State sequence prediction in imprecise hidden Markov models

Jasper De Bock & Gert de Cooman

27 July 2011

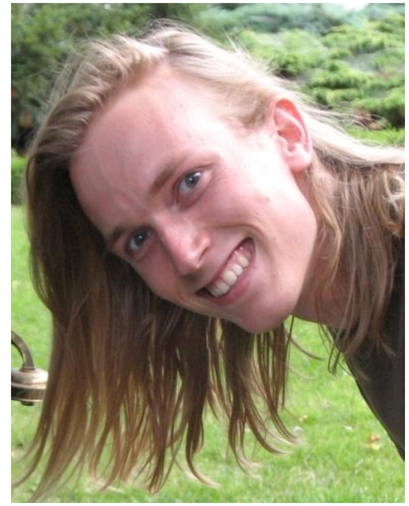
Jasper De Bock & Gert de Cooman

Jasper De Bock & Gert de Cooman

UNIVERSITEIT GENT

Research group

SYSTeMS



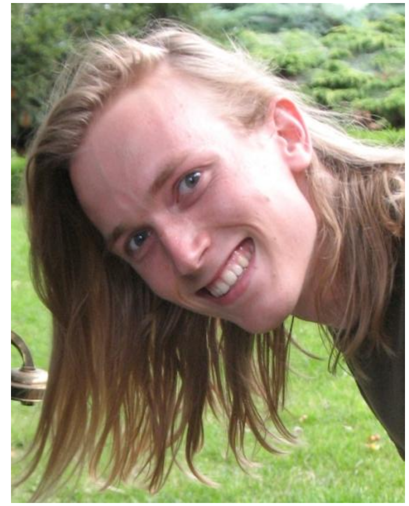
Jasper De Bock

Gert de Cooman

UNIVERSITEIT GENT

Research group

SYSTeMS



Jasper De Bock

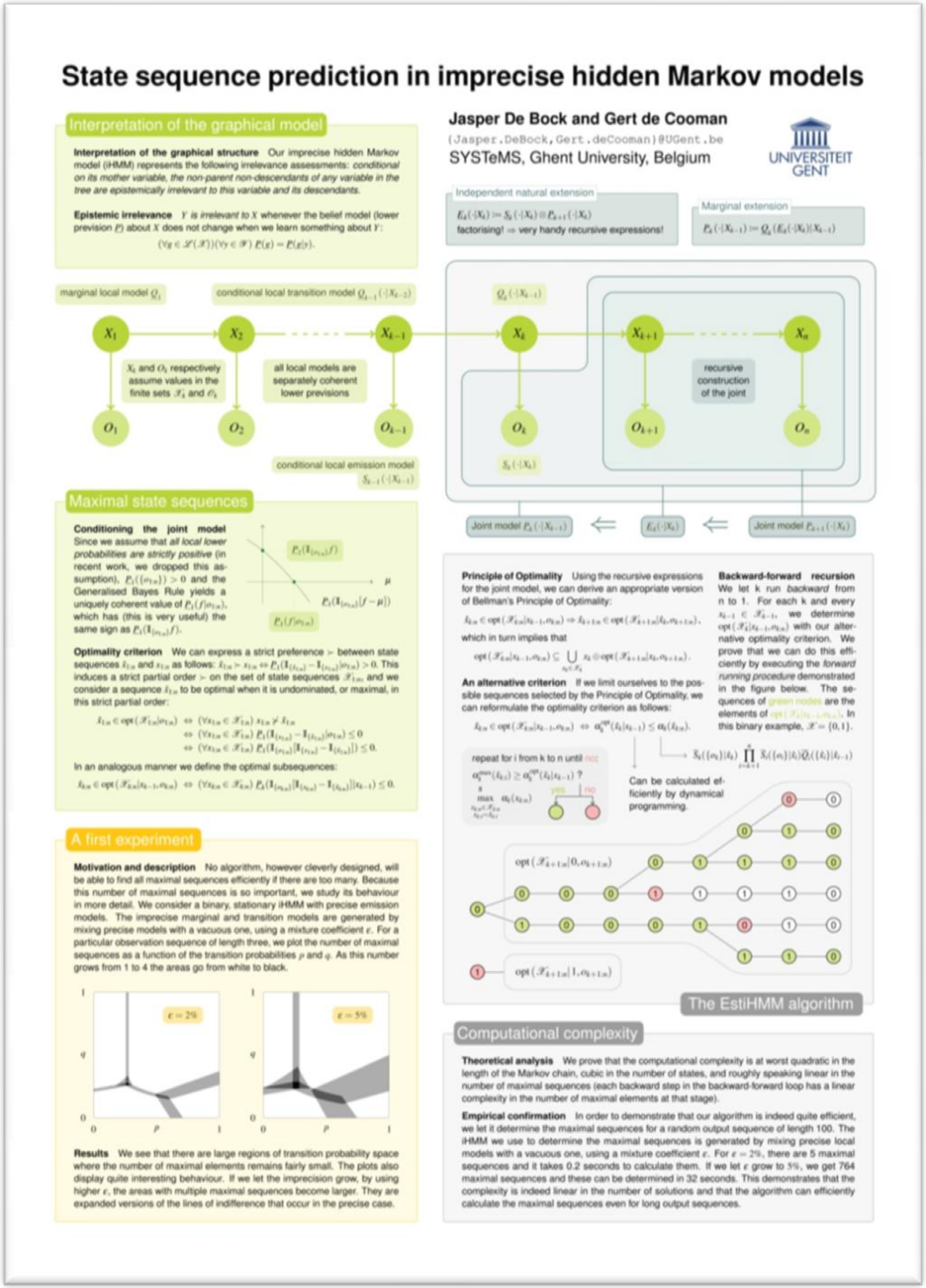
Gert de Cooman

Filip Hermans

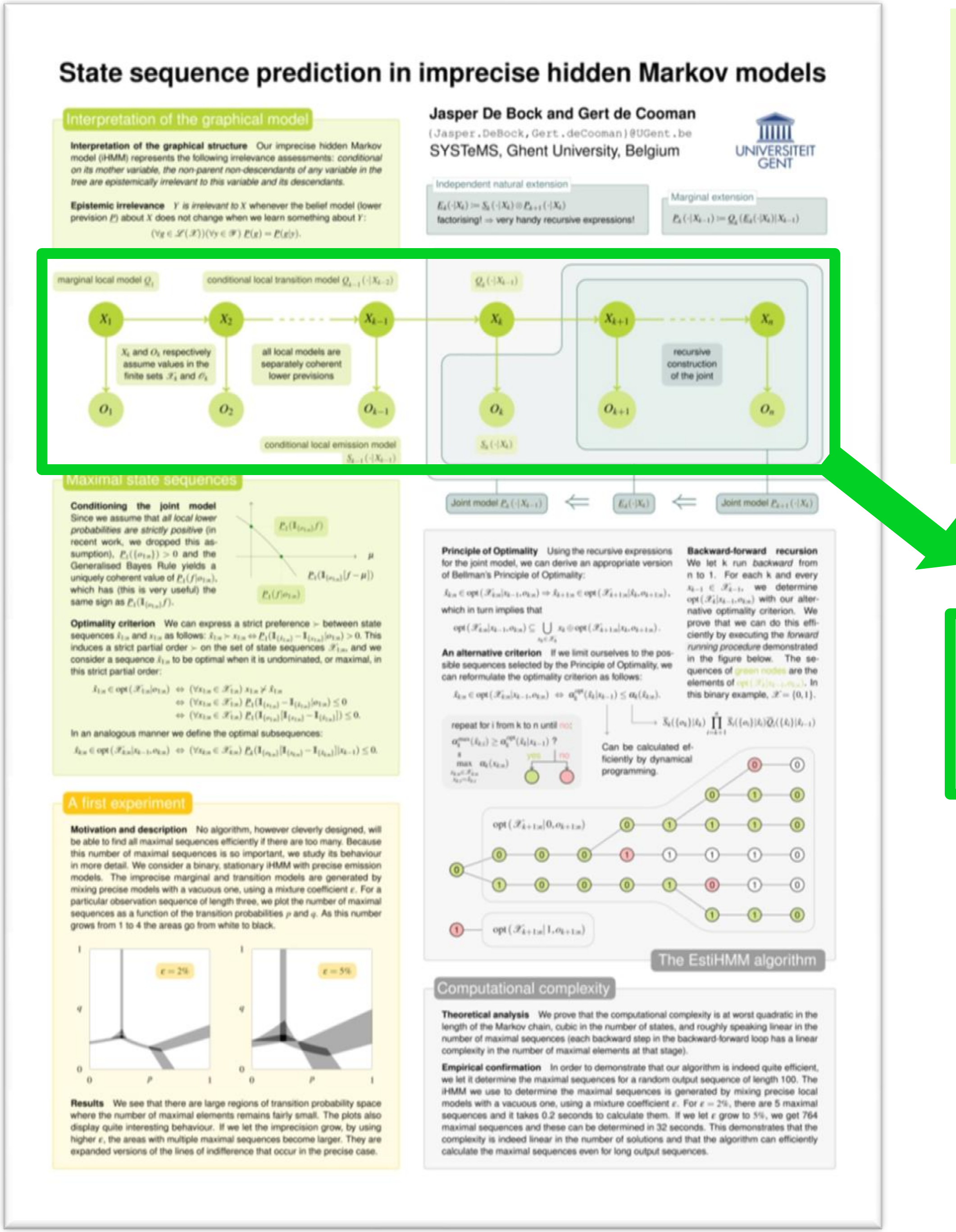
Erik Quaeghebeur

Keivan Shariatmadar

Arthur Van Camp



