Supplementary Material: An Inverse QSAR Method Based on Linear Regression and Integer Programming

Jianshen Zhu[1], Naveed Ahmed Azam[1], Kazuya Haraguchi[1], Liang Zhao[2], Hiroshi Nagamochi[1] and Tatsuya Akutsu[3]

1. Department of Applied Mathematics and Physics, Kyoto University, Kyoto 606-8501, Japan

2. Graduate School of Advanced Integrated Studies in Human Survavibility (Shishu-Kan), Kyoto University, Kyoto 606-8306, Japan

3. Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji 611-0011, Japan

1 A Full Description of Descriptors

Associated with the two functions α and β in a chemical graph $\mathbb{C} = (H, \alpha, \beta)$, we introduce functions ac : $V(E) \to (\Lambda \setminus \{\mathtt{H}\}) \times (\Lambda \setminus \{\mathtt{H}\}) \times [1, 3]$, cs : $V(E) \to (\Lambda \setminus \{\mathtt{H}\}) \times [1, 6]$ and ec : $V(E) \to ((\Lambda \setminus \{\mathtt{H}\}) \times [1, 6]) \times ((\Lambda \setminus \{\mathtt{H}\}) \times [1, 6]) \times [1, 3]$ in the following.

To represent a feature of the exterior of $\mathbb{C}$, a chemical rooted tree in $\mathcal{T}(\mathbb{C})$ is called a *fringe-configuration* of $\mathbb{C}$.

We also represent leaf-edges in the exterior of $\mathbb{C}$. For a leaf-edge $uv \in E(\langle\mathbb{C}\rangle)$ with $\deg_{\langle\mathbb{C}\rangle}(u) = 1$, we define the *adjacency-configuration* of e to be an ordered tuple $(\alpha(u), \alpha(v), \beta(uv))$. Define

$$\Gamma^{\mathrm{lf}}_{\mathrm{ac}} \triangleq \{(\mathtt{a}, \mathtt{b}, m) \mid \mathtt{a}, \mathtt{b} \in \Lambda, m \in [1, \min\{\mathrm{val}(\mathtt{a}), \mathrm{val}(\mathtt{b})\}]\}$$

as a set of possible adjacency-configurations for leaf-edges.

To represent a feature of an interior-vertex $v \in V^{\mathrm{int}}(\mathbb{C})$ such that $\alpha(v) = \mathtt{a}$ and $\deg_{\langle\mathbb{C}\rangle}(v) = d$ (i.e., the number of non-hydrogen atoms adjacent to v is d) in a chemical graph $\mathbb{C} = (H, \alpha, \beta)$, we use a pair $(\mathtt{a}, d) \in (\Lambda \setminus \{\mathtt{H}\}) \times [1, 4]$, which we call the *chemical symbol* $\mathrm{cs}(v)$ of the vertex v. We treat $(\mathtt{a}, d)$ as a single symbol $\mathtt{a}d$, and define Λ_{dg} to be the set of all chemical symbols $\mu = \mathtt{a}d \in (\Lambda \setminus \{\mathtt{H}\}) \times [1, 4]$.

We define a method for featuring interior-edges as follows. Let $e = uv \in E^{\mathrm{int}}(\mathbb{C})$ be an interior-edge $e = uv \in E^{\mathrm{int}}(\mathbb{C})$ such that $\alpha(u) = \mathtt{a}$, $\alpha(v) = \mathtt{b}$ and $\beta(e) = m$ in a chemical graph $\mathbb{C} = (H, \alpha, \beta)$. To feature this edge e, we use a tuple $(\mathtt{a}, \mathtt{b}, m) \in (\Lambda \setminus \{\mathtt{H}\}) \times (\Lambda \setminus \{\mathtt{H}\}) \times [1, 3]$, which we call the *adjacency-configuration* $\mathrm{ac}(e)$ of the edge e. We introduce a total order $<$ over the elements in Λ to distinguish between $(\mathtt{a}, \mathtt{b}, m)$ and $(\mathtt{b}, \mathtt{a}, m)$ $(\mathtt{a} \neq \mathtt{b})$ notationally. For a tuple $\nu = (\mathtt{a}, \mathtt{b}, m)$, let $\overline{\nu}$ denote the tuple $(\mathtt{b}, \mathtt{a}, m)$.

