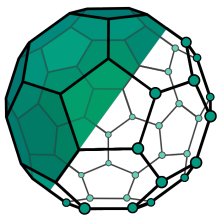
# A Polyhedral Journey Through Extremal Chemical Graphs

Hadrien Mélot

Algorithms Lab, Computer Science Dept
Faculty of Science, University of Mons, Belgium

Computers in Scientific Discovery 11
Kranjska Gora, Slovenia, May 4-8, 2026

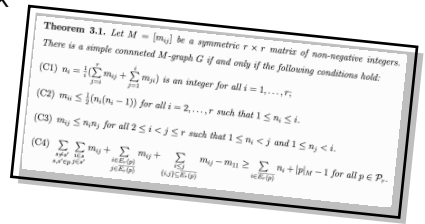


UMONS
University of Mons

# In Memoriam: Pierre Hansen (1940 - 2025)

This talk is dedicated to the memory of **Pierre Hansen**.

- One of the founding fathers of the CSD conference series (with Patrick Fowler)
- My former PhD advisor and a great mentor
- A brilliant mind in OR and Mathematical Chemistry
- Co-author of a result (Hansen et al., 2017) playing a central role in this talk



**Theorem 3.1.** *Let $M = [m_{ij}]$ be a symmetric $r \times r$ matrix of non-negative integers. There is a simple connneted $M$-graph $G$ if and only if the following conditions hold:*

(C1) $n_i = \frac{1}{2}(\sum_{j=1}^{r} m_{ij} + \sum_{j=1}^{i} m_{jj})$ *is an integer for all* $i = 1, \ldots, r$;

(C2) $m_{ii} \leq \frac{1}{2}(n_i(n_i - 1))$ *for all* $i = 2, \ldots, r$ *such that* $1 \leq n_i \leq i$.

(C3) $m_{ij} \leq n_i n_j$ *for all* $2 \leq i < j \leq r$ *such that* $1 \leq n_i < j$ *and* $1 \leq n_j < i$.

(C4) $\sum_{\substack{s,s' \in p \\ s,s' \in p}} \sum_{\substack{t,t' \in p \\ \{t,t'\} \neq \{s,s'\}}} m_{tt'} + \sum_{\substack{i \in E_r(p) \\ j \in E_r(p)}} m_{ij} + \sum_{\substack{i<j \\ \{i,j\} \subseteq E_r(p)}} m_{ij} - m_{11} \geq \sum_{i \in E_r(p)} n_i + |p|_M - 1$ *for all* $p \in \mathcal{P}_r$.

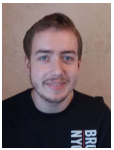
# In Memoriam: Pierre Hansen (1940 - 2025)

- **Key Roles:** Tenured Professor at HEC Montréal (since 1990) and Director of GERAD (1996–2001)
- **Mentorship:** Supervised 46 Master's, PhD, and postdocs
- **Recognition:**
  - EURO Gold Medal (1986)
  - Fellow of the Royal Society of Canada (1999)
  - Posthumously appointed Professor Emeritus by HEC Montréal (2025)

# In Memoriam: Pierre Hansen (1940 - 2025)

- **Astounding Volume:** Over 400 peer-reviewed scientific papers
- **Global Impact:** 50,000 citations with an h-index of 92 (Google Scholar, April 2026)
- **Multidisciplinary Vision:** Bridged gaps between CS, economics, chemistry, and AI
  - Co-invented **Variable Neighborhood Search (VNS)**
  - **AutoGraphiX (AGX)**: automated discovery in graph theory
  - **Data Mining:** Proved NP-hardness of Euclidean sum-of-squares clustering
  - **Math. Chemistry:** e.g., problems regarding graph energy

This talk is based on three papers, representing joint work with:



Sébastien Bonte
UMONS

Gauvain Devillez
UMONS

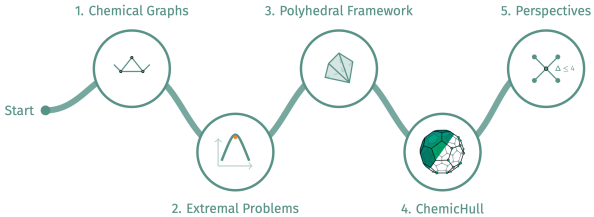
Valentin Dusollier
UMONS

Alain Hertz
Polytechnique Montréal

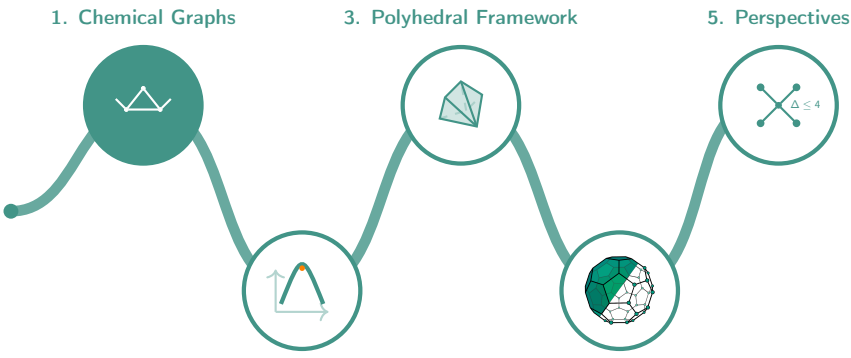
David Schindl
Haute Ecole de Gestion de Genève

# Outline: Embarking on our Polyhedral Journey

1. Chemical Graphs & Topological Indices
2. Extremal Problems & Discoveries for 33 Indices
3. The Polyhedral Framework
4. ChemicHull: From Theory to Software
5. Perspectives



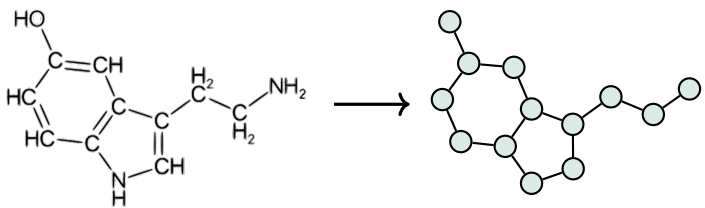
# Chemical Graphs & Topological Indices



# Graphs in Chemistry

Graphs are widely used in mathematical chemistry to abstract the structure of molecules:

- all or some atoms are represented by the **vertices**,
- the chemical bonds between them by the **edges**.



# What is a Chemical Graph?

According to Patrick Fowler[1], the definition of a chemical graph depends on the context:

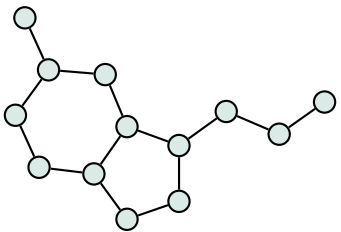
- To model skeletons of saturated hydrocarbons (like alkanes), the **maximum degree** is $\Delta \leq 4$.
- To model conjugated $\pi$ systems (alkenes, polyenes, benzenoids, fullerenes), a carbon atom participates in at most three single bonds, so $\Delta \leq 3$.

[1] Personal communication.

# Chemical Graphs in this Talk

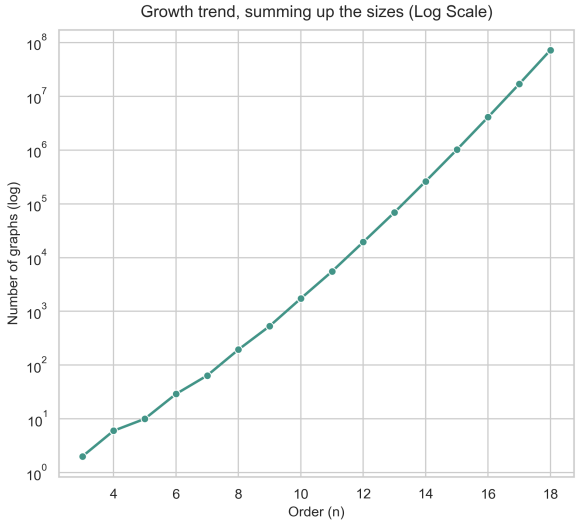
**Working Definition**

In this talk, a **chemical graph** is a simple, undirected, connected graph with maximum degree $\Delta \leq 3$.



