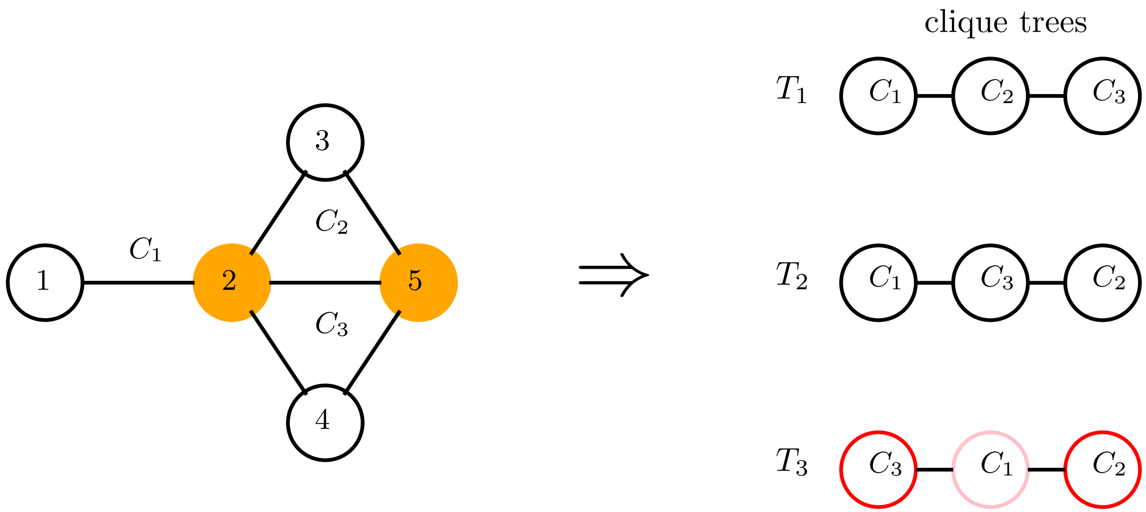
$$C \cap C' \subseteq C'' \rightarrow \text{ junction property}$$



# Clique tree

- A clique tree $T = (\mathcal{C}, \mathcal{E})$
  $\forall C$, $\forall C' \in \mathcal{C}$ and any $C'' \in \mathcal{C}$ on the unique path in $T$ between $C$ and $C'$

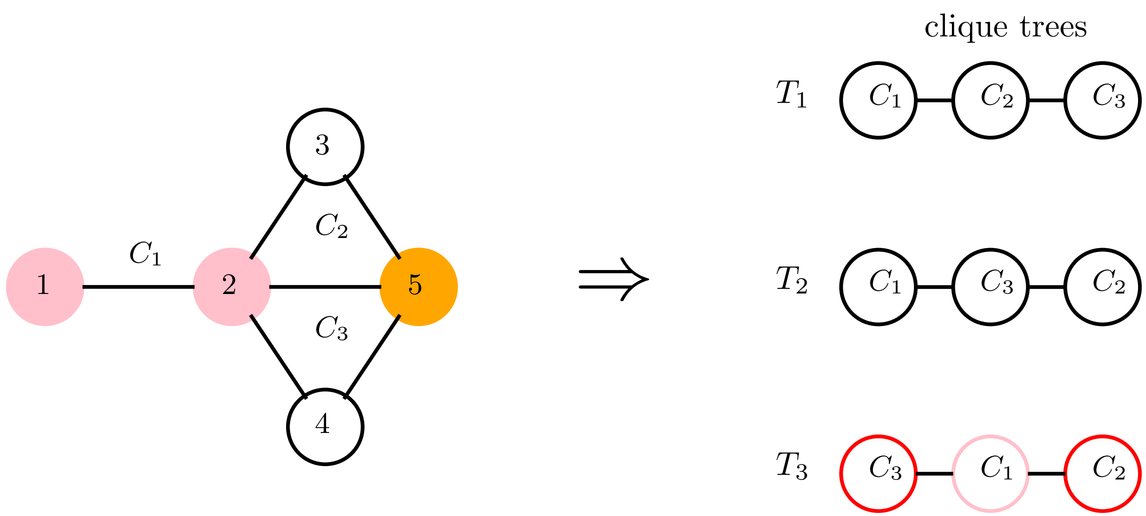
$$C \cap C' \subseteq C'' \rightarrow \text{ junction property}$$



# Clique tree

- A clique tree $T = (\mathcal{C}, \mathcal{E})$
  $\forall C$, $\forall C' \in \mathcal{C}$ and any $C'' \in \mathcal{C}$ on the unique path in $T$ between $C$ and $C'$

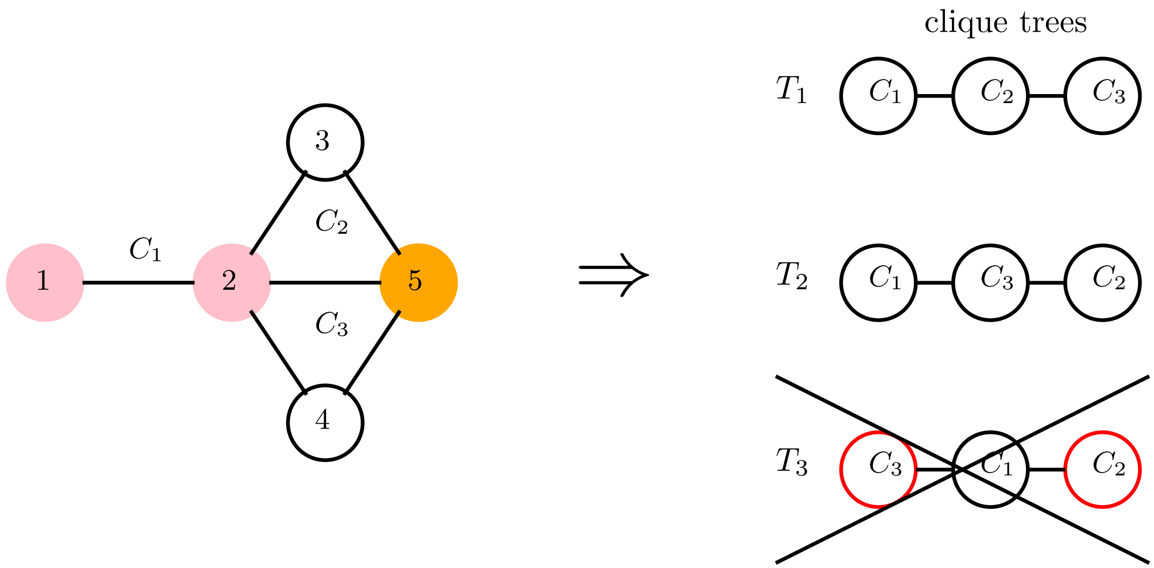
$$C \cap C' \subseteq C'' \rightarrow$$ junction property