Let $e = uv \in E^{\mathrm{int}}(\mathbb{C})$ be an interior-edge $e = uv \in E^{\mathrm{int}}(\mathbb{C})$ such that $\mathrm{cs}(u) = \mu$, $\mathrm{cs}(v) = \mu'$ and $\beta(e) = m$ in a chemical graph $\mathbb{C} = (H, \alpha, \beta)$. To feature this edge e, we use a tuple $(\mu, \mu', m) \in \Lambda_{\mathrm{dg}} \times \Lambda_{\mathrm{dg}} \times [1, 3]$, which we call the *edge-configuration* $\mathrm{ec}(e)$ of the edge e. We introduce a total order $<$ over the elements in Λ_{dg} to distinguish between (μ, μ', m) and (μ', μ, m) $(\mu \neq \mu')$ notationally. For a tuple $\gamma = (\mu, \mu', m)$, let $\overline{\gamma}$ denote the tuple (μ', μ, m).

Let π be a chemical property for which we will construct a prediction function η from a feature vector $f(\mathbb{C})$ of a chemical graph $\mathbb{C}$ to a predicted value $y \in \mathbb{R}$ for the chemical property of $\mathbb{C}$.

We first choose a set Λ of chemical elements and then collect a data set D_π of chemical compounds C whose chemical elements belong to Λ, where we regard D_π as a set of chemical graphs $\mathbb{C}$ that represent the chemical compounds C in D_π. To define the interior/exterior of chemical graphs $\mathbb{C} \in D_\pi$, we next choose a branch-parameter ρ, where we recommend $\rho = 2$.

Let $\Lambda^{\mathrm{int}}(D_\pi) \subseteq \Lambda$ (resp., $\Lambda^{\mathrm{ex}}(D_\pi) \subseteq \Lambda$) denote the set of chemical elements used in the set $V^{\mathrm{int}}(\mathbb{C})$ of interior-vertices (resp., the set $V^{\mathrm{ex}}(\mathbb{C})$ of exterior-vertices) of $\mathbb{C}$ over all chemical graphs $\mathbb{C} \in D_\pi$, and $\Gamma^{\mathrm{int}}(D_\pi)$

denote the set of edge-configurations used in the set $E^{\rm int}(\mathbb{C})$ of interior-edges in $\mathbb{C}$ over all chemical graphs $\mathbb{C} \in D_\pi$. Let $\mathcal{F}(D_\pi)$ denote the set of chemical rooted trees ψ r-isomorphic to a chemical rooted tree in $\mathcal{T}(\mathbb{C})$ over all chemical graphs $\mathbb{C} \in D_\pi$, where possibly a chemical rooted tree $\psi \in \mathcal{F}(D_\pi)$ consists of a single chemical element $\mathtt{a} \in \Lambda \setminus \{\mathtt{H}\}$.

We define an integer encoding of a finite set A of elements to be a bijection $\sigma : A \to [1, |A|]$, where we denote by $[A]$ the set $[1, |A|]$ of integers. Introduce an integer coding of each of the sets $\Lambda^{\rm int}(D_\pi)$, $\Lambda^{\rm ex}(D_\pi)$, $\Gamma^{\rm int}(D_\pi)$ and $\mathcal{F}(D_\pi)$. Let $[\mathtt{a}]^{\rm int}$ (resp., $[\mathtt{a}]^{\rm ex}$) denote the coded integer of an element $\mathtt{a} \in \Lambda^{\rm int}(D_\pi)$ (resp., $\mathtt{a} \in \Lambda^{\rm ex}(D_\pi)$), $[\gamma]$ denote the coded integer of an element γ in $\Gamma^{\rm int}(D_\pi)$ and $[\psi]$ denote an element ψ in $\mathcal{F}(D_\pi)$.

Over 99% of chemical compounds $\mathbb{C}$ with up to 100 non-hydrogen atoms in PubChem have degree at most 4 in the hydrogen-suppressed graph $\langle \mathbb{C} \rangle$. We assume that a chemical graph $\mathbb{C}$ treated in this paper satisfies $\deg_{\langle \mathbb{C} \rangle}(v) \leq 4$ in the hydrogen-suppressed graph $\langle \mathbb{C} \rangle$.