Even with these constraints severely restricting the graph space, the number of chemical graphs grows exponentially.

# Exponential Growth of Chemical Graphs



# Topological Indices

- A **topological index** (or molecular descriptor) is a specific graph invariant designed to quantify the topology of the underlying molecule.
- These indices are used in QSAR and QSPR studies.

**QSAR** (Quantitative Structure-**Activity** Relationship):
Predicts biological effects (e.g., drug efficacy, toxicity)

**QSPR** (Quantitative Structure-**Property** Relationship):
Predicts physicochemical properties (e.g., boiling point, density, solubility)

# Degree-Based Topological Indices

**Definition**

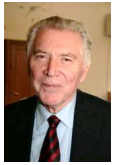
Let $G$ be a graph with maximum degree $\Delta$. A **degree-based topological index** $\mathcal{I}$ is a graph invariant of the form

$$\mathcal{I}(G) = \sum_{1 \le i \le j \le \Delta} c_{ij}\, m_{ij}$$

where:

- $m_{ij}$ is the number of $ij$-edges in $G$, *i.e.* edges connecting vertices of degrees $i$ and $j$.
- $c_{ij}$ is a weight depending only on the degrees $i$ and $j$.

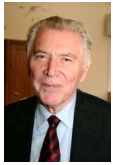
# Example: The Randić Index



One example of a degree-based topological index is the Randić index $R$. This is the first such index, introduced in 1975 by Milan Randić, and remains one of the most widely studied.

Its weights are defined as: $c_{ij} = \frac{1}{\sqrt{i \cdot j}}$ $\Rightarrow (c_{ij})_{1 \le i \le j \le 3} = \begin{pmatrix} 1 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ & \frac{1}{2} & \frac{1}{\sqrt{6}} \\ & & \frac{1}{3} \end{pmatrix}$

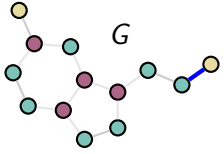
# Example: The Randić Index



One example of a degree-based topological index is the Randić index $R$. This is the first such index, introduced in 1975 by Milan Randić, and remains one of the most widely studied.

Its weights are defined as: $c_{ij} = \frac{1}{\sqrt{i \cdot j}}$ $\Rightarrow (c_{ij})_{1 \le i \le j \le 3} = \begin{pmatrix} 1 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ & \frac{1}{2} & \frac{1}{\sqrt{6}} \\ & & \frac{1}{3} \end{pmatrix}$

deg 1
deg 2
deg 3

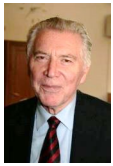



$m_{12} = 1$

$R(G) = 1 \cdot \frac{1}{\sqrt{2}}$

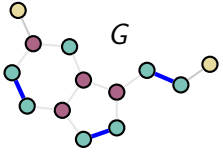
# Example: The Randić Index

One example of a degree-based topological index is the Randić index $R$. This is the first such index, introduced in 1975 by Milan Randić, and remains one of the most widely studied.



Its weights are defined as: $c_{ij} = \frac{1}{\sqrt{i \cdot j}}$ $\Rightarrow (c_{ij})_{1 \le i \le j \le 3} = \begin{pmatrix} 1 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ & \frac{1}{2} & \frac{1}{\sqrt{6}} \\ & & \frac{1}{3} \end{pmatrix}$

deg 1
deg 2
deg 3



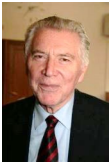

$m_{12} = 1$

$m_{22} = 3$

$R(G) = 1 \cdot \frac{1}{\sqrt{2}} + 3 \cdot \frac{1}{2}$

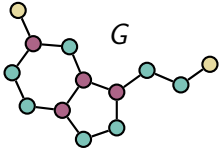
# Example: The Randić Index

One example of a degree-based topological index is the Randić index $R$. This is the first such index, introduced in 1975 by Milan Randić, and remains one of the most widely studied.



Its weights are defined as: $c_{ij} = \frac{1}{\sqrt{i \cdot j}}$ $\Rightarrow (c_{ij})_{1 \le i \le j \le 3} = \begin{pmatrix} 1 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ & \frac{1}{2} & \frac{1}{\sqrt{6}} \\ & & \frac{1}{3} \end{pmatrix}$

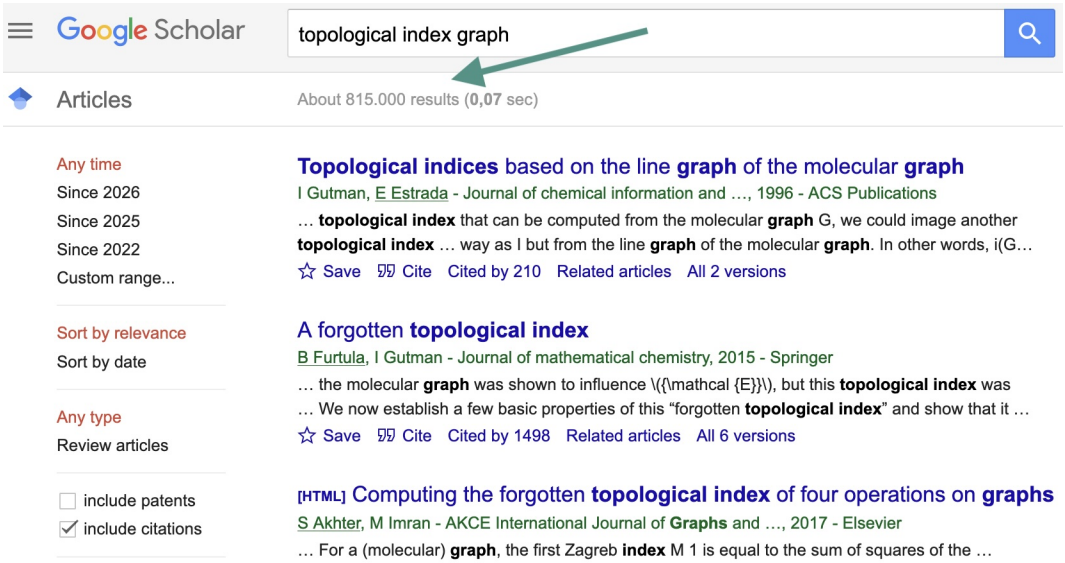
deg 1
deg 2
deg 3




$m_{12} = 1 \quad m_{23} = 7$
$m_{13} = 1 \quad m_{33} = 2$
$m_{22} = 3 \quad m_{11} = 0$

$$R(G) = 1 \cdot \frac{1}{\sqrt{2}} + 1 \cdot \frac{1}{\sqrt{3}} + 3 \cdot \frac{1}{2} + 7 \cdot \frac{1}{\sqrt{6}} + 2 \cdot \frac{1}{3} \approx 6.309$$