# Clique tree

- A clique tree $T = (\mathcal{C}, \mathcal{E})$
  $\forall C$, $\forall C' \in \mathcal{C}$ and any $C'' \in \mathcal{C}$ on the unique path in $T$ between $C$ and $C'$

$$C \cap C' \subseteq C'' \rightarrow$$ junction property




$T_3$ is not a clique tree

## Properties of clique trees

- A graph is chordal $\Leftrightarrow$ there exists a clique tree
- For $C$ and $C'$ s.t. $(C, C') \in \mathcal{E}$, $C \cap C' \in \mathcal{S}$
  $\longrightarrow$ every edge corresponds to m.v.s.
- $\mathcal{T}$ : the set of clique trees for a graph $G$
- In general clique trees are not uniquely defined (e.g. Hara and Takemura(2006))

# Properties of clique trees

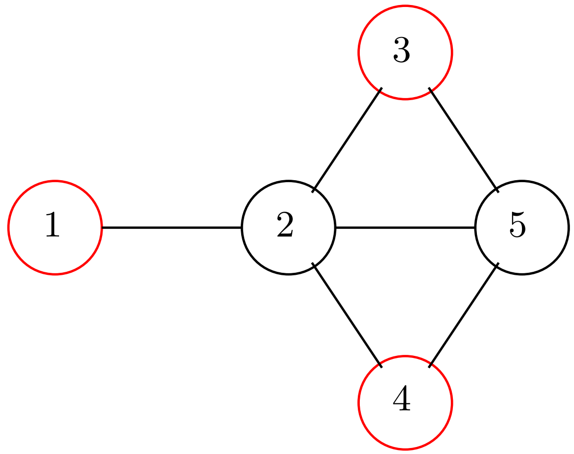
- Clique trees appear in many practical problem
  - Local computation of the marginal distribution in Baysian network (Junction tree algorithm)
  - Efficient computation of the MLE in the statistical model determined by “not” chordal graph
  - Calculation of a Markov basis for a decomposable model (Dobra(2004))

# Simplicial vertex and p.e.o.

- Simplicial vertex

  $v \in V$ s.t. $N_G(v)$ induces a clique

  (where $N_G(v)$ denotes the open adjacency set of $v$)



# Properties of simplicial vertex

For a maximal clique $C \in \mathcal{C}$, we denote by

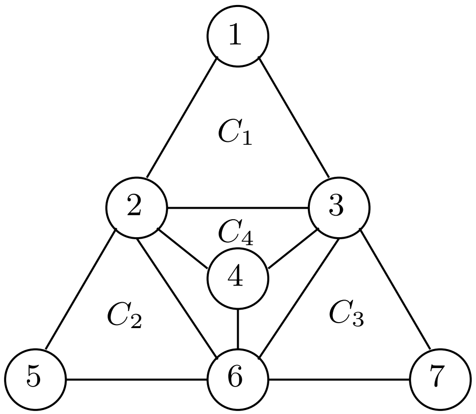
- $\mathrm{Simp}(C)$ : the set of simplicial vertices in $C$
- $\mathrm{Sep}(C)$ : the set of non-simplicial vertices in $C$
- $C = \mathrm{Simp}(C) \cup \mathrm{Sep}(C)$ $\longrightarrow$ disjoint union.
- a simplicial clique : $C \in \mathcal{C}$ s.t. $\mathrm{Simp}(C) \neq \emptyset$