State sequence prediction in imprecise hidden Markov models

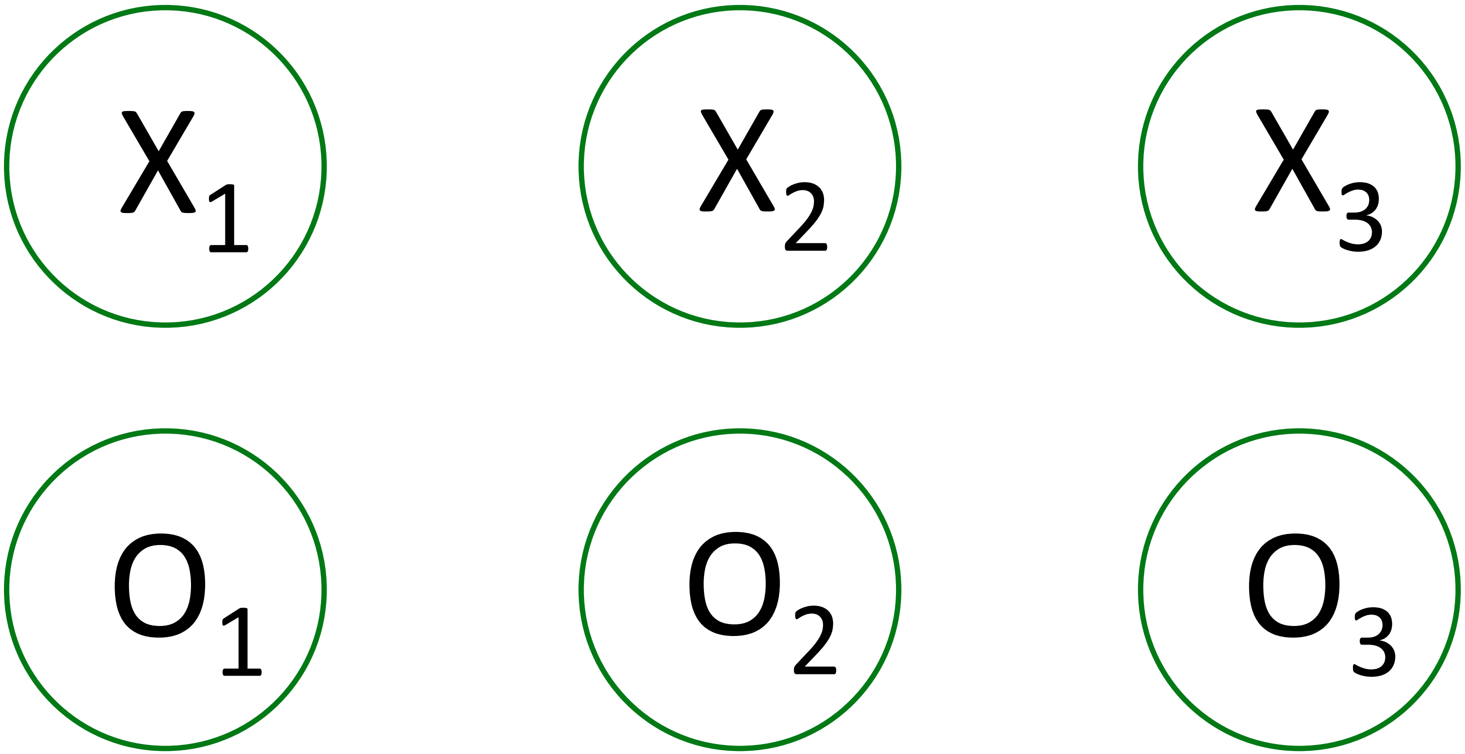


State sequence prediction in imprecise hidden Markov models

The imprecise hidden Markov model

Imprecise hidden Markov model

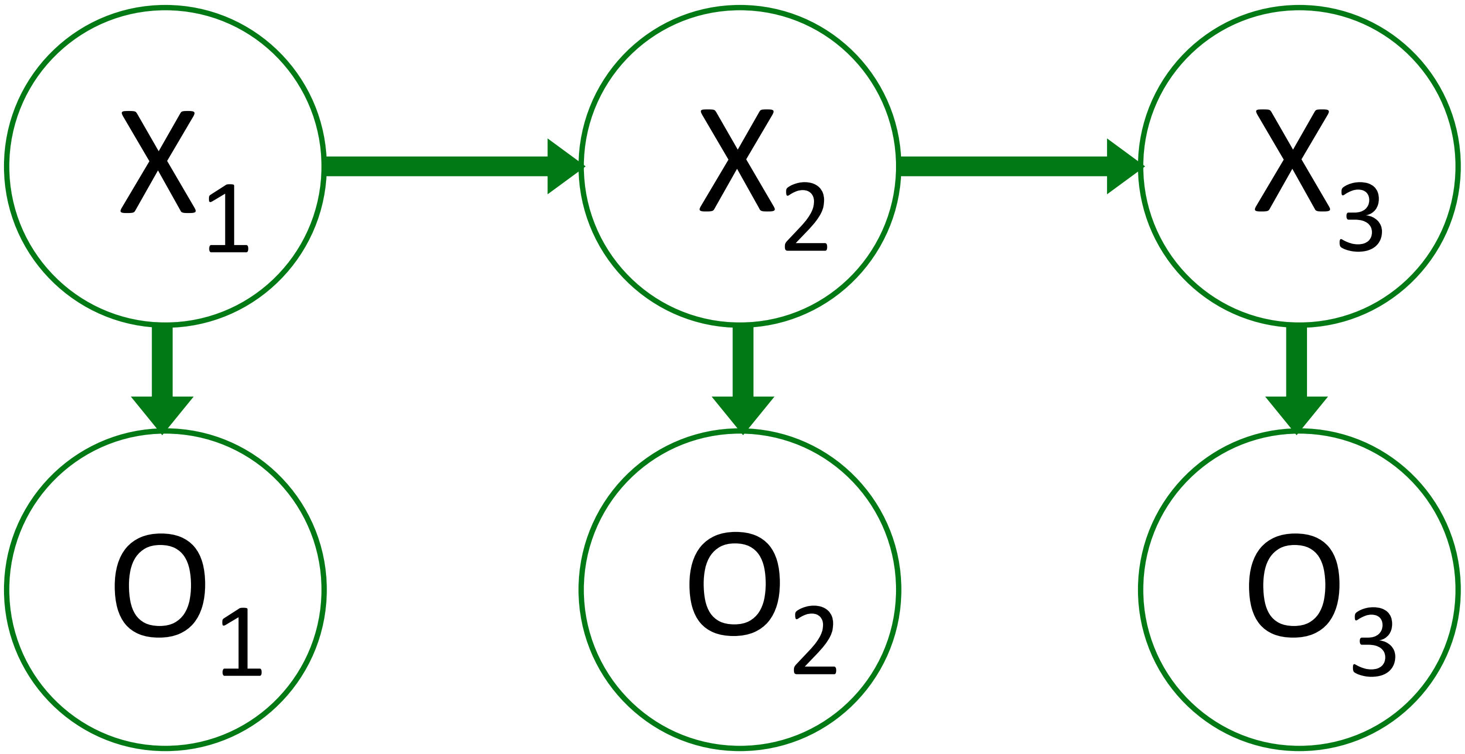
A sequence of hidden state variables



A sequence of observable variables

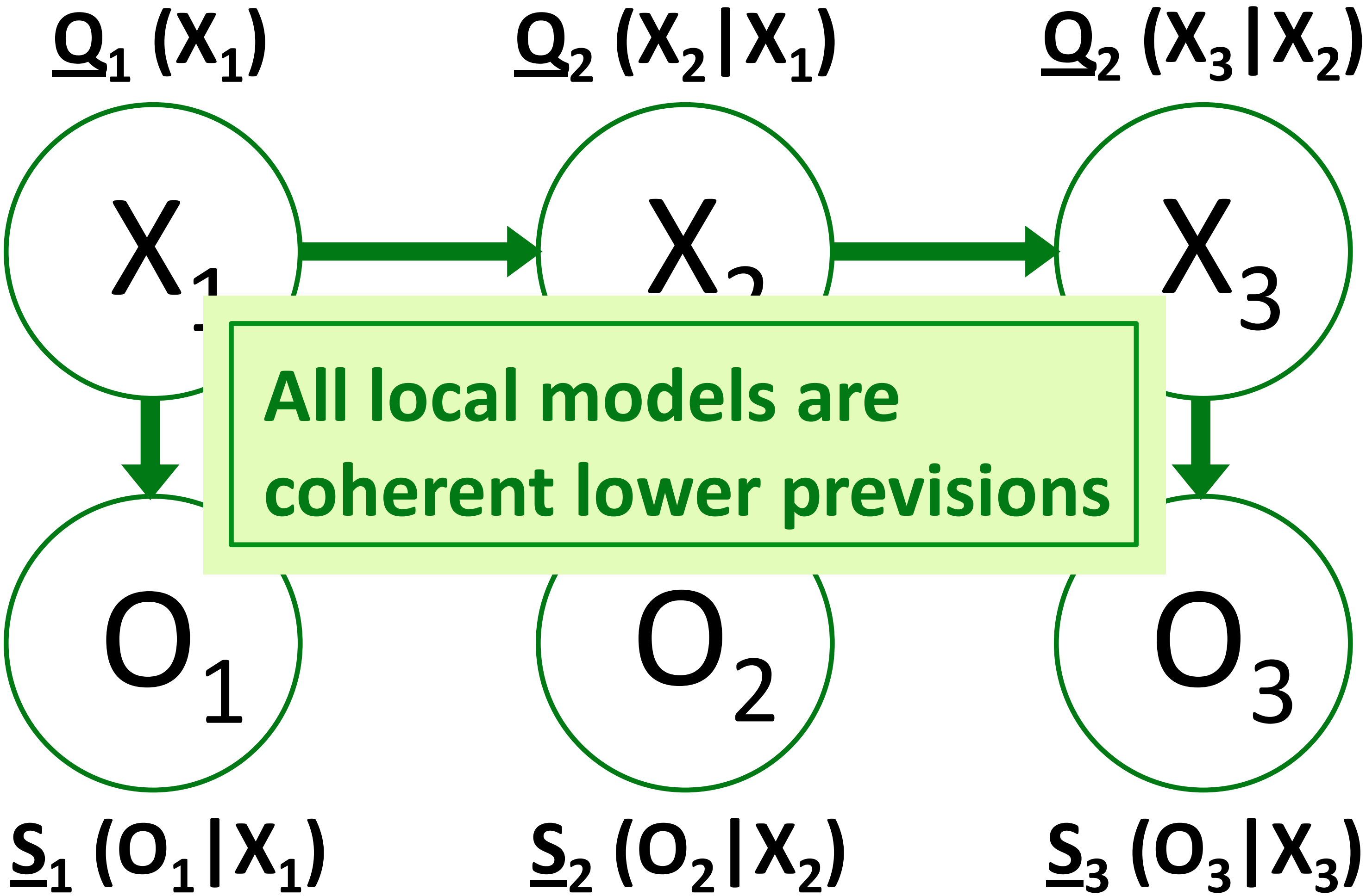
Imprecise hidden Markov model

A sequence of hidden state variables



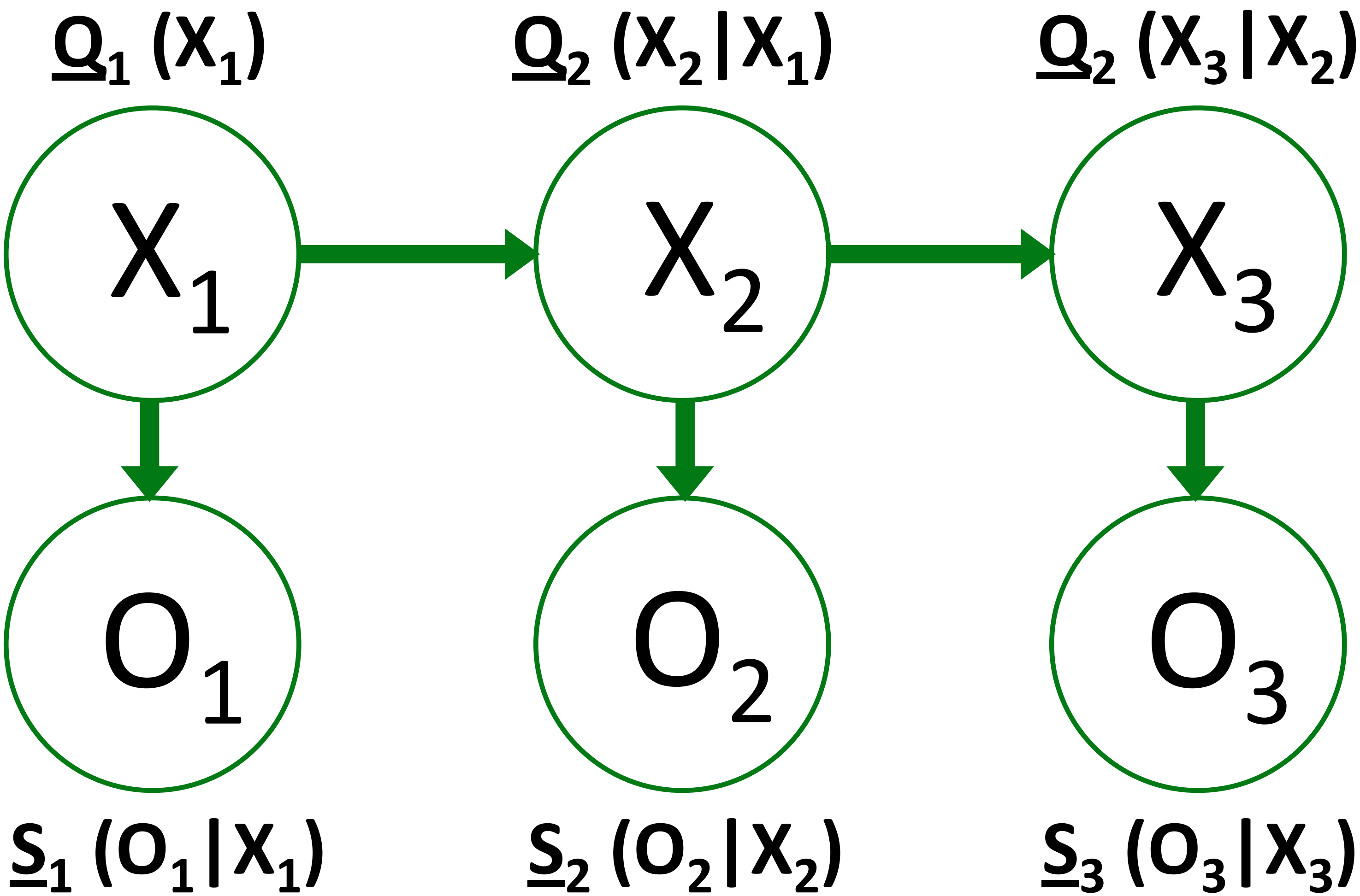
A sequence of observable variables

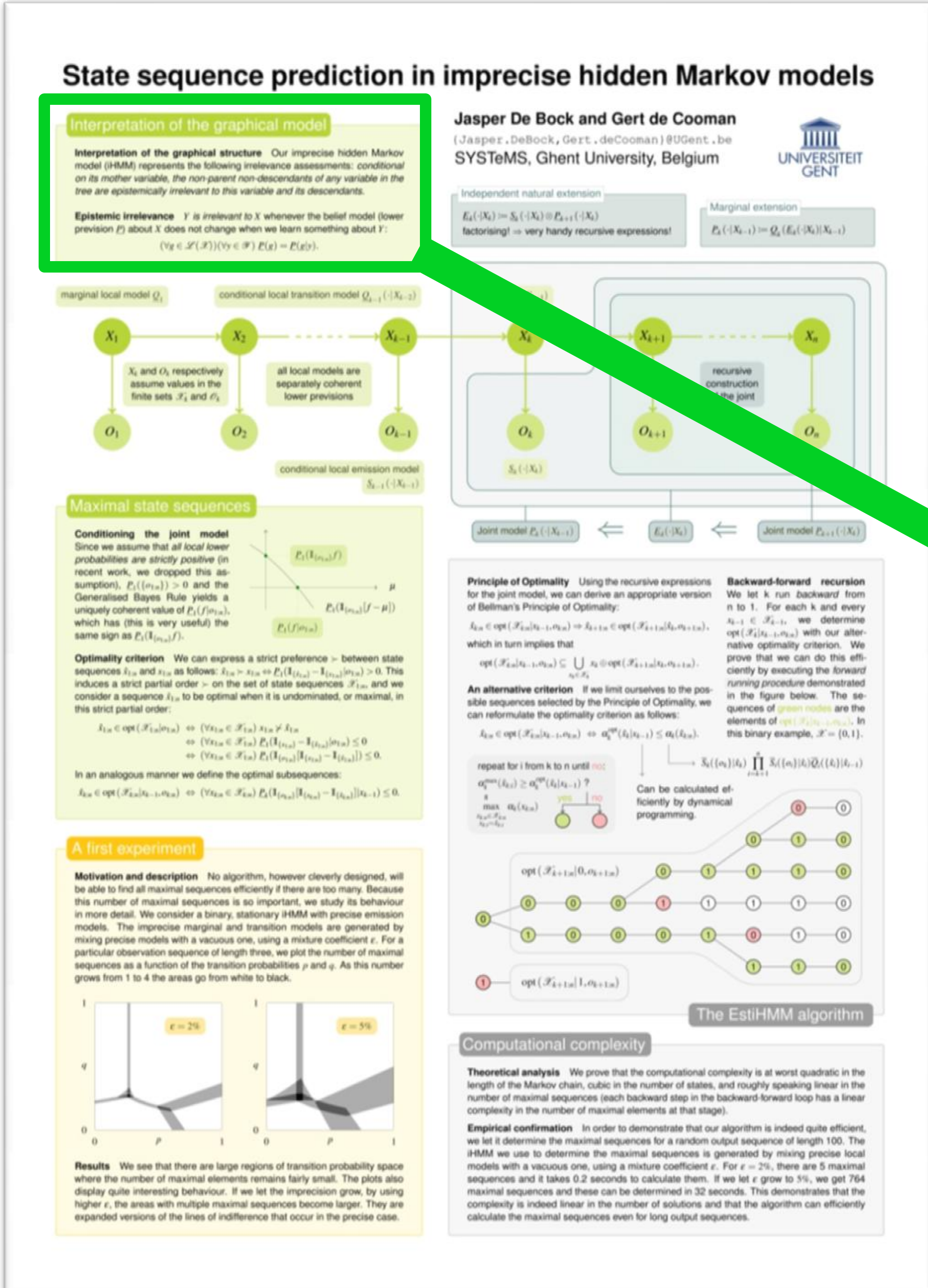
Imprecise hidden Markov model



