

# Inférence de la structure de dépendance à partir d'événements

F. Harlé<sup>1</sup>, F. Chatelain<sup>2</sup>, C. Gouy-Pailler<sup>1</sup>, S. Achard<sup>2</sup>

<sup>1</sup>CEA, LIST, Laboratoire d'Analyse de Données et Intelligence des Systèmes, France

<sup>2</sup>Univ. Grenoble Alpes, GIPSA-Lab, France  
CNRS, GIPSA-Lab, France

Réunion AIGM, Grenoble  
30 juin 2015



UNIVERSITÉ DE  
GRENOBLE



# Plan

- 1 Introduction
- 2 Recherche du modèle causal
  - Principe
  - Calcul des factorisations
  - Sélection de la factorisation
  - Classes d'équivalence
  - Nombre d'arêtes
- 3 Estimation de la loi de probabilité jointe
  - Loi connue
  - Loi inconnue
  - Algorithme
- 4 Résultats
  - Simulations
  - Application sur des données de consommation électrique
- 5 Conclusions et perspectives

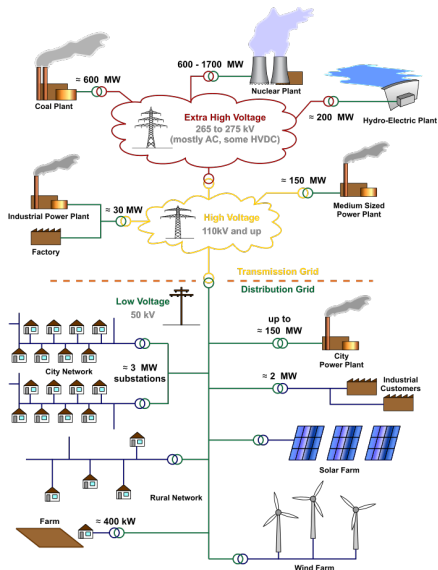
# Plan

- 1 Introduction
- 2 Recherche du modèle causal
  - Principe
  - Calcul des factorisations
  - Sélection de la factorisation
  - Classes d'équivalence
  - Nombre d'arêtes
- 3 Estimation de la loi de probabilité jointe
  - Loi connue
  - Loi inconnue
  - Algorithme
- 4 Résultats
  - Simulations
  - Application sur des données de consommation électrique
- 5 Conclusions et perspectives

## Contexte



## Indicateurs financiers

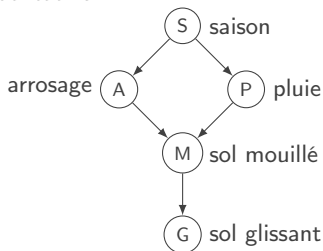


## Réseau électrique

# Représentations des relations de causalité [Pea00]

## Approche graphique : qualitative

$$G = (V, E)$$



### Intérêts :

- intuitive, adaptée à la conception humaine d'un système
- explicite, symbolique, informations pertinentes sur les relations de causalité
- compact

### Limites :

- pas d'information quantitative
- statique

# Représentations des relations de causalité [Pea00]

## Approche probabiliste : quantitative

variable		valeur prise
saison	S	printemps, été, automne, hiver
pluie	P	oui, non
arrosage	A	on, off
sol mouillé	M	oui, non
sol glissant	G	oui, non

## Probabilité jointe

$$P(v) = P(S = s, P = p, A = a, M = m, G = g)$$

Notation :  $P(x)$  pour  $P(X = x)$

## Intérêts :

- formalisme mathématique
- loi de probabilité des variables
- incertitudes possibles

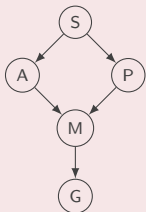
## Limites :

- stockage pour de nombreuses variables
- relations entre les variables non visibles

## Décomposition et compatibilité

### Décomposition de la loi de probabilité jointe

Variables  $X_{1,j}, \dots, X_{n,j}$  **parents markoviens** de  $X_j$  : l'ensemble minimal des prédécesseurs de  $X_j$  qui rendent  $X_j$  indépendant de tous ses autres prédécesseurs.



**Décomposition** de  $P(v)$  à l'aide des parents markoviens de chaque variable

$$P(v) = P(s, p, a, m, g)$$

$$P(v) = P(s)P(a|s)P(p|s)P(m|a, p)P(g|m).$$

Unicité de l'ensemble des parents Markoviens si  $P(v)$  est strictement positive.

### Compatibilité au sens de Markov

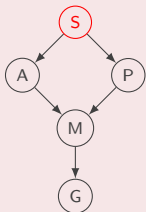
Si une loi de probabilité  $P$  admet une factorisation d'après les parents markoviens des variables relatifs au graphe acyclique et dirigé  $G$ ,  $P$  et  $G$  sont dits **compatibles**.

Pour un ordre donné des variables, si  $P(v)$  est strictement positive, le graphe  $G$  compatible est unique.

## Décomposition et compatibilité

### Décomposition de la loi de probabilité jointe

Variables  $X_{1,j}, \dots, X_{n,j}$  **parents markoviens** de  $X_j$  : l'ensemble minimal des prédécesseurs de  $X_j$  qui rendent  $X_j$  indépendant de tous ses autres prédécesseurs.



**Décomposition** de  $P(v)$  à l'aide des parents markoviens de chaque variable

$$P(v) = P(s, p, a, m, g)$$

$$P(v) = P(s)P(a|s)P(p|s)P(m|a, p)P(g|m).$$

Unicité de l'ensemble des parents Markoviens si  $P(v)$  est strictement positive.

### Compatibilité au sens de Markov

Si une loi de probabilité  $P$  admet une factorisation d'après les parents markoviens des variables relatifs au graphe acyclique et dirigé  $G$ ,  $P$  et  $G$  sont dits **compatibles**.

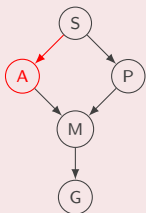
Pour un ordre donné des variables, si  $P(v)$  est strictement positive, le graphe  $G$  compatible est unique.



## Décomposition et compatibilité

### Décomposition de la loi de probabilité jointe

Variables  $X_{1,j}, \dots, X_{n,j}$  **parents markoviens** de  $X_j$  : l'ensemble minimal des prédécesseurs de  $X_j$  qui rendent  $X_j$  indépendant de tous ses autres prédécesseurs.



**Décomposition** de  $P(v)$  à l'aide des parents markoviens de chaque variable

$$P(v) = P(s, p, a, m, g)$$

$$P(v) = P(s)P(a|s)P(p|s)P(m|a, p)P(g|m).$$

Unicité de l'ensemble des parents Markoviens si  $P(v)$  est strictement positive.

### Compatibilité au sens de Markov

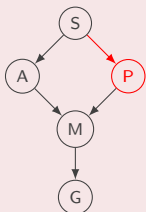
Si une loi de probabilité  $P$  admet une factorisation d'après les parents markoviens des variables relatifs au graphe acyclique et dirigé  $G$ ,  $P$  et  $G$  sont dits **compatibles**.

Pour un ordre donné des variables, si  $P(v)$  est strictement positive, le graphe  $G$  compatible est unique.

# Décomposition et compatibilité

## Décomposition de la loi de probabilité jointe

Variables  $X_{1,j}, \dots, X_{n,j}$  **parents markoviens** de  $X_j$  : l'ensemble minimal des prédécesseurs de  $X_j$  qui rendent  $X_j$  indépendant de tous ses autres prédécesseurs.



**Décomposition** de  $P(v)$  à l'aide des parents markoviens de chaque variable

$$P(v) = P(s, p, a, m, g)$$

$$P(v) = P(s)P(a|s)P(p|s)P(m|a, p)P(g|m).$$

Unicité de l'ensemble des parents Markoviens si  $P(v)$  est strictement positive.