# Boundary cliques and endpoints in clique trees

**Definition 1** (Boundary clique(Shibata(1988)))**.**
*A simplicial clique $C$ is a boundary clique if*

$$\exists C' \in \mathcal{C} \quad s.t. \quad \mathrm{Sep}(C) = C \cap C'.$$

*Then we call $C'$ dominant clique for $C$.*

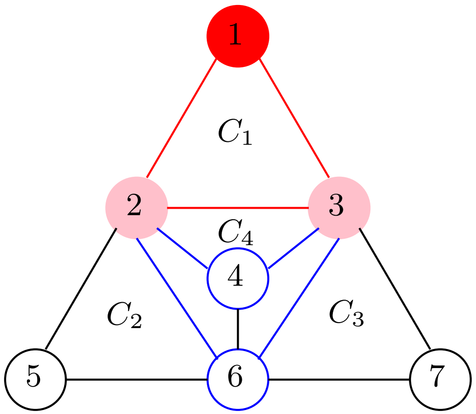


# Boundary cliques and endpoints in clique trees

**Definition 1** (Boundary clique(Shibata(1988)))**.**
*A simplicial clique $C$ is a boundary clique if*

$$\exists C' \in \mathcal{C} \quad s.t. \quad \mathrm{Sep}(C) = C \cap C'.$$

*Then we call $C'$ dominant clique for $C$.*




Boundary clique : $C_1$

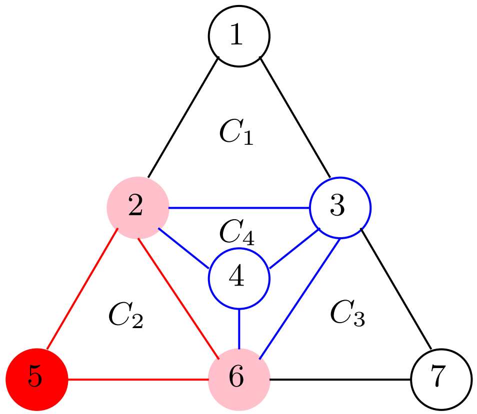
Dominant clique : $C_4$

# Boundary cliques and endpoints in clique trees

**Definition 1** (Boundary clique(Shibata(1988)))**.**
*A simplicial clique $C$ is a boundary clique if*

$$\exists C' \in \mathcal{C} \quad s.t. \quad \mathrm{Sep}(C) = C \cap C'.$$

*Then we call $C'$ dominant clique for $C$.*



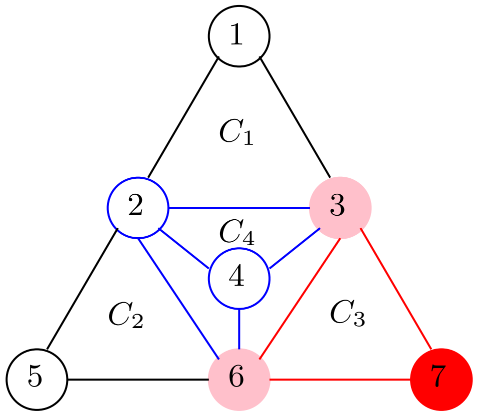

Boundary clique : $C_2$
Dominant clique : $C_4$

# Boundary cliques and endpoints in clique trees

**Definition 1** (Boundary clique(Shibata(1988)))**.**
*A simplicial clique $C$ is a boundary clique if*

$$\exists C' \in \mathcal{C} \quad s.t. \quad \mathrm{Sep}(C) = C \cap C'.$$

*Then we call $C'$ dominant clique for $C$.*



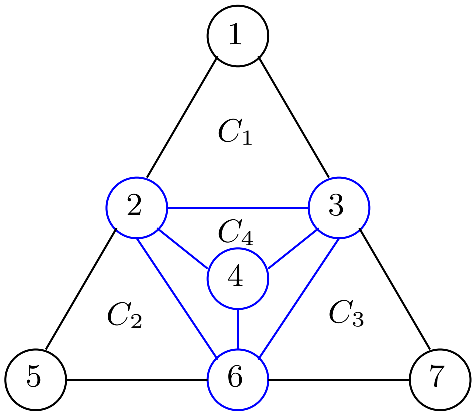

Boundary clique : $C_3$
Dominant clique : $C_4$

# Boundary cliques and endpoints in clique trees

**Definition 1** (Boundary clique(Shibata(1988))). *A simplicial clique $C$ is a boundary clique if*

$$\exists C' \in \mathcal{C} \quad s.t. \quad \mathrm{Sep}(C) = C \cap C'.$$

*Then we call $C'$ dominant clique for $C$.*



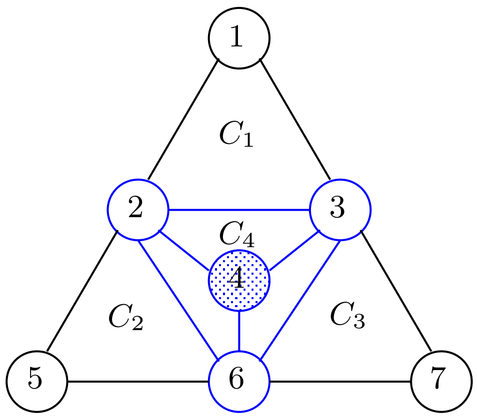

Dominant clique for all boundary cliques

# Boundary cliques and endpoints in clique trees

**Definition 1** (Boundary clique(Shibata(1988)))**.**
*A simplicial clique* $C$ *is a boundary clique if*

$$\exists C' \in \mathcal{C} \quad s.t. \quad \mathrm{Sep}(C) = C \cap C'.$$

*Then we call* $C'$ *dominant clique for* $C$.



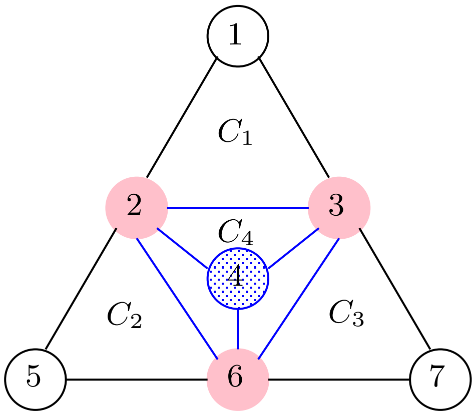

$\{4\}$ is a simplicial vertex, however,

# Boundary cliques and endpoints in clique trees

**Definition 1** (Boundary clique(Shibata(1988)))**.**
*A simplicial clique $C$ is a* boundary clique *if*

$$\exists C' \in \mathcal{C} \quad s.t. \quad \mathrm{Sep}(C) = C \cap C'.$$

*Then we call $C'$* dominant clique *for $C$.*



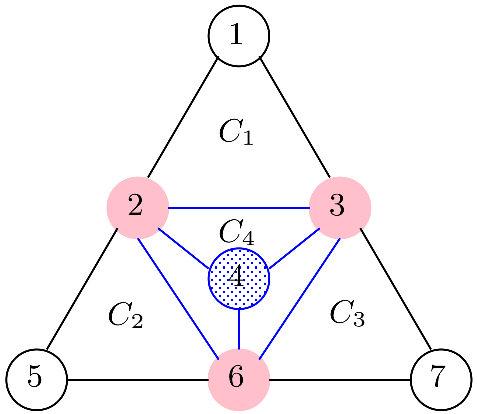

$\{4\}$ is a simplicial vertex, however,
there exists no dominant clique for this maximal clique

# Boundary cliques and endpoints in clique trees

**Definition 1** (Boundary clique(Shibata(1988))). *A simplicial clique $C$ is a boundary clique if*

$$\exists C' \in \mathcal{C} \quad s.t. \quad \mathrm{Sep}(C) = C \cap C'.$$

*Then we call $C'$ dominant clique for $C$.*




$\{4\}$ : weakly simplicial vertex

# Characterization of endpoints of clique tree

**Theorem**

$C$ is a boundary clique
$\Leftrightarrow \exists T \in \mathcal{T}$ s.t. $C \in \mathcal{C}$ is an endpoint of $T$

- $\Leftarrow$ : Shibata(1988)
- $\Rightarrow$ : Hara and Takemura(2006)

# Characterization of minimal Markov bases

- We consider Markov bases for decomposable models from the view point of minimality in terms of chordal graphs
  1. notations and terminologies
  2. a complete description of minimal Markov bases in terms of chordal graphs
  3. a necessary and sufficient condition that the minimal MB is unique
  4. The minimality of Dobra's MB

# Contingency table

- A survey on health care

  A. Drinking habit
  1. everyday, 2. a few times a week, 3. less than a times a week

  B. Smoking : 1. yes, 2. no

  C. Breakfast : 1. bread, 2. rice

  D. Exercise habit : 1. yes, 2. no

| | | $A_1$ | | $A_2$ | | $A_3$ | |
|---|---|---|---|---|---|---|---|
| | | $B_1$ | $B_2$ | $B_1$ | $B_2$ | $B_1$ | $B_2$ |
| $C_1$ | $D_1$ | 3 | 7 | 4 | 5 | 1 | 3 |
| | $D_2$ | 34 | 3 | 16 | 14 | 6 | 22 |
| $C_2$ | $D_1$ | 13 | 24 | 16 | 15 | 4 | 15 |
| | $D_2$ | 28 | 7 | 18 | 9 | 12 | 16 |

# Notation

- $\Delta$ : the set of variables $\Leftrightarrow \{A, B, C, D\}$
- $I_\delta$ : the number of levels of $\delta \in \Delta$
  $\Leftrightarrow I_A = 3,\ I_B = I_C = I_D = 2$
- $\mathcal{I}_\delta$ : the set of levels of $\delta$
  $\Leftrightarrow \mathcal{I}_A = \{A_1, A_2, A_3\}$
- a cell $\boldsymbol{i} \in \mathcal{I} = \prod_{\delta \in \Delta} \mathcal{I}_\delta$
  $\Leftrightarrow \boldsymbol{i} = (i_A, i_B, i_C, i_D) = (A_i, B_j, C_k, D_l)$
- $n(\boldsymbol{i})$ : the frequency of $\boldsymbol{i}$