In our model, we use an integer $\mathrm{mass}^*(\mathtt{a}) = \lfloor 10 \cdot \mathrm{mass}(\mathtt{a}) \rfloor$, for each $\mathtt{a} \in \Lambda$.

We define the *feature vector* $f(\mathbb{C})$ of a chemical graph $\mathbb{C} = (H, \alpha, \beta) \in D_\pi$ to be a vector that consists of the following non-negative integer descriptors $\mathrm{dcp}_i(\mathbb{C})$, $i \in [1, K]$, where $K = 14 + |\Lambda^{\rm int}(D_\pi)| + |\Lambda^{\rm ex}(D_\pi)| + |\Gamma^{\rm int}(D_\pi)| + |\mathcal{F}(D_\pi)| + |\Gamma^{\rm lf}_{\rm ac}|$.

1. $\mathrm{dcp}_1(\mathbb{C})$: the number $|V(H)| - |V_{\mathtt{H}}|$ of non-hydrogen atoms in $\mathbb{C}$.

2. $\mathrm{dcp}_2(\mathbb{C})$: the rank $\mathrm{r}(\mathbb{C})$ of $\mathbb{C}$.

3. $\mathrm{dcp}_3(\mathbb{C})$: the number $|V^{\rm int}(\mathbb{C})|$ of interior-vertices in $\mathbb{C}$.

4. $\mathrm{dcp}_4(\mathbb{C})$: the average $\overline{\mathrm{ms}}(\mathbb{C})$ of mass* over all atoms in $\mathbb{C}$;
   i.e., $\overline{\mathrm{ms}}(\mathbb{C}) \triangleq \frac{1}{|V(H)|} \sum_{v \in V(H)} \mathrm{mass}^*(\alpha(v))$.

5. $\mathrm{dcp}_i(\mathbb{C})$, $i = 4 + d, d \in [1, 4]$: the number $\mathrm{dg}^{\overline{\mathtt{H}}}_d(\mathbb{C})$ of non-hydrogen vertices $v \in V(H) \setminus V_{\mathtt{H}}$ of degree $\deg_{\langle \mathbb{C} \rangle}(v) = d$ in the hydrogen-suppressed chemical graph $\langle \mathbb{C} \rangle$.

6. $\mathrm{dcp}_i(\mathbb{C})$, $i = 8 + d, d \in [1, 4]$: the number $\mathrm{dg}^{\rm int}_d(\mathbb{C})$ of interior-vertices of interior-degree $\deg_{\mathbb{C}^{\rm int}}(v) = d$ in the interior $\mathbb{C}^{\rm int} = (V^{\rm int}(\mathbb{C}), E^{\rm int}(\mathbb{C}))$ of $\mathbb{C}$.

7. $\mathrm{dcp}_i(\mathbb{C})$, $i = 12 + m$, $m \in [2, 3]$: the number $\mathrm{bd}^{\rm int}_m(\mathbb{C})$ of interior-edges with bond multiplicity m in $\mathbb{C}$; i.e., $\mathrm{bd}^{\rm int}_m(\mathbb{C}) \triangleq \{e \in E^{\rm int}(\mathbb{C}) \mid \beta(e) = m\}$.

8. $\mathrm{dcp}_i(\mathbb{C})$, $i = 14 + [\mathtt{a}]^{\rm int}$, $\mathtt{a} \in \Lambda^{\rm int}(D_\pi)$: the frequency $\mathrm{na}^{\rm int}_{\mathtt{a}}(\mathbb{C}) = |V_{\mathtt{a}}(\mathbb{C}) \cap V^{\rm int}(\mathbb{C})|$ of chemical element $\mathtt{a}$ in the set $V^{\rm int}(\mathbb{C})$ of interior-vertices in $\mathbb{C}$.

9. $\mathrm{dcp}_i(\mathbb{C})$, $i = 14 + |\Lambda^{\rm int}(D_\pi)| + [\mathtt{a}]^{\rm ex}$, $\mathtt{a} \in \Lambda^{\rm ex}(D_\pi)$: the frequency $\mathrm{na}^{\rm ex}_{\mathtt{a}}(\mathbb{C}) = |V_{\mathtt{a}}(\mathbb{C}) \cap V^{\rm ex}(\mathbb{C})|$ of chemical element $\mathtt{a}$ in the set $V^{\rm ex}(\mathbb{C})$ of exterior-vertices in $\mathbb{C}$.

