Mixed-Integer Programming for Cycle Detection in Non-reversible Markov Processes

Isabel Beckenbach  Leon Eifler  Ambros Gleixner  Jakob Witzig
Mathematical Optimization, Zuse Institute Berlin

Konstantin Fackeldey  Andreas Grever  Marcus Weber
Computational Molecular Design, Zuse Institute Berlin

Aussois 2018
Motivation

- Many chemical and biological processes can be interpreted as a
  Markov State Models
Motivation

- Many chemical and biological processes can be interpreted as a Markov State Models.

- After simulation (e.g., Monte-Carlo) a discrete set of states \( S = \{1, \ldots, n\} \) and a transition matrix \( P \) describe the underlying Markov process.
Motivation

- Many chemical and biological processes can be interpreted as a **Markov State Models**

- After simulation (e.g. Monte-Carlo) a discrete set of states $S = \{1, \ldots, n\}$ and a transition matrix $P$ describe the underlying Markov process

- A clustering of the states can be used to analyze the process
Problem Description

- Let $P \in \mathbb{R}^{n \times n}$ a transition matrix of a Markov process with stationary distribution $\pi \in \mathbb{R}^n$.

- A process is called **reversible**, if for all states $i$ and $j$ holds:

  $$\pi_i p_{ij} = \pi_j p_{ji}$$
Problem Description

- Let $P \in \mathbb{R}^{n \times n}$ a transition matrix of a Markov process with stationary distribution $\pi \in \mathbb{R}^n$.
- A process is called reversible, if for all states $i$ and $j$ holds:

$$\pi_i p_{ij} = \pi_j p_{ji}$$

Goal

Find clustering $C_1, \ldots, C_m$, s.t. the process between two consecutive clusters is non-reversible and within each cluster coherent.
Problem Description

▶ Let $P \in \mathbb{R}^{n \times n}$ a transition matrix of a Markov process with stationary distribution $\pi \in \mathbb{R}^n$.

▶ A process is called reversible, if for all states $i$ and $j$ holds:

$$\pi_i p_{ij} = \pi_j p_{ji}$$

Goal  Find clustering $C_1, \ldots, C_m$, s.t. the process between two consecutive clusters is non-reversible and within each cluster coherent.
Problem Description

- Let $P \in \mathbb{R}^{n \times n}$ a transition matrix of a Markov process with stationary distribution $\pi \in \mathbb{R}^n$.

- A process is called **reversible**, if for all states $i$ and $j$ holds:

  $$\pi_i p_{ij} = \pi_j p_{ji}$$

**Goal**  Find clustering $C_1, \ldots, C_m$, s.t. the process between two consecutive clusters is **non-reversible** and within each cluster **coherent**.
Why is this interesting in practice?

- If a process is reversible the leading eigenvalues can be used to find a meaningful clustering.


Why is this interesting in practice?

- If a process is **reversible** the leading eigenvalues can be used to find a meaningful clustering.

- If a process is **non-reversible** some of the leading eigenvalues are complex and close to the unit circle\(^1\) \(\Rightarrow\) only meaningful interpretation for dominant cycles, but not in general.

---


Why is this interesting in practice?

- If a process is **reversible** the leading eigenvalues can be used to find a meaningful clustering.

- If a process is **non-reversible** some of the leading eigenvalues are complex and close to the unit circle\(^1\) ⇒ only meaningful interpretation for dominant cycles, but not in general.

State-of-the-art methods for non-reversible processes are, e.g.,

- Schur-decomposition\(^2\)

- “Fuzzy” clustering (i.e. assign fractions of states to clusters) and apply a rounding heuristic

but none of them can prove optimality or has some guarantee for the solution quality.