# Notation

- a marginal cell for $D \subset \Delta$ :

  $\boldsymbol{i}_D \in \mathcal{I}_D = \prod_{\delta \in D} \mathcal{I}_\delta \Leftrightarrow \boldsymbol{i}_{AB} = (A_i, B_j)$

- a marginal frequency of $i_D$ :

  $n_D(\boldsymbol{i}_D) = \sum_{\boldsymbol{i}_{D^c} \in \mathcal{I}_{D^c}} n(\boldsymbol{i}_D, \boldsymbol{i}_{D^c})$

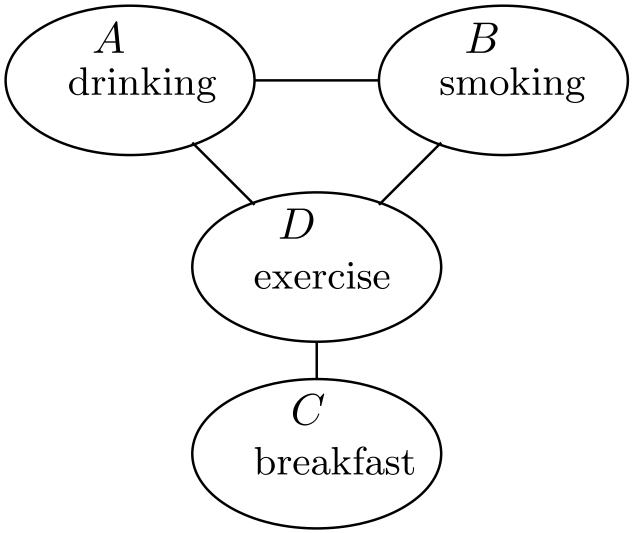
| | | $A_1$ | | $A_2$ | | $A_3$ | | |
|---|---|---|---|---|---|---|---|---|
| | | $B_1$ | $B_2$ | $B_1$ | $B_2$ | $B_1$ | $B_2$ | $\boldsymbol{i}_{CD}$ |
| $C_1$ | $D_1$ | 3 | 7 | 4 | 5 | 1 | 3 | 23 |
| | $D_2$ | 34 | 3 | 16 | 14 | 6 | 22 | 95 |
| $C_2$ | $D_1$ | 13 | 24 | 16 | 15 | 4 | 15 | 87 |
| | $D_2$ | 28 | 7 | 18 | 9 | 12 | 16 | 90 |
| | $\boldsymbol{i}_{AB}$ | 78 | 41 | 54 | 43 | 23 | 56 | 295 |

# Notations and terminologies

- the generating class :
  $\mathcal{D} = \{D_1, \ldots, D_r\}$, $D_r \subset \Delta$ satisfying
  - no inclusion relation among $D_j$'s
  - $\bigcup_{j=1}^{r} D_j = \Delta$
- $D_i \in \mathcal{D}$ : a generating set
- the independence graph for $\mathcal{D}$:
  $\mathcal{G}^{\mathcal{D}} = (\Delta, E^{\mathcal{D}})$
  $(\delta, \delta') \in E^{\mathcal{D}} \Leftrightarrow \delta, \delta' \in \exists D \in \mathcal{D}$

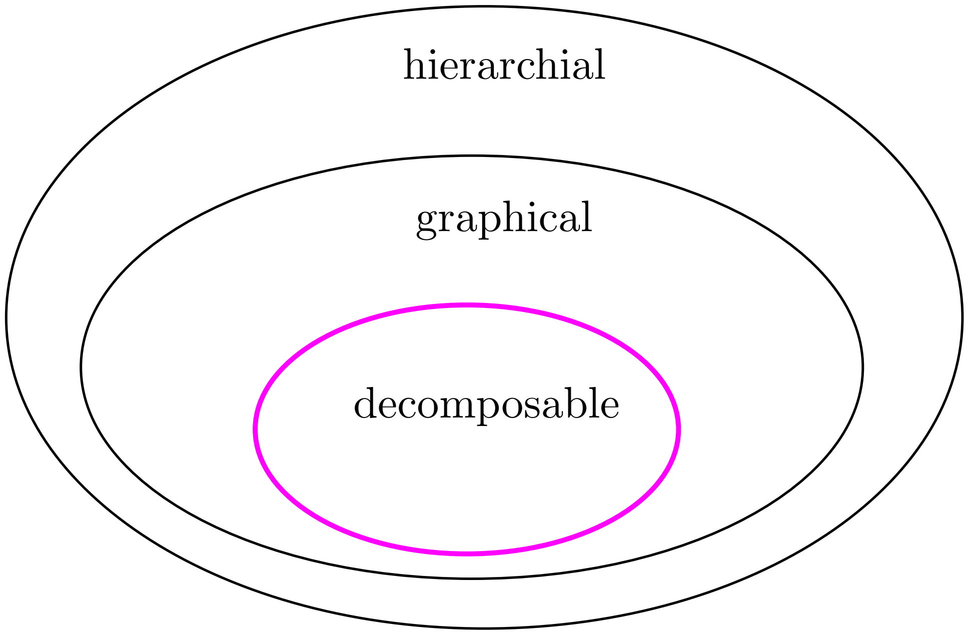
# Notations and terminologies

- $\mathcal{G}^{\mathcal{D}}$ with $\mathcal{D} = \{\{A, B, D\}, \{C, D\}\}$



# Decomposable models

- A decomposable model :
  $\mathcal{D}$ is the set of maximal clique of a chordal graph



# Notations and terminologies

- $\boldsymbol{b} = \{n_D(\boldsymbol{i}_D) \in \mathcal{I}_D, D \in \mathcal{D}\}$
- $\boldsymbol{b} = A\boldsymbol{n}$, $\quad A$ : the incidence matrix for $\mathcal{D}$
- $\boldsymbol{b}$-fiber : $\mathcal{F}_{\boldsymbol{b}} = \{\boldsymbol{n} \geq 0 \mid A\boldsymbol{n} = \boldsymbol{b}\}$
- $\mathcal{B}_{\text{nd}} = \{\boldsymbol{b} \mid |\mathcal{F}_{\boldsymbol{b}}| \geq 2\}$

# Degree two fibers

- $\boldsymbol{b} = \{n_D(\boldsymbol{i}_D) \in \mathcal{I}_D, D \in \mathcal{D}\}$
- deg $\boldsymbol{b}$ :

$$\deg \boldsymbol{b} = \sum_{D \in \mathcal{D}} \sum_{\boldsymbol{i}_D \in \mathcal{I}_D} n_D(\boldsymbol{i}_D)$$

- a degree two fiber : $\mathcal{F}_{\boldsymbol{b}}$ s.t. deg $\boldsymbol{b} = 2$