## Compatibilité au sens de Markov

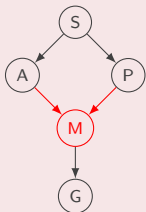
Si une loi de probabilité  $P$  admet une factorisation d'après les parents markoviens des variables relatifs au graphe acyclique et dirigé  $G$ ,  $P$  et  $G$  sont dits **compatibles**.

Pour un ordre donné des variables, si  $P(v)$  est strictement positive, le graphe  $G$  compatible est unique.

## Décomposition et compatibilité

### Décomposition de la loi de probabilité jointe

Variables  $X_{1,j}, \dots, X_{n,j}$  **parents markoviens** de  $X_j$  : l'ensemble minimal des prédécesseurs de  $X_j$  qui rendent  $X_j$  indépendant de tous ses autres prédécesseurs.



**Décomposition** de  $P(v)$  à l'aide des parents markoviens de chaque variable

$$P(v) = P(s, p, a, m, g)$$

$$P(v) = P(s)P(a|s)P(p|s)P(m|a, p)P(g|m).$$

Unicité de l'ensemble des parents Markoviens si  $P(v)$  est strictement positive.

### Compatibilité au sens de Markov

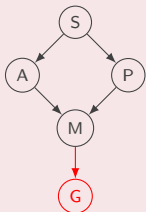
Si une loi de probabilité  $P$  admet une factorisation d'après les parents markoviens des variables relatifs au graphe acyclique et dirigé  $G$ ,  $P$  et  $G$  sont dits **compatibles**.

Pour un ordre donné des variables, si  $P(v)$  est strictement positive, le graphe  $G$  compatible est unique.

## Décomposition et compatibilité

### Décomposition de la loi de probabilité jointe

Variables  $X_{1,j}, \dots, X_{n,j}$  **parents markoviens** de  $X_j$  : l'ensemble minimal des prédécesseurs de  $X_j$  qui rendent  $X_j$  indépendant de tous ses autres prédécesseurs.



**Décomposition** de  $P(v)$  à l'aide des parents markoviens de chaque variable

$$P(v) = P(s, p, a, m, g)$$

$$P(v) = P(s)P(a|s)P(p|s)P(m|a, p)P(g|m).$$

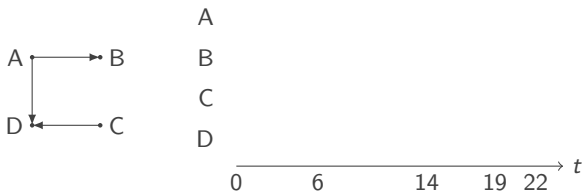
Unicité de l'ensemble des parents Markoviens si  $P(v)$  est strictement positive.

### Compatibilité au sens de Markov

Si une loi de probabilité  $P$  admet une factorisation d'après les parents markoviens des variables relatifs au graphe acyclique et dirigé  $G$ ,  $P$  et  $G$  sont dits **compatibles**.

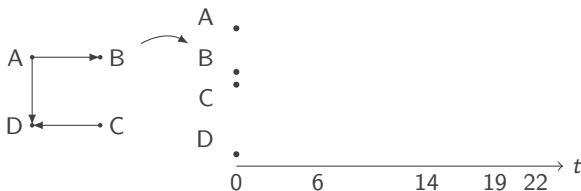
Pour un ordre donné des variables, si  $P(v)$  est strictement positive, le graphe  $G$  compatible est unique.

# Problématique



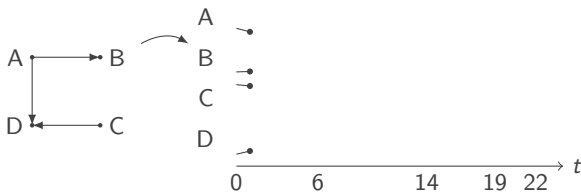
- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

# Problématique



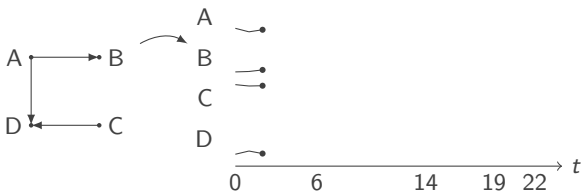
- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

# Problématique



- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

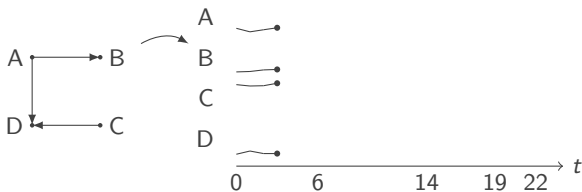
# Problématique



- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

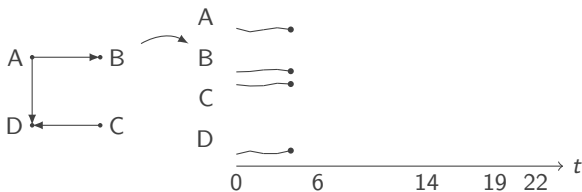


# Problématique



- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

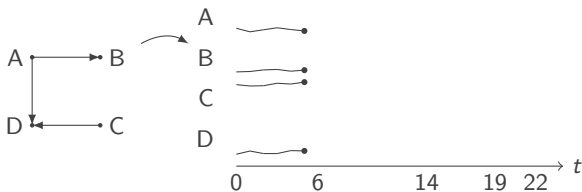
# Problématique



- Point de départ :

- séries temporelles multivariées, structure du système inconnue ;
- événements qui impactent les variables dépendantes.

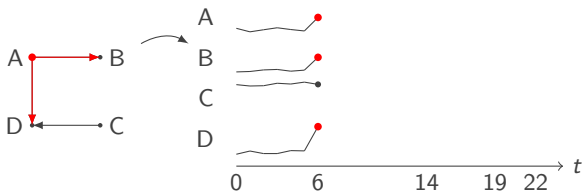
# Problématique



- Point de départ :

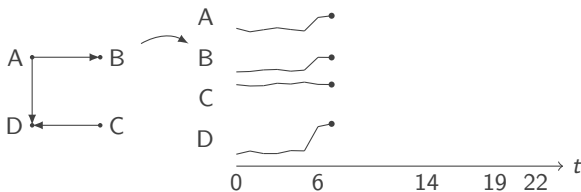
- séries temporelles multivariées, structure du système inconnue ;
- événements qui impactent les variables dépendantes.

# Problématique



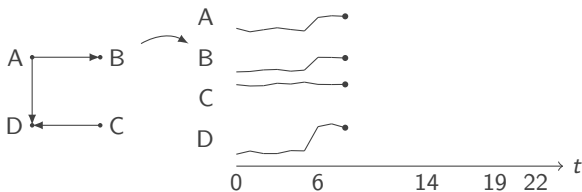
- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

# Problématique



- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

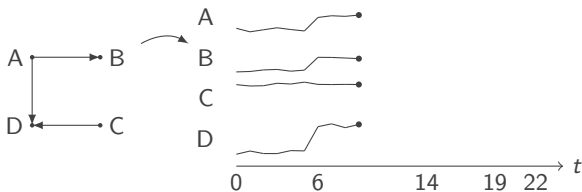
# Problématique



- Point de départ :

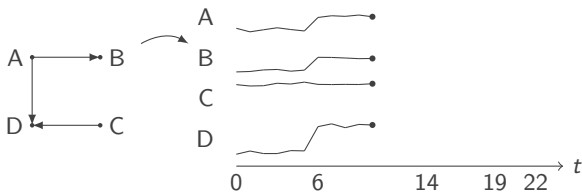
- séries temporelles multivariées, structure du système inconnue ;
- événements qui impactent les variables dépendantes.