# And Randić is not alone. . .



# A Multitude of Degree-Based Topological Indices

| Name | Short name | $c_{ij}$ |
|---|---|---|
| Atom-bond connectivity index | ABC | $\sqrt{\frac{i+j-2}{ij}}$ |
| Atom-bond sum-connectivity index | ABSC | $\sqrt{\frac{i+j-2}{i+j}}$ |
| Albertson index | Albertson | $\lvert i-j\rvert$ |
| Arithmetic-geometric index | AG | $\frac{i+j}{2\sqrt{ij}}$ |
| Difference between AG and GA | AG-GA | $\frac{i+j}{2\sqrt{ij}} - \frac{2\sqrt{ij}}{i+j}$ |
| Extended index | Extended | $\frac{1}{2}(\frac{i}{j} + \frac{j}{i})$ |
| Forgotten index | Forgotten | $i^2 + j^2$ |
| Geometric-arithmetic index | GA | $\frac{2\sqrt{ij}}{i+j}$ |
| First Gourava index | Gourava1 | $i + j + ij$ |
| Second Gourava index | Gourava2 | $(i + j)ij$ |
| First hyper-Gourava index | hGourava1 | $(i + j + ij)^2$ |
| Second hyper-Gourava index | hGourava2 | $((i + j)ij)^2$ |
| Gourava sum-connectivity index | GouravaSC | $\frac{1}{\sqrt{i+j+ij}}$ |
| Gourava product-connectivity index | GouravaPC | $\sqrt{ij(i + j)}$ |
| Harmonic index | Harmonic | $\frac{2}{i+j}$ |
| Inverse degree index | InvDeg | $i^{-2} + j^{-2}$ |
| Inverse sum of degree index | InvSumDeg | $\frac{ij}{i+j}$ |
| Randić index | Randić | $\frac{1}{\sqrt{ij}}$ |
| Reciprocal Randić index | rRandić | $\sqrt{ij}$ |
| Sigma index | Sigma | $(i - j)^2$ |
| Sombor index | Sombor | $\sqrt{i^2 + j^2}$ |
| Reduced Sombor index | rSombor | $\sqrt{(i - 1)^2 + (j - 1)^2}$ |
| Sum connectivity index | SumConn | $\frac{1}{\sqrt{i+j}}$ |
| Reciprocal sum connectivity index | rSumConn | $\sqrt{i + j}$ |
| First Zagreb index | Zagreb1 | $i + j$ |
| Second Zagreb index | Zagreb2 | $ij$ |
| Augmented Zagreb index | aZagreb | $(\frac{ij}{i+j-2})^3$ |
| First hyper-Zagreb index | hZagreb1 | $(i + j)^2$ |
| Second hyper-Zagreb index | hZagreb2 | $(ij)^2$ |
| Nat. log. of the mult. sum Zagreb index | lnZagreb1 | $\ln(i + j)$ |
| Nat. log. of the first mult. Zagreb index | lnZagreb2 | $2(\frac{\ln(i)}{i} + \frac{\ln(j)}{j})$ |
| Nat. log. of the second mult. Zagreb index | lnZagreb3 | $\ln(i) + \ln(j)$ |
| Modified first Zagreb index | mZagreb | $i^{-3} + j^{-3}$ |

# A Multitude of Degree-Based Topological Indices

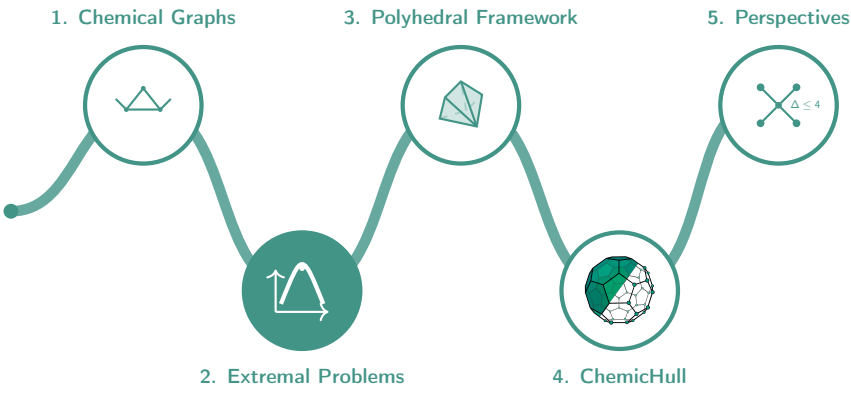
According to Ivan Gutman (2013)[2]:

*"Today there are far too many such descriptors and a firm criterion to stop or slow down their proliferation seems to be lacking."*

**Observation:** More than a decade later, this trend has persisted, with novel degree-based descriptors appearing regularly in the literature.

[2] I. Gutman. Degree-Based Topological Indices. *Croatica Chemica Acta* 86 (2013), pp. 351–361.

# Extremal Problems & Discoveries for 33 Indices



# Bounding indices

## Our Goals:

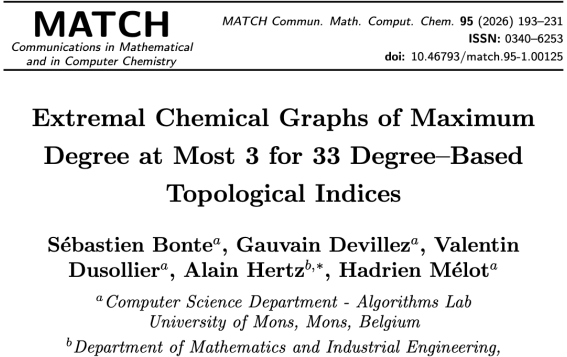
- Let $\mathcal{I}$ be a degree-based topological index, find
  - a tight **upper bound** (**max** problem) and
  - a tight **lower bound** (**min** problem)

  on $\mathcal{I}$ among chemical graphs with fixed order $n$ and size $m$
- Characterize the **extremal graphs** reaching these bounds
- Based on these bounds for various indices, are they really **distinct from an extremal point of view**?

# A Multitude of Indices

In a recent paper (MATCH, 2026), we listed **33** common degree-based indices.

For each of the 33 indices, we aim to solve the **max** and **min** extremal problems.



**MATCH**
*Communications in Mathematical and in Computer Chemistry*

*MATCH Commun. Math. Comput. Chem.* **95** (2026) 193–231
**ISSN:** 0340–6253
**doi:** 10.46793/match.95-1.00125

**Extremal Chemical Graphs of Maximum Degree at Most 3 for 33 Degree–Based Topological Indices**

**Sébastien Bonte[a], Gauvain Devillez[a], Valentin Dusollier[a], Alain Hertz[b,*], Hadrien Mélot[a]**

[a] *Computer Science Department - Algorithms Lab University of Mons, Mons, Belgium*

[b] *Department of Mathematics and Industrial Engineering, Polytechnique Montréal - Gerad, Montréal, Canada*

# Exploration with PHOEG

To tackle this, we started with the help of a computer using PHOEG (see **Dusollier**'s talk on Thursday).

For small graphs ($n \leq 5$), we can easily check the 20 existing chemical graphs (and find the extremal ones).

As we started exploring these problems for $n \geq 6$, interesting patterns emerged.

# Family $F_1$ of chemical graphs

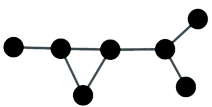
**Definition**

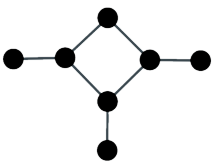
Let $n \geq 6$, $F_1$ is the set of chemical graphs with the following $m_{ij}$ values:

| $m_{12}$ | $m_{13}$ | $m_{22}$ | $m_{23}$ | $m_{33}$ | |
|---|---|---|---|---|---|
| 0 | $\frac{3n-2m}{2}$ | 0 | 0 | $\frac{4m-3n}{2}$ | if $n$ if even |
| 0 | $\frac{3n-2m-1}{2}$ | 0 | 2 | $\frac{4m-3n-3}{2}$ | if $n$ if odd |

**Example:** If $n = 7$ and $m = 7$, then $m_{13} = 3, m_{23} = 2, m_{33} = 2$.

There exist 2 graphs $\in F_1$ of order 7 and size 7:





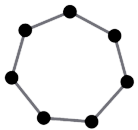
# Family $F_2$ of chemical graphs

**Definition**

$F_2$ is the set of chemical graphs with the following numbers $m_{ij}$ values:

| $m_{12}$ | $m_{13}$ | $m_{22}$ | $m_{23}$ | $m_{33}$ | |
|---|---|---|---|---|---|
| 2 | 0 | $m-2$ | 0 | 0 | if $m = n-1$ |
| 0 | 0 | $m$ | 0 | 0 | if $m = n$ |
| 0 | 0 | $m-5$ | 4 | 1 | if $m = n+1$ |
| 0 | 0 | $3n-2m-1$ | 2 | $3m-3n-1$ | if $n+1 < m \leq \frac{3n-3}{2}$ |

**Example:** If $n = 7$ and $m = 7$, then $m_{22} = 7$. There is only one graph $\in F_2$ of order 7 and size 7: the cycle $C_7$.