10. $\mathrm{dcp}_i(\mathbb{C})$, $i = 14 + |\Lambda^{\rm int}(D_\pi)| + |\Lambda^{\rm ex}(D_\pi)| + [\gamma]$, $\gamma \in \Gamma^{\rm int}(D_\pi)$: the frequency $\mathrm{ec}_\gamma(\mathbb{C})$ of edge-configuration γ in the set $E^{\rm int}(\mathbb{C})$ of interior-edges in $\mathbb{C}$.

11. $\mathrm{dcp}_i(\mathbb{C})$, $i = 14 + |\Lambda^{\rm int}(D_\pi)| + |\Lambda^{\rm ex}(D_\pi)| + |\Gamma^{\rm int}(D_\pi)| + [\psi]$, $\psi \in \mathcal{F}(D_\pi)$: the frequency $\mathrm{fc}_\psi(\mathbb{C})$ of fringe-configuration ψ in the set of ρ-fringe-trees in $\mathbb{C}$.

12. $\mathrm{dcp}_i(\mathbb{C})$, $i = 14 + |\Lambda^{\rm int}(D_\pi)| + |\Lambda^{\rm ex}(D_\pi)| + |\Gamma^{\rm int}(D_\pi)| + |\mathcal{F}(D_\pi)| + [\nu]$, $\nu \in \Gamma^{\rm lf}_{\rm ac}$: the frequency $\mathrm{ac}^{\rm lf}_\nu(\mathbb{C})$ of adjacency-configuration ν in the set of leaf-edges in $\langle \mathbb{C} \rangle$.

2 Specifying Target Chemical Graphs

Given a prediction function η and a target value $y^* \in \mathbb{R}$, we call a chemical graph $\mathbb{C}^*$ such that $\eta(x^*) = y^*$ for the feature vector $x^* = f(\mathbb{C}^*)$ a *target chemical graph*. This section presents a set of rules for specifying topological substructure of a target chemical graph in a flexible way in Stage 4.

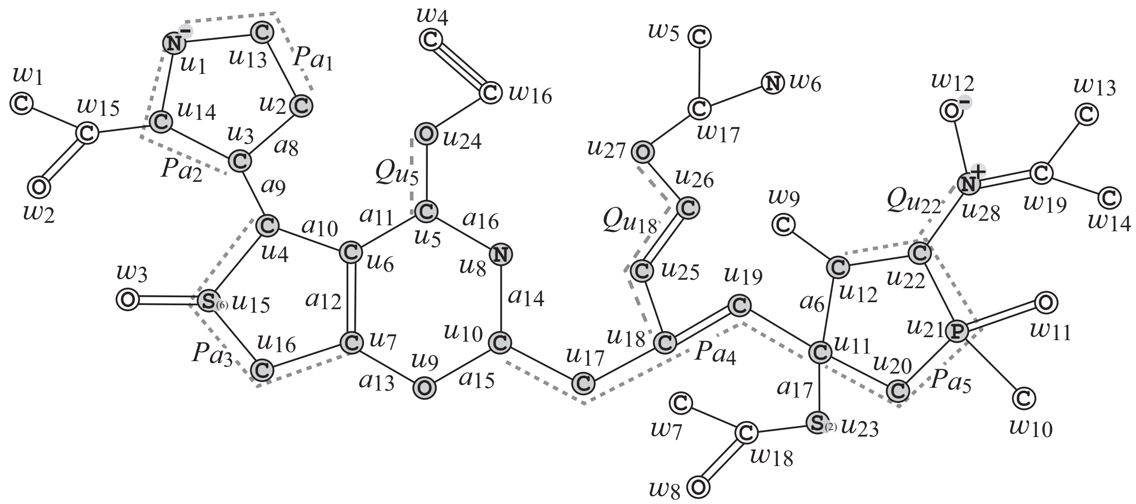


Figure 1: An illustration of a hydrogen-suppressed chemical graph $\langle\mathbb{C}\rangle$ obtained from a chemical graph $\mathbb{C}$ with $\mathrm{r}(\mathbb{C}) = 4$ by removing all the hydrogens, where for $\rho = 2$, $V^{\mathrm{ex}}(\mathbb{C}) = \{w_i \mid i \in [1, 19]\}$ and $V^{\mathrm{int}}(\mathbb{C}) = \{u_i \mid i \in [1, 28]\}$.