---


Our Goal

1. Find a clustering \( C_1, \ldots, C_m, m \in \mathbb{N} \) given.

2. Maximize the net-flow between consecutive clusters and the coherence within each cluster.
Our Goal

1. Find a clustering \( C_1, \ldots, C_m, m \in \mathbb{N} \) given.

2. Maximize the net-flow between consecutive clusters and the coherence within each cluster.

3. Should be applicable for all non-reversible processes with stationary distribution.
Our Goal

1. Find a clustering $C_1, \ldots, C_m, m \in \mathbb{N}$ given.

2. Maximize the net-flow between consecutive clusters and the coherence within each cluster.

3. Should be applicable for all non-reversible processes with stationary distribution.

4. We want to be able to either prove optimality or to measure the solution quality.
Cycle-Clustering MINLP

\[
\begin{align*}
\text{max } & \text{ net-flow } + \alpha \cdot \sum_{t \in C} \text{coherence of } C_t \\
\text{s.t. } & \sum_{t \in C} x_{it} = 1 \quad \forall i \in S \\
& \sum_{i \in S} x_{it} \geq 1 \quad \forall t \in C \\
\end{align*}
\]

\(x_{it} \in \{0, 1\} \quad \forall i \in S, \forall t \in C\)

- Set of clusters \(C\)
- Discrete set of states \(S\)
Cycle-Clustering MINLP

\[
\begin{align*}
\text{max} & \quad \text{net-flow} + \alpha \cdot \sum_{t \in C} \text{coherence of } C_t \\
\text{s.t.} & \quad \sum_{t \in C} x_{it} = 1 \quad \forall i \in S \\
& \quad \sum_{i \in S} x_{it} \geq 1 \quad \forall t \in C
\end{align*}
\]

- Set of clusters \( C \)
- Discrete set of states \( S \)
- Binary variable \( x_{it} = 1 \iff \text{state } i \text{ is assigned to cluster } t \)
Cycle-Clustering MINLP

\[
\begin{align*}
\text{max} & \quad \sum_{t \in C} \epsilon_t + \alpha \cdot \sum_{t \in C} \text{coherence of } C_t \\
\text{s.t.} & \quad \sum_{t \in C} x_{it} = 1 \quad \forall i \in S \\
& \quad \sum_{i \in S} x_{it} \geq 1 \quad \forall t \in C \\
& \quad \epsilon_t = \sum_{i \neq j \in S} \pi_{ij} (x_{it} x_{j\phi(t)} - x_{i\phi(t)} x_{jt}) \quad \forall t \in C \\
x_{it} \in \{0, 1\} \quad \forall i \in S, \forall t \in C \\
\epsilon_t \in \mathbb{R}_{\geq 0} \quad \forall t \in C
\end{align*}
\]

- Set of clusters \( C \)
- Discrete set of states \( S \)
- Binary variable \( x_{it} = 1 \iff \text{state } i \text{ is assigned to cluster } t \)
- Cont. variable \( \epsilon_t \) representing the net-flow between \( C_t \) and \( C_{\phi(t)} \)
Cycle-Clustering MINLP

\[ \text{max} \sum_{t \in C} \epsilon_t + \alpha \cdot \sum_{t \in C} \sum_{i,j \in S} \pi_{ij} x_{it} x_{jt} \]

s.t.
\[ \sum_{t \in C} x_{it} = 1 \quad \forall i \in S \]  
\[ \sum_{i \in S} x_{it} \geq 1 \quad \forall t \in C \]  
\[ \epsilon_t = \sum_{i \neq j \in S} \pi_{ij} (x_{it} x_{j\phi(t)} - x_{i\phi(t)} x_{jt}) \quad \forall t \in C \]  
\[ x_{it} \in \{0, 1\} \quad \forall i \in S, \forall t \in C \]  
\[ \epsilon_t \in \mathbb{R}_{\geq 0} \quad \forall t \in C \]

- Set of clusters \( C \)
- Discrete set of states \( S \)
- Binary variable \( x_{it} = 1 \) \( \iff \) state \( i \) is assigned to cluster \( t \)
- Cont. variable \( \epsilon_t \) representing the net-flow between \( C_t \) and \( C_{\phi(t)} \)
Linearizing the MINLP-Formulation