# A Surprising Discovery

**Theorem**

*$F_1 \cup F_2$ is the set of extremal chemical graphs for **13 out of the 33** indices, solving both* **min** *and* **max** *problems.*

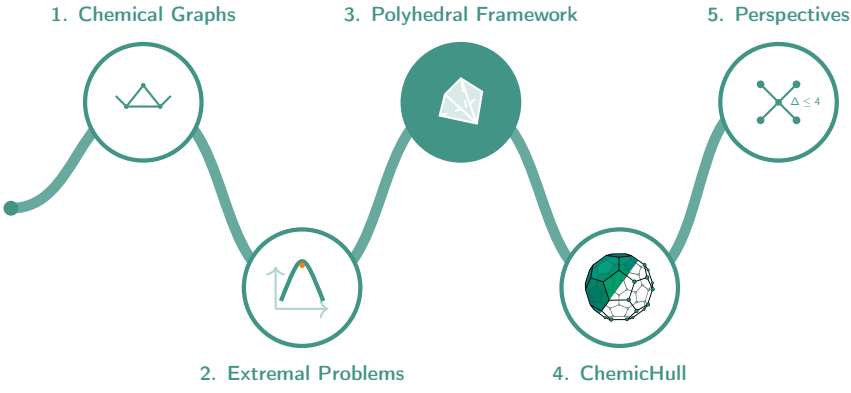
*These indices are ABSC, AG, AG-GA, Extended, GA, GouravaSC, Harmonic, Randić, Sombor, rSombor, SumConn, rSumConn and lnZagreb1.*

**Key idea in the proof:** all these indices have common conditions on their $c_{ij}$ values, that are sufficient to prove this result using graph transformations.

# 100% of problems solved for 33 indices

- **13 out of the 33 (39%)** indices share the exact same set of optimal graphs ($F_1 \cup F_2$)
- Only **5 structural families** (which we named $F_1$ to $F_5$) are sufficient to fully characterize the extremal graphs for **29 out of 33 (88%)** indices (both for **min** and **max** problems)
- Some additional families allow to solve the 4 remaining indices
- **Why so little diversity?** The answer will be polyhedral

# The Polyhedral Framework



# From Graphs to Points

To optimize degree-based indices, only the edge distribution ($m_{ij}$ values) matters.

Given two integers $n$ and $m$, optimizing $\mathcal{I}(G)$ among chemical graphs = optimizing the linear function:

$$\sum_{1 \le i \le j \le 3} c_{ij} m_{ij}$$

where a vector $(m_{11}, m_{12}, m_{13}, m_{22}, m_{23}, m_{33})$ is $(n, m)$-*feasible*.

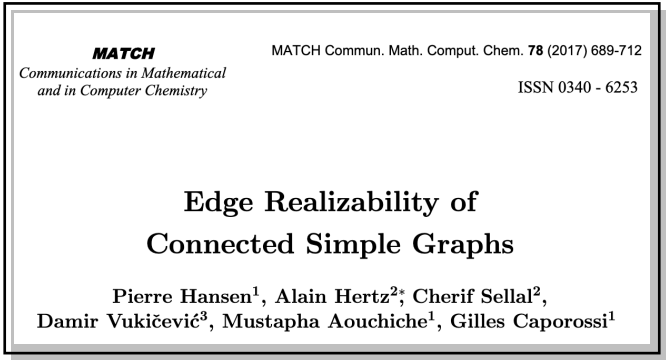
**Definition**

A vector is $(n, m)$-*feasible* if there exists at least one chemical graph of $n$ vertices and $m$ edges, having exactly $m_{ij}$ edges between vertices of degree $i$ and $j$.

# Characterizing Feasibility (Hansen *et al.*, 2017)

But how to ensure that a vector $(m_{ij})$ is $(n, m)$-feasible ?

Thanks to **Hansen *et al.*(2017)**, we have a complete set of necessary and sufficient conditions linking $n, m$, the number $n_i$ of vertices of degree $i$, and the edge counts $m_{ij}$.



**MATCH**
*Communications in Mathematical and in Computer Chemistry*

MATCH Commun. Math. Comput. Chem. **78** (2017) 689-712

ISSN 0340 - 6253

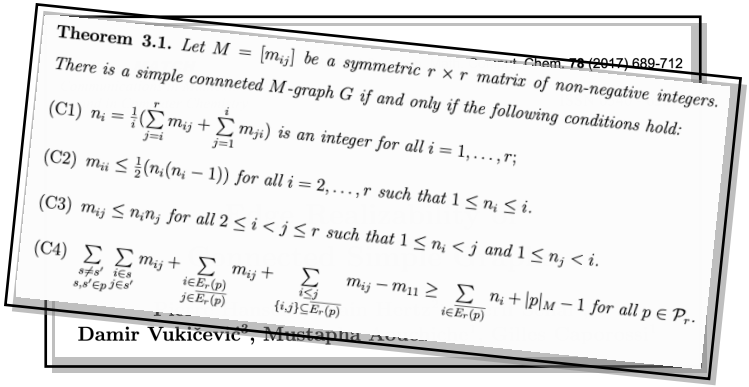
**Edge Realizability of Connected Simple Graphs**

**Pierre Hansen[1], Alain Hertz[2,*], Cherif Sellal[2], Damir Vukičević[3], Mustapha Aouchiche[1], Gilles Caporossi[1]**

# Characterizing Feasibility (Hansen *et al.*, 2017)

But how to ensure that a vector $(m_{ij})$ is $(n, m)$-feasible ?

Thanks to **Hansen *et al.*(2017)**, we have a complete set of necessary and sufficient conditions linking $n, m$, the number $n_i$ of vertices of degree $i$, and the edge counts $m_{ij}$.



**Theorem 3.1.** *Let $M = [m_{ij}]$ be a symmetric $r \times r$ matrix of non-negative integers. There is a simple connneted $M$-graph $G$ if and only if the following conditions hold:*

(C1) $n_i = \frac{1}{i}(\sum_{j=i}^{r} m_{ij} + \sum_{j=1}^{i} m_{ji})$ *is an integer for all* $i = 1, \ldots, r$;

(C2) $m_{ii} \leq \frac{1}{2}(n_i(n_i - 1))$ *for all* $i = 2, \ldots, r$ *such that* $1 \leq n_i \leq i$.

(C3) $m_{ij} \leq n_i n_j$ *for all* $2 \leq i < j \leq r$ *such that* $1 \leq n_i < j$ *and* $1 \leq n_j < i$.

(C4) $\sum_{\substack{s \neq s' \\ s,s' \in p}} \sum_{\substack{i \in s \\ j \in s'}} m_{ij} + \sum_{\substack{i \in \overline{E_r(p)} \\ j \in \overline{E_r(p)}}} m_{ij} + \sum_{\substack{i<j \\ \{i,j\} \subseteq \overline{E_r(p)}}} m_{ij} - m_{11} \geq \sum_{i \in \overline{E_r(p)}} n_i + |p|_M - 1$ *for all* $p \in \mathcal{P}_r$.

Chem. 78 (2017) 689-712

Damir Vukičević[3], Mustapha

# Integer Program (IP)

Restricting Hansen *et al.*'s characterization to chemical graphs, our problem reduces to optimizing a linear function subject to integer linearised constraints:

**Integer Program (IP)**

**optimize** $\sum_{1 \le i \le j \le 3} c_{ij} m_{ij}$

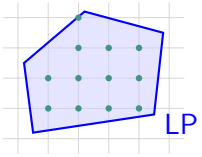
**subject to** Hansen *et al.*'s Linearised Constraints (16 variables, 28 equations/inequalities)

$m_{ij} \geq 0$ and integer $\quad \forall i \leq j$

# Linear Relaxation vs. Convex Hull

**Limits of this IP:**

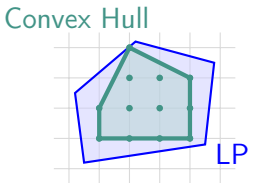
- Its **Linear Relaxation (LP)** is not tight and can have thousands of **fractional** extreme points (e.g., 11648 points for $n = 14, m = 13$)
- Need solvers and / or IP methods (e.g., Branch & Cut)



# Linear Relaxation vs. Convex Hull

**Limits of this IP:**

- Its **Linear Relaxation (LP)** is not tight and can have thousands of **fractional** extreme points (e.g., 11648 points for $n = 14, m = 13$)
- Need solvers and / or IP methods (e.g., Branch & Cut)




**Our Goal:** Characterize the **Convex Hull** of the valid integer points:

- If we have it, we don't need solvers
- The optimum **must** lie on an integer vertex
- Hopefully, it could provides an explanation for the low diversity of extremal graphs

# Dimensionality Reduction (6D to 3D)

To find this convex hull, we can first drastically simplify the space. For fixed $n$ and $m$, the feasible region is actually of lower dimension since:

- $m_{11} = 0$ if $n \geq 3$,
- $m = m_{12} + m_{13} + m_{22} + m_{23} + m_{33}$,
- $n = \frac{3}{2}m_{12} + \frac{4}{3}m_{13} + m_{22} + \frac{5}{6}m_{23} + \frac{2}{3}m_{33}$.