# Problématique



- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

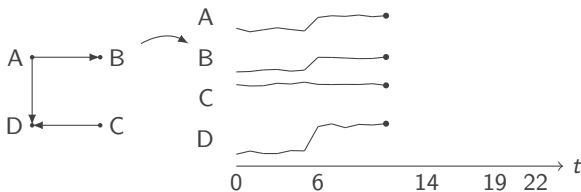
# Problématique



- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

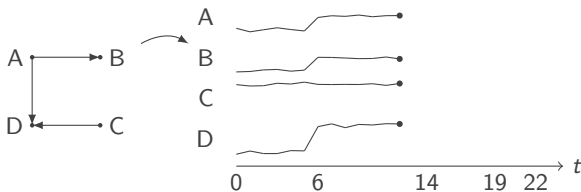


# Problématique



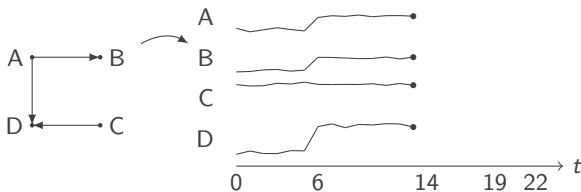
- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

# Problématique



- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

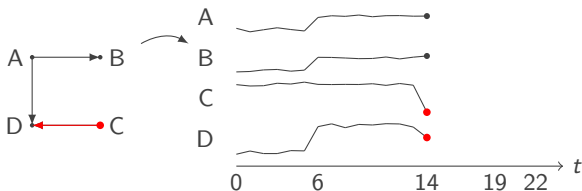
# Problématique



- Point de départ :

- séries temporelles multivariées, structure du système inconnue ;
- événements qui impactent les variables dépendantes.

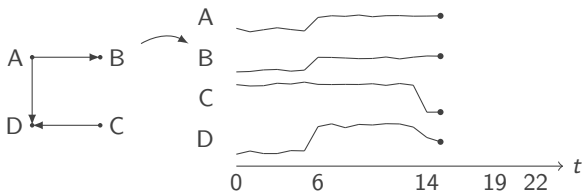
# Problématique



- Point de départ :

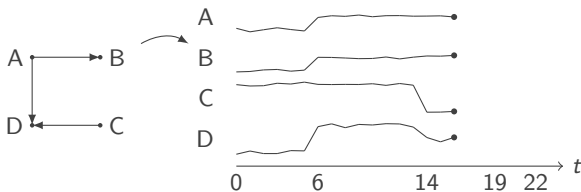
- séries temporelles multivariées, structure du système inconnue ;
- événements qui impactent les variables dépendantes.

# Problématique



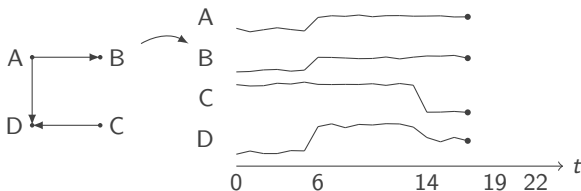
- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

# Problématique



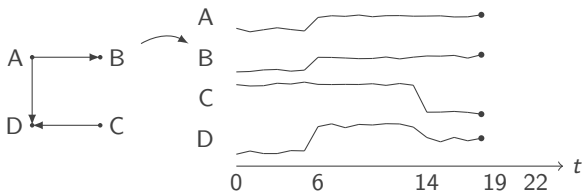
- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

# Problématique



- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

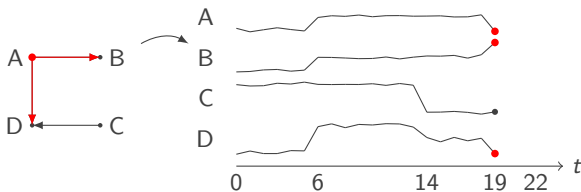
# Problématique



- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

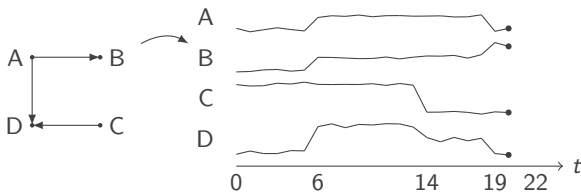


# Problématique



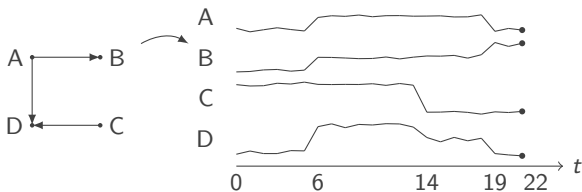
- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

# Problématique



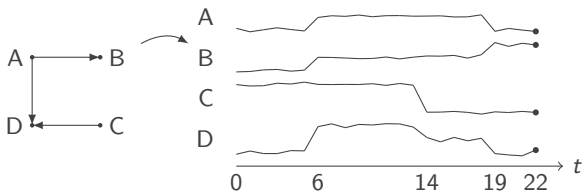
- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

# Problématique



- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.

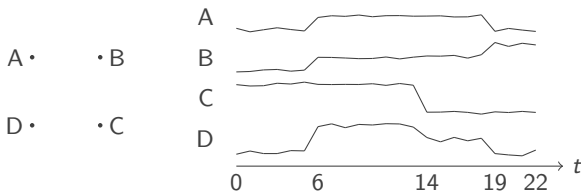
# Problématique



- Point de départ :

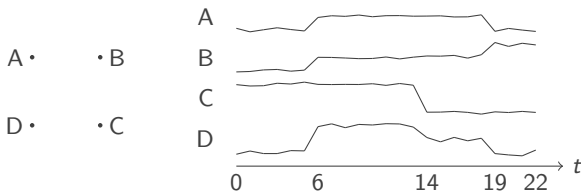
- séries temporelles multivariées, structure du système inconnue ;
- événements qui impactent les variables dépendantes.

# Problématique



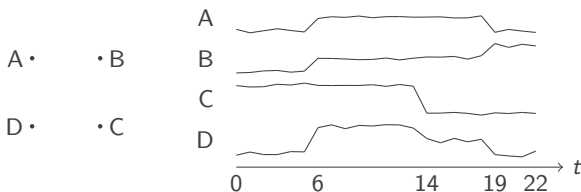
- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.
- Objectif : estimation de la structure de dépendance.

# Problématique



- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.
- Objectif : estimation de la structure de dépendance.

# Problématique



- Point de départ :
  - séries temporelles multivariées, structure du système inconnue ;
  - événements qui impactent les variables dépendantes.
- Objectif : estimation de la structure de dépendance.
- Hypothèses :
  - événements = ruptures dans les valeurs moyennes des séries ;
  - dépendances entre les variables dans l'apparition ou non de ruptures [TP01] ;
  - les changements dus à un événement sont simultanés.

- Objectif : recherche du graphe  $G$
- Principe : trouver la factorisation  $Q$  qui correspond à la probabilité jointe

$$P(r_1, \dots, r_N) = \prod_j P(r_j | pa_j)$$

→ **Comment trouver la factorisation ?**

→ Comment estimer  $P(E)$  ?

→ Comment trouver le nombre d'arêtes ?



# Plan

- 1 Introduction
- 2 Recherche du modèle causal
  - Principe
  - Calcul des factorisations
  - Sélection de la factorisation
  - Classes d'équivalence
  - Nombre d'arêtes
- 3 Estimation de la loi de probabilité jointe
  - Loi connue
  - Loi inconnue
  - Algorithme
- 4 Résultats
  - Simulations
  - Application sur des données de consommation électrique
- 5 Conclusions et perspectives

# Principe et notations

$$P(r_1, \dots, r_N) = \prod_j P(r_j | pa_j)$$

Principe :

- ➊ **générer toutes les combinaisons** de variables et de parents markoviens qui correspondent à une factorisation possible  $Q_1, Q_2, \dots, Q_D$
- ➋ **comparer la probabilité jointe**  $P$  avec les factorisations  $Q_1, \dots, Q_D$ , pour les observations  $r_1, \dots, r_N$
- ➌ **construire le graphe**  $G_d$  compatible avec la meilleure factorisation  $Q_d$ .