# Primitive move

- a move $\boldsymbol{z}$ : an integer array s.t. $\dim \boldsymbol{n} = \dim \boldsymbol{z}$ satisfying $A\boldsymbol{z} = 0$
- $\boldsymbol{z} = \boldsymbol{z}^+ - \boldsymbol{z}^-$, $\boldsymbol{z}^+ \geq 0$, $\boldsymbol{z}^- \geq 0$
- $\boldsymbol{z}^+ = \{z^+(\boldsymbol{i})\}_{\boldsymbol{i}\in\mathcal{I}}, \quad \boldsymbol{z}^- = \{z^-(\boldsymbol{i})\}_{\boldsymbol{i}\in\mathcal{I}}$
- Total frequencies of $\boldsymbol{z}^+$ and $\boldsymbol{z}^-$ are the same $z = \sum_{\boldsymbol{i}\in\mathcal{I}} z^+(\boldsymbol{i}) = \sum_{\boldsymbol{i}\in\mathcal{I}} z^-(\boldsymbol{i})$
- primitive move : a move $\boldsymbol{z}$ with $z = 2$

# Markov bases of decomposable models

- Dobra(2003), Hoşten and Sullivant(2002)
  For decomposable models there exists a Markov basis consisting of primitive moves
- Takemura and Aoki(2004)
  The union of sets of moves which connect each degree two fiber is a MB

$$\Downarrow$$

The clarification of the structure of deg 2 fibers $\mathcal{F}_{\boldsymbol{b}}$ s.t. $|\mathcal{F}_{\boldsymbol{b}}| \geq 2$ is essential

# Degree 2 fibers

- There exist at most two levels with positive one dimensional marginal
  $\Rightarrow$ it suffices to consider the table with
  $$I_\delta = 2, \quad \delta \in \Delta, \; \mathcal{I} = \{0,1\}^{|\Delta|}$$
- $\mathcal{I}_\delta = (i_\delta, i_\delta^*) = (0, 1), \quad i_\delta^* = 1 - i_\delta$
- $\boldsymbol{i}_{\Delta'}^* = (i_\delta^*)_{\{\delta \in \Delta'\}}$

# Nondegenerate variables

- $\boldsymbol{b}$ s.t. $\deg \boldsymbol{b} = 2$ is given
- $\delta \in \Delta$ is degenerate :
  there exist a unique levels $i_\delta$ s.t. $n_\delta(i_\delta) = 2$
  i.e. $n_\delta(i_\delta^*) = 0$
- $\delta \in \Delta$ is nondegenerate :
  $n_\delta(i_\delta) = n_\delta(i_\delta^*) = 1$
- $\bar{\Delta}_{\boldsymbol{b}}$ : the set of nondegenerate variable
- $\mathcal{G}(\bar{\Delta}_{\boldsymbol{b}})$ : The subgraph of $\mathcal{G}^{\mathcal{D}}$ induced by $\bar{\Delta}_{\boldsymbol{b}}$

# Structure of degree 2 fibers

**Theorem** (Hara, Aoki and Takemura(2007))

$\mathcal{F}_{\boldsymbol{b}}$ : a degree two fiber s.t. $\bar{\Delta}_{\boldsymbol{b}} \neq \emptyset$
$c(\boldsymbol{b})$ : # of connected components of $\mathcal{G}(\bar{\Delta}_{\boldsymbol{b}})$
Then

$$|\mathcal{F}_{\boldsymbol{b}}| = 2^{c(\boldsymbol{b})-1}$$

- This theorem holds for general hierarchical models
- $\forall \Delta' \subset \Delta$ s.t. $\mathcal{G}^{\mathcal{D}}(\Delta')$ is disconnected
  $\exists \boldsymbol{b} \in \mathcal{B}_{\text{nd}}$ s.t. $\bar{\Delta}_{\boldsymbol{b}} = \Delta'$

# Notations

- $(\boldsymbol{i})(\boldsymbol{j})$ : the element of degree 2 fiber s.t.

$$n(\boldsymbol{i}') = \begin{cases} 1 & \text{if } \boldsymbol{i}' = \boldsymbol{i} \text{ or } \boldsymbol{i}' = \boldsymbol{j} \\ 0 & \text{otherwise} \end{cases}$$

- Example

| | |
|---|---|
| 1 | 0 |
| 0 | 1 |

(00)(11)

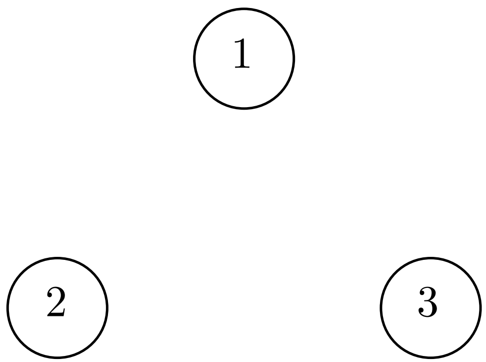
| | |
|---|---|
| 0 | 1 |
| 1 | 0 |

(01)(10)

# Example : 3way complete independence model

- $\mathcal{D} = \{\{1\},\{2\},\{3\}\}$
- Let $\boldsymbol{b} \in \mathcal{B}_{\text{nd}}$ satisfy $\bar{\Delta}_{\boldsymbol{b}} = \{1,2,3\}$
- $\bar{\Delta}_{\boldsymbol{b}} = \{1,2,3\}$

$$\Rightarrow \quad c(\boldsymbol{b}) = 3$$
$$\Rightarrow \quad |\mathcal{F}_{\boldsymbol{b}}| = 2^{(3-1)} = 4$$

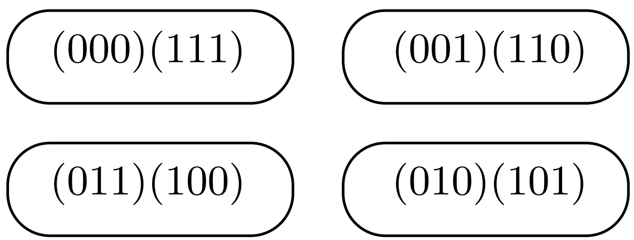


# Example : 3way complete independence model

- $\mathcal{D} = \{\{1\}, \{2\}, \{3\}\}$
- Let $\boldsymbol{b} \in \mathcal{B}_{\text{nd}}$ satisfy $\bar{\Delta}_{\boldsymbol{b}} = \{1, 2, 3\}$
- $\bar{\Delta}_{\boldsymbol{b}} = \{1, 2, 3\}$
$$\Rightarrow \quad c(\boldsymbol{b}) = 3$$
$$\Rightarrow \quad |\mathcal{F}_{\boldsymbol{b}}| = 2^{(3-1)} = 4$$
- $\mathcal{F}_{\boldsymbol{b}} = \{(000)(111), (001)(110), (010)(101), (011)(100)\}$