# Dimensionality Reduction (6D to 3D)

To find this convex hull, we can first drastically simplify the space. For fixed $n$ and $m$, the feasible region is actually of lower dimension since:

- $m_{11} = 0$ if $n \geq 3$,
- $m_{22} = 6n - 5m - 4m_{12} - 3m_{13} + m_{33}$,
- $m_{23} = 6m - 6n + 3m_{12} + 2m_{13} - 2m_{33}$.

Points associated to chemical graphs or order $n$ and size $m$ are **fully described by only 3 variables**: $m_{12}$, $m_{13}$ and $m_{33}$.

# Dimensionality Reduction (6D to 3D)

To find this convex hull, we can first drastically simplify the space. For fixed $n$ and $m$, the feasible region is actually of lower dimension since:

- $m_{11} = 0$ if $n \geq 3$,
- $m_{22} = 6n - 5m - 4m_{12} - 3m_{13} + m_{33}$,
- $m_{23} = 6m - 6n + 3m_{12} + 2m_{13} - 2m_{33}$.

Points associated to chemical graphs or order $n$ and size $m$ are **fully described by only 3 variables**: $m_{12}$, $m_{13}$ and $m_{33}$.

By substituting $m_{22}$ and $m_{23}$, the objective function reduces to:

$$\sum_{1 \leq i \leq j \leq 3} c_{ij} m_{ij} = \left( c'_{12} m_{12} + c'_{13} m_{13} + c'_{33} m_{33} \right) + \text{constant (dep. on } n \text{ and } m)$$

# Dimensionality Reduction (6D to 3D)

To find this convex hull, we can first drastically simplify the space. For fixed $n$ and $m$, the feasible region is actually of lower dimension since:

- $m_{11} = 0$ if $n \geq 3$,
- $m_{22} = 6n - 5m - 4m_{12} - 3m_{13} + m_{33}$,
- $m_{23} = 6m - 6n + 3m_{12} + 2m_{13} - 2m_{33}$.

Points associated to chemical graphs or order $n$ and size $m$ are **fully described by only 3 variables**: $m_{12}$, $m_{13}$ and $m_{33}$.

By substituting $m_{22}$ and $m_{23}$, the objective function reduces to:

$$\sum_{1 \leq i \leq j \leq 3} c_{ij} m_{ij} = \left( c'_{12} m_{12} + c'_{13} m_{13} + c'_{33} m_{33} \right) + \text{constant (dep. on } n \text{ and } m)$$

**Our problem is now purely 3-dimensional!**

# Example: The Reduced Randić Index ($\widehat{R}$)

To illustrate, let's apply this dimensionality reduction to the Randić index. By substituting $m_{22}$ and $m_{23}$ with their expressions in terms of $n, m, m_{12}, m_{13}$, and $m_{33}$, the original index $R$ transforms into a new function $\widehat{R}$:

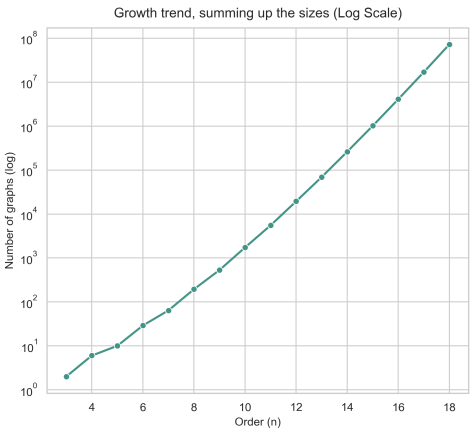
$$\widehat{R}(m_{ij}) = \frac{\sqrt{6}+\sqrt{2}-4}{2} m_{12} + \frac{2\sqrt{6}+2\sqrt{3}-9}{3} m_{13} + \frac{-2\sqrt{6}+5}{6} m_{33}.$$

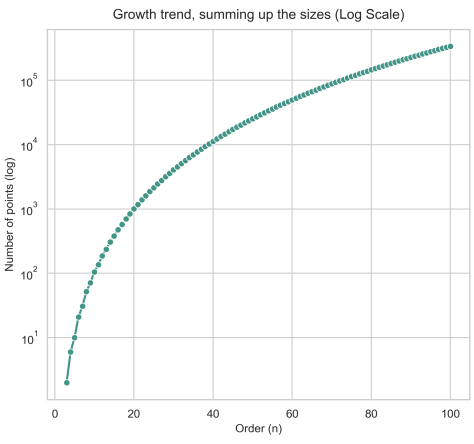
**Equivalence**

Optimizing the reduced function $\widehat{R}(m_{ij})$ in 3D is **strictly equivalent** to optimizing the original Randić index $R(G)$, as they only differ by a constant depending purely on $n$ and $m$.

# Growth of points vs. Graphs

Reducing the problem to 3D points drastically mitigates the combinatorial explosion of chemical graphs, though the growth of these points remains substantial.



# Methodology: Conjecturing the Convex Hulls

To understand the general structure of the feasible region, we proceeded in three main steps:

1. **Computation:** We computed the exact convex hulls of points for small values of $(n, m)$. The $(n, m)$-feasible points were determined using Hansen et al.'s linearised constraints.
2. **Pattern Recognition:** By analyzing these small cases, we observed a strict periodicity in the polytope descriptions. This allowed us to conjecture and generalise the equations of the facets and the coordinates of the vertices for arbitrary values of $(n, m)$.
3. **Formal Proof:** Building upon these generalized periodic patterns, we were ultimately able to determine and rigorously prove the exact convex hull for **any** $(n, m)$.

# Convex Hull for any $(n, m)$

We were able to **fully characterize the convex hull** of this 3D feasible region **for any** $n$ **and** $m$:

# Convex Hull for any $(n, m)$

We were able to **fully characterize the convex hull** of this 3D feasible region **for any** $n$ **and** $m$:

- We established **96 families of polytopes** that perfectly cover all possible values of $n$ and $m$.

# Convex Hull for any $(n, m)$

We were able to **fully characterize the convex hull** of this 3D feasible region **for any** $n$ **and** $m$:

- We established **96 families of polytopes** that perfectly cover all possible values of $n$ and $m$.
- Surprisingly, the polyhedral complexity is strictly bounded! It has **at most 10 facets** and **16 vertices**, meaning the **complexity doesn't grow with** $n$ **and** $m$.

# Finding Optimal Points

Because every topological index translates to a linear function in our 3D space, the optima **necessarily lie on the boundary** of the convex hull.

**Example:** If $n = 15$ and $m = 18$, the convex hull has exactly 6 vertices. To optimize $\widehat{R}$, we just evaluate it at these 6 points:

| **Vertices** |
|---|
| $(0, 0, 0)$ |
| $(0, 4, 12)$ |
| $(1, 3, 13)$ |
| $(3, 0, 12)$ |
| $(0, 0, 8)$ |
| $(1, 0, 10)$ |

# Finding Optimal Points

Because every topological index translates to a linear function in our 3D space, the optima **necessarily lie on the boundary** of the convex hull.