Notations :

- $R_j$  **variable indicatrice** d'un événement dans le signal  $j$

$$R_j = \begin{cases} 1 & \text{s'il y a un événement dans le signal } j \\ 0 & \text{sinon,} \end{cases}$$

- $\epsilon = (r_1, \dots, r_N)$  **configuration** des indicatrices,  $2^N$  combinaisons de 0 et de 1
- $P_{\epsilon_k} = P(R_1 = r_1, \dots, R_N = r_N)$  probabilité jointe,  $P(E) = (P_{\epsilon_0}, \dots, P_{\epsilon_L})$
- $Q_d$  produit des probabilités qui composent la  $d$ -ième **factorisation** possible
- $G_d$  **graphe** compatible avec la factorisation  $Q_d$

# Calcul des factorisations

# Calcul des factorisations

Probabilité jointe

$$P(E) = (P_{\epsilon_0}, \dots, P_{\epsilon_L})$$

$$\sum_{k=0}^L P_{\epsilon_k} = 1$$

$$P_{\epsilon_0} = P(R_1 = 0, \dots, R_N = 0)$$

$$P_{\epsilon_L} = P(R_1 = 1, \dots, R_N = 1)$$

...

# Calcul des factorisations

Probabilité jointe

$$P(E) = (P_{\epsilon_0}, \dots, P_{\epsilon_L}) \quad \sum_{k=0}^L P_{\epsilon_k} = 1$$

$$P_{\epsilon_0} = P(R_1 = 0, \dots, R_N = 0)$$

$$\dots$$

$$P_{\epsilon_L} = P(R_1 = 1, \dots, R_N = 1)$$

Liste des factorisations

$$Q_1(R) = P(R_1)P(R_2)P(R_3)$$

...

$$Q_D(R) = P(R_1)P(R_2|R_3)P(R_3|R_1)$$

# Calcul des factorisations

Probabilité jointe

$$P(E) = (P_{\epsilon_0}, \dots, P_{\epsilon_L}) \quad \sum_{k=0}^L P_{\epsilon_k} = 1$$

$$P_{\epsilon_0} = P(R_1 = 0, \dots, R_N = 0)$$

$$\dots$$

$$P_{\epsilon_L} = P(R_1 = 1, \dots, R_N = 1)$$

Liste des factorisations

$$Q_1(R) = P(R_1)P(R_2)P(R_3)$$

...

$$Q_D(R) = P(R_1)P(R_2|R_3)P(R_3|R_1)$$

Probabilités marginales et conditionnelles

$$P(R_1 = 0) = P_{000} + P_{001} + P_{010} + P_{011}$$

$$P(R_1 = 1) = P_{100} + P_{101} + P_{110} + P_{111}$$

...

$$P(R_N = 1) = P_{001} + P_{011} + P_{101} + P_{111}$$

$$P(R_2 = 0|R_3 = 0) = \frac{P_{000} + P_{100}}{P_{000} + P_{010} + P_{100} + P_{110}}$$

$$P(R_2 = 1|R_3 = 0) = \frac{P_{010} + P_{110}}{P_{000} + P_{010} + P_{100} + P_{110}}$$

...

# Calcul des factorisations

Probabilité jointe

$$P(E) = (P_{\epsilon_0}, \dots, P_{\epsilon_L}) \quad \sum_{k=0}^L P_{\epsilon_k} = 1$$

$$P_{\epsilon_0} = P(R_1 = 0, \dots, R_N = 0)$$

$$\dots$$

$$P_{\epsilon_L} = P(R_1 = 1, \dots, R_N = 1)$$

Liste des factorisations

$$Q_1(R) = P(R_1)P(R_2)P(R_3)$$

...

$$Q_D(R) = P(R_1)P(R_2|R_3)P(R_3|R_1)$$

Probabilités marginales et conditionnelles

$$P(R_1 = 0) = P_{000} + P_{001} + P_{010} + P_{011}$$

$$P(R_1 = 1) = P_{100} + P_{101} + P_{110} + P_{111}$$

...

$$P(R_N = 1) = P_{001} + P_{011} + P_{101} + P_{111}$$

$$P(R_2 = 0|R_3 = 0) = \frac{P_{000} + P_{100}}{P_{000} + P_{010} + P_{100} + P_{110}}$$

$$P(R_2 = 1|R_3 = 0) = \frac{P_{010} + P_{110}}{P_{000} + P_{010} + P_{100} + P_{110}}$$

...

Calcul des factorisations

$$Q_1(0, 0, 0) = P(R_1 = 0)P(R_2 = 0)P(R_3 = 0)$$

...

$$Q_D(1, 1, 1) = P(R_1 = 1)P(R_2 = 1|R_3 = 1)$$

$$P(R_3 = 1|R_1 = 1)$$

## Sélection de la factorisation

- probabilité jointe de référence

$$P(E) = (P_{\epsilon_0}, \dots, P_{\epsilon_L})$$

- $d$ -ième factorisation,  $1 \leq d \leq D$ , pour toutes les configurations  $\epsilon_0, \dots, \epsilon_L$

$$Q_d(E) = \prod_j P(r_j | pa_{d,j})$$

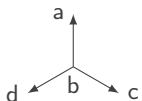
- comparaison des  $Q_d(E)$  avec  $P(E)$  : divergence de Kullback-Leibler

$$D_{KL,d} = D_{KL}(P(E) || Q_d(E)) = \sum_{k=0}^L P(\epsilon_k) \log \frac{P(\epsilon_k)}{Q_d(\epsilon_k)}$$

- sélection de la factorisation  $\hat{Q}_s(E)$ , telle que  $D_{KL,s} = \min_d (D_{KL,d})$
- structure de dépendance : graphe  $G_s$  compatible avec  $\hat{Q}_s(E)$



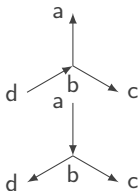
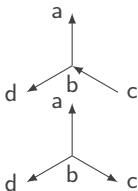
## Classes d'équivalence [TV91, Fry90]



$$P(a, b, c, d) = P(a|b)P(b)P(c|b)P(d|b)$$

Parents markoviens et ordre des variables inconnus.

graphes équivalents



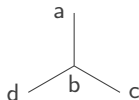
décompositions équivalentes

squelette

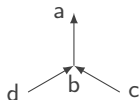
$$P(a)P(b|a)P(c|b)P(d|b)$$

$$P(a|b)P(b|c)P(c)P(d|b)$$

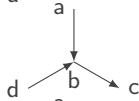
$$P(a|b)P(c|b)P(b|d)P(d)$$



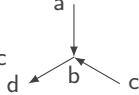
graphes non équivalents  
avec structure en v



$$P(a|b)P(b|c, d)P(c)P(d)$$



$$P(a)P(b|a, d)P(c|b)P(d)$$



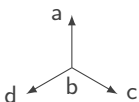
$$P(a)P(b|a, c)P(c)P(d|b)$$

## Recherche de la classe d'équivalence du modèle causal

critère graphique : même squelette et mêmes structures en v, pattern  $G_p$  [TV91].

# Nombre d'arêtes

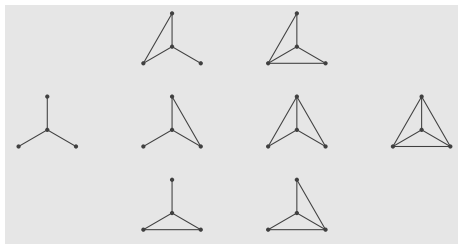
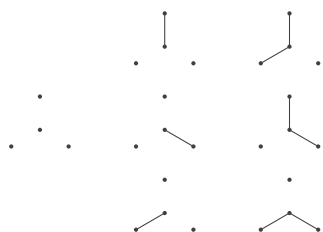
Graphe réel et décomposition associée



$$\begin{aligned}
 P(a, b, c, d) &= P(a|b)P(b)P(c|b)P(d|b) \\
 &= P(a|b, c)P(b)P(c|b)P(d|b) \\
 &= P(a|b, c, d)P(b)P(c|b)P(d|b)
 \end{aligned}$$

...