- automatic linearization by SCIP / textbook linearization

\[ E := \{ (i, j) \in S \times S : i \neq j, \pi_i p_{ij} + \pi_j p_{ji} > 0 \} \]

- Binary variables \( y_{ij} = 1 \) ⇔ states \( i \) and \( j \) are in the same cluster, \( \forall i, j \in E \):

- Binary variables \( z_{ij} = 1 \) ⇔ states \( i \in C_t \) and \( j \in C_{\varphi(t)} \), \( \forall i, j \in E \):
Linearizing the MINLP-Formulation

- automatic linearization by SCIP / textbook linearization
- compact linearization for binary quadratic problems (L. Liberti, 2007)
Linearizing the MINLP-Formulation

- automatic linearization by SCIP / textbook linearization

- compact linearization for binary quadratic problems (L. Liberti, 2007)

- problem specific linearization which uses the cyclic structure
Linearizing the MINLP-Formulation

- automatic linearization by SCIP / textbook linearization
  -

- compact linearization for binary quadratic problems (L. Liberti, 2007)
  -

- problem specific linearization which uses the cyclic structure
  -

- Set of relevant transitions: \( E := \{(i,j) \in S \times S : i \neq j, \ \pi_i p_{ij} + \pi_j p_{ji} > 0\} \)

- Binary variables \( y_{ij} = 1 \Leftrightarrow \) states \( i \) and \( j \) are in the same cluster, \( \forall i, j \in E : i < j \)

- Binary variables \( z_{ij} = 1 \Leftrightarrow \) states \( i \in C_t \) and \( j \in C_{\phi(t)} \), \( \forall i, j \in E \)
MIP Model

\[\begin{align*}
\text{max} \quad & \sum_{(i,j) \in E} z_{ij} (\pi_i p_{ij} - \pi_j p_{ji}) + \alpha \sum_{(i,j) \in E: i < j} y_{ij} (\pi_i p_{ij} + \pi_j p_{ji}) \\
\text{s.t.} \quad & \sum_{t \in C} x_{it} = 1 \quad \forall i \in S \\
& \sum_{i \in S} x_{it} \geq 1 \quad \forall t \in C \\
& x_{it} + x_{jt} - y_{ij} + z_{ij} - x_{j\phi(t)} - x_{i\phi^{-1}(t)} \leq 1 \quad \forall t \in C, (i,j) \in E \\
& x_{it} + x_{j\phi(t)} - z_{ij} + y_{ij} - x_{jt} - x_{i\phi(t)} \leq 1 \quad \forall t \in C, (i,j) \in E \\
& y_{ij} + z_{ij} + z_{ji} \leq 1 \quad \forall (i,j) \in E: i < j \\
& x_{it} \in \{0, 1\} \quad \forall i \in S, \forall t \in C \\
& y_{ij}, z_{ij}, z_{ji} \in \{0, 1\} \quad \forall (i,j) \in E: i < j
\end{align*}\]
MIP Model

\[
\begin{align*}
\text{max} & \quad \sum_{(i,j) \in E} z_{ij} (\pi_i p_{ij} - \pi_j p_{ji}) + \alpha \sum_{(i,j) \in E : i < j} y_{ij} (\pi_i p_{ij} + \pi_j p_{ji}) \\
\text{s.t.} & \quad \sum_{t \in C} x_{it} = 1 \quad \forall i \in S \\
& \quad \sum_{i \in S} x_{it} \geq 1 \quad \forall t \in C \\
& \quad x_{it} + x_{jt} - y_{ij} + z_{ij} - x_{j\phi(t)} - x_{i\phi^{-1}(t)} \leq 1 \quad \forall t \in C, (i, j) \in E \\
& \quad x_{it} + x_{j\phi(t)} - z_{ij} + y_{ij} - x_{jt} - x_{i\phi(t)} \leq 1 \quad \forall t \in C, (i, j) \in E \\
& \quad y_{ij} + z_{ij} + z_{ji} \leq 1 \quad \forall (i, j) \in E : i < j \\
& \quad x_{it} \in \{0, 1\} \quad \forall i \in S, \forall t \in C \\
& \quad y_{ij}, z_{ij}, z_{ji} \in \{0, 1\} \quad \forall (i, j) \in E : i < j
\end{align*}
\]
MIP Model