**Example:** If $n = 15$ and $m = 18$, the convex hull has exactly 6 vertices. To optimize $\widehat{R}$, we just evaluate it at these 6 points:

| **Vertices** | $\widehat{R}$ |
|---|---|
| $(0, 0, 0)$ | $0$ |
| $(0, 4, 12)$ | $\approx -0.2225$ |
| $(1, 3, 13)$ | $\approx -0.1677$ |
| $(3, 0, 12)$ | $\approx -0.0024$ |
| $(0, 0, 8)$ | $\approx 0.1346$ |
| $(1, 0, 10)$ | $\approx 0.1002$ |

$$\widehat{R}(m_{ij}) = \frac{\sqrt{6}+\sqrt{2}-4}{2} m_{12} + \frac{2\sqrt{6}+2\sqrt{3}-9}{3} m_{13} + \frac{-2\sqrt{6}+5}{6} m_{33}$$

# Finding Optimal Points

Because every topological index translates to a linear function in our 3D space, the optima **necessarily lie on the boundary** of the convex hull.

**Example:** If $n = 15$ and $m = 18$, the convex hull has exactly 6 vertices. To optimize $\widehat{R}$, we just evaluate it at these 6 points:

| **Vertices** | $\widehat{R}$ | |
|---|---|---|
| $(0, 0, 0)$ | $0$ | |
| $(0, 4, 12)$ | $\approx -0.2225$ | $\leftarrow$ **min** |
| $(1, 3, 13)$ | $\approx -0.1677$ | |
| $(3, 0, 12)$ | $\approx -0.0024$ | |
| $(0, 0, 8)$ | $\approx 0.1346$ | $\leftarrow$ **max** |
| $(1, 0, 10)$ | $\approx 0.1002$ | |

$$\widehat{R}(m_{ij}) = \frac{\sqrt{6}+\sqrt{2}-4}{2} m_{12} + \frac{2\sqrt{6}+2\sqrt{3}-9}{3} m_{13} + \frac{-2\sqrt{6}+5}{6} m_{33}$$

## Remarks

- This gives have a beautiful **mathematical answer** to the low diversity of indices (there are a very few number of possible optimal points)

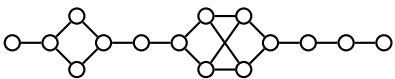
# Remarks

- This gives have a beautiful **mathematical answer** to the low diversity of indices (there are a very few number of possible optimal points)
- One can instantly find the optimal points **without any solver**

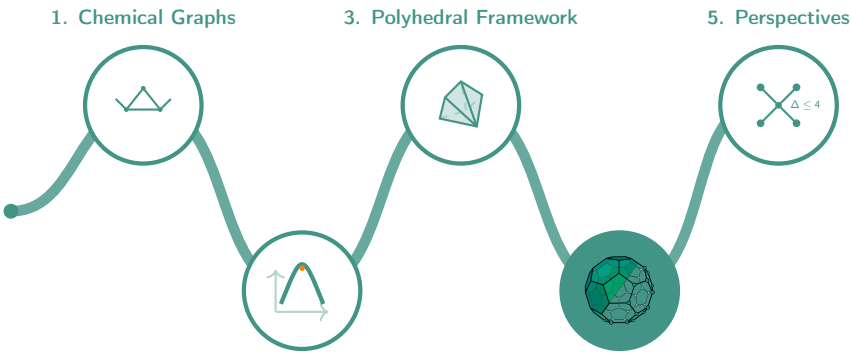
# Remarks

- This gives have a beautiful **mathematical answer** to the low diversity of indices (there are a very few number of possible optimal points)
- One can instantly find the optimal points **without any solver**
- Chemical graphs that map to **inner points** of the polytope are **never optimal** for any degree-based index (even not yet invented)

  **Example:** With $n = 15$ and $m = 18$, the point $(1, 1, 8)$ is an inner point. Therefore, no matter which is your favorite degree-based index, this graph will never be extremal:



# ChemicHull: From Theory to Software



# Motivation for ChemicHull

The polyhedral description is mathematically beautiful and solves the problem entirely.

However, how can chemists or graph theorists easily

- find optimal points for any given degree-based index?
- find a graph (molecule) from an (optimal) point?
- explore the polytopes?

# Motivation of ChemicHull

Of course, one can use the 31-page paper and its tables

Complete polyhedral description of chemical graphs of maximum degree at most 3

Valentin Dusollier[1], Sébastien Bonte[1], Gauvain Devillez[1], Alain Hertz[2], Hadrien Mélot[1], David Schindl[3]

[1] Computer Science Department - Algorithms Lab
University of Mons, Mons, Belgium

[2] Department of Mathematics and Industrial Engineering
Polytechnique Montréal - Gerad, Montréal, Canada
Corresponding author, alain.hertz@gerad.ca

[3] Haute Ecole de Gestion de Genève
University of Applied Sciences Western Switzerland, Genève, Switzerland

June 25, 2025

*In memory of Pierre Hansen.*

# Motivation of ChemicHull

Of course, one can use the 31-page paper and its tables

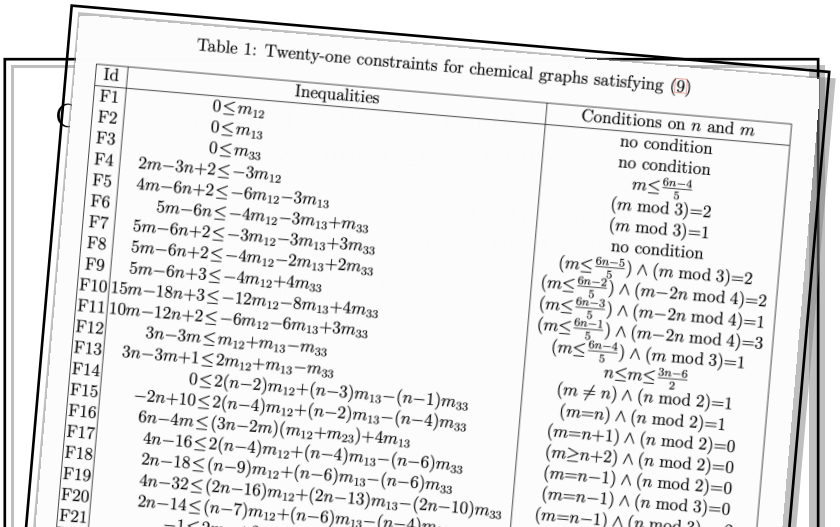


Table 1: Twenty-one constraints for chemical graphs satisfying (9)

| Id | Inequalities | Conditions on $n$ and $m$ |
|---|---|---|
| F1 | $0 \leq m_{12}$ | no condition |
| F2 | $0 \leq m_{13}$ | no condition |
| F3 | $0 \leq m_{33}$ | $m \leq \frac{6n-4}{5}$ |
| F4 | $2m-3n+2 \leq -3m_{12}$ | $(m \bmod 3)=2$ |
| F5 | $4m-6n+2 \leq -6m_{12}-3m_{13}$ | $(m \bmod 3)=1$ |
| F6 | $5m-6n \leq -4m_{12}-3m_{13}+m_{33}$ | no condition |
| F7 | $5m-6n+2 \leq -3m_{12}-3m_{13}+3m_{33}$ | $(m \leq \frac{6n-5}{5}) \wedge (m \bmod 3)=2$ |
| F8 | $5m-6n+2 \leq -4m_{12}-2m_{13}+2m_{33}$ | $(m \leq \frac{6n-2}{5}) \wedge (m-2n \bmod 4)=2$ |
| F9 | $5m-6n+3 \leq -4m_{12}+4m_{33}$ | $(m \leq \frac{6n-3}{5}) \wedge (m-2n \bmod 4)=1$ |
| F10 | $15m-18n+3 \leq -12m_{12}-8m_{13}+4m_{33}$ | $(m \leq \frac{6n-1}{5}) \wedge (m-2n \bmod 4)=3$ |
| F11 | $10m-12n+2 \leq -6m_{12}-6m_{13}+3m_{33}$ | $(m \leq \frac{6n-4}{5}) \wedge (m \bmod 3)=1$ |
| F12 | $3n-3m \leq m_{12}+m_{13}-m_{33}$ | $n \leq m \leq \frac{3n-6}{2}$ |
| F13 | $3n-3m+1 \leq 2m_{12}+m_{13}-m_{33}$ | $(m \neq n) \wedge (n \bmod 2)=1$ |
| F14 | $0 \leq 2(n-2)m_{12}+(n-3)m_{13}-(n-1)m_{33}$ | $(m=n) \wedge (n \bmod 2)=1$ |
| F15 | $-2n+10 \leq 2(n-4)m_{12}+(n-2)m_{13}-(n-4)m_{33}$ | $(m=n+1) \wedge (n \bmod 2)=0$ |
| F16 | $6n-4m \leq (3n-2m)(m_{12}+m_{23})+4m_{13}$ | $(m \geq n+2) \wedge (n \bmod 2)=0$ |
| F17 | $4n-16 \leq 2(n-4)m_{12}+(n-4)m_{13}-(n-6)m_{33}$ | $(m=n-1) \wedge (n \bmod 2)=0$ |
| F18 | $2n-18 \leq (n-9)m_{12}+(n-6)m_{13}-(n-6)m_{33}$ | $(m=n-1) \wedge (n \bmod 3)=0$ |
| F19 | $4n-32 \leq (2n-16)m_{12}+(2n-13)m_{13}-(2n-10)m_{33}$ | $(m=n-1) \wedge (n \bmod 3)=0$ |
| F20 | $2n-14 \leq (n-7)m_{12}+(n-6)m_{13}-(n-4)m_{33}$ | |
| F21 | | |