All local models are coherent lower previsions

Imprecise hidden Markov model

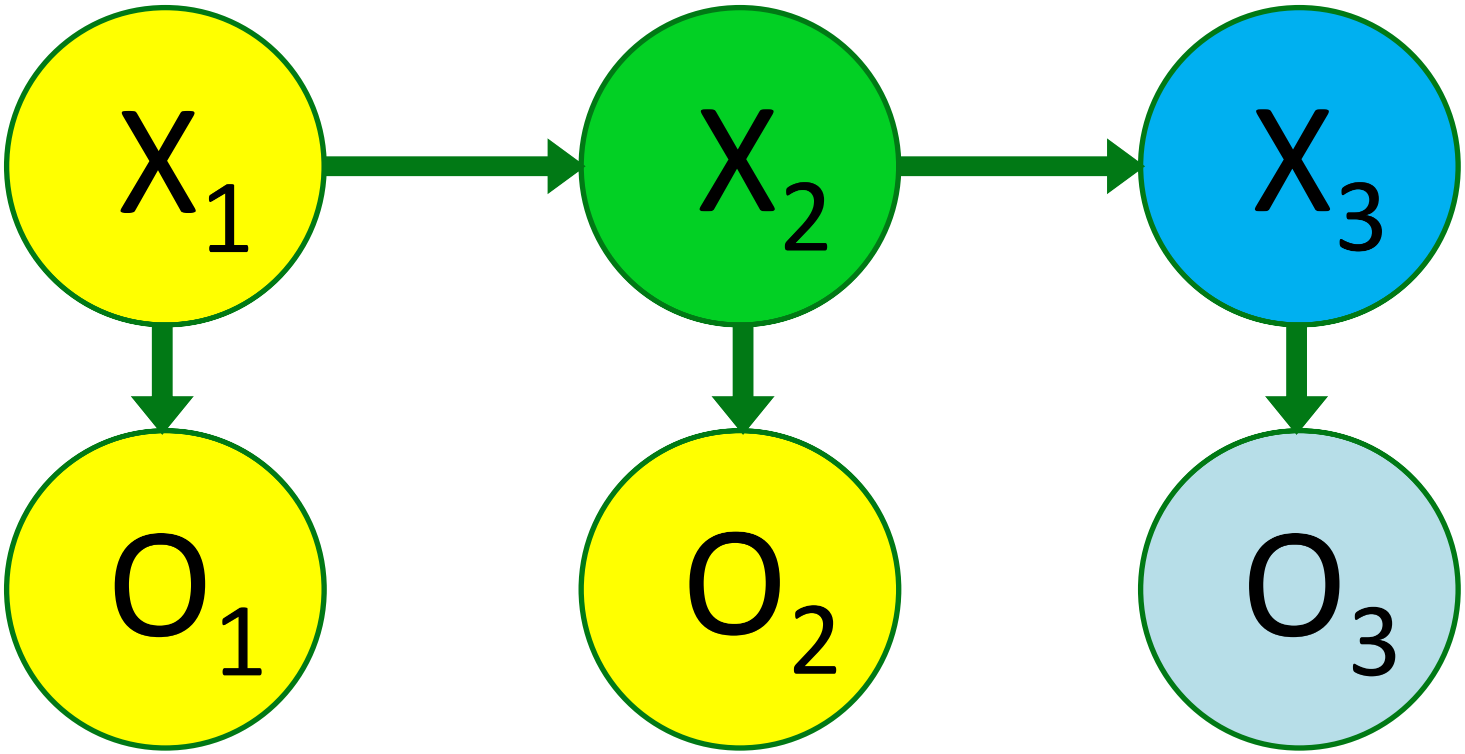




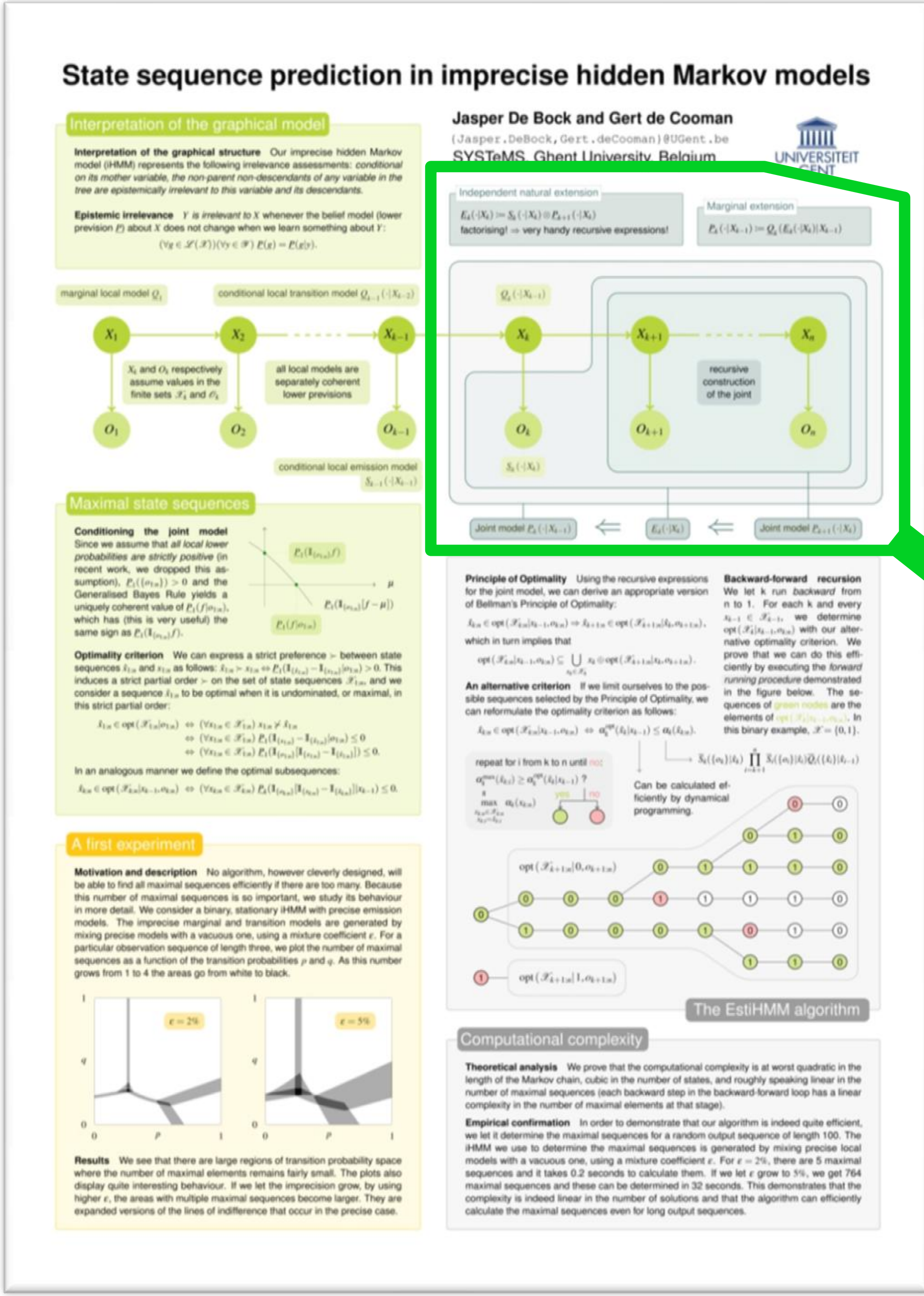
State sequence prediction in imprecise hidden Markov models

Epistemic Irrelevance

Epistemic irrelevance



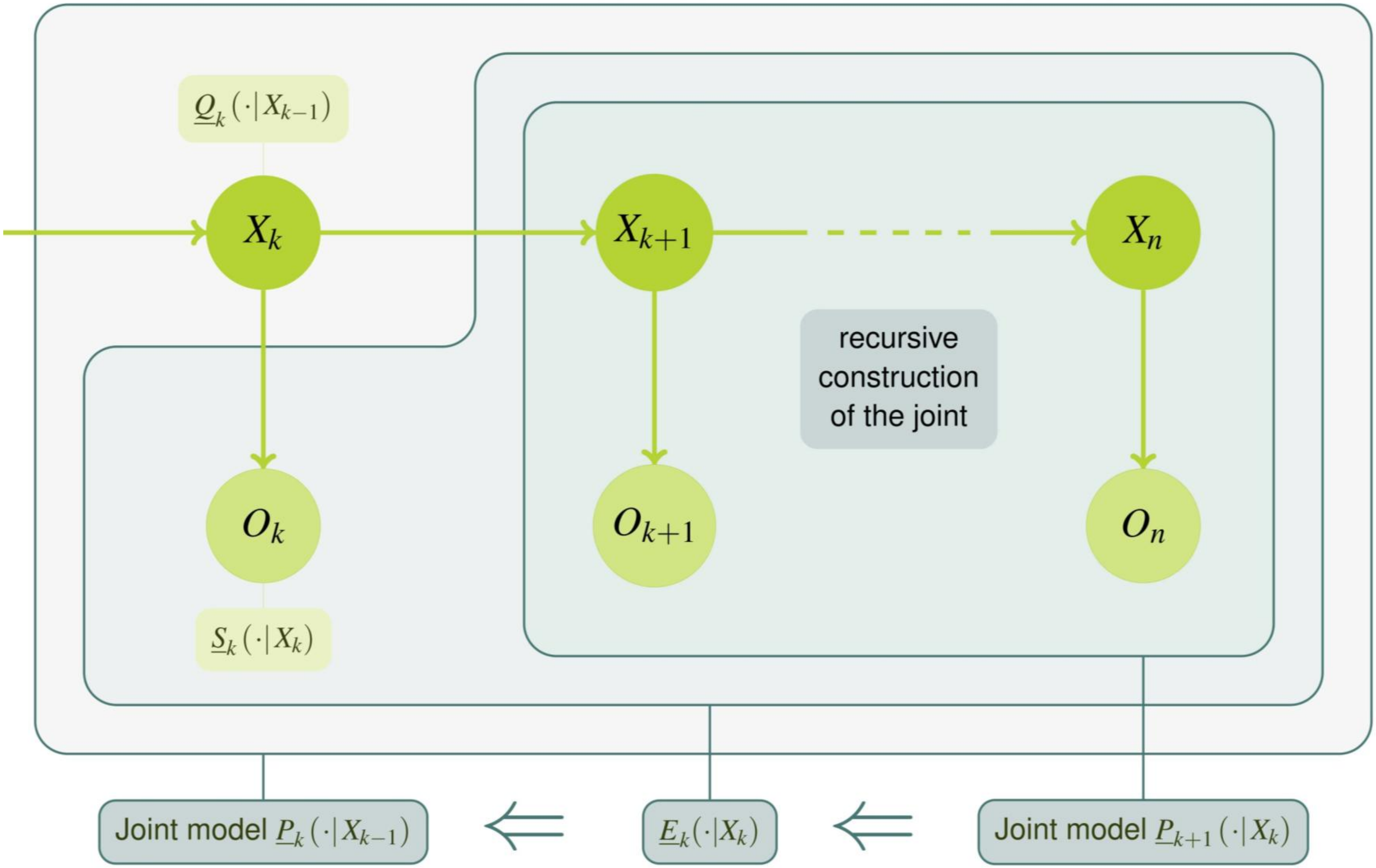
Conditional on its **mother variable**, the **non-parent non-descendants** of any **variable** in the tree are epistemically irrelevant to this **variable** and its **descendants**



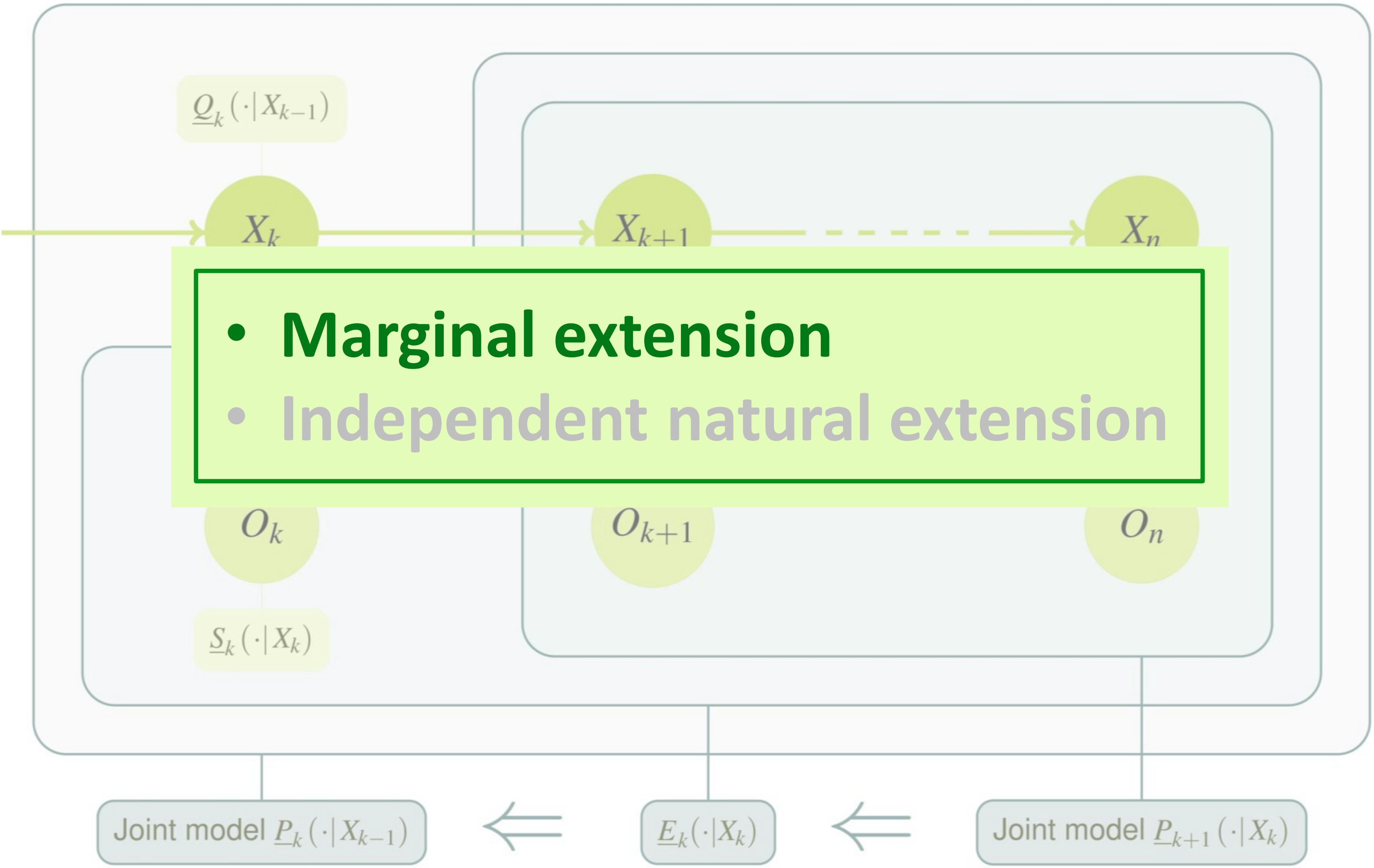
State sequence prediction in imprecise hidden Markov models