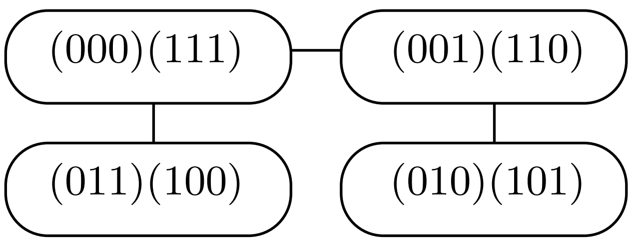
# Construction of a minimal MB

- For each fiber $\mathcal{F}_b$



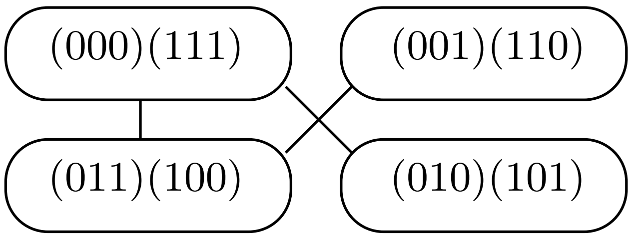
# Construction of a minimal MB

- For each fiber $\mathcal{F}_{\boldsymbol{b}}$
- Construct a tree $\mathcal{T}_{\boldsymbol{b}} = (\mathcal{F}_{\boldsymbol{b}}, \mathcal{M}^0_{\boldsymbol{b}})$
  $\mathcal{M}^0_{\boldsymbol{b}}$ : a set of edges $\Leftrightarrow$ a set of moves



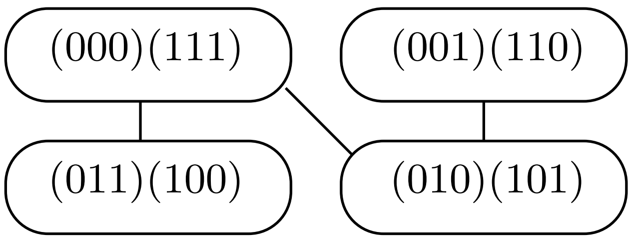
# Construction of a minimal MB

- For each fiber $\mathcal{F}_{\boldsymbol{b}}$
- Construct a tree $\mathcal{T}_{\boldsymbol{b}} = (\mathcal{F}_{\boldsymbol{b}}, \mathcal{M}^0_{\boldsymbol{b}})$
  $\mathcal{M}^0_{\boldsymbol{b}}$ : a set of edges $\Leftrightarrow$ a set of moves



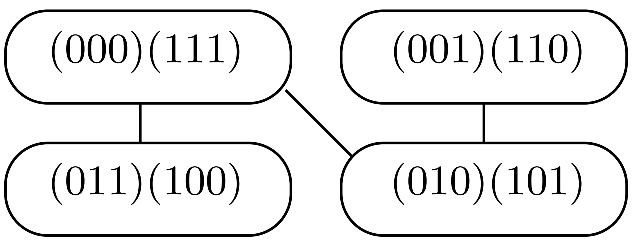
# Construction of a minimal MB

- For each fiber $\mathcal{F}_{\boldsymbol{b}}$
- Construct a tree $\mathcal{T}_{\boldsymbol{b}} = (\mathcal{F}_{\boldsymbol{b}}, \mathcal{M}^0_{\boldsymbol{b}})$
  $\mathcal{M}^0_{\boldsymbol{b}}$ : a set of edges $\Leftrightarrow$ a set of moves



# Construction of a minimal MB

- For each fiber $\mathcal{F}_{\boldsymbol{b}}$
- Construct a tree $\mathcal{T}_{\boldsymbol{b}} = (\mathcal{F}_{\boldsymbol{b}}, \mathcal{M}^0_{\boldsymbol{b}})$
  $\mathcal{M}^0_{\boldsymbol{b}}$ : a set of edges $\Leftrightarrow$ a set of moves
- Dobra(2003), Takemura and Aoki(2004)
  $\mathcal{M}^0 = \bigcup_{\boldsymbol{b} \in \mathcal{B}_{\text{nd}}} \mathcal{M}^0_{\boldsymbol{b}} \quad \Rightarrow \quad$ a minimal MB



# The uniqueness of minimal MB

- $|\mathcal{F}_{\boldsymbol{b}}| = 2^{c(\boldsymbol{b})-1}$
- $c(\boldsymbol{b}) \geq 3 \quad \Rightarrow \quad \mathcal{T}_{\boldsymbol{b}}$ is not uniquely defined
  $\Rightarrow$ minimal MB are not unique
- In the case of decomposable models

> $c(\boldsymbol{b}) = 2$ for all $\boldsymbol{b} \in \mathcal{B}_{\mathrm{nd}}$
> $\Leftrightarrow$ minimal MB is unique

# The unique minimal Markov basis

**Lemma**(HAT(2007))

A decomposable model has the unique minimal MB if and only if the number of connected components in any induced subgraphs of $\mathcal{G}^{\mathcal{D}}$ is less than three.

# In terms of chordal graphs ...

**Theorem**(HAT(2007))

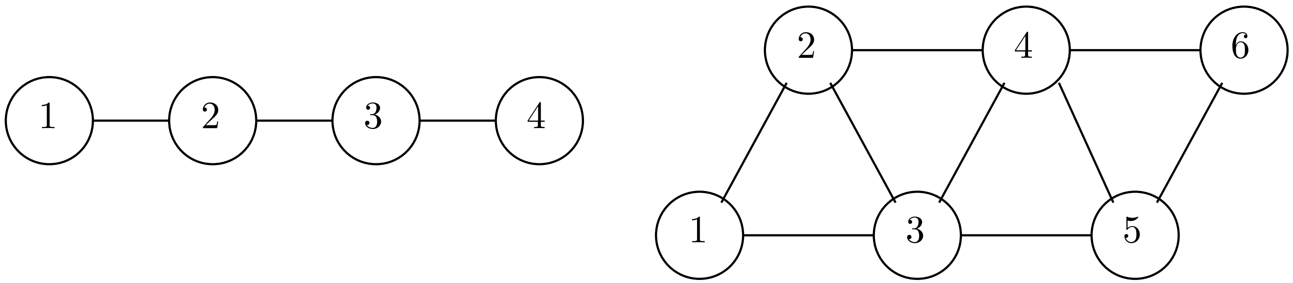
For a decomposable model, there exists the unique minimal MB if and only if

- $\mathcal{G}^{\mathcal{D}}$ has only two boundary cliques $D$, $D'$
- $D$ and $D'$ satisfy $D'' \subset D \cup D'$ for all $D'' \in \mathcal{D}$.