# Motivation of ChemicHull

Of course, one can use the 31-page paper and its tables

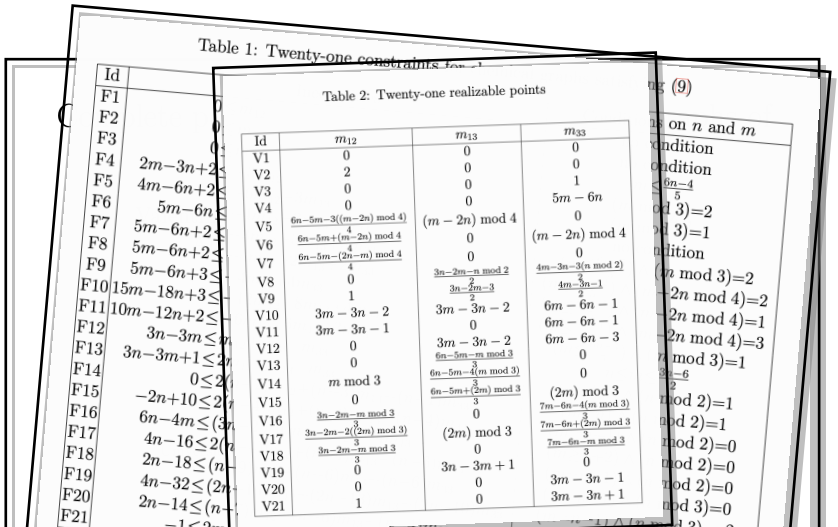


Table 1: Twenty-one constraints

| Id | |
|---|---|
| F1 | |
| F2 | |
| F3 | |
| F4 | $2m-3n+2$ |
| F5 | $4m-6n+2$ |
| F6 | $5m-6n$ |
| F7 | $5m-6n+2$ |
| F8 | $5m-6n+2$ |
| F9 | $5m-6n+3$ |
| F10 | $15m-18n+3$ |
| F11 | $10m-12n+2$ |
| F12 | $3n-3m$ |
| F13 | $3n-3m+1$ |
| F14 | $0$ |
| F15 | $-2n+10$ |
| F16 | $6n-4m$ |
| F17 | $4n-16$ |
| F18 | $2n-18$ |
| F19 | $4n-32$ |
| F20 | $2n-14$ |
| F21 | |

Table 2: Twenty-one realizable points

| Id | $m_{12}$ | $m_{13}$ | $m_{33}$ |
|---|---|---|---|
| V1 | 0 | 0 | 0 |
| V2 | 2 | 0 | 0 |
| V3 | 0 | 0 | 1 |
| V4 | 0 | 0 | $5m-6n$ |
| V5 | $\frac{6n-5m-3((m-2n) \bmod 4)}{4}$ | $(m-2n) \bmod 4$ | 0 |
| V6 | $\frac{6n-5m+(m-2n) \bmod 4}{4}$ | 0 | $(m-2n) \bmod 4$ |
| V7 | $\frac{6n-5m-(2n-m) \bmod 4}{4}$ | 0 | 0 |
| V8 | 0 | $\frac{3n-2m-n \bmod 2}{2}$ | $\frac{4m-3n-3(n \bmod 2)}{2}$ |
| V9 | 1 | $\frac{3n-2m-3}{2}$ | $\frac{4m-3n-1}{2}$ |
| V10 | $3m-3n-2$ | $3m-3n-2$ | $6m-6n-1$ |
| V11 | $3m-3n-1$ | 0 | $6m-6n-1$ |
| V12 | 0 | $3m-3n-2$ | $6m-6n-3$ |
| V13 | 0 | $\frac{6n-5m-m \bmod 3}{3}$ | 0 |
| V14 | $m \bmod 3$ | $\frac{6n-5m-4(m \bmod 3)}{3}$ | 0 |
| V15 | 0 | $\frac{6n-5m+(2m) \bmod 3}{3}$ | $(2m) \bmod 3$ |
| V16 | $\frac{3n-2m-m \bmod 3}{3}$ | 0 | $\frac{7m-6n-4(m \bmod 3)}{3}$ |
| V17 | $\frac{3n-2m-2(2m) \bmod 3)}{3}$ | $(2m) \bmod 3$ | $\frac{7m-6n+(2m) \bmod 3}{3}$ |
| V18 | $\frac{3n-2m-m \bmod 3}{3}$ | 0 | $\frac{7m-6n-m \bmod 3}{3}$ |
| V19 | 0 | $3n-3m+1$ | 0 |
| V20 | 0 | 0 | $3m-3n-1$ |
| V21 | 1 | 0 | $3m-3n+1$ |

(9)

on $n$ and $m$

ndition

ndition

$\frac{6n-4}{5}$

d 3)=2

d 3)=1

dition

($n$ mod 3)=2

$2n$ mod 4)=2

$2n$ mod 4)=1

$2n$ mod 4)=3

mod 3)=1

$\frac{-6}{2}$

od 2)=1

od 2)=1

od 2)=0

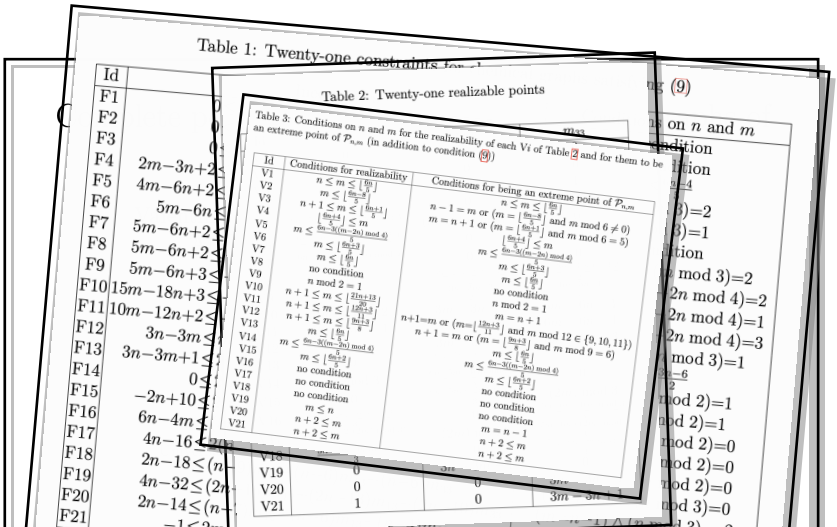
od 2)=0

od 2)=0

od 3)=0

# Motivation of ChemicHull

Of course, one can use the 31-page paper and its tables

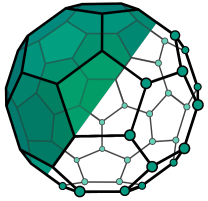


# Motivation of ChemicHull

Of course, one can use the 31-page paper and its tables, **but it is not practical.**

# ChemicHull

Another solution is to dive into ChemicHull[1], the companion website of these results.