Nombre d'arêtes inconnu



0

1

2

3

4

5

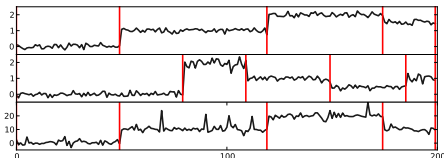
6

nb arêtes

Recherche d'une classe d'équivalence pour chaque nombre d'arêtes.

# Récapitulatif

**Objectif** : estimer  $G$  à partir de l'observation des effets d'événements dans  $N$  signaux



**Variables** : indicateurs des événements,  $R_j = 0$  ou  $1$  dans le signal  $j$

$G$  compatible avec une factorisation  $Q_d$  de la probabilité jointe  $P$

→ recherche de  $Q_d$ , qui correspond à  $P$ .

$$P_{\epsilon_0} = P(R_1 = 0, \dots, R_N = 0) \quad Q_{d,\epsilon_0} = \prod_j P(R_j = 0 | PA_{d,j} = (0, \dots, 0))$$

$$P_{\epsilon_L} = P(R_1 = 1, \dots, R_N = 1) \quad Q_{d,\epsilon_L} = \prod_j P(R_j = 1 | PA_{d,j} = (1, \dots, 1))$$

Meilleure factorisation : celle qui minimise  $D_{KL}(P || Q_d)$

**Résultats** :

- classe d'équivalence de  $Q_d$ : graphe non dirigé  $G_{p,d}$ , sauf pour les structures en  $v$
- un graphe  $G_{p,d,u}$  par nombre d'arêtes  $u$

- Objectif : recherche du graphe  $G$
- Principe : trouver la factorisation  $Q$  qui correspond à la probabilité jointe

$$P(r_1, \dots, r_N) = \prod_j P(r_j | p_{aj})$$

→ Comment trouver la factorisation ?

→ **Comment obtenir la loi de probabilité jointe  $P$  ?**

→ Comment trouver le nombre d'arêtes ?

# Plan

- 1 Introduction
- 2 Recherche du modèle causal
  - Principe
  - Calcul des factorisations
  - Sélection de la factorisation
  - Classes d'équivalence
  - Nombre d'arêtes
- 3 Estimation de la loi de probabilité jointe
  - Loi connue
  - Loi inconnue
  - Algorithme
- 4 Résultats
  - Simulations
  - Application sur des données de consommation électrique
- 5 Conclusions et perspectives

# Loi de la probabilité jointe connue

## Estimation par Monte Carlo

Loi de probabilité jointe  $P(E)$ , configurations  $(\epsilon_0, \dots, \epsilon_L)$ ,  $L - 1 = 2^N$ , liste des  $D$  classes d'équivalence pour les graphes de  $N$  nœuds

**pour**  $i \leftarrow 1, M$  **faire**

    échantillonner  $P_i(E) = (P_{i,\epsilon_0}, \dots, P_{i,\epsilon_L})$  d'après  $P(E)$

    calculer les probabilités conditionnelles  $P_{c,i}$  entre les variables  $(R_1, \dots, R_N)$

**pour**  $d \leftarrow 1, D$  **faire**

        calculer  $Q_{d,i}(E)$  le produit des  $P_{c,i}$  d'une factorisation de  $d$

        calculer  $D_{KL,d,i} = D_{KL}(P_i(E) || Q_{d,i}(E))$

**fin**

**fin**

**pour**  $d \leftarrow 1, D$  **faire**

    calculer  $D_{KL,d} = \frac{1}{M} \sum_{i=1}^M D_{KL,d,i}$

**fin**

**pour**  $u \leftarrow 1, U = N(N - 1)/2 - 1$  **faire**

    trouver  $S_u = \{d | G_d \text{ a } u \text{ arêtes}, 1 \leq d \leq D\}$

    trouver  $s_u$  tel que  $D_{KL,s_u} = \min_{s \in S_u} (D_{KL,s})$

**fin**

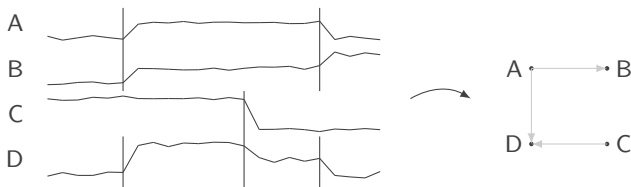
**Résultat** : Ensemble  $\{G_{s_u}\}$  des graphes des classes ( $s_u$ ) de  $u$  arêtes

# Loi de la probabilité jointe inconnue

Estimation de la loi de probabilité jointe pour échantillonner  $P = (P_{\epsilon_0}, \dots, P_{\epsilon_L})$

Observations : série temporelle multivariée  $\mathbf{X}$  ( $N \times M$ )

$x_{j,i}$  mutuellement indépendantes



Détection conjointe des événements pour estimer les probabilités  $P_{\epsilon_k}$ .

Événements : ruptures dans la moyenne, simultanés sur les variables dépendantes.

Paramètres à estimer :

- matrice des indicatrices  $\mathbf{R}$  ( $N \times M$ )
- probabilités  $P_{\epsilon_k} = P(R_1 = r_1, \dots, R_N = r_N)$

D'après le théorème de Bayes :

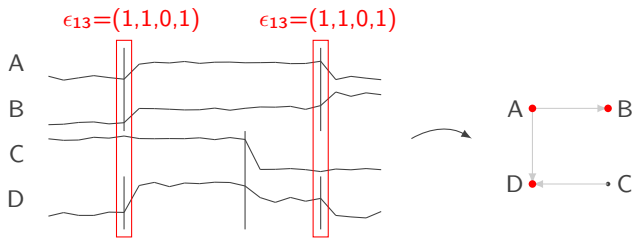
$$f(\mathbf{R}, P | \mathbf{X}) \propto L(\mathbf{X} | \mathbf{R}) f(\mathbf{R}, P)$$

# Loi de la probabilité jointe inconnue

Estimation de la loi de probabilité jointe pour échantillonner  $P = (P_{\epsilon_0}, \dots, P_{\epsilon_L})$

Observations : série temporelle multivariée  $\mathbf{X}$  ( $N \times M$ )

$x_{j,i}$  mutuellement indépendantes



Détection conjointe des événements pour estimer les probabilités  $P_{\epsilon_k}$ .

Événements : ruptures dans la moyenne, simultanés sur les variables dépendantes.

Paramètres à estimer :

- matrice des indicatrices  $\mathbf{R}$  ( $N \times M$ )
- probabilités  $P_{\epsilon_k} = P(R_1 = r_1, \dots, R_N = r_N)$

D'après le théorème de Bayes :

$$f(\mathbf{R}, P | \mathbf{X}) \propto L(\mathbf{X} | \mathbf{R}) f(\mathbf{R}, P)$$

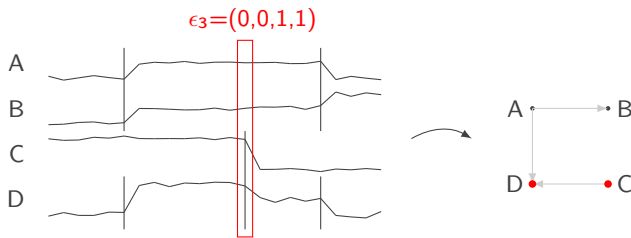


# Loi de la probabilité jointe inconnue

Estimation de la loi de probabilité jointe pour échantillonner  $P = (P_{\epsilon_0}, \dots, P_{\epsilon_L})$

Observations : série temporelle multivariée  $\mathbf{X}$  ( $N \times M$ )

$x_{j,i}$  mutuellement indépendantes



Détection conjointe des événements pour estimer les probabilités  $P_{\epsilon_k}$ .

Événements : ruptures dans la moyenne, simultanés sur les variables dépendantes.