# Example

- $\mathcal{G}^{\mathcal{D}}$ with

$$\mathcal{D} = \{\{1,\ldots,r{-}1\},\{2,\ldots,r\},\ldots,\{r,\ldots,2r{-}2\}\}$$



# Markov basis for two maximal clique model

- $\mathcal{D} = \{D_1, D_2\}$
- $S = D_1 \cap D_2$ : the minimal vertex separator
- $D'_1 = D_1 \setminus S$, $D'_2 = D_2 \setminus S$
- $\mathcal{D}$ has the unique minimal MB $\mathcal{M}^0(D_1, D_2)$

$$\mathcal{M}^0(D_1, D_2)$$
$$= \{(i_{D'_1}\ j_S\ i_{D'_2})(i^*_{D'_1}\ j_S\ i^*_{D'_2}) - (i_{D'_1}\ j_S\ i^*_{D'_2})(i^*_{D'_1}\ j_S\ i_{D'_2}) \mid i_{D'_j} \in \mathcal{I}_{D'_j}, j_S \in \mathcal{I}_S\}$$

# Dobra's construction of Markov bases

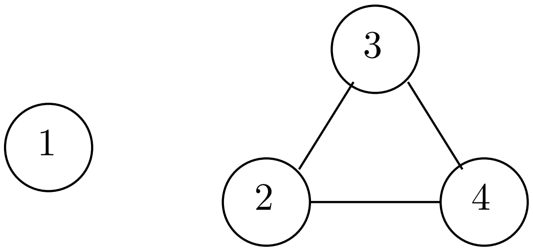
- $\mathcal{D} = \{\{1\}, \{2\}, \{3\}, \{4\}\}$, $D_j = \{j\}$

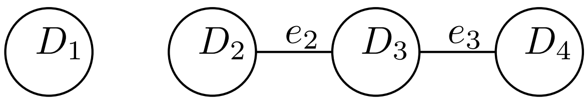
# Dobra's construction of Markov bases

- $\mathcal{D} = \{\{1\}, \{2\}, \{3\}, \{4\}\}$, $D_j = \{j\}$
- remove $e_1$
  $V_{e_1} = \{1\}$, $V'_{e_1} = \{2, 3, 4\}$

# Dobra's construction of Markov bases

- $\mathcal{D} = \{\{1\}, \{2\}, \{3\}, \{4\}\}$, $D_j = \{j\}$
- remove $e_1$
  $V_{e_1} = \{1\}$, $V'_{e_1} = \{2, 3, 4\}$




$$\Downarrow$$

$$\mathcal{M}^0(V_{e_1}, V'_{e_1})$$

$\mathcal{M}^0(V_{e_1}, V'_{e_1})$ : The set of primitive moves of the model $\mathcal{D}' = \{V_{e_1}, V'_{e_1}\}$

# Dobra's construction of Markov bases

- $\mathcal{D} = \{\{1\}, \{2\}, \{3\}, \{4\}\}$, $D_j = \{j\}$
- remove $e_2$
  $V_{e_2} = \{1, 2\}$, $V'_{e_2} = \{3, 4\}$

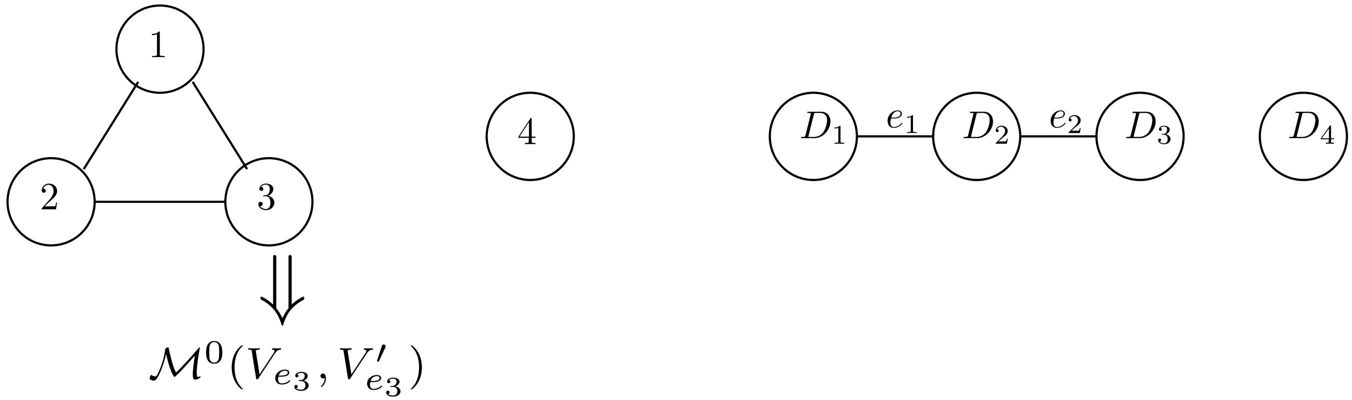

$$\Downarrow$$

$$\mathcal{M}^0(V_{e_2}, V'_{e_2})$$

$$\mathcal{M}^0(V_{e_1}, V'_{e_1}) \cup \mathcal{M}^0(V_{e_2}, V'_{e_2})$$

# Dobra's construction of Markov bases

- $\mathcal{D} = \{\{1\}, \{2\}, \{3\}, \{4\}\}$, $D_j = \{j\}$
- remove $e_3$
  $V_{e_3} = \{1, 2, 3\}$, $V'_{e_3} = \{4\}$




$\mathcal{M}^0(V_{e_1}, V'_{e_1}) \cup \mathcal{M}^0(V_{e_2}, V'_{e_2}) \cup \mathcal{M}^0(V_{e_3}, V'_{e_3}) \Rightarrow$ Markov basis

# Dobra's construction of Markov bases

- $\mathcal{D}$ : A decomposable model
  $\mathcal{T} = (\mathcal{D}, \mathcal{E})$ : A clique tree of $\mathcal{G}^{\mathcal{D}}$
- $\mathcal{T}_e = (\mathcal{D}_e, \mathcal{E}_e)$, $\mathcal{T}_e' = (\mathcal{D}_e', \mathcal{E}_e')$
  two subtree of $\mathcal{T}$ by removing $e \in \mathcal{E}$
- $V_e = \bigcup_{D \in \mathcal{D}_e} D$, $\quad V_e' = \bigcup_{D \in \mathcal{D}_e'} D$

$$\mathcal{M}^{\mathcal{T}} = \bigcup_e \mathcal{M}^0(V_e, V_e') \quad \Rightarrow \quad \text{Markov basis}$$

# Minimality of Dobra's Markov bases

**Theorem(HAT(2007))**

Dobra's MB $\mathcal{M}^{\mathcal{T}}$ is minimal if and only if the decomposable model has the unique minimal MB

- We can show that $\mathcal{M}^{\mathcal{T}}$ includes at least 4 moves for every $\mathcal{F}_{\boldsymbol{b}}$ s.t. $|\mathcal{F}_{\boldsymbol{b}}| = 4$