Recursive construction of a joint model for the imprecise hidden Markov model

Recursive construction of a joint model



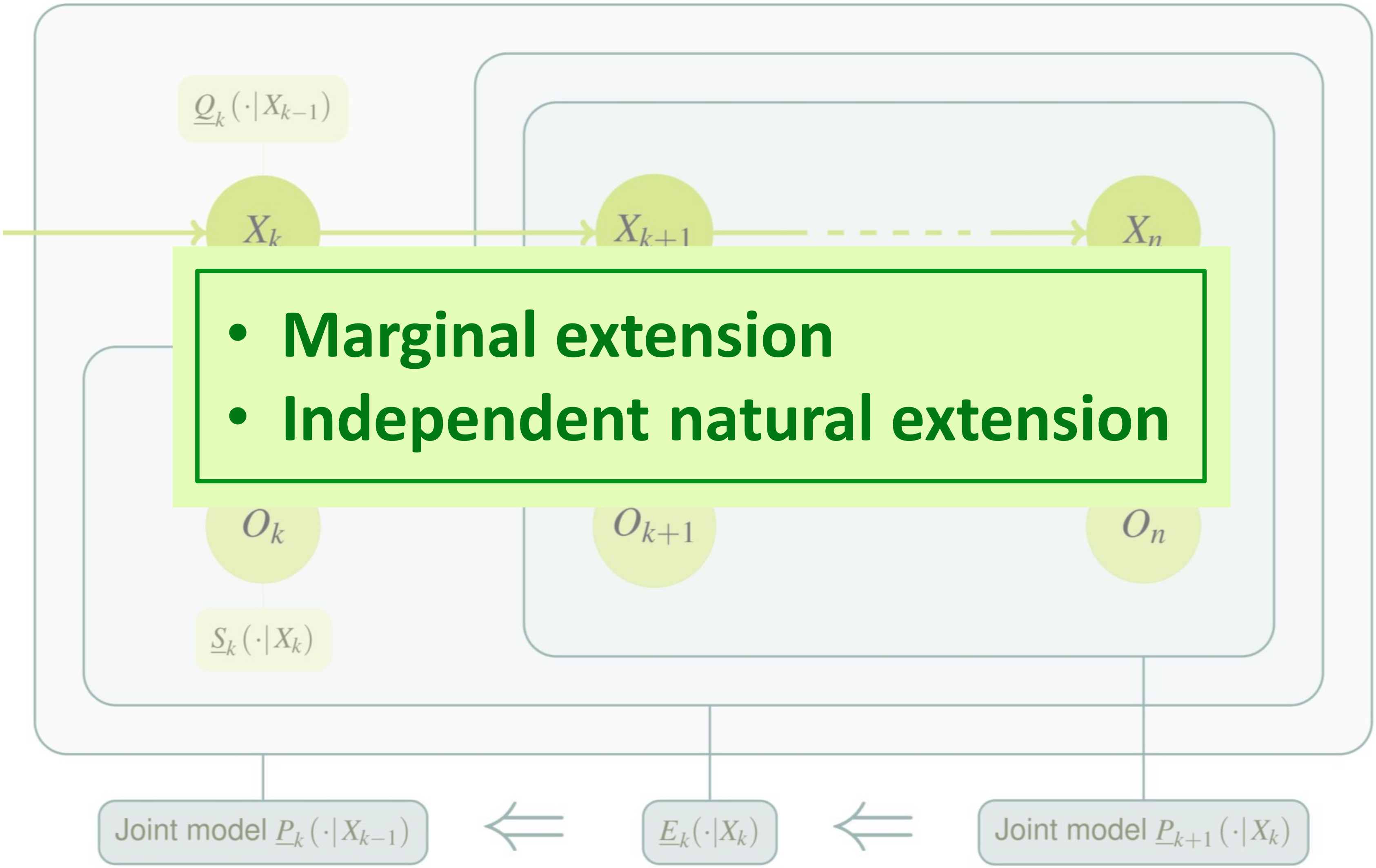
Recursive construction of a joint model

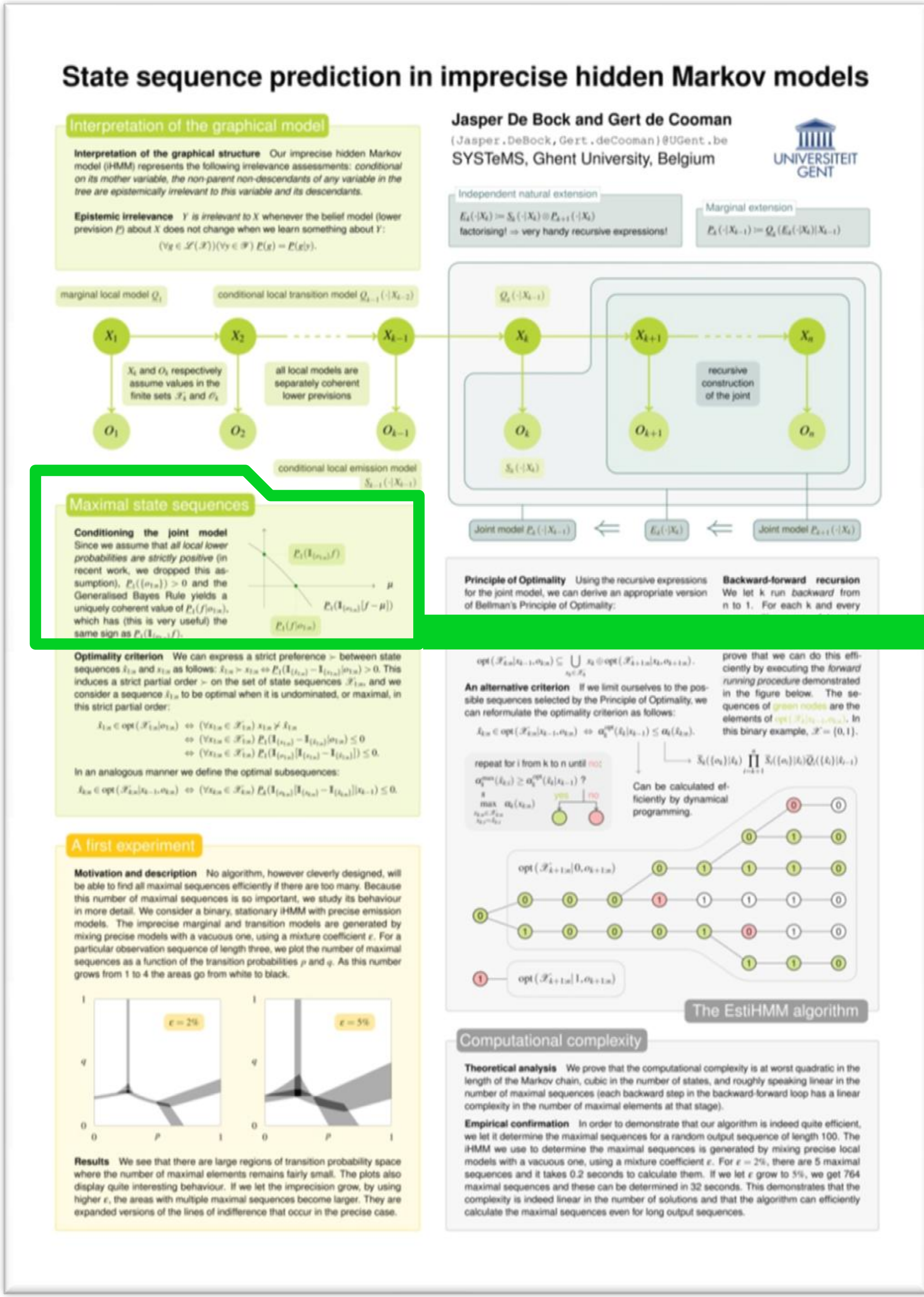


- **Marginal extension**
- Independent natural extension

Recursive construction of a joint model

- **Marginal extension**
- **Independent natural extension**

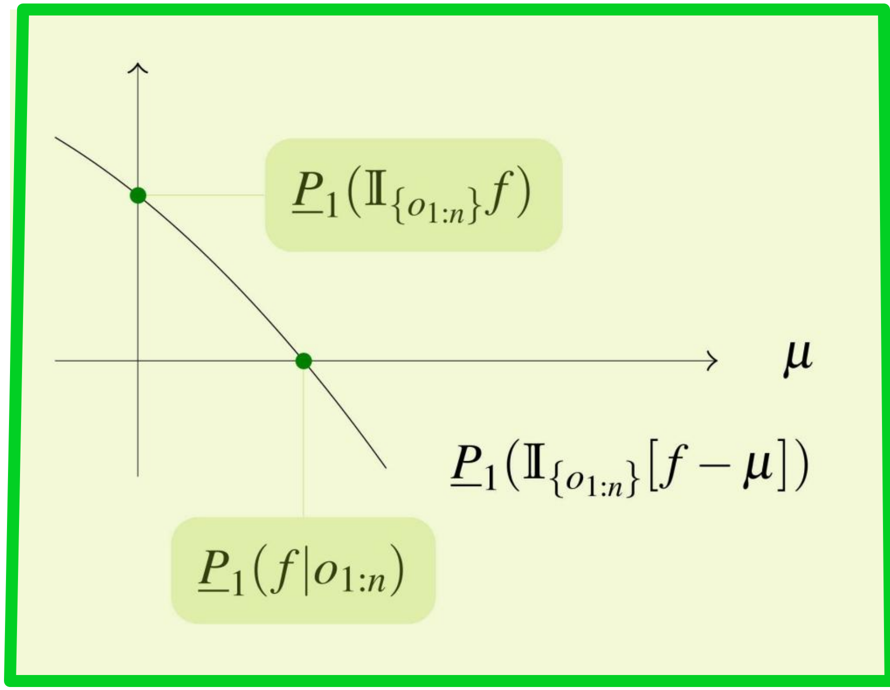




State sequence prediction in imprecise hidden Markov models

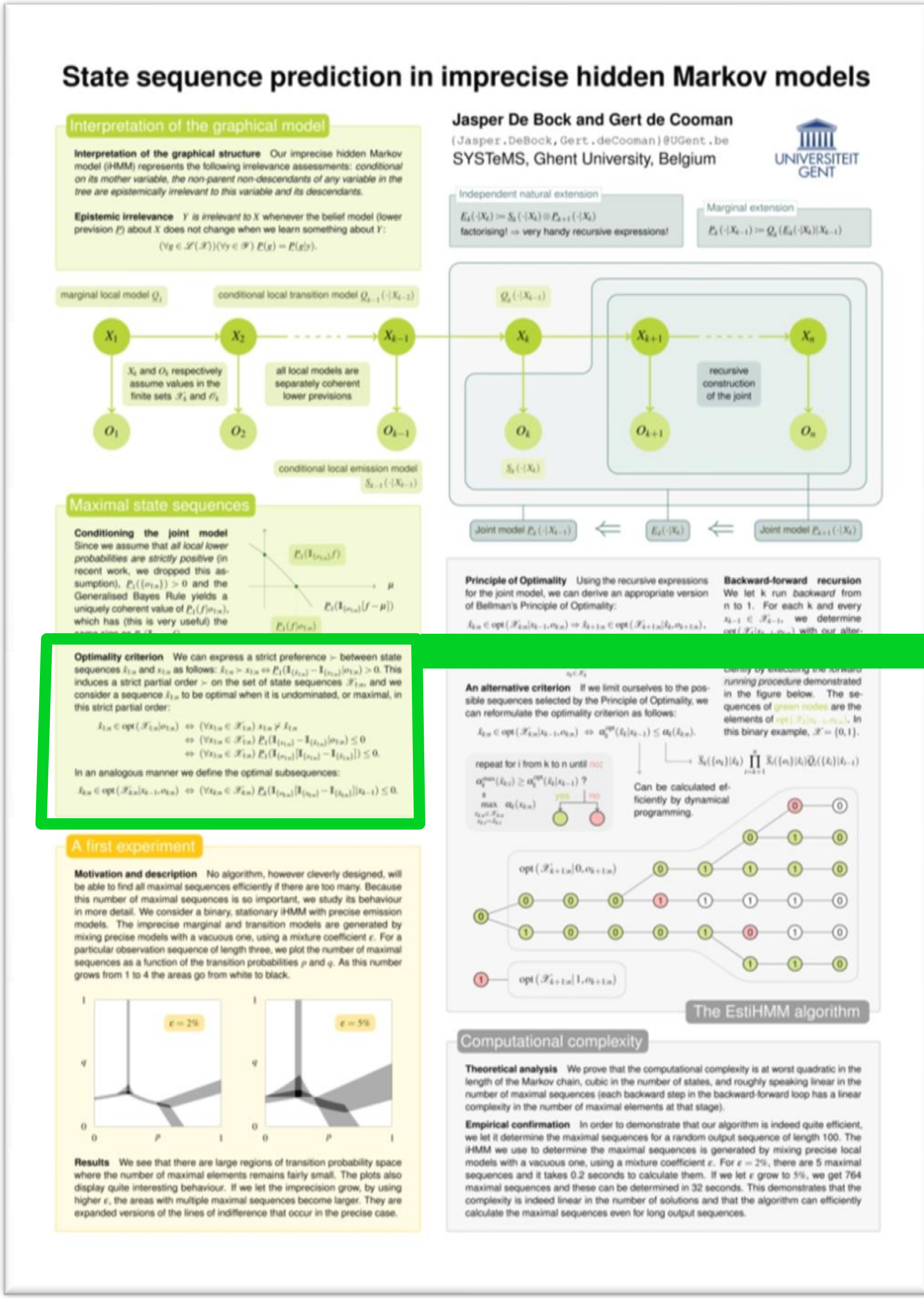
Conditioning the model on the observations

State sequence prediction in imprecise hidden Markov models



Conditioning the model on the observations

Generalised Bayes rule:
An extension of the Bayes rule to imprecise probabilities



State sequence prediction in imprecise hidden Markov models

Maximal state sequences

State sequence prediction in imprecise hidden Markov models

We predict the state sequence by calculating a set of optimal sequences

Notion of optimality: **maximality**

Strict partial ordening:

$$\hat{x}_{1:n} \succ x_{1:n} \Leftrightarrow \underline{P}_1(\mathbb{I}_{\{\hat{x}_{1:n}\}} - \mathbb{I}_{\{x_{1:n}\}} | o_{1:n}) > 0.$$

Maximal state sequences:

$$\begin{aligned}
&\hat{x}_{1:n} \in \operatorname{opt}(\mathscr{X}_{1:n} | o_{1:n}) \\
\Leftrightarrow\ & (\forall x_{1:n} \in \mathscr{X}_{1:n})\ x_{1:n} \not\succ \hat{x}_{1:n} \\
\Leftrightarrow\ & (\forall x_{1:n} \in \mathscr{X}_{1:n})\ \underline{P}_1(\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}} | o_{1:n}) \leq 0
\end{aligned}$$

Maximal state sequences

State sequence prediction in imprecise hidden Markov models

We predict the state sequence by calculating a set of optimal sequences

Notion of optimality: **maximality**

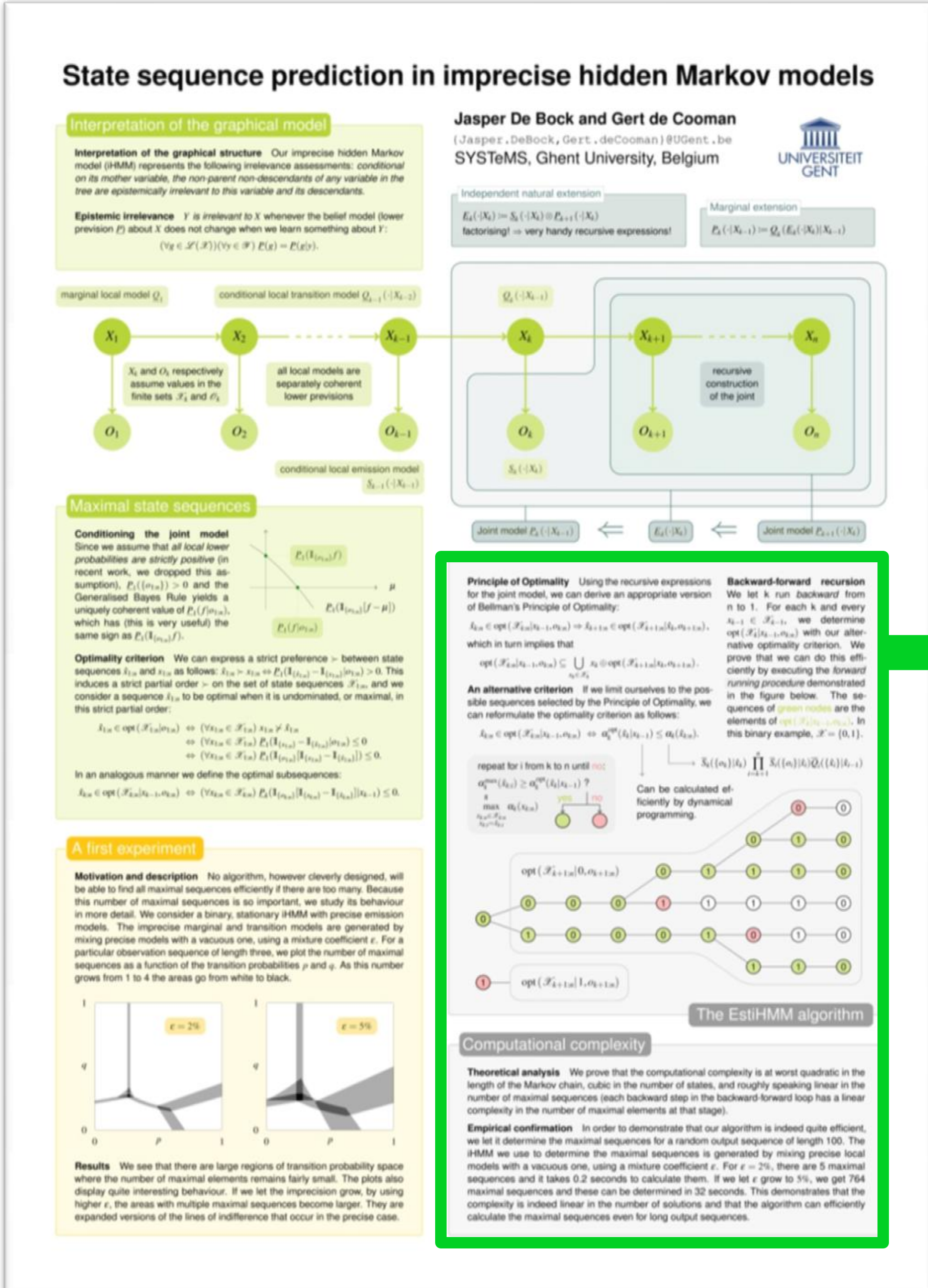
Strict partial ordening:

$$\hat{x}_{1:n} \succ x_{1:n} \Leftrightarrow \underline{P}_1(\mathbb{I}_{\{\hat{x}_{1:n}\}} - \mathbb{I}_{\{x_{1:n}\}} | o_{1:n}) > 0.$$

Maximal state sequences:

$$\begin{aligned}
&\hat{x}_{1:n} \in \operatorname{opt}(\mathscr{X}_{1:n} | o_{1:n}) \\
\Leftrightarrow\ & (\forall x_{1:n} \in \mathscr{X}_{1:n})\ x_{1:n} \not\succ \hat{x}_{1:n} \\
\Leftrightarrow\ & (\forall x_{1:n} \in \mathscr{X}_{1:n})\ \underline{P}_1(\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}} | o_{1:n}) \leq 0 \\
\Leftrightarrow\ & (\forall x_{1:n} \in \mathscr{X}_{1:n})\ \underline{P}_1(\mathbb{I}_{\{o_{1:n}\}}[\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}}]) \leq 0.
\end{aligned}$$

→ **Maximal state sequences**



State sequence prediction in imprecise hidden Markov models

EstiHMM:

an efficient algorithm to determine the maximal state sequences in an imprecise hidden Markov model

State sequence prediction in imprecise hidden Markov models

Principle of Optimality Using the recursive expressions for the joint model, we can derive an appropriate version of Bellman's Principle of Optimality:

$$\hat{x}_{k:n} \in \mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}\right) \Rightarrow \hat{x}_{k+1:n} \in \mathrm{opt}\left(\mathscr{X}_{k+1:n}|\hat{x}_k, o_{k+1:n}\right),$$

which in turn implies that

$$\mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}\right) \subseteq \bigcup_{x_k \in \mathscr{X}_k} x_k \oplus \mathrm{opt}\left(\mathscr{X}_{k+1:n}|x_k, o_{k+1:n}\right).$$