\[
\begin{align*}
\text{max} & \quad \sum_{(i,j) \in E} z_{ij}(\pi_i p_{ij} - \pi_j p_{ji}) + \alpha \sum_{(i,j) \in E: i < j} y_{ij}(\pi_i p_{ij} + \pi_j p_{ji}) \\
\text{s.t.} & \quad \sum_{t \in C} x_{it} = 1 \quad \forall i \in S \\
& \quad \sum_{i \in S} x_{it} \geq 1 \quad \forall t \in C \\
& \quad x_{it} + x_{jt} - y_{ij} + z_{ij} - x_{i\phi(t)} - x_{i\phi^{-1}(t)} \leq 1 \quad \forall t \in C, (i,j) \in E \\
& \quad x_{it} + x_{j\phi(t)} - z_{ij} + y_{ij} - x_{jt} - x_{i\phi(t)} \leq 1 \quad \forall t \in C, (i,j) \in E \\
& \quad y_{ij} + z_{ij} + z_{ji} \leq 1 \quad \forall (i,j) \in E: i < j \\
& \quad x_{it} \in \{0, 1\} \quad \forall i \in S, \forall t \in C \\
& \quad y_{ij}, z_{ij}, z_{ji} \in \{0, 1\} \quad \forall (i,j) \in E: i < j
\end{align*}
\]
MIP Model

\[
\begin{align*}
\text{max} & \sum_{(i,j) \in E} z_{ij}(\pi_ip_{ij} - \pi_jp_{ji}) + \alpha \sum_{(i,j) \in E : i < j} y_{ij}(\pi_ip_{ij} + \pi_jp_{ji}) \\
s.t. & \quad \sum_{t \in C} x_{it} = 1 \quad \forall i \in S \\
& \quad \sum_{i \in S} x_{it} \geq 1 \quad \forall t \in C \\
& \quad x_{it} + x_{jt} - y_{ij} + z_{ij} - x_{j\phi(t)} - x_{i\phi^{-1}(t)} \leq 1 \quad \forall t \in C, (i, j) \in E \\
& \quad x_{it} + x_{j\phi(t)} - z_{ij} + y_{ij} - x_{jt} - x_{i\phi(t)} \leq 1 \quad \forall t \in C, (i, j) \in E \\
& \quad y_{ij} + z_{ij} + z_{ji} \leq 1 \quad \forall (i, j) \in E : i < j \\
& \quad x_{it} \in \{0, 1\} \quad \forall i \in S, \forall t \in C \\
& \quad y_{ij}, z_{ij}, z_{ji} \in \{0, 1\} \quad \forall (i, j) \in E : i < j
\end{align*}
\]
Facets of the Linearization

Theorem
Let \((i, j) \in E, t \in C, \text{ and } m \geq 4\). Then the inequalities

\[
\begin{align*}
\xi_it + \xi_jt - \gamma_{ij} + \zeta_{ij} - \xi_{j\phi(t)} - \xi_{i\phi^{-1}(t)} & \leq 1 \quad \text{and} \\
\xi_it + \xi_{j\phi(t)} - \zeta_{ij} + \gamma_{ij} - \xi_jt - \xi_{i\phi(t)} & \leq 1
\end{align*}
\]

are facet-defining for the cycle-clustering polytope (CCP).
Facets of the Linearization

Theorem

Let $(i, j) \in E$, $t \in C$, and $m \geq 4$. Then the inequalities

\[ x_{it} + x_{jt} - y_{ij} + z_{ij} = x_{j\phi(t)} - x_{i\phi(t)} \leq 1 \quad \text{and} \quad x_{it} + x_{j\phi(t)} - y_{ij} + y_{ji} - x_{jt} - x_{i\phi(t)} \leq 1 \]

are facet-defining for the cycle-clustering polytope (CCP).