Paramètres à estimer :

- matrice des indicatrices  $\mathbf{R}$  ( $N \times M$ )
- probabilités  $P_{\epsilon_k} = P(R_1 = r_1, \dots, R_N = r_N)$

D'après le théorème de Bayes :

$$f(\mathbf{R}, P | \mathbf{X}) \propto L(\mathbf{X} | \mathbf{R}) f(\mathbf{R}, P)$$

# Estimation de $P(E)$ par une méthode d'inférence Bayésienne

$$f(\mathbf{R}, P|\mathbf{X}) \propto L(\mathbf{X}|\mathbf{R})f(\mathbf{R}, P)$$

- Choix d'un modèle pour la fonction de vraisemblance  $L(\mathbf{X}|\mathbf{R})$  [HCGPA14].
- Matrice des indicateurs de ruptures :

à l'instant  $i$  :

$$\mathbf{R} = \begin{pmatrix} \dots & 0\dots 1\dots 0\dots 1\dots 0\dots \\ \dots & 0\dots 0\dots 0\dots 1\dots 0\dots \\ \dots & 1\dots 0\dots 0\dots 0\dots 0\dots \\ \dots & 0\dots 1\dots 0\dots 0\dots 1\dots \end{pmatrix}$$

$$\begin{aligned} R_i &= (R_{1,i}, R_{2,i}, R_{3,i}, R_{4,i}) \\ &= (1, 1, 0, 0) \\ &= \epsilon_{12} \end{aligned}$$

- Relation de causalité : un changement dans la variable  $j$  affecte instantanément les variables qui dépendent de  $j$ .
- $P_{\epsilon_k}$  est la probabilité d'observer la configuration  $\epsilon_k$  dans  $\mathbf{R}$ .
- $S_k$  est le nombre de fois que  $\epsilon_k$  est dans  $\mathbf{R}$ .
- Les colonnes de  $\mathbf{R}$  sont supposées a priori indépendantes :  $f(\mathbf{R}) = \prod_{i=1}^N f(R_i)$ ,  
d'où  $f(\mathbf{R}|P) = \prod_{0 \leq k \leq L} P_{\epsilon_k}^{S_k}$ .
- $P = (P_{\epsilon_k})_{0 \leq k \leq L}$  suit la loi de Dirichlet  $\mathcal{D}_L(d)$  [DTD07].

Finalement :

$$f(\mathbf{R}, P) \propto \prod_{0 \leq k \leq L} P_{\epsilon_k}^{S_k + d_k - 1}$$

# Algorithme

$$f(R, P|X) \propto L(X|R) \prod_{0 \leq k \leq L} P_{\epsilon_k}^{S_k + d_k - 1}$$

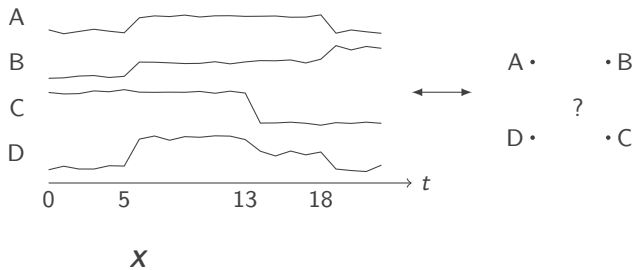
- Méthode de Monte Carlo par Chaîne de Markov,  $V$  itérations.
- Échantillonneur de Gibbs.
- Échantillonnage des paramètres
  - $R$  pour la segmentation
  - $P$  pour les probabilités des configurations

**Données** : signal multivarié  $X$   
**Résultat** : estimation de  $R$  et  $P$

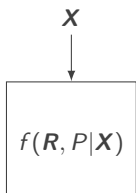
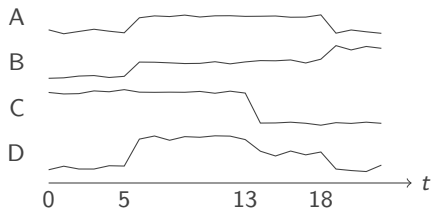
```

pour  $v \leftarrow 1, V$  faire
  |
  | pour  $i \leftarrow 1, M$  faire
  | |
  | | échantillonner  $R_i^{(v)}$ 
  | | calculer les  $S_k$  pour  $R^{(v)}$ 
  | | échantillonner  $\hat{P}^{(v)}$ 
  | |
  | | fin
  |
  | fin
  
```

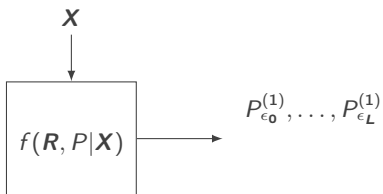
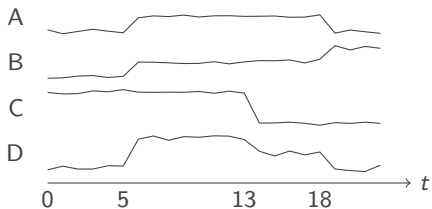
# Récapitulatif



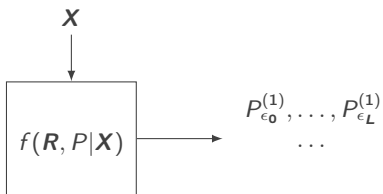
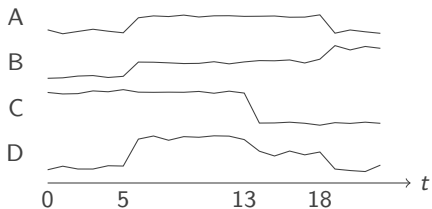
# Récapitulatif



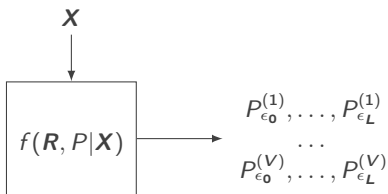
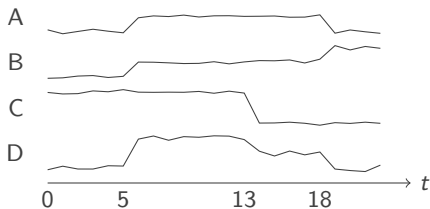
# Récapitulatif



# Récapitulatif

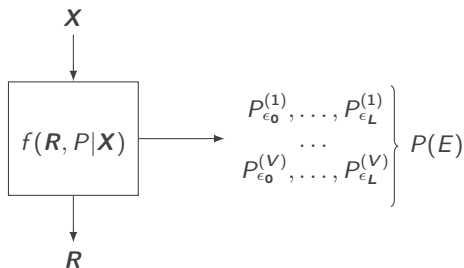
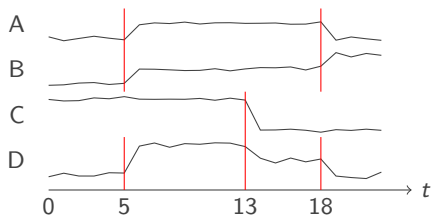


# Récapitulatif

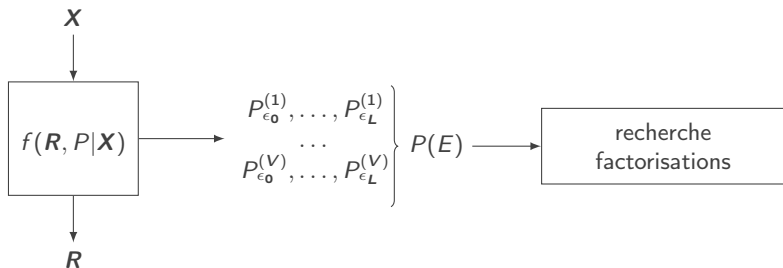
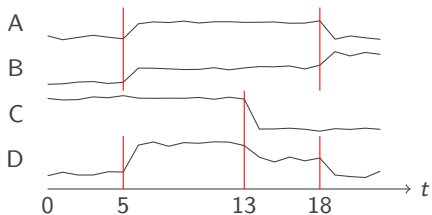