We first describe how to reduce a chemical graph $\mathbb{C} = (H, \alpha, \beta)$ into an abstract form based on which our specification rules will be defined. To illustrate the reduction process, we use the chemical graph $\mathbb{C} = (H, \alpha, \beta)$ such that $\langle\mathbb{C}\rangle$ is given in Figure 1.

R1 **Removal of all ρ-fringe-trees:** The interior $H^{\mathrm{int}} = (V^{\mathrm{int}}(\mathbb{C}), E^{\mathrm{int}}(\mathbb{C}))$ of $\mathbb{C}$ is obtained by removing the non-root vertices of each ρ-fringe-trees $\mathbb{C}[u] \in \mathcal{T}(\mathbb{C}), u \in V^{\mathrm{int}}(\mathbb{C})$. Figure 2 illustrates the interior H^{int} of chemical graph $\mathbb{C}$ with $\rho = 2$ in Figure 1.

R2 **Removal of some leaf paths:** We call a u, v-path Q in H^{int} a *leaf path* if vertex v is a leaf-vertex of H^{int} and the degree of each internal vertex of Q in H^{int} is 2, where we regard that Q is rooted at vertex u. A connected subgraph S of the interior H^{int} of $\mathbb{C}$ is called a *cyclical-base* if S is obtained from H by removing the vertices in $V(Q_u) \setminus \{u\}, u \in X$ for a subset X of interior-vertices and a set $\{Q_u \mid u \in X\}$ of leaf u, v-paths Q_u such that no two paths Q_u and $Q_{u'}$ share a vertex. Figure 3(a) illustrates a cyclical-base $S = H^{\mathrm{int}} - \bigcup_{u \in X}(V(Q_u) \setminus \{u\})$ of the interior H^{int} for a set $\{Q_{u_5} = (u_5, u_{24}), Q_{u_{18}} = (u_{18}, u_{25}, u_{26}, u_{27}), Q_{u_{22}} = (u_{22}, u_{28})\}$ of leaf paths in Figure 2.

R3 **Contraction of some pure paths:** A path in S is called *pure* if each internal vertex of the path is of degree 2. Choose a set $\mathcal{P}$ of several pure paths in S so that no two paths share vertices except for their end-vertices. A graph S' is called a *contraction* of a graph S (with respect to $\mathcal{P}$) if S' is obtained from S by replacing each pure u, v-path with a single edge $a = uv$, where S' may contain multiple edges between the same pair of adjacent vertices. Figure 3(b) illustrates a contraction S' obtained from the chemical graph S by contracting each uv-path $P_a \in \mathcal{P}$ into a new edge $a = uv$, where $a_1 = u_1u_2, a_2 = u_1u_3, a_3 = u_4u_7, a_4 = u_{10}u_{11}$ and $a_5 = u_{11}u_{12}$ and $\mathcal{P} = \{P_{a_1} = (u_1, u_{13}, u_2), P_{a_2} = (u_1, u_{14}, u_3), P_{a_3} = (u_4, u_{15}, u_{16}, u_7), P_{a_4} = (u_{10}, u_{17}, u_{18}, u_{19}, u_{11}), P_{a_5} = (u_{11}, u_{20}, u_{21}, u_{22}, u_{12})\}$ of pure paths in Figure 3(a).