very technical proof

Follows the proof technique of Chopra and Rao (1993).
Facets of the Linearization

Theorem

Let \((i, j) \in E, t \in C, \) and \(m \geq 4\). Then the inequalities

\[
\begin{align*}
    x_{it} + x_{jt} - y_{ij} + z_{ij} - x_{i \phi(t)} - x_{j \phi(t)} & \leq 1 \\
    x_{it} + x_{j \phi(t)} - y_{ij} + y_{jj} - x_{jt} - x_{i \phi(t)} & \leq 1
\end{align*}
\]

are facet-defining for the cycle-clustering polytope (CCP).

Proof idea

1. Let \(\hat{a}x + \hat{b}y + \hat{c}z \leq \hat{\delta}\) facet-defining with

\[
\{(x, y, z) \in CCP \mid ax + by + cz = \delta\} \subseteq \{(x, y, z) \in CCP \mid \hat{a}x + \hat{b}y + \hat{c}z = \hat{\delta}\}
\]

2. Compare coefficients and show that \(ax + by + cz \leq \delta\) is a multiple of \(\hat{a}x + \hat{b}y + \hat{c}z \leq \hat{\delta}\)

Follows the proof technique of Chopra and Rao (1993).
1. various types of triangle inequalities
   ▶ special case: $m = 3$
   ▶ general case: $m \geq 4$ (motivated by S. Chopra and M. R. Rao. (1993); M. Grötschel and Y. Wakabayashi (1990))

2. subtour and path inequalities
   ▶ motivated by TSP subtour elimination inequalities

3. partitioning inequalities
   ▶ modification of M. Grötschel and Y. Wakabayashi (1990)
1. **various types of triangle inequalities**
   - special case: $m = 3$
   - general case: $m \geq 4$ (motivated by S. Chopra and M. R. Rao. (1993); M. Grötschel and Y. Wakabayashi (1990))

2. **subtour and path inequalities**
   - motivated by TSP subtour elimination inequalities

3. **partitioning inequalities**
   - modification of M. Grötschel and Y. Wakabayashi (1990)
Simple Triangle Inequalities

special case $m = 3$:

$$z_{ij} + z_{jk} - z_{ki} \leq 1 \quad \forall (i, j), (j, k), (k, i) \in E$$

holds in general:

$$y_{ij} + y_{jk} - y_{ki} \leq 1 \quad \forall (i, j), (j, k), (k, i) \in E$$
Triangle Inequalities Involving 2 Clusters

\[ k \text{ is in the successor cluster of } i \]

\[ k \text{ is in the predecessor cluster of } i \]
Triangle Inequalities Involving 2 Clusters

$k$ is in the successor cluster of $i$

$y_{ij} + z_{ik} - z_{jk} \leq 1$

$k$ is in the predecessor cluster of $i$

$y_{ij} + z_{ki} - z_{kj} \leq 1$
Facet-Defining Triangle Inequalities

A combination of

- $y_{ij} + y_{jk} - y_{ik} \leq 1$
- $y_{ij} + z_{ik} - z_{jk} \leq 1$
- $y_{ij} + z_{ki} - z_{kj} \leq 1$

yield a facet-defining inequality.
Facet-Defining Triangle Inequalities

A combination of

- $y_{ij} + y_{jk} - y_{ik} \leq 1$
- $y_{ij} + z_{ik} - z_{jk} \leq 1$
- $y_{ij} + z_{ki} - z_{kj} \leq 1$

yield a facet-defining inequality.

Theorem

Let $i, j, k \in S$ with $(i, j), (j, k), (i, k) \in E$. Then the triangle inequality

$$y_{ij} + y_{jk} - y_{ik} + 0.5( z_{ij} + z_{ji} + z_{jk} + z_{kj} - z_{ik} - z_{ki} ) \leq 1$$

is facet-defining for the cycle clustering polytope.
Triangle Inequalities Involving 2 Clusters (cont.)

\[ z_{ij} + z_{ik} - y_{jk} \leq 1 \]
\[ z_{ji} + z_{ki} - y_{jk} \leq 1 \]
Facet-Defining Triangle Inequalities (cont.)