An alternative criterion If we limit ourselves to the possible sequences selected by the Principle of Optimality, we can reformulate the optimality criterion as follows:

$$\hat{x}_{k:n} \in \mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}\right) \Leftrightarrow \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1}) \leq \alpha_k(\hat{x}_{k:n}).$$

Backward-forward recursion We let k run *backward* from n to 1. For each k and every $x_{k-1} \in \mathscr{X}_{k-1}$, we determine $\mathrm{opt}\left(\mathscr{X}_k|x_{k-1}, o_{k:n}\right)$ with our alternative optimality criterion. We prove that we can do this efficiently by executing the *forward running procedure* demonstrated in the figure below. The sequences of green nodes are the elements of $\mathrm{opt}\left(\mathscr{X}_k|x_{k-1}, o_{k:n}\right)$. In this binary example, $\mathscr{X} = \{0, 1\}$.

$$\overline{S}_k(\{o_k\}|\hat{x}_k)\prod_{i=k+1}^{n} \overline{S}_i(\{o_i\}|\hat{x}_i)\overline{Q}_i(\{\hat{x}_i\}|\hat{x}_{i-1})$$

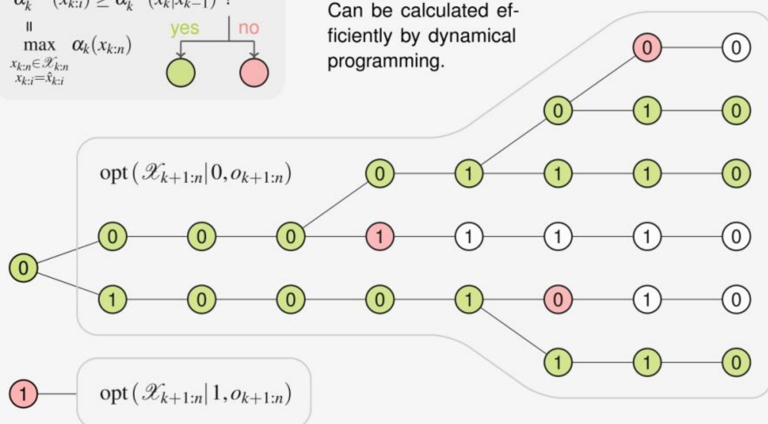
Can be calculated efficiently by dynamical programming.

repeat for i from k to n until no:

$$\alpha_k^{\max}(\hat{x}_{k:i}) \geq \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1}) \ ?$$

$$\alpha_k^{\max}(\hat{x}_{k:i}) = \max_{\substack{x_{k:n}\in\mathscr{X}_{k:n}\\ x_{k:i}=\hat{x}_{k:i}}} \alpha_k(x_{k:n})$$

yes | no



The EstiHMM algorithm

Computational complexity

Theoretical analysis We prove that the computational complexity is at worst quadratic in the length of the Markov chain, cubic in the number of states, and roughly speaking linear in the number of maximal sequences (each backward step in the backward-forward loop has a linear complexity in the number of maximal elements at that stage).

Empirical confirmation In order to demonstrate that our algorithm is indeed quite efficient, we let it determine the maximal sequences for a random output sequence of length 100. The iHMM we use to determine the maximal sequences is generated by mixing precise local models with a vacuous one, using a mixture coefficient ε. For $\varepsilon = 2\%$, there are 5 maximal sequences and it takes 0.2 seconds to calculate them. If we let ε grow to 5%, we get 764 maximal sequences and these can be determined in 32 seconds. This demonstrates that the complexity is indeed linear in the number of solutions and that the algorithm can efficiently calculate the maximal sequences even for long output sequences.

EstiHMM:

an efficient algorithm to determine the maximal state sequences in an imprecise hidden Markov model

EstiHMM: an efficient algorithm to determine the maximal sequences

Principle of Optimality Using the recursive expressions for the joint model, we can derive an appropriate version of Bellman's Principle of Optimality:

$$\hat{x}_{k:n} \in \mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}\right) \Rightarrow \hat{x}_{k+1:n} \in \mathrm{opt}\left(\mathscr{X}_{k+1:n}|\hat{x}_k, o_{k+1:n}\right),$$

which in turn implies that

$$\mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}\right) \subseteq \bigcup_{x_k \in \mathscr{X}_k} x_k \oplus \mathrm{opt}\left(\mathscr{X}_{k+1:n}|x_k, o_{k+1:n}\right).$$

An alternative criterion If we limit ourselves to the possible sequences selected by the Principle of Optimality, we can reformulate the optimality criterion as follows:

$$\hat{x}_{k:n} \in \mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}\right) \Leftrightarrow \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1}) \le \alpha_k(\hat{x}_{k:n}).$$

$$\overline{S}_k(\{o_k\}|\hat{x}_k) \prod_{i=k+1}^{n} \overline{S}_i(\{o_i\}|\hat{x}_i)\overline{Q}_i(\{\hat{x}_i\}|\hat{x}_{i-1})$$

repeat for i from k to n until no:

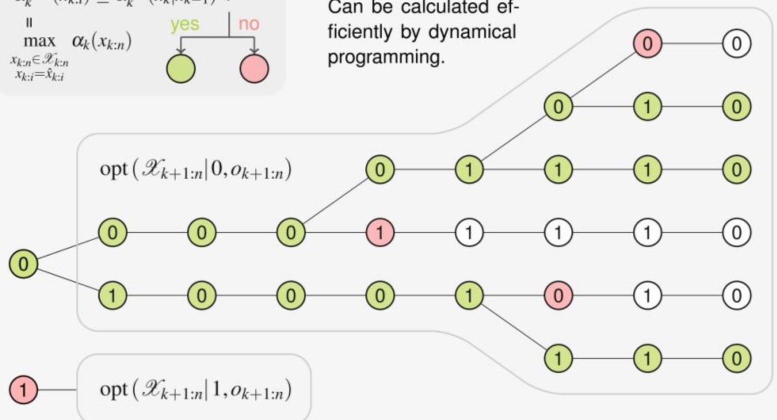
$\alpha_k^{\max}(\hat{x}_{k:i}) \ge \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1})$?

$= \max_{\substack{x_{k:n} \in \mathscr{X}_{k:n} \\ x_{k:i} = \hat{x}_{k:i}}} \alpha_k(x_{k:n})$

yes | no

Can be calculated efficiently by dynamical programming.

Backward-forward recursion We let k run *backward* from n to 1. For each k and every $x_{k-1} \in \mathscr{X}_{k-1}$, we determine $\mathrm{opt}\left(\mathscr{X}_k|x_{k-1}, o_{k:n}\right)$ with our alternative optimality criterion. We prove that we can do this efficiently by executing the *forward running procedure* demonstrated in the figure below. The sequences of green nodes are the elements of $\mathrm{opt}\left(\mathscr{X}_k|x_{k-1}, o_{k:n}\right)$. In this binary example, $\mathscr{X} = \{0, 1\}$.



The EstiHMM algorithm

Computational complexity

Theoretical analysis We prove that the computational complexity is at worst quadratic in the length of the Markov chain, cubic in the number of states, and roughly speaking linear in the number of maximal sequences (each backward step in the backward-forward loop has a linear complexity in the number of maximal elements at that stage).

Empirical confirmation In order to demonstrate that our algorithm is indeed quite efficient, we let it determine the maximal sequences for a random output sequence of length 100. The iHMM we use to determine the maximal sequences is generated by mixing precise local models with a vacuous one, using a mixture coefficient ε. For $\varepsilon = 2\%$, there are 5 maximal sequences and it takes 0.2 seconds to calculate them. If we let ε grow to 5%, we get 764 maximal sequences and these can be determined in 32 seconds. This demonstrates that the complexity is indeed linear in the number of solutions and that the algorithm can efficiently calculate the maximal sequences even for long output sequences.

Principle of Optimality Using the recursive expressions for the joint model, we can derive an appropriate version of Bellman's Principle of Optimality:

$$\hat{x}_{k:n} \in \mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1},o_{k:n}\right) \Rightarrow \hat{x}_{k+1:n} \in \mathrm{opt}\left(\mathscr{X}_{k+1:n}|\hat{x}_k,o_{k+1:n}\right),$$

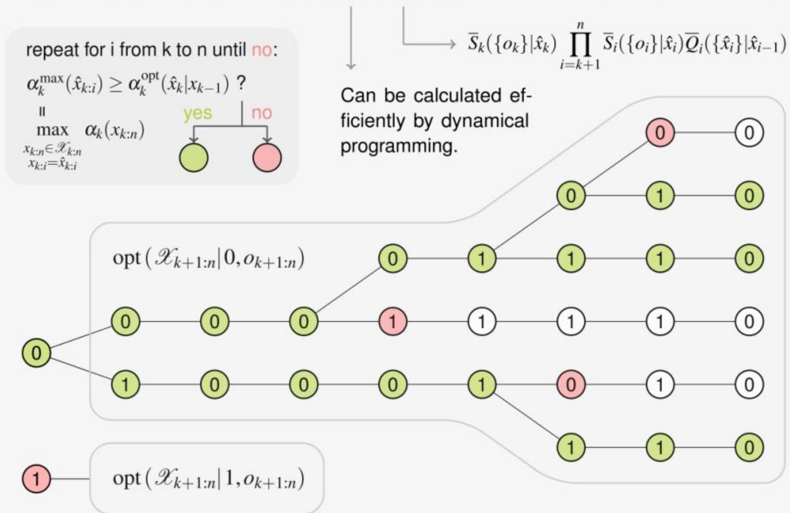
which in turn implies that

$$\mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1},o_{k:n}\right) \subseteq \bigcup_{x_k\in\mathscr{X}_k} x_k \oplus \mathrm{opt}\left(\mathscr{X}_{k+1:n}|x_k,o_{k+1:n}\right).$$

An alternative criterion If we limit ourselves to the possible sequences selected by the Principle of Optimality, we can reformulate the optimality criterion as follows:

$$\hat{x}_{k:n} \in \mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1},o_{k:n}\right) \Leftrightarrow \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1}) \leq \alpha_k(\hat{x}_{k:n}).$$

Backward-forward recursion We let k run *backward* from n to 1. For each k and every $x_{k-1} \in \mathscr{X}_{k-1}$, we determine $\mathrm{opt}\left(\mathscr{X}_k|x_{k-1},o_{k:n}\right)$ with our alternative optimality criterion. We prove that we can do this efficiently by executing the *forward running procedure* demonstrated in the figure below. The sequences of green nodes are the elements of $\mathrm{opt}\left(\mathscr{X}_k|x_{k-1},o_{k:n}\right)$. In this binary example, $\mathscr{X}=\{0,1\}$.



The EstiHMM algorithm

Computational complexity

Theoretical analysis We prove that the computational complexity is at worst quadratic in the length of the Markov chain, cubic in the number of states, and roughly speaking linear in the number of maximal sequences (each backward step in the backward-forward loop has a linear complexity in the number of maximal elements at that stage).

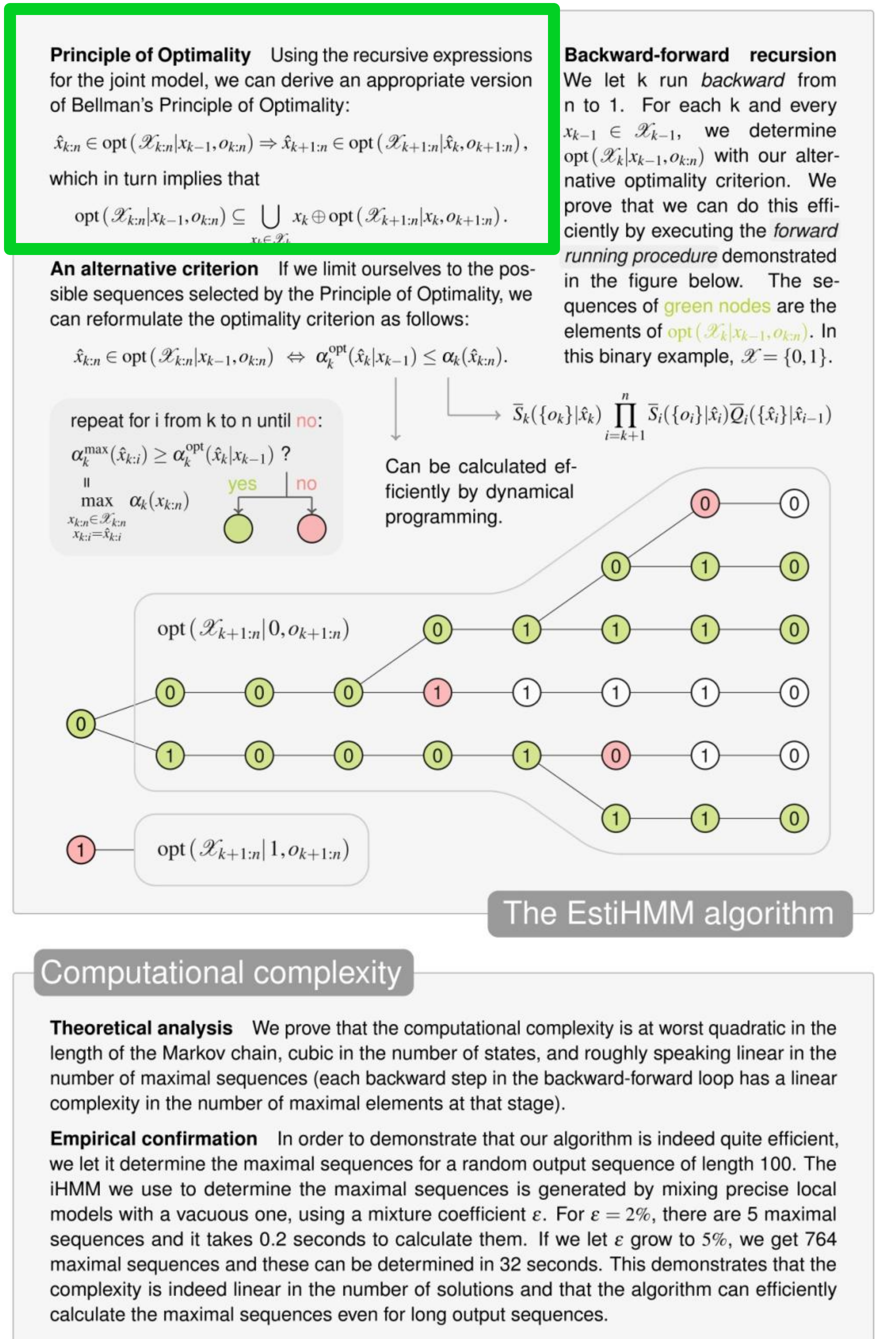
Empirical confirmation In order to demonstrate that our algorithm is indeed quite efficient, we let it determine the maximal sequences for a random output sequence of length 100. The iHMM we use to determine the maximal sequences is generated by mixing precise local models with a vacuous one, using a mixture coefficient ε. For $\varepsilon = 2\%$, there are 5 maximal sequences and it takes 0.2 seconds to calculate them. If we let ε grow to 5%, we get 764 maximal sequences and these can be determined in 32 seconds. This demonstrates that the complexity is indeed linear in the number of solutions and that the algorithm can efficiently calculate the maximal sequences even for long output sequences.