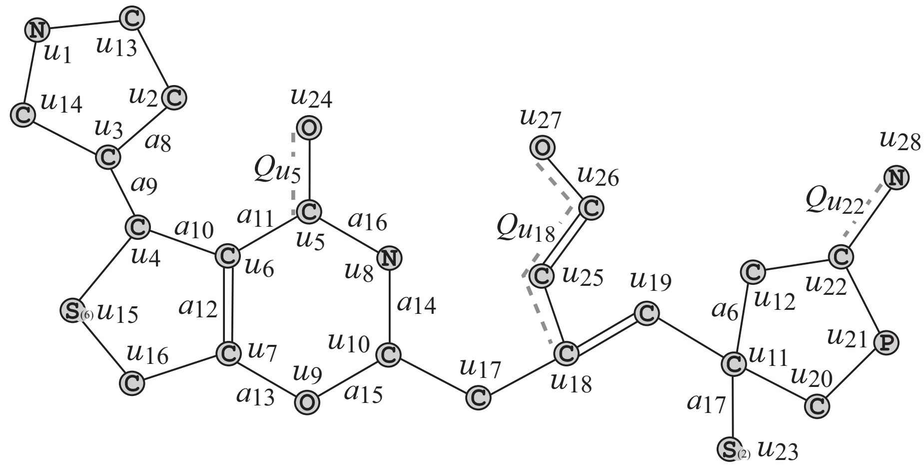
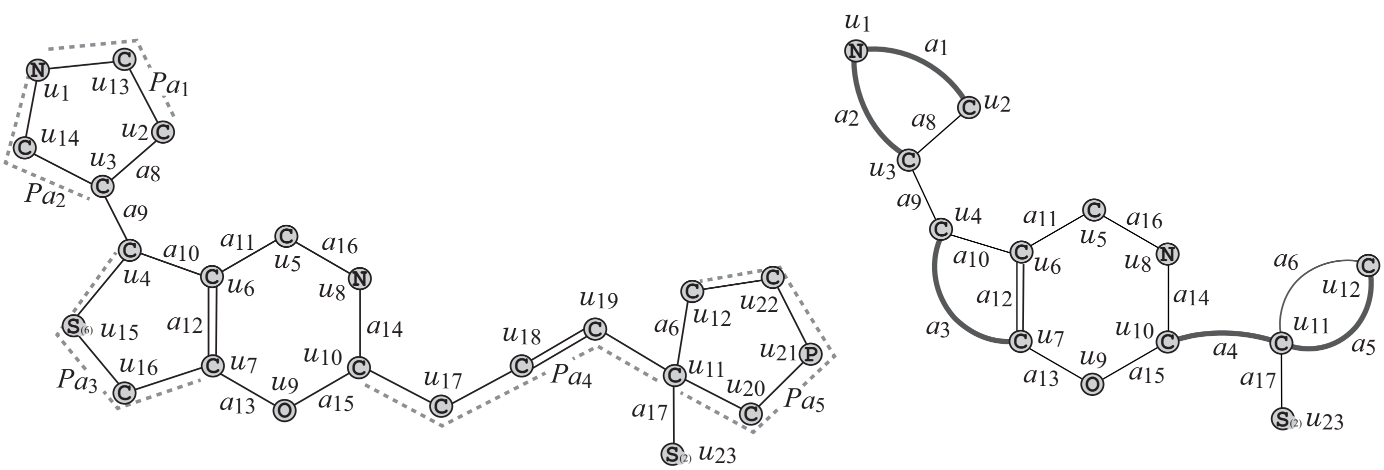


Figure 2: The interior $H^{\rm int}$ of chemical graph $\mathbb{C}$ with $\langle\mathbb{C}\rangle$ in Figure 1 for $\rho = 2$.



(a) A cyclical-base S of $H^{\rm int}$

(b) A contraction S' of S

Figure 3: (a) A cyclical-base $S = H^{\rm int} - \bigcup_{u \in \{u_5, u_{18}, u_{22}\}} (V(Q_u) \setminus \{u\})$ of the interior $H^{\rm int}$ in Figure 2; (b) A contraction S' of S for a pure path set $\mathcal{P} = \{P_{a_1}, P_{a_2}, \ldots, P_{a_5}\}$ in (a), where a new edge obtained by contracting a pure path is depicted with a thick line.