Theorem

Let \( i, j, k \in S \) with \( (i, j), (j, k), (i, k) \in E \).

- If \( m > 4 \), then
  \[
  z_{ij} + z_{ik} - y_{jk} \leq 1 \quad \text{and} \quad z_{ji} + z_{ki} - y_{jk} \leq 1
  \]
  are facet-defining for the cycle clustering polytope.

- If \( m = 4 \), then
  \[
  z_{ij} + z_{ik} - 2y_{jk} - (z_{jk} + z_{kj} + z_{ji} + z_{ki}) \leq 0
  \]
  is facet-defining for the cycle clustering polytope.
Theorem
Let $i, j, k \in S$ with $(i, j), (j, k), (i, k) \in E$.

- If $m > 4$, then
  \[
  z_{ij} + z_{ik} - y_{jk} \leq 1 \quad \text{and} \quad z_{ji} + z_{ki} - y_{jk} \leq 1
  \]
  are facet-defining for the cycle clustering polytope.

- If $m = 4$, then
  \[
  z_{ij} + z_{ik} - 2y_{jk} - (z_{jk} + z_{kj} + z_{ji} + z_{ki}) \leq 0
  \]
  is facet-defining for the cycle clustering polytope.
Facet-Defining Triangle Inequalities (cont.)

**Theorem**

Let $i, j, k \in S$ with $(i, j), (j, k), (i, k) \in E$.

- If $m > 4$, then
  
  $$z_{ij} + z_{ik} - y_{jk} \leq 1 \quad \text{and} \quad z_{ji} + z_{ki} - y_{jk} \leq 1$$

  are facet-defining for the cycle clustering polytope.

- If $m = 4$, then
  
  $$z_{ij} + z_{ik} - 2y_{jk} - (z_{jk} + z_{kj} + z_{ji} + z_{ki}) \leq 0$$

  is facet-defining for the cycle clustering polytope.
Facet-Defining Triangle Inequalities (cont.)

Theorem
Let $i, j, k \in S$ with $(i, j), (j, k), (i, k) \in E$.

- If $m > 4$, then
  
  $$z_{ij} + z_{ik} - y_{jk} \leq 1 \quad \text{and} \quad z_{ji} + z_{ki} - y_{jk} \leq 1$$

  are facet-defining for the cycle clustering polytope.

- If $m = 4$, then
  
  $$z_{ij} + z_{ik} - 2y_{jk} - (z_{jk} + z_{kj} + z_{ji} + z_{ki}) \leq 0$$

  is facet-defining for the cycle clustering polytope.
Theorem

Let $i, j, k \in S$ with $(i, j), (j, k), (i, k) \in E$.

- If $m > 4$, then
  $$z_{ij} + z_{ik} - y_{jk} \leq 1 \quad \text{and} \quad z_{ji} + z_{ki} - y_{jk} \leq 1$$
  are facet-defining for the cycle clustering polytope.

- If $m = 4$, then
  $$z_{ij} + z_{ik} - 2y_{jk} - (z_{jk} + z_{kj} + z_{ji} + z_{ki}) \leq 0$$
  is facet-defining for the cycle clustering polytope.
Theorem

Let \( i, j, k \in S \) with \((i, j), (j, k), (i, k) \in E\).

- If \( m > 4 \), then
  \[
  z_{ij} + z_{ik} - y_{jk} \leq 1 \quad \text{and} \quad z_{ji} + z_{ki} - y_{jk} \leq 1
  \]
  are facet-defining for the cycle clustering polytope.

- If \( m = 4 \), then
  \[
  z_{ij} + z_{ik} - 2y_{jk} - (z_{jk} + z_{kj} + z_{ji} + z_{ki}) \leq 0
  \]
  is facet-defining for the cycle clustering polytope.
Theorem
Let \( i, j, k \in S \) with \((i, j), (j, k), (i, k) \in E\).