EstiHMM: an efficient algorithm to determine the maximal sequences

- Principle of optimality

EstiHMM: an efficient algorithm to determine the maximal sequences

- **Principle of optimality**



Principle of Optimality Using the recursive expressions for the joint model, we can derive an appropriate version of Bellman's Principle of Optimality:

$$\hat{x}_{k:n} \in \operatorname{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \Rightarrow \hat{x}_{k+1:n} \in \operatorname{opt}(\mathscr{X}_{k+1:n}|\hat{x}_k, o_{k+1:n}),$$

which in turn implies that

$$\operatorname{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \subseteq \bigcup_{x_k \in \mathscr{X}_k} x_k \oplus \operatorname{opt}(\mathscr{X}_{k+1:n}|x_k, o_{k+1:n}).$$

An alternative criterion If we limit ourselves to the possible sequences selected by the Principle of Optimality, we can reformulate the optimality criterion as follows:

$$\hat{x}_{k:n} \in \operatorname{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \Leftrightarrow \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1}) \leq \alpha_k(\hat{x}_{k:n}).$$

$$\overline{S}_k(\{o_k\}|\hat{x}_k) \prod_{i=k+1}^{n} \overline{S}_i(\{o_i\}|\hat{x}_i)\overline{Q}_i(\{\hat{x}_i\}|\hat{x}_{i-1})$$

repeat for i from k to n until no:

$$\alpha_k^{\max}(\hat{x}_{k:i}) \geq \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1}) \text{ ?}$$

$$\| \quad \max_{\substack{x_{k:n} \in \mathscr{X}_{k:n} \\ x_{k:i} = \hat{x}_{k:i}}} \alpha_k(x_{k:n})$$

yes no

Can be calculated efficiently by dynamical programming.

Backward-forward recursion We let k run *backward* from n to 1. For each k and every $x_{k-1} \in \mathscr{X}_{k-1}$, we determine $\operatorname{opt}(\mathscr{X}_k|x_{k-1}, o_{k:n})$ with our alternative optimality criterion. We prove that we can do this efficiently by executing the *forward running procedure* demonstrated in the figure below. The sequences of green nodes are the elements of $\operatorname{opt}(\mathscr{X}_k|x_{k-1}, o_{k:n})$. In this binary example, $\mathscr{X} = \{0, 1\}$.

$\operatorname{opt}(\mathscr{X}_{k+1:n}|0, o_{k+1:n})$

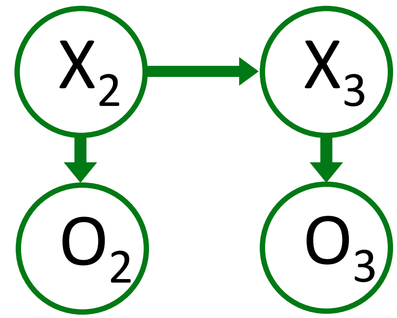
$\operatorname{opt}(\mathscr{X}_{k+1:n}|1, o_{k+1:n})$

The EstiHMM algorithm

Computational complexity

Theoretical analysis We prove that the computational complexity is at worst quadratic in the length of the Markov chain, cubic in the number of states, and roughly speaking linear in the number of maximal sequences (each backward step in the backward-forward loop has a linear complexity in the number of maximal elements at that stage).

Empirical confirmation In order to demonstrate that our algorithm is indeed quite efficient, we let it determine the maximal sequences for a random output sequence of length 100. The iHMM we use to determine the maximal sequences is generated by mixing precise local models with a vacuous one, using a mixture coefficient ε. For $\varepsilon = 2\%$, there are 5 maximal sequences and it takes 0.2 seconds to calculate them. If we let ε grow to 5%, we get 764 maximal sequences and these can be determined in 32 seconds. This demonstrates that the complexity is indeed linear in the number of solutions and that the algorithm can efficiently calculate the maximal sequences even for long output sequences.



EstiHMM: an efficient algorithm to determine the maximal sequences

- Principle of optimality

Principle of Optimality Using the recursive expressions for the joint model, we can derive an appropriate version of Bellman's Principle of Optimality:

$$\hat{x}_{k:n} \in \mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}\right) \Rightarrow \hat{x}_{k+1:n} \in \mathrm{opt}\left(\mathscr{X}_{k+1:n}|\hat{x}_k, o_{k+1:n}\right),$$

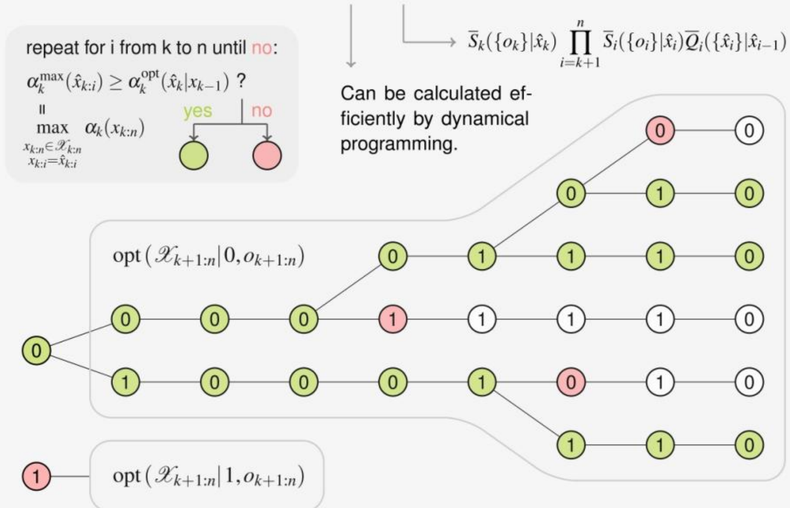
which in turn implies that

$$\mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}\right) \subseteq \bigcup_{x_k \in \mathscr{X}_k} x_k \oplus \mathrm{opt}\left(\mathscr{X}_{k+1:n}|x_k, o_{k+1:n}\right).$$

An alternative criterion If we limit ourselves to the possible sequences selected by the Principle of Optimality, we can reformulate the optimality criterion as follows:

$$\hat{x}_{k:n} \in \mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}\right) \Leftrightarrow \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1}) \leq \alpha_k(\hat{x}_{k:n}).$$

Backward-forward recursion We let k run *backward* from n to 1. For each k and every $x_{k-1} \in \mathscr{X}_{k-1}$, we determine $\mathrm{opt}\left(\mathscr{X}_k|x_{k-1}, o_{k:n}\right)$ with our alternative optimality criterion. We prove that we can do this efficiently by executing the *forward running procedure* demonstrated in the figure below. The sequences of green nodes are the elements of $\mathrm{opt}\left(\mathscr{X}_k|x_{k-1}, o_{k:n}\right)$. In this binary example, $\mathscr{X} = \{0, 1\}$.

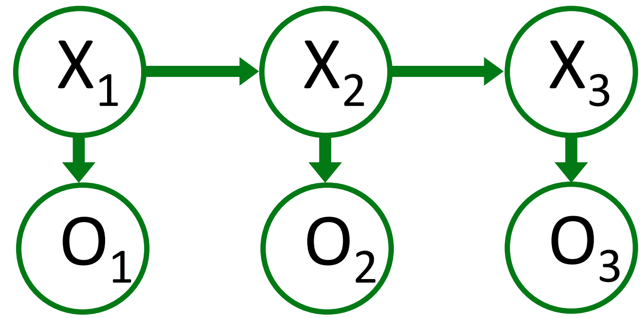


The EstiHMM algorithm

Computational complexity

Theoretical analysis We prove that the computational complexity is at worst quadratic in the length of the Markov chain, cubic in the number of states, and roughly speaking linear in the number of maximal sequences (each backward step in the backward-forward loop has a linear complexity in the number of maximal elements at that stage).

Empirical confirmation In order to demonstrate that our algorithm is indeed quite efficient, we let it determine the maximal sequences for a random output sequence of length 100. The iHMM we use to determine the maximal sequences is generated by mixing precise local models with a vacuous one, using a mixture coefficient ε. For $\varepsilon = 2\%$, there are 5 maximal sequences and it takes 0.2 seconds to calculate them. If we let ε grow to 5%, we get 764 maximal sequences and these can be determined in 32 seconds. This demonstrates that the complexity is indeed linear in the number of solutions and that the algorithm can efficiently calculate the maximal sequences even for long output sequences.



Principle of Optimality Using the recursive expressions for the joint model, we can derive an appropriate version of Bellman's Principle of Optimality:

$$\hat{x}_{k:n} \in \mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1},o_{k:n}\right) \Rightarrow \hat{x}_{k+1:n} \in \mathrm{opt}\left(\mathscr{X}_{k+1:n}|\hat{x}_k,o_{k+1:n}\right),$$

which in turn implies that

$$\mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1},o_{k:n}\right) \subseteq \bigcup x_k \oplus \mathrm{opt}\left(\mathscr{X}_{k+1:n}|x_k,o_{k+1:n}\right).$$

An alternative criterion If we limit ourselves to the possible sequences selected by the Principle of Optimality, we can reformulate the optimality criterion as follows:

$$\hat{x}_{k:n} \in \mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1},o_{k:n}\right) \Leftrightarrow \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1}) \leq \alpha_k(\hat{x}_{k:n}).$$

repeat for i from k to n until no:

$\alpha_k^{\max}(\hat{x}_{k:i}) \geq \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1})$?

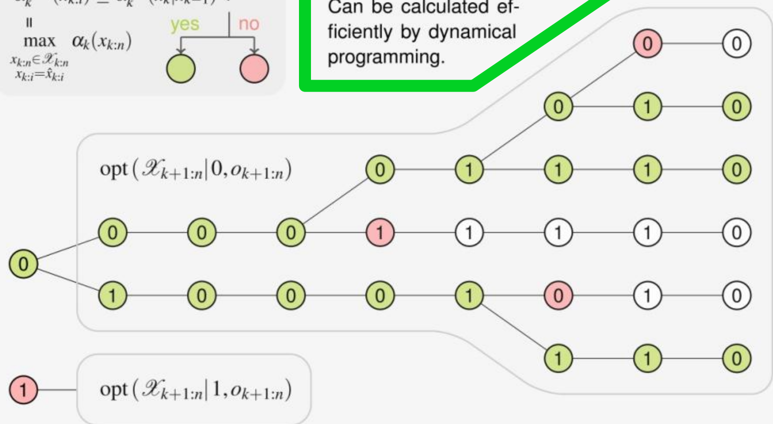
$\alpha_k^{\max}(\hat{x}_{k:i}) = \max_{\substack{x_{k:n}\in\mathscr{X}_{k:n}\\ x_{k:i}=\hat{x}_{k:i}}} \alpha_k(x_{k:n})$

yes / no

$$\overline{S}_k(\{o_k\}|\hat{x}_k)\prod_{i=k+1}^{n}\overline{S}_i(\{o_i\}|\hat{x}_i)\overline{Q}_i(\{\hat{x}_i\}|\hat{x}_{i-1})$$

Can be calculated efficiently by dynamical programming.

Backward-forward recursion We let k run *backward* from n to 1. For each k and every $x_{k-1} \in \mathscr{X}_{k-1}$, we determine $\mathrm{opt}\left(\mathscr{X}_k|x_{k-1},o_{k:n}\right)$ with our alternative optimality criterion. We prove that we can do this efficiently by executing the *forward running procedure* demonstrated in the figure below. The sequences of green nodes are the elements of $\mathrm{opt}\left(\mathscr{X}_k|x_{k-1},o_{k:n}\right)$. In this binary example, $\mathscr{X} = \{0,1\}$.



The EstiHMM algorithm

Computational complexity

Theoretical analysis We prove that the computational complexity is at worst quadratic in the length of the Markov chain, cubic in the number of states, and roughly speaking linear in the number of maximal sequences (each backward step in the backward-forward loop has a linear complexity in the number of maximal elements at that stage).

Empirical confirmation In order to demonstrate that our algorithm is indeed quite efficient, we let it determine the maximal sequences for a random output sequence of length 100. The iHMM we use to determine the maximal sequences is generated by mixing precise local models with a vacuous one, using a mixture coefficient ε. For $\varepsilon = 2\%$, there are 5 maximal sequences and it takes 0.2 seconds to calculate them. If we let ε grow to 5%, we get 764 maximal sequences and these can be determined in 32 seconds. This demonstrates that the complexity is indeed linear in the number of solutions and that the algorithm can efficiently calculate the maximal sequences even for long output sequences.

EstiHMM: an efficient algorithm to determine the maximal sequences

- Principle of optimality
- **Deriving an alternative optimality criterion**

$$\hat{x}_{1:n} \in \mathrm{opt}\left(\mathscr{X}_{1:n}|o_{1:n}\right)$$

$$\Updownarrow$$

$$(\forall x_{1:n} \in \mathscr{X}_{1:n})\ \underline{P}_1\left(\mathbb{I}_{\{o_{1:n}\}}\left[\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}}\right]\right) \leq 0.$$

EstiHMM: an efficient algorithm to determine the maximal sequences

- Principle of optimality
- **Deriving an alternative optimality criterion**

$$\hat{x}_{1:n} \in \operatorname{opt}\left(\mathscr{X}_{1:n} | o_{1:n}\right)$$

$$\Updownarrow$$

$$\cancel{(\forall x_{1:n} \in \mathscr{X}_{1:n})\ \underline{P}_1\left(\mathbb{I}_{\{o_{1:n}\}}\left[\mathbb{I}_{\{x_{1:n}\}} - \mathbb{I}_{\{\hat{x}_{1:n}\}}\right]\right) \leq 0.}$$

$$\alpha_k^{\mathrm{opt}}\left(\hat{x}_k | x_{k-1}\right) \leq \alpha_k\left(\hat{x}_{k:n}\right).$$

Principle of Optimality Using the recursive expressions for the joint model, we can derive an appropriate version of Bellman's Principle of Optimality:

$$\hat{x}_{k:n} \in \operatorname{opt}\left(\mathscr{X}_{k:n} | x_{k-1}, o_{k:n}\right) \Rightarrow \hat{x}_{k+1:n} \in \operatorname{opt}\left(\mathscr{X}_{k+1:n} | \hat{x}_k, o_{k+1:n}\right),$$

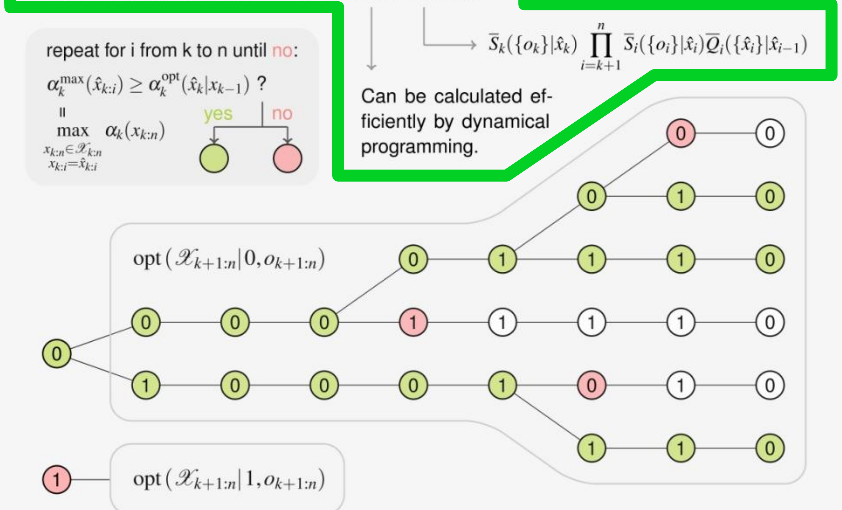
which in turn implies that

$$\operatorname{opt}\left(\mathscr{X}_{k:n} | x_{k-1}, o_{k:n}\right) \subseteq \bigcup x_k \oplus \operatorname{opt}\left(\mathscr{X}_{k+1:n} | x_k, o_{k+1:n}\right).$$

An alternative criterion If we limit ourselves to the possible sequences selected by the Principle of Optimality, we can reformulate the optimality criterion as follows:

$$\hat{x}_{k:n} \in \operatorname{opt}\left(\mathscr{X}_{k:n} | x_{k-1}, o_{k:n}\right) \Leftrightarrow \alpha_k^{\mathrm{opt}}\left(\hat{x}_k | x_{k-1}\right) \leq \alpha_k\left(\hat{x}_{k:n}\right).$$

Backward-forward recursion We let k run *backward* from n to 1. For each k and every $x_{k-1} \in \mathscr{X}_{k-1}$, we determine $\operatorname{opt}\left(\mathscr{X}_k | x_{k-1}, o_{k:n}\right)$ with our alternative optimality criterion. We prove that we can do this efficiently by executing the *forward running procedure* demonstrated in the figure below. The sequences of green nodes are the elements of $\operatorname{opt}\left(\mathscr{X}_k | x_{k-1}, o_{k:n}\right)$. In this binary example, $\mathscr{X} = \{0, 1\}$.