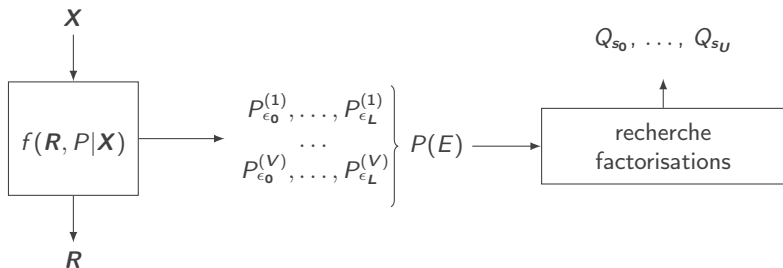
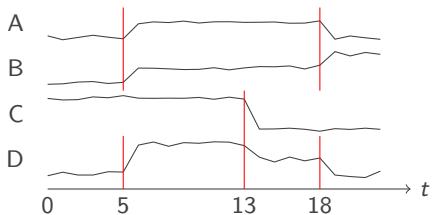
# Récapitulatif



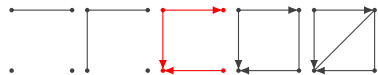
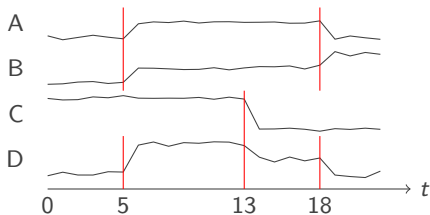
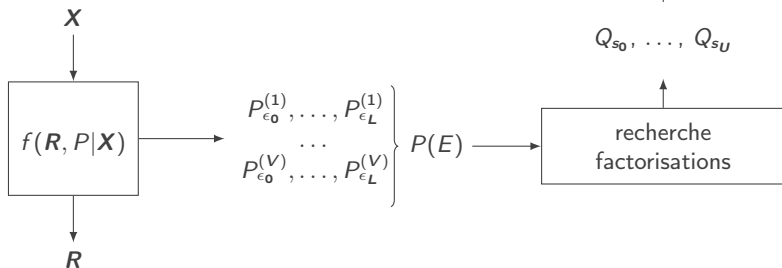
# Récapitulatif



# Récapitulatif



# Récapitulatif

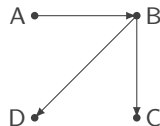

 $G_{p,s_1}, \dots, G_{p,s_U}$ 
 $Q_{s_0}, \dots, Q_{s_U}$ 


# Plan

- 1 Introduction
- 2 Recherche du modèle causal
  - Principe
  - Calcul des factorisations
  - Sélection de la factorisation
  - Classes d'équivalence
  - Nombre d'arêtes
- 3 Estimation de la loi de probabilité jointe
  - Loi connue
  - Loi inconnue
  - Algorithme
- 4 Résultats
  - Simulations
  - Application sur des données de consommation électrique
- 5 Conclusions et perspectives

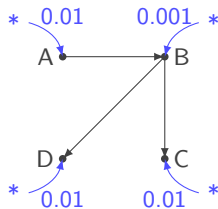
# Simulations : génération des données

- 4 signaux de 1000 points



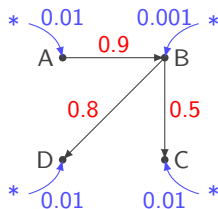
# Simulations : génération des données

- 4 signaux de 1000 points
- événements spontanés sur  $A$ ,  $B$ ,  $C$  et  $D$



# Simulations : génération des données

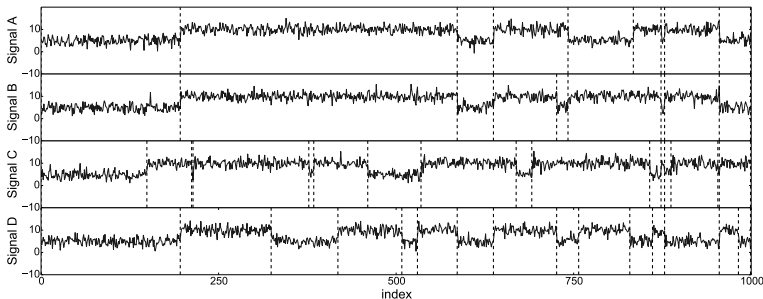
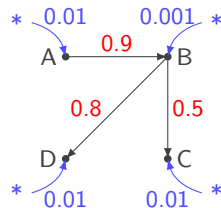
- 4 signaux de 1000 points
- événements spontanés sur  $A$ ,  $B$ ,  $C$  et  $D$
- propagation instantanée des événements entre les variables dépendantes



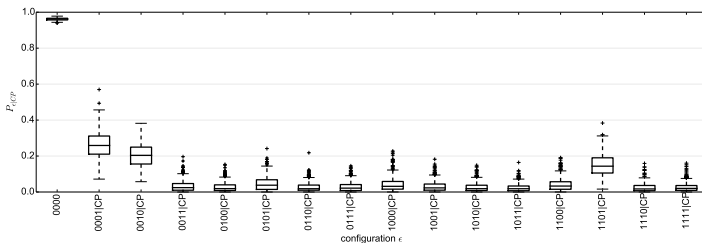
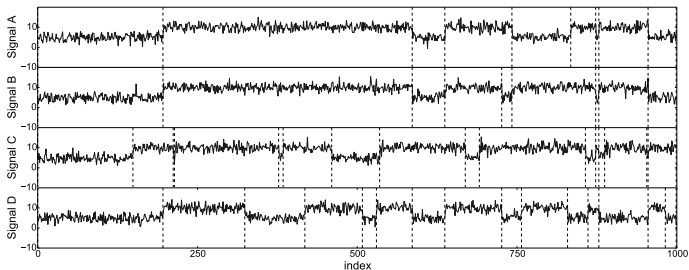
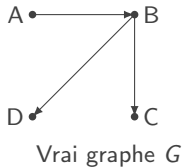


# Simulations : génération des données

- 4 signaux de 1000 points
- événements spontanés sur  $A$ ,  $B$ ,  $C$  et  $D$
- propagation instantanée des événements entre les variables dépendantes

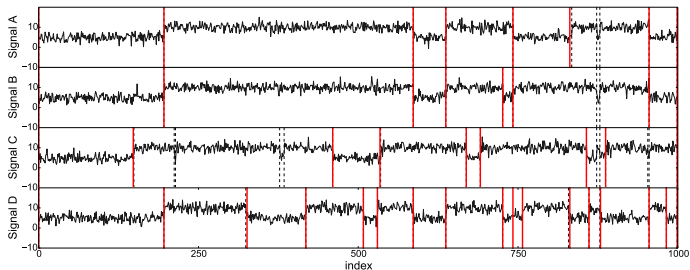
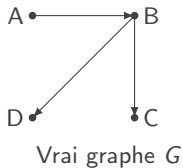


# Simulations : estimation de la loi jointe

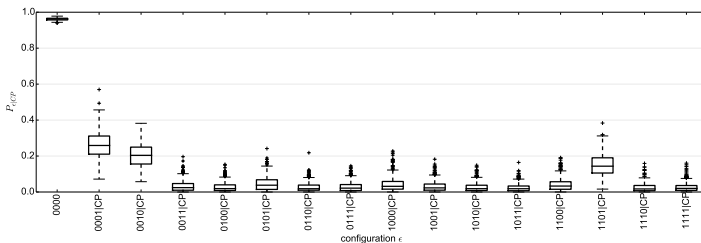


Estimation de  $P$

# Simulations : estimation de la loi jointe

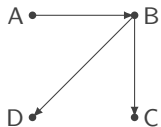


## Estimation de $R$

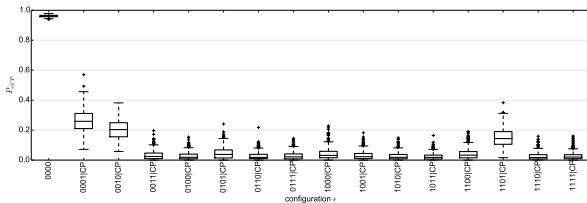


## Estimation de $P$

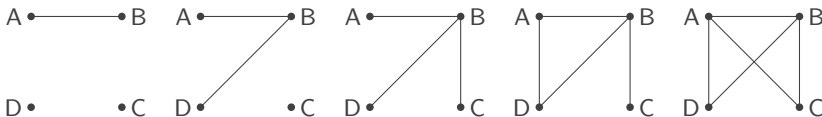
# Simulations : estimation du graphe



Vrai graphe  $G$



Estimation de  $P$



Estimations de  $G$

- Objectif : recherche du graphe  $G$
- Principe : trouver la factorisation  $Q$  qui correspond à la probabilité jointe

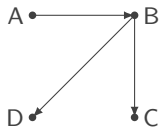
$$P(r_1, \dots, r_N) = \prod_j P(r_j | p_{aj})$$

→ Comment trouver la factorisation ?