# Invariant Markov bases

- Suppose that move $z$ is included in a Markov basis.
- Then we also use a move where the levels of the moves are interchanged (invariance w.r.t. permutations of levels)

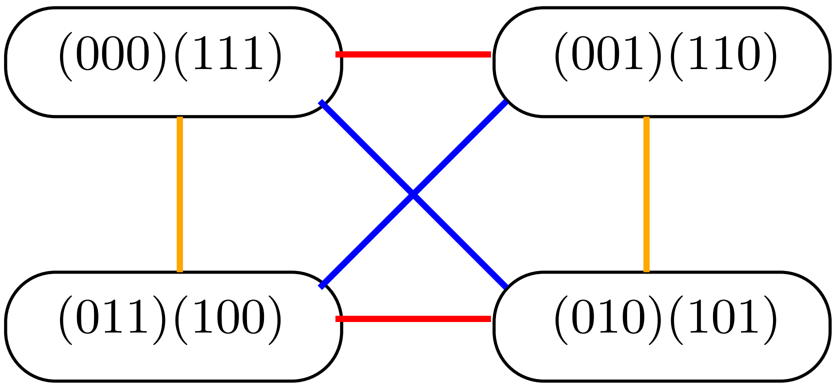
$$\boxed{\begin{matrix} +1 & -1 & 0 \\ -1 & +1 & 0 \\ 0 & 0 & 0 \end{matrix}} \Rightarrow \boxed{\begin{matrix} +1 & -1 & 0 \\ 0 & 0 & 0 \\ -1 & +1 & 0 \end{matrix}} \Rightarrow \boxed{\begin{matrix} +1 & 0 & -1 \\ -1 & 0 & +1 \\ 0 & 0 & 0 \end{matrix}}$$

# Invariant Markov bases

**Theorem(HAT(2007))**

Minimal invariant Markov basis corresponds to a vector space basis of the vector space $\{0,1\}^{c-1}$ over $\mathrm{GF}(2)$ (in each relevant fiber)

3-way complete independence case:

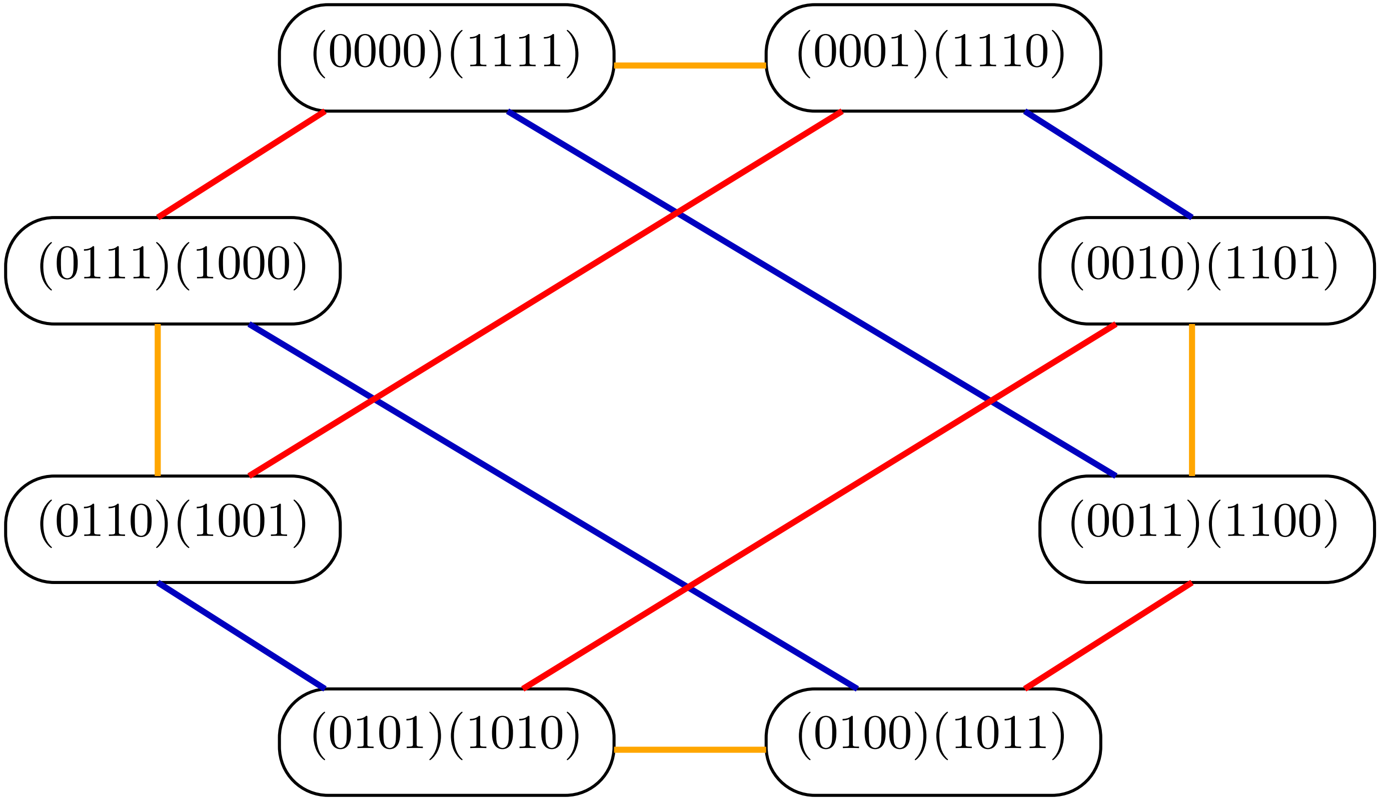


# Invariant Markov bases

- 3 orbits
- 2 orbits are enough to connect this fiber
- Corresponds to linear bases $\{(1,0),(0,1)\}$, $\{(1,0),(1,1)\}$, $\{(0,0),(1,1)\}$ of the vector space $\{0,1\}^2$.

# Invariant Markov bases

Four way case:



# Invariant Markov bases

- 3 orbits are enough (out of 7)
- In this picture we chose $\{(0,0,1),(0,1,1),$ $(1,1,1)\}$ as a linear basis of $\{0,1\}^3$.

# Conclusion

- We gave a brief review on chordal graphs and their clique trees
- We considered Markov bases from the viewpoint of minimality
  - a complete description of minimal MB in decomposable models
  - a necessary and sufficient condition that the unique minimal MB exists
  - the minimality of Dobra's Markov bases

# References


- Hara, H., Aoki, S and Takemura,A.(2007). Fibers of sample size two of hierarchical models and Markov bases of decomposable models for contingency tables. `arvix:math.ST/0701429`.
- Hara, H., Takemura, A.(2006). Boundary cliques, clique trees and perfect sequences of maximal cliques of a chordal graph. `arvix:cs.DM/0607055`.