The EstiHMM algorithm

Computational complexity

Theoretical analysis We prove that the computational complexity is at worst quadratic in the length of the Markov chain, cubic in the number of states, and roughly speaking linear in the number of maximal sequences (each backward step in the backward-forward loop has a linear complexity in the number of maximal elements at that stage).

Empirical confirmation In order to demonstrate that our algorithm is indeed quite efficient, we let it determine the maximal sequences for a random output sequence of length 100. The iHMM we use to determine the maximal sequences is generated by mixing precise local models with a vacuous one, using a mixture coefficient ε. For $\varepsilon = 2\%$, there are 5 maximal sequences and it takes 0.2 seconds to calculate them. If we let ε grow to 5%, we get 764 maximal sequences and these can be determined in 32 seconds. This demonstrates that the complexity is indeed linear in the number of solutions and that the algorithm can efficiently calculate the maximal sequences even for long output sequences.

Principle of Optimality Using the recursive expressions for the joint model, we can derive an appropriate version of Bellman's Principle of Optimality:

$$\hat{x}_{k:n} \in \operatorname{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \Rightarrow \hat{x}_{k+1:n} \in \operatorname{opt}(\mathscr{X}_{k+1:n}|\hat{x}_k, o_{k+1:n}),$$

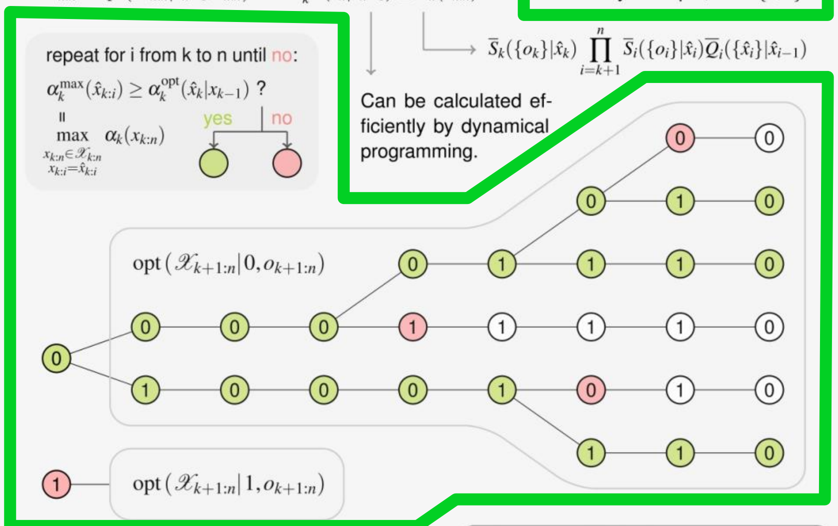
which in turn implies that

$$\operatorname{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \subseteq \bigcup_{x_k \in \mathscr{X}_k} x_k \oplus \operatorname{opt}(\mathscr{X}_{k+1:n}|x_k, o_{k+1:n}).$$

An alternative criterion If we limit ourselves to the possible sequences selected by the Principle of Optimality, we can reformulate the optimality criterion as follows:

$$\hat{x}_{k:n} \in \operatorname{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \Leftrightarrow \alpha_k^{\text{opt}}(\hat{x}_k|x_{k-1}) \leq \alpha_k(\hat{x}_{k:n}).$$

Backward-forward recursion We let k run *backward* from n to 1. For each k and every $x_{k-1} \in \mathscr{X}_{k-1}$, we determine $\operatorname{opt}(\mathscr{X}_k|x_{k-1}, o_{k:n})$ with our alternative optimality criterion. We prove that we can do this efficiently by executing the *forward running procedure* demonstrated in the figure below. The sequences of green nodes are the elements of $\operatorname{opt}(\mathscr{X}_k|x_{k-1}, o_{k:n})$. In this binary example, $\mathscr{X} = \{0, 1\}$.



The EstiHMM algorithm

Computational complexity

Theoretical analysis We prove that the computational complexity is at worst quadratic in the length of the Markov chain, cubic in the number of states, and roughly speaking linear in the number of maximal sequences (each backward step in the backward-forward loop has a linear complexity in the number of maximal elements at that stage).

Empirical confirmation In order to demonstrate that our algorithm is indeed quite efficient, we let it determine the maximal sequences for a random output sequence of length 100. The iHMM we use to determine the maximal sequences is generated by mixing precise local models with a vacuous one, using a mixture coefficient ε. For $\varepsilon = 2\%$, there are 5 maximal sequences and it takes 0.2 seconds to calculate them. If we let ε grow to 5%, we get 764 maximal sequences and these can be determined in 32 seconds. This demonstrates that the complexity is indeed linear in the number of solutions and that the algorithm can efficiently calculate the maximal sequences even for long output sequences.

EstiHMM: an efficient algorithm to determine the maximal sequences

- Principle of optimality
- Deriving an alternative optimality criterion
- **A recursive approach**

EstiHMM: an efficient algorithm to determine the maximal sequences

- Principle of optimality
- Deriving an alternative optimality criterion
- A recursive approach

Complexity

Principle of Optimality Using the recursive expressions for the joint model, we can derive an appropriate version of Bellman's Principle of Optimality:

$$\hat{x}_{k:n} \in \mathrm{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \Rightarrow \hat{x}_{k+1:n} \in \mathrm{opt}(\mathscr{X}_{k+1:n}|\hat{x}_k, o_{k+1:n}),$$

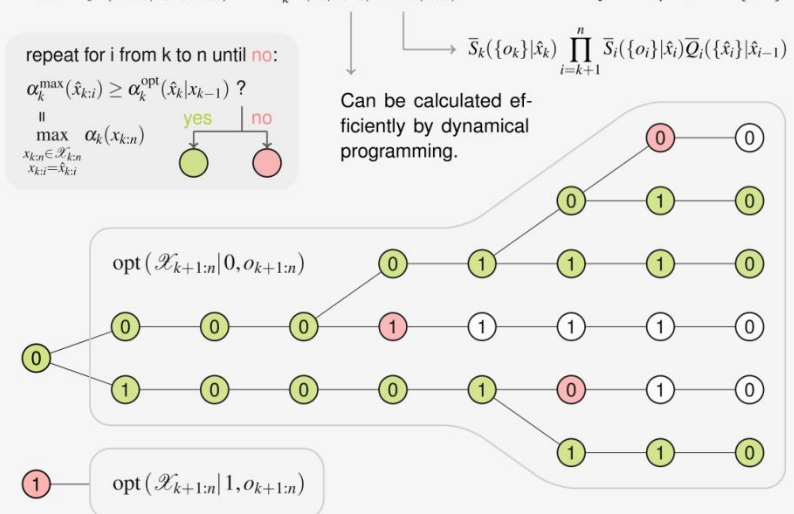
which in turn implies that

$$\mathrm{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \subseteq \bigcup_{x_k \in \mathscr{X}_k} x_k \oplus \mathrm{opt}(\mathscr{X}_{k+1:n}|x_k, o_{k+1:n}).$$

An alternative criterion If we limit ourselves to the possible sequences selected by the Principle of Optimality, we can reformulate the optimality criterion as follows:

$$\hat{x}_{k:n} \in \mathrm{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \Leftrightarrow \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1}) \le \alpha_k(\hat{x}_{k:n}).$$

Backward-forward recursion We let k run *backward* from n to 1. For each k and every $x_{k-1} \in \mathscr{X}_{k-1}$, we determine $\mathrm{opt}(\mathscr{X}_k|x_{k-1}, o_{k:n})$ with our alternative optimality criterion. We prove that we can do this efficiently by executing the *forward running procedure* demonstrated in the figure below. The sequences of green nodes are the elements of $\mathrm{opt}(\mathscr{X}_k|x_{k-1}, o_{k:n})$. In this binary example, $\mathscr{X} = \{0, 1\}$.



The EstiHMM algorithm

Computational complexity

Theoretical analysis We prove that the computational complexity is at worst quadratic in the length of the Markov chain, cubic in the number of states, and roughly speaking linear in the number of maximal sequences (each backward step in the backward-forward loop has a linear complexity in the number of maximal elements at that stage).

Empirical confirmation In order to demonstrate that our algorithm is indeed quite efficient, we let it determine the maximal sequences for a random output sequence of length 100. The iHMM we use to determine the maximal sequences is generated by mixing precise local models with a vacuous one, using a mixture coefficient ε. For $\varepsilon = 2\%$, there are 5 maximal sequences and it takes 0.2 seconds to calculate them. If we let ε grow to 5%, we get 764 maximal sequences and these can be determined in 32 seconds. This demonstrates that the complexity is indeed linear in the number of solutions and that the algorithm can efficiently calculate the maximal sequences even for long output sequences.

EstiHMM: an efficient algorithm to determine the maximal sequences

- Principle of optimality
- Deriving an alternative optimality criterion
- A recursive approach

Complexity

- **Theoretical analysis**

Principle of Optimality Using the recursive expressions for the joint model, we can derive an appropriate version of Bellman's Principle of Optimality:

$$\hat{x}_{k:n} \in \mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}\right) \Rightarrow \hat{x}_{k+1:n} \in \mathrm{opt}\left(\mathscr{X}_{k+1:n}|\hat{x}_k, o_{k+1:n}\right),$$

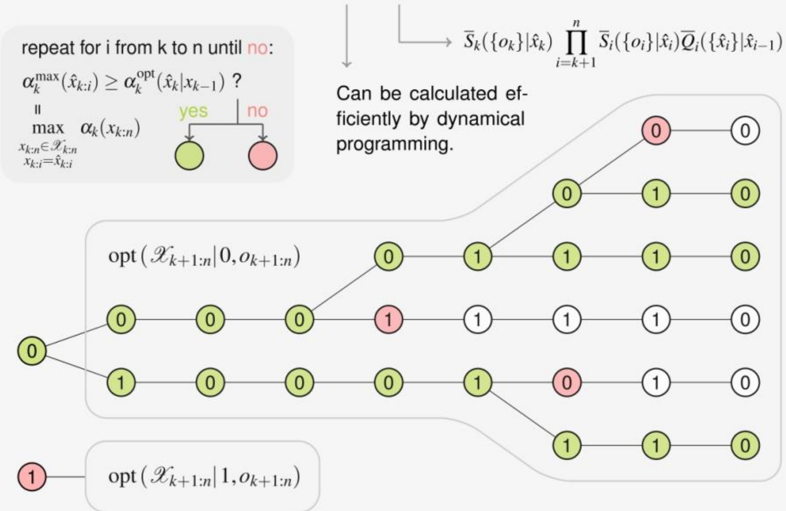
which in turn implies that

$$\mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}\right) \subseteq \bigcup_{x_k \in \mathscr{X}_k} x_k \oplus \mathrm{opt}\left(\mathscr{X}_{k+1:n}|x_k, o_{k+1:n}\right).$$

An alternative criterion If we limit ourselves to the possible sequences selected by the Principle of Optimality, we can reformulate the optimality criterion as follows:

$$\hat{x}_{k:n} \in \mathrm{opt}\left(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}\right) \Leftrightarrow \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1}) \leq \alpha_k(\hat{x}_{k:n}).$$

Backward-forward recursion We let k run *backward* from n to 1. For each k and every $x_{k-1} \in \mathscr{X}_{k-1}$, we determine $\mathrm{opt}\left(\mathscr{X}_k|x_{k-1}, o_{k:n}\right)$ with our alternative optimality criterion. We prove that we can do this efficiently by executing the *forward running procedure* demonstrated in the figure below. The sequences of green nodes are the elements of $\mathrm{opt}\left(\mathscr{X}_k|x_{k-1}, o_{k:n}\right)$. In this binary example, $\mathscr{X} = \{0, 1\}$.



The EstiHMM algorithm

Computational complexity

Theoretical analysis We prove that the computational complexity is at worst quadratic in the length of the Markov chain, cubic in the number of states, and roughly speaking linear in the number of maximal sequences (each backward step in the backward-forward loop has a linear complexity in the number of maximal elements at that stage).

Empirical confirmation In order to demonstrate that our algorithm is indeed quite efficient, we let it determine the maximal sequences for a random output sequence of length 100. The iHMM we use to determine the maximal sequences is generated by mixing precise local models with a vacuous one, using a mixture coefficient ε. For $\varepsilon = 2\%$, there are 5 maximal sequences and it takes 0.2 seconds to calculate them. If we let ε grow to 5%, we get 764 maximal sequences and these can be determined in 32 seconds. This demonstrates that the complexity is indeed linear in the number of solutions and that the algorithm can efficiently calculate the maximal sequences even for long output sequences.

EstiHMM: an efficient algorithm to determine the maximal sequences

Principle of Optimality Using the recursive expressions for the joint model, we can derive an appropriate version of Bellman's Principle of Optimality:

$$\hat{x}_{k:n} \in \operatorname{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \Rightarrow \hat{x}_{k+1:n} \in \operatorname{opt}(\mathscr{X}_{k+1:n}|\hat{x}_k, o_{k+1:n}),$$

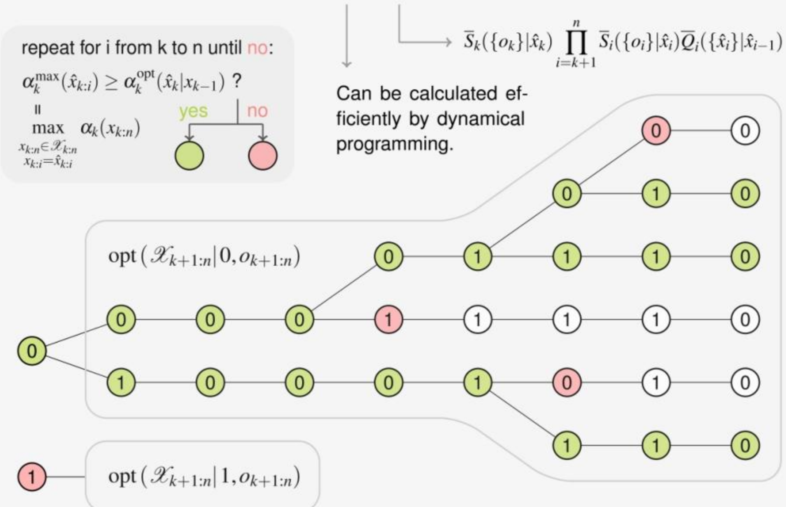
which in turn implies that

$$\operatorname{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \subseteq \bigcup_{x_k \in \mathscr{X}_k} x_k \oplus \operatorname{opt}(\mathscr{X}_{k+1:n}|x_k, o_{k+1:n}).$$

An alternative criterion If we limit ourselves to the possible sequences selected by the Principle of Optimality, we can reformulate the optimality criterion as follows:

$$\hat{x}_{k:n} \in \operatorname{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \Leftrightarrow \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1}) \leq \alpha_k(\hat{x}_{k:n}).$$

Backward-forward recursion We let k run *backward* from n to 1. For each k and every $x_{k-1} \in \mathscr{X}_{k-1}$, we determine $\operatorname{opt}(\mathscr{X}_k|x_{k-1}, o_{k:n})$ with our alternative optimality criterion. We prove that we can do this efficiently by executing the *forward running procedure* demonstrated in the figure below. The sequences of green nodes are the elements of $\operatorname{opt}(\mathscr{X}_k|x_{k-1}, o_{k:n})$. In this binary example, $\mathscr{X} = \{0, 1\}$.



The EstiHMM algorithm

Computational complexity

Theoretical analysis We prove that the computational complexity is at worst quadratic in the length of the Markov chain, cubic in the number of states, and roughly speaking linear in the number of maximal sequences (each backward step in the backward-forward loop has a linear complexity in the number of maximal elements at that stage).