- If \( m > 4 \), then
  \[
  z_{ij} + z_{ik} - y_{jk} \leq 1 \quad \text{and} \quad z_{ji} + z_{ki} - y_{jk} \leq 1
  \]
  are facet-defining for the cycle clustering polytope.

- If \( m = 4 \), then
  \[
  z_{ij} + z_{ik} - 2y_{jk} - (z_{jk} + z_{kj} + z_{ji} + z_{ki}) \leq 0
  \]
  is facet-defining for the cycle clustering polytope.
Facet-Defining Triangle Inequalities (cont.)

Theorem
Let $i, j, k \in S$ with $(i, j), (j, k), (i, k) \in E$.

- If $m > 4$, then
  \[
  z_{ij} + z_{ik} - y_{jk} \leq 1 \quad \text{and} \quad z_{ji} + z_{ki} - y_{jk} \leq 1
  \]
  are facet-defining for the cycle clustering polytope.

- If $m = 4$, then
  \[
  z_{ij} + z_{ik} - 2y_{jk} - (z_{jk} + z_{kj} + z_{ji} + z_{ki}) \leq 0
  \]
  is facet-defining for the cycle clustering polytope.
Valid Inequalities

1. various types of triangle inequalities
   ▶ special case: \( m = 3 \)
   ▶ general case: \( m \geq 4 \) (motivated by S. Chopra and M. R. Rao. (1993); M. Grötschel and Y. Wakabayashi (1990))

2. subtour and path inequalities
   ▶ motivated by TSP subtour elimination inequalities

3. partitioning inequalities
   ▶ modification of M. Grötschel and Y. Wakabayashi (1990)
Subtour Elimination Inequalities

- Let $K = \{(i_1, i_2), (i_2, i_3), \ldots, (i_{s-1}, i_s), (i_s, i_1)\} \subset E$ with $1 < |K| < m$
- $U \subset K$, $|U| \leq |K| - 1$
Let $K = \{(i_1, i_2), (i_2, i_3), \ldots, (i_{s-1}, i_s), (i_s, i_1)\} \subset E$ with $1 < |K| < m$

- $U \subset K$, $|U| \leq |K| - 1$

$$\sum_{(i, j) \in K} z_{ij} \leq |K| - 1$$
Subtour Elimination Inequalities

- Let $K = \{(i_1, i_2), (i_2, i_3), \ldots, (i_{s-1}, i_s), (i_s, i_1)\} \subset E$ with $1 < |K| < m$
- $U \subset K$, $|U| \leq |K| - 1$

\[ \sum_{(i, j) \in K} z_{ij} \leq |K| - 1 \]
\[ \sum_{(i, j) \in K} z_{ij} + \sum_{(i, j) \in U} y_{ij} \leq |K| - 1 \]
Subtour Elimination Inequalities

- Let $K = \{(i_1, i_2), (i_2, i_3), \ldots, (i_{s-1}, i_s), (i_s, i_1)\} \subset E$ with $1 < |K| < m$
- $U \subset K$, $|U| \leq |K| - 1$

\[
\sum_{(i,j) \in K} z_{ij} \leq |K| - 1
\]

\[
\sum_{(i,j) \in K} z_{ij} + \sum_{(i,j) \in U} y_{ij} \leq |K| - 1
\]

- none of them is facet-defining
- but can be separated in polynomial time
Valid Inequalities

1. Various types of triangle inequalities
   - Special case: $m = 3$
   - General case: $m \geq 4$ (motivated by S. Chopra and M. R. Rao. (1993); M. Grötschel and Y. Wakabayashi (1990))

2. Subtour and path inequalities
   - Motivated by TSP subtour elimination inequalities

3. Partitioning inequalities
   - Modification of M. Grötschel and Y. Wakabayashi (1990)
Partition Inequalities

Let $A, B \subset S$, $A \cap B = \emptyset$. The partitioning inequality
\[
\sum_{i \in A, j \in B} z_{ij} - \sum_{i < j \in A} y_{ij} - \sum_{i < j \in B} y_{ij} \leq \min\{|A|, |B|\}
\]
(4)
is valid for the cycle clustering polytope.