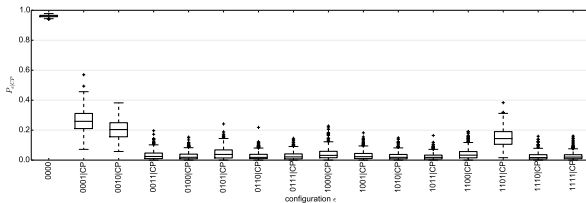
→ Comment obtenir la loi de probabilité jointe  $P$  ?

→ **Comment trouver le nombre d'arêtes ?**

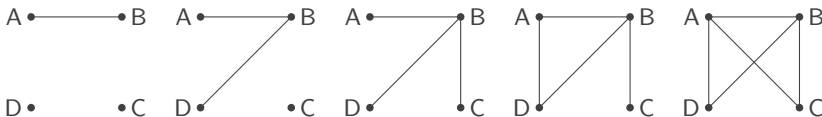
# Simulations : estimation du graphe



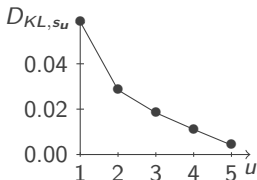
Vrai graphe  $G$



Estimation de  $P$



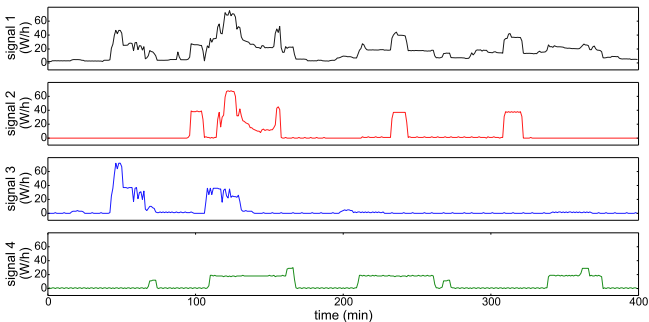
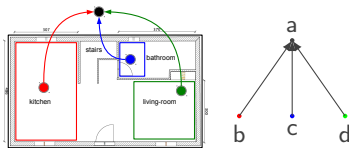
Estimations de  $G$



$D_{KL}(P||Q_{s_u})$  en fonction du nombre d'arêtes

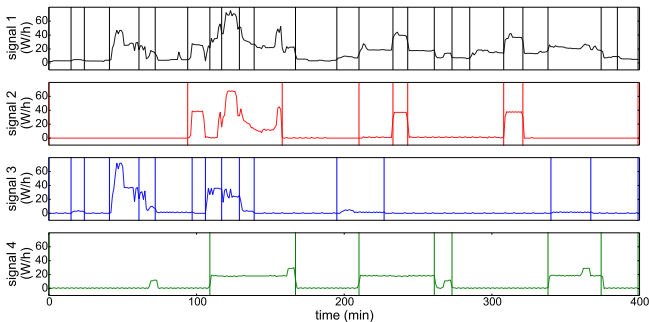
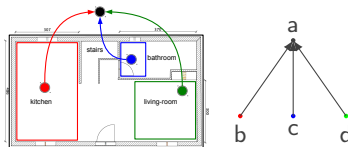
# Application sur des données de consommation électrique

## Structure de dépendance sous-jacente



# Application sur des données de consommation électrique

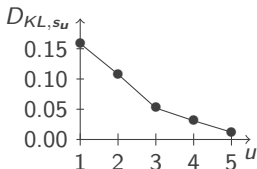
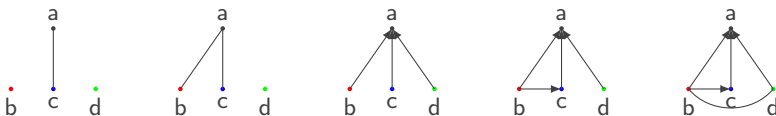
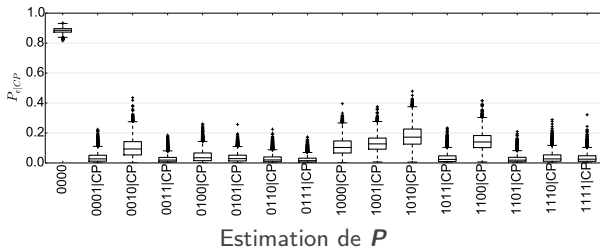
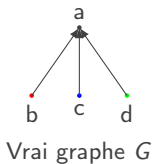
## Structure de dépendance sous-jacente



## Estimation de $R$



# Application sur des données de consommation électrique



$D_{KL}(P||Q_{s_u})$  en fonction du nombre d'arêtes

# Plan

- 1 Introduction
- 2 Recherche du modèle causal
  - Principe
  - Calcul des factorisations
  - Sélection de la factorisation
  - Classes d'équivalence
  - Nombre d'arêtes
- 3 Estimation de la loi de probabilité jointe
  - Loi connue
  - Loi inconnue
  - Algorithme
- 4 Résultats
  - Simulations
  - Application sur des données de consommation électrique
- 5 Conclusions et perspectives

# Conclusions

## Méthode

- estimation de la structure de dépendance à partir d'événements
- détection conjointe des événements
- estimation de la loi jointe des probabilités des événements
- sélection de la factorisation qui minimise  $D_{KL}(P||Q_{su})$

## Avantages

- peu d'a priori sur les données
- estimation de  $G$  via la propagation d'événements
- estimation de la probabilité jointe

## Limites

- ruptures simultanées, fréquence suffisante
- lenteur et complexité : traiter un nombre raisonnable de signaux
- estimation d'une classe de modèles équivalents
- une solution par nombre d'arêtes possible

# Perspectives

- intégration d'un test d'indépendance pour réduire le nombre de factorisations
- recherche d'un critère pour l'estimation du nombre d'arêtes
- recherche d'un critère pour déterminer les nœuds racines
- intégration de variables cachées
- modification de la fonction de vraisemblance pour traiter d'autres types d'événements
- application sur d'autres données réelles

Merci pour votre attention !  
Des questions ?

## Références

- [DTD07] N. Dobigeon, J.-Y. Tournet, and M. Davy, *Joint segmentation of piecewise constant autoregressive processes by using a hierarchical model and a bayesian sampling approach*, Signal Processing, IEEE Transactions on 55 (2007), no. 4, 1251–1263.
- [Fry90] Morten Frydenberg, *The chain graph markov property*, Scandinavian Journal of Statistics (1990), 333–353.
- [HCGPA14] Flore Harlé, Florent Chatelain, Cédric Gouy-Pailler, and Sophie Achard, *Rank-based multiple change-point detection in multivariate time series*, Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European, IEEE, 2014, pp. 1337–1341.
- [Pea00] Judea Pearl, *Causality: models, reasoning and inference*, vol. 29, Cambridge Univ Press, 2000.
- [TP01] Jin Tian and Judea Pearl, *Causal discovery from changes*, Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence, Morgan Kaufmann Publishers Inc., 2001, pp. 512–521.
- [TV91] J. Pearl T. Verma, *Equivalence and synthesis of causal models*, vol. 6, Elsevier, 1991, pp. 255–323.