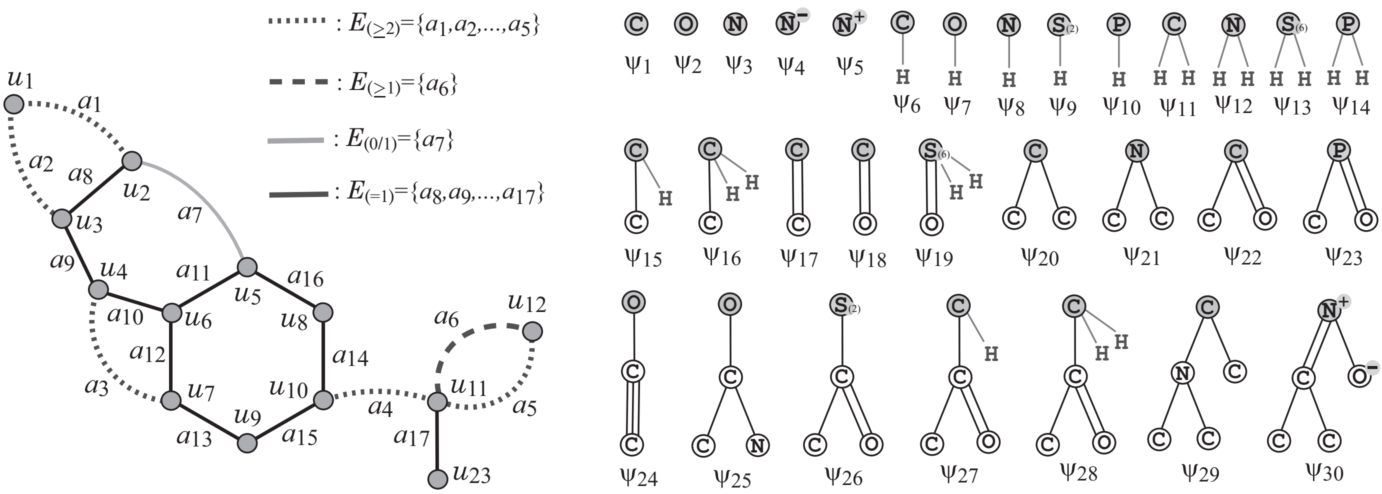
We will define a set of rules so that a chemical graph can be obtained from a graph (called a seed graph in the next section) by applying processes R3 to R1 in a reverse way. We specify topological substructures of a target chemical graph with a tuple $(G_{\rm C}, \sigma_{\rm int}, \sigma_{\rm ce})$ called a *target specification* defined under the set of the following rules.

Seed Graph

A *seed graph* $G_{\rm C} = (V_{\rm C}, E_{\rm C})$ is defined to be a graph (possibly with multiple edges) such that the edge set $E_{\rm C}$ consists of four sets $E_{(\geq 2)}$, $E_{(\geq 1)}$, $E_{(0/1)}$ and $E_{(=1)}$, where each of them can be empty. A seed graph plays a role of the most abstract form S' in R3. Figure 4(a) illustrates an example of a seed graph $G_{\rm C}$ with $\mathrm{r}(G_{\rm C}) = 5$, where $V_{\rm C} = \{u_1, u_2, \ldots, u_{12}, u_{23}\}$, $E_{(\geq 2)} = \{a_1, a_2, \ldots, a_5\}$, $E_{(\geq 1)} = \{a_6\}$, $E_{(0/1)} = \{a_7\}$ and $E_{(=1)} = \{a_8, a_9, \ldots, a_{16}\}$.

A *subdivision* S of $G_{\rm C}$ is a graph constructed from a seed graph $G_{\rm C}$ according to the following rules:

- Each edge $e = uv \in E_{(\geq 2)}$ is replaced with a u,v-path P_e of length at least 2;



(a) A seed graph $G_\mathrm{C}=(V_\mathrm{C},E_\mathrm{C})$ (b) A set $\mathcal{F}$ of chemical rooted trees

Figure 4: (a) An illustration of a seed graph G_C with $\mathrm{r}(G_\mathrm{C})=5$ where the vertices in V_C are depicted with gray circles, the edges in $E_{(\geq 2)}$ are depicted with dotted lines, the edges in $E_{(\geq 1)}$ are depicted with dashed lines, the edges in $E_{(0/1)}$ are depicted with gray bold lines and the edges in $E_{(=1)}$ are depicted with black solid lines; (b) A set $\mathcal{F}=\{\psi_1,\psi_2,\ldots,\psi_{30}\}\subseteq\mathcal{F}(D_\pi)$ of 30 chemical rooted trees $\psi_i, i\in[1,30]$, where the root of each tree is depicted with a gray circle, where the hydrogens attached to non-root vertices are omitted in the figure.