Theorem
The partitioning inequalities of type (4) are facet-defining if $m > 4$ and $||A| - |B|| = 1$. 
Partition Inequalities

Let $A, B \subset S$, $A \cap B = \emptyset$. The partitioning inequality

$$\sum_{i \in A, j \in B} z_{ij} - \sum_{i < j \in A} y_{ij} - \sum_{i < j \in B} y_{ij} \leq \min\{|A|, |B|\}$$

is valid for the cycle clustering polytope.
Partition Inequalities

Let $A, B \subset S$, $A \cap B = \emptyset$. The partitioning inequality

$$\sum_{i \in A, j \in B} z_{ij} - \sum_{i < j \in A} y_{ij} - \sum_{i < j \in B} y_{ij} \leq \min\{|A|, |B|\}$$

(4)

is valid for the cycle clustering polytope.

**Theorem**

The partitioning inequalities of type (4) are facet-defining if $m > 4$ and $||A| - |B|| = 1$. 
How do the inequalities perform in practice?

Our test instances:

1. Artificial potentials with 3, 4 or 6 minima

2. Repressilator (synthetic genetic regulatory network)

3. Hindmarsh-Rose (simulate membrane potential of cells in a human heart)
Only one slide left before you see some numbers...
## Computational Results: Impact of Inequalities

<table>
<thead>
<tr>
<th>Inequ.</th>
<th>time [s]</th>
<th>nodes</th>
<th>dual integral</th>
</tr>
</thead>
<tbody>
<tr>
<td>default</td>
<td>130.6</td>
<td>256.6</td>
<td>1213.3</td>
</tr>
<tr>
<td>triangle</td>
<td>77.6</td>
<td>28.8</td>
<td>474.1</td>
</tr>
<tr>
<td>subtour</td>
<td>122.6</td>
<td>174.3</td>
<td>1036.4</td>
</tr>
<tr>
<td>partition</td>
<td>122.6</td>
<td>121.6</td>
<td>933.1</td>
</tr>
<tr>
<td>all</td>
<td>68.3</td>
<td>15.4</td>
<td>396.6</td>
</tr>
</tbody>
</table>

Test set of 176 instances with 50 – 250 states. Time limit 1h.
Mixed Integer Programming is more than cutting planes...

Beside of three different type of cutting planes we designed primal heuristics to improve the primal side as well:

- greedy heuristic assigning states to clusters

- exchange heuristic that swaps states between clusters (even if it gets worse in between)

- problem-tailored LP rounding heuristic
Mixed Integer Programming is more than cutting planes...

Beside of three different type of cutting planes we designed primal heuristics to improve the primal side as well:

- greedy heuristic assigning states to clusters \(-32\%\) primal integral
- exchange heuristic that swaps states between clusters (even if it gets worse in between) \(-53\%\) primal integral
- problem-tailored LP rounding heuristic \(-32\%\) primal integral
Now, put it all together

<table>
<thead>
<tr>
<th>Solver</th>
<th>solved</th>
<th>time [s]</th>
<th>primal int</th>
<th>dual int</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCIP</td>
<td>113</td>
<td>182.4</td>
<td>3274.8</td>
<td>1794.0</td>
</tr>
<tr>
<td>CC-SCIP</td>
<td>130</td>
<td>64.7</td>
<td>82.9</td>
<td>327.1</td>
</tr>
</tbody>
</table>

*(out of competition)*

| Gurobi    | 118    | 153.1    | 1236.6     | 1524.9   |

Test set of 176 instances with 50 – 250 states. Time limit 1h.
Conclusion

Summary

▶ First MIP model for detecting cyclic clustering of non-reversible Markov State Models

▶ Various (facet-defining) valid inequalities for the cycle-clustering polytope

▶ Very brief description of problem-tailored primal heuristics

References


▶ Methods: Master thesis of Leon Eifler, publication coming soon...
Thank you for your attention.

Questions?