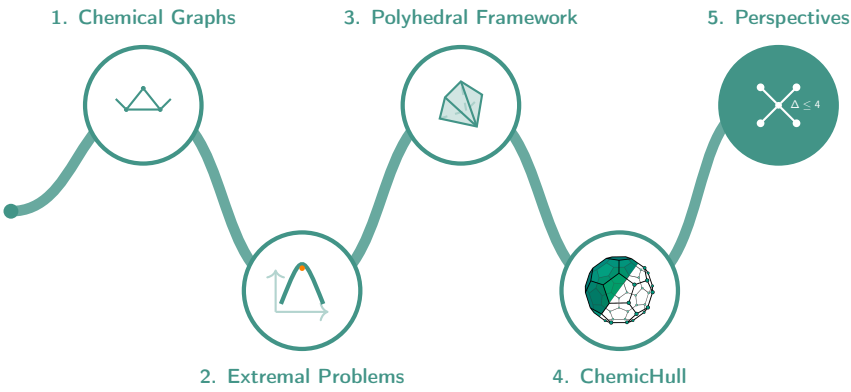


`https://phoeg.umons.ac.be/chemichull`

**Short Live Demonstration**

[1] Bonte et al., ChemicHull: an online tool for determining extremal chemical graphs of maximum degree at most 3 for any degree-based topological indices. Accepted in *Discrete Mathematical Chemistry*, 2026.

# Perspectives



# From a Point to a Graph

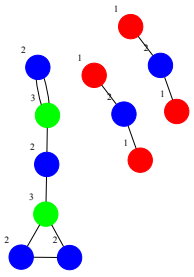
**How to get a graph (molecule) from a point?** ChemicHull uses a constructive algorithm from Hansen *et al.* (2017) to build **one** representative graph step by step.

# From a Point to a Graph

**How to get a graph (molecule) from a point?** ChemicHull uses a constructive algorithm from Hansen *et al.* (2017) to build **one** representative graph step by step.

**Example:** point $(4, 0, 0)$ (when $n = 12$ and $m = 11$)

$$\implies m_{12} = 4, m_{13} = 0, m_{33} = 0, m_{22} = 1, m_{23} = 6, n_1 = 4, n_2 = 6, n_3 = 2$$

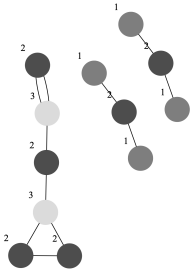


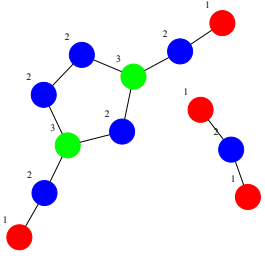
# From a Point to a Graph

**How to get a graph (molecule) from a point?** ChemicHull uses a constructive algorithm from Hansen *et al.* (2017) to build **one** representative graph step by step.

**Example:** point $(4, 0, 0)$ (when $n = 12$ and $m = 11$)

$$\implies m_{12} = 4, m_{13} = 0, m_{33} = 0, m_{22} = 1, m_{23} = 6, n_1 = 4, n_2 = 6, n_3 = 2$$



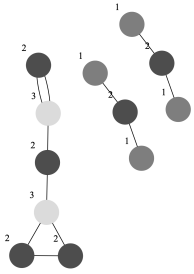


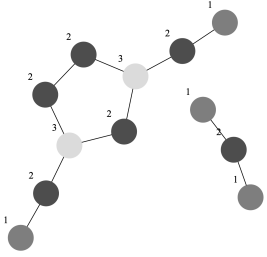
# From a Point to a Graph

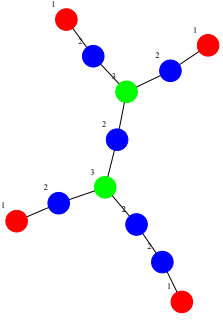
**How to get a graph (molecule) from a point?** ChemicHull uses a constructive algorithm from Hansen *et al.* (2017) to build **one** representative graph step by step.

**Example:** point $(4, 0, 0)$ (when $n = 12$ and $m = 11$)

$$\implies m_{12} = 4, m_{13} = 0, m_{33} = 0, m_{22} = 1, m_{23} = 6, n_1 = 4, n_2 = 6, n_3 = 2$$







# Perspective 1: from a Point to all Corresponding Graphs

Given $(n, m)$, multiple non-isomorphic graphs can share the exact same $(m_{ij})$ vector.

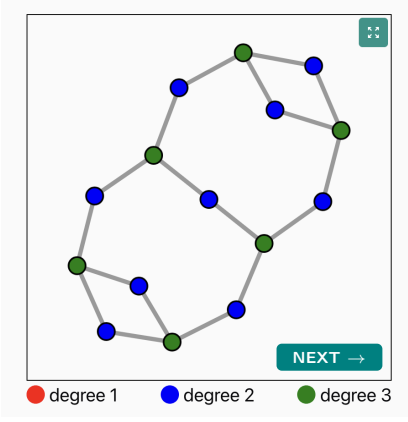
**Example:** Both chemical graphs of order 6 and size 6 below map to the point $(1, 0, 0)$

**Open Problem (Enumerating all graphs from a point)**

Given $n$, $m$, $m_{12}$, $m_{13}$, $m_{33}$ values, how to enumerate **all** corresponding chemical graphs?

# Perspective 1: from a Point to all Corresponding Graphs

Such a generator could be useful on ChemicHull to allow users to scroll among all optimal molecules



# Perspective 1: from a Point to all Corresponding Graphs

However this generator should yield graphs by small batches since the number of graphs for a point can be huge.

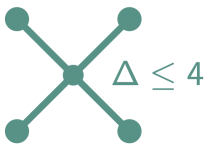
**Example:** Consider the point $(0, 0, m)$ implying all edges link vertices of degree 3. Generating all graphs for this single point is thus equivalent to enumerating **all cubic graphs** of a given order (this number grows exponentially).

# Perspective 1: from a Point to all Corresponding Graphs

However this generator should yield graphs by small batches since the number of graphs for a point can be huge.

I dont' know... maybe some people in this room can help?

# Perspective 2: $\Delta \leq 4$




What if we want to model saturated hydrocarbons (like alkanes)? As mentioned at the beginning, we must extend our maximum degree from $\Delta \leq 3$ to $\Delta \leq 4$.

**Problem ($\Delta \leq 4$)**

Given $(n, m)$, can we describe the convex hull of chemical graphs with $\Delta \leq 4$ (in a $m_{ij}$ space)?

**Remark:** solving this problem could have a huge impact in Mathematical Chemistry (much more literature with **this definition** of chemical graphs).

# Perspective 2: $\Delta \leq 4$

This seemingly innocent change

- adds **four** new variables: $m_{14}, m_{24}, m_{34}, m_{44}$
- induces a 7D space increasing significantly the complexity

However,

- we are currently working to compute (parametrized) convex hulls in a fully automated way
- preliminary results suggest a limited number of vertices in the convex hulls; we conjecture a constant bound, as observed for $\Delta \leq 3$

# Perspective 2: $\Delta \leq 4$

**Example:** $n = 77,\ m = 76$

| | $\Delta \leq 3$ | $\Delta \leq 4$ |
|---|---|---|
| # of graphs | $6.16 \times 10^{25}$ *(a)* | $5.21 \times 10^{29}$ *(b)* |
| # of distinct points | 6884 | $\simeq$ 30 million *(c)* |
| # of Convex Hull vertices | 13 | 216 *(c)* |

*(a)* OEIS A000672
*(b)* OEIS A000602
*(c)* Estimation

# Conclusion

- What started as a surprising observation of few families of extremal graphs for 33 indices led to a complete polyhedral explanation
- Extremal problems for any degree-based topological index (for $\Delta \leq 3$) are now fundamentally solved and unified
- ChemicHull makes this results accessible to chemists and graph theorists instantly
- Perspectives: enumerating all graphs from a point and cracking the $\Delta \leq 4$ combinatorial explosion

**Thank you for your attention!**