Empirical confirmation In order to demonstrate that our algorithm is indeed quite efficient, we let it determine the maximal sequences for a random output sequence of length 100. The iHMM we use to determine the maximal sequences is generated by mixing precise local models with a vacuous one, using a mixture coefficient ε. For $\varepsilon = 2\%$, there are 5 maximal sequences and it takes 0.2 seconds to calculate them. If we let ε grow to 5%, we get 764 maximal sequences and these can be determined in 32 seconds. This demonstrates that the complexity is indeed linear in the number of solutions and that the algorithm can efficiently calculate the maximal sequences even for long output sequences.

- Principle of optimality
- Deriving an alternative optimality criterion
- A recursive approach

Complexity

- Theoretical analysis
- **Linear in the number of maximal sequences**

Principle of Optimality Using the recursive expressions for the joint model, we can derive an appropriate version of Bellman's Principle of Optimality:

$$\hat{x}_{k:n} \in \operatorname{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \Rightarrow \hat{x}_{k+1:n} \in \operatorname{opt}(\mathscr{X}_{k+1:n}|\hat{x}_k, o_{k+1:n}),$$

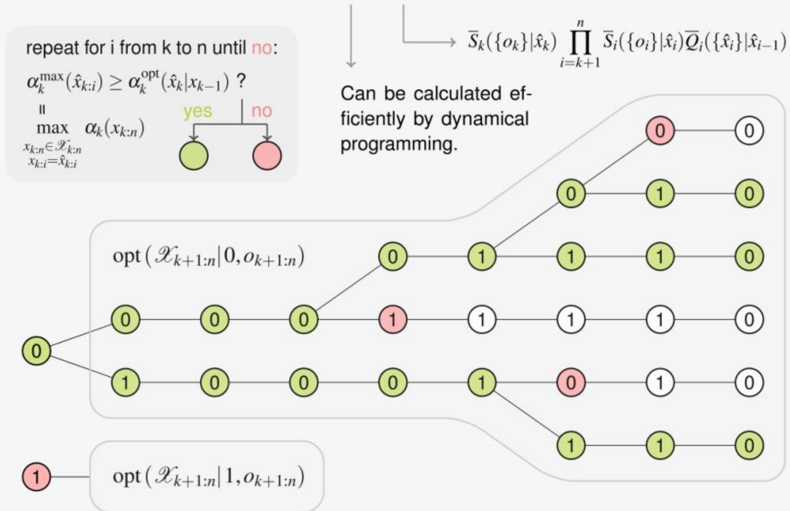
which in turn implies that

$$\operatorname{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \subseteq \bigcup_{x_k \in \mathscr{X}_k} x_k \oplus \operatorname{opt}(\mathscr{X}_{k+1:n}|x_k, o_{k+1:n}).$$

An alternative criterion If we limit ourselves to the possible sequences selected by the Principle of Optimality, we can reformulate the optimality criterion as follows:

$$\hat{x}_{k:n} \in \operatorname{opt}(\mathscr{X}_{k:n}|x_{k-1}, o_{k:n}) \Leftrightarrow \alpha_k^{\mathrm{opt}}(\hat{x}_k|x_{k-1}) \leq \alpha_k(\hat{x}_{k:n}).$$

Backward-forward recursion We let k run *backward* from n to 1. For each k and every $x_{k-1} \in \mathscr{X}_{k-1}$, we determine $\operatorname{opt}(\mathscr{X}_k|x_{k-1}, o_{k:n})$ with our alternative optimality criterion. We prove that we can do this efficiently by executing the *forward running procedure* demonstrated in the figure below. The sequences of green nodes are the elements of $\operatorname{opt}(\mathscr{X}_k|x_{k-1}, o_{k:n})$. In this binary example, $\mathscr{X} = \{0, 1\}$.



The EstiHMM algorithm

Computational complexity

Theoretical analysis We prove that the computational complexity is at worst quadratic in the length of the Markov chain, cubic in the number of states, and roughly speaking linear in the number of maximal sequences (each backward step in the backward-forward loop has a linear complexity in the number of maximal elements at that stage).

Empirical confirmation In order to demonstrate that our algorithm is indeed quite efficient, we let it determine the maximal sequences for a random output sequence of length 100. The iHMM we use to determine the maximal sequences is generated by mixing precise local models with a vacuous one, using a mixture coefficient ε. For $\varepsilon = 2\%$, there are 5 maximal sequences and it takes 0.2 seconds to calculate them. If we let ε grow to 5%, we get 764 maximal sequences and these can be determined in 32 seconds. This demonstrates that the complexity is indeed linear in the number of solutions and that the algorithm can efficiently calculate the maximal sequences even for long output sequences.

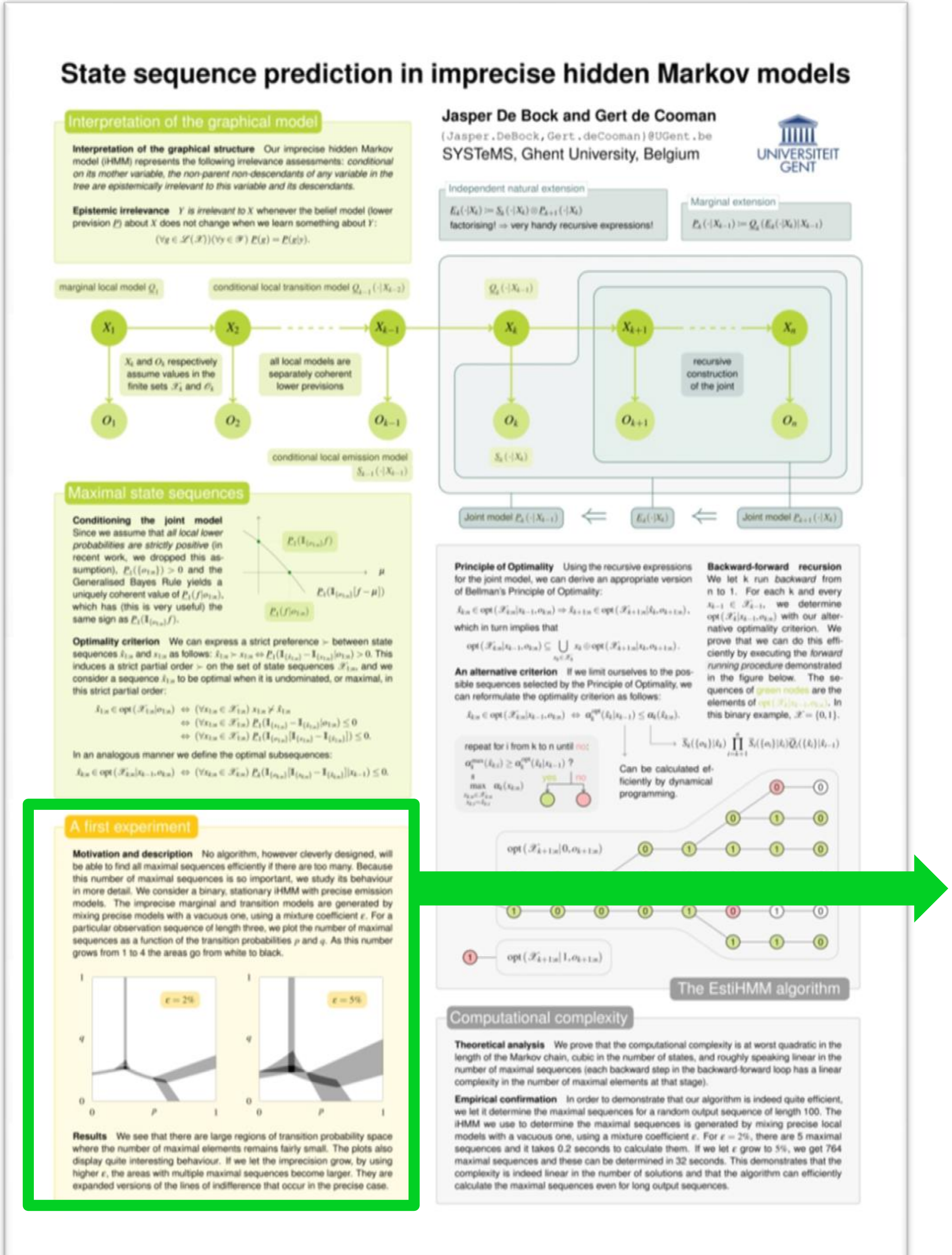
EstiHMM: an efficient algorithm to determine the maximal sequences

- Principle of optimality
- Deriving an alternative optimality criterion
- A recursive approach

Complexity

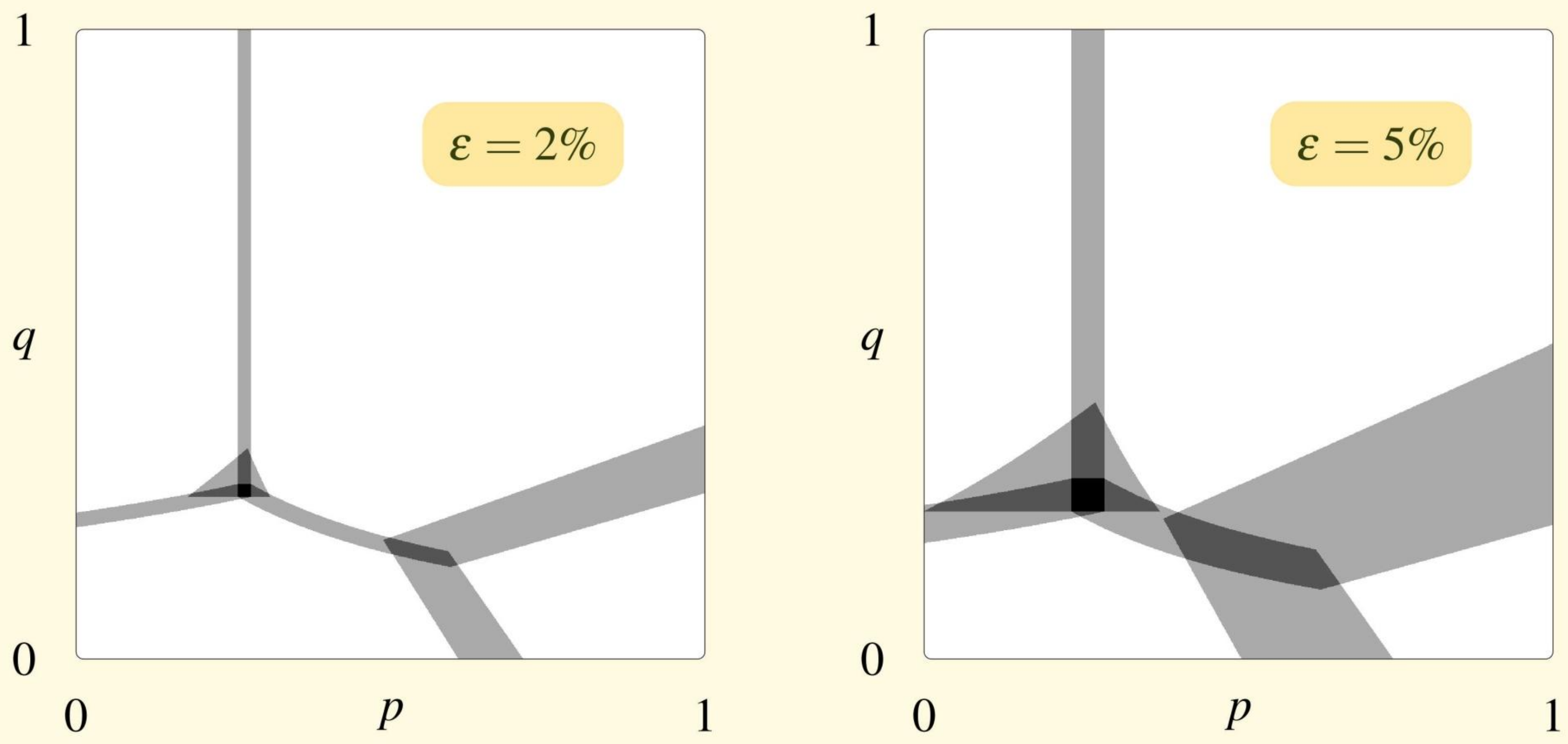
- Theoretical analysis
- Linear in the number of maximal sequences
- Empirical confirmation

State sequence prediction in imprecise hidden Markov models

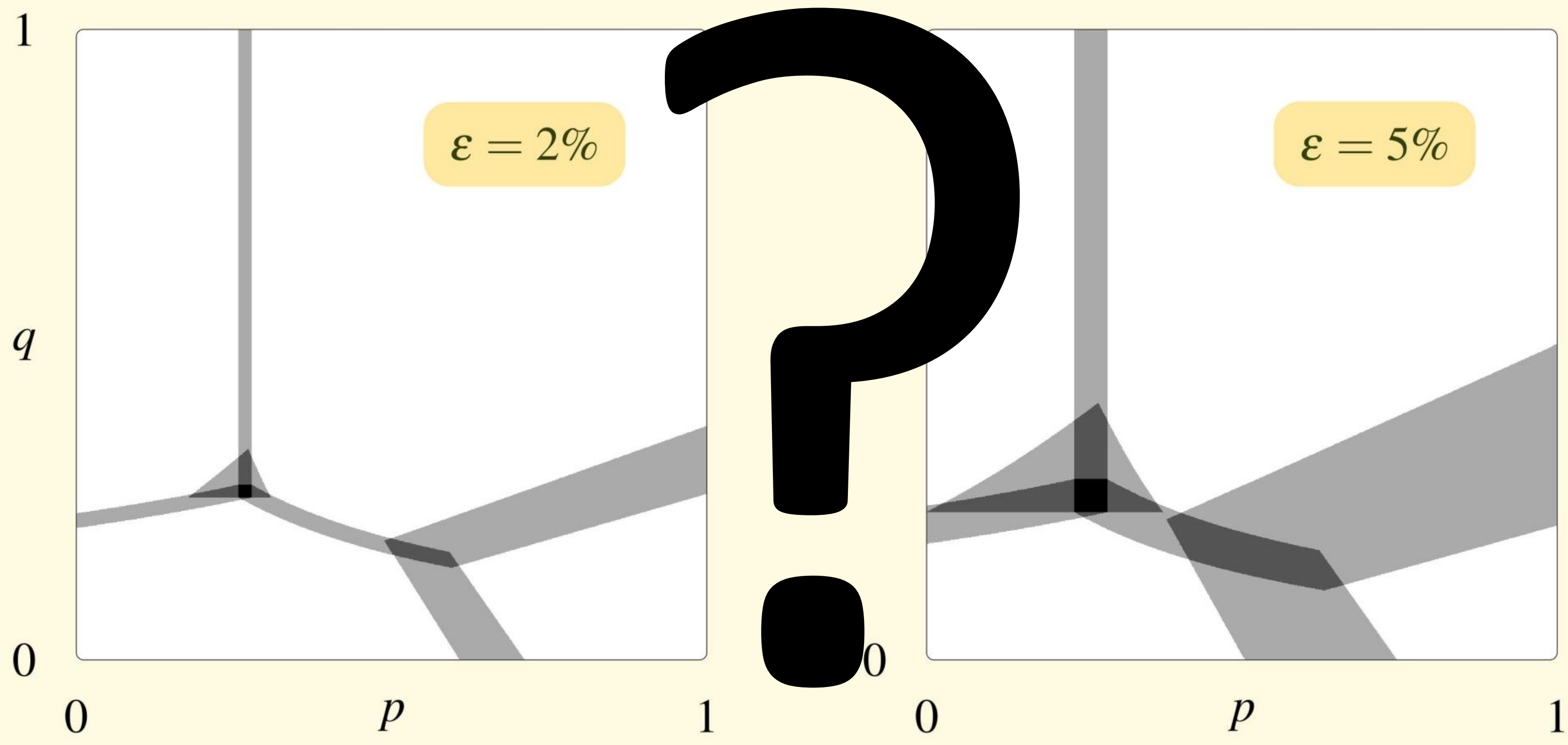


A first experiment

A first experiment



A first experiment



A first experiment