- Each edge $e=uv\in E_{(\geq 1)}$ is replaced with a u,v-path P_e of length at least 1 (equivalently e is directly used or replaced with a u,v-path P_e of length at least 2);
- Each edge $e\in E_{(0/1)}$ is either used or discarded, where $E_{(0/1)}$ is required to be chosen as a non-separating edge subset of $E(G_\mathrm{C})$ since otherwise the connectivity of a final chemical graph $\mathbb{C}$ is not guaranteed; $\mathrm{r}(\mathbb{C})=\mathrm{r}(G_\mathrm{C})-|E'|$ holds for a subset $E'\subseteq E_{(0/1)}$ of edges discarded in a final chemical graph $\mathbb{C}$; and
- Each edge $e\in E_{(=1)}$ is always used directly.

We allow a possible elimination of edges in $E_{(0/1)}$ as an optional rule in constructing a target chemical graph from a seed graph, even though such an operation has not been included in the process R3. A subdivision S plays a role of a cyclical-base in R2. A target chemical graph $\mathbb{C}=(H,\alpha,\beta)$ will contain S as a subgraph of the interior H^int of $\mathbb{C}$.

Interior-specification

A graph H^* that serves as the interior H^int of a target chemical graph $\mathbb{C}$ will be constructed as follows. First construct a subdivision S of a seed graph G_C by replacing each edge $e=uu'\in E_{(\geq 2)}\cup E_{(\geq 1)}$ with a pure u,u'-path P_e. Next construct a supergraph H^* of S by attaching a leaf path Q_v at each vertex $v\in V_\mathrm{C}$ or at an internal vertex $v\in V(P_e)\setminus\{u,u'\}$ of each pure u,u'-path P_e for some edge $e=uu'\in E_{(\geq 2)}\cup E_{(\geq 1)}$, where possibly $Q_v=(v), E(Q_v)=\emptyset$ (i.e., we do not attach any new edges to v). We introduce the following rules for specifying the size of H^*, the length $|E(P_e)|$ of a pure path P_e, the length $|E(Q_v)|$ of a leaf path Q_v, the number of leaf paths Q_v and a bond-multiplicity of each interior-edge, where we call the set of prescribed constants an *interior-specification* σ_int:

- Lower and upper bounds $\mathrm{n}^\mathrm{int}_\mathrm{LB}, \mathrm{n}^\mathrm{int}_\mathrm{UB}\in\mathbb{Z}_+$ on the number of interior-vertices of a target chemical graph $\mathbb{C}$.

- For each edge $e = uu' \in E_{(\geq 2)} \cup E_{(\geq 1)}$,

  a lower bound $\ell_{\rm LB}(e)$ and an upper bound $\ell_{\rm UB}(e)$ on the length $|E(P_e)|$ of a pure u, u'-path P_e. (For a notational convenience, set $\ell_{\rm LB}(e) := 0$, $\ell_{\rm UB}(e) := 1$, $e \in E_{(0/1)}$ and $\ell_{\rm LB}(e) := 1$, $\ell_{\rm UB}(e) := 1$, $e \in E_{(=1)}$.)

  a lower bound $\mathrm{bl}_{\rm LB}(e)$ and an upper bound $\mathrm{bl}_{\rm UB}(e)$ on the number of leaf paths Q_v attached at internal vertices v of a pure u, u'-path P_e.

  a lower bound $\mathrm{ch}_{\rm LB}(e)$ and an upper bound $\mathrm{ch}_{\rm UB}(e)$ on the maximum length $|E(Q_v)|$ of a leaf path Q_v attached at an internal vertex $v \in V(P_e) \setminus \{u, u'\}$ of a pure u, u'-path P_e.

- For each vertex $v \in V_{\rm C}$,

  a lower bound $\mathrm{ch}_{\rm LB}(v)$ and an upper bound $\mathrm{ch}_{\rm UB}(v)$ on the number of leaf paths Q_v attached to v, where $0 \leq \mathrm{ch}_{\rm LB}(v) \leq \mathrm{ch}_{\rm UB}(v) \leq 1$.

  a lower bound $\mathrm{ch}_{\rm LB}(v)$ and an upper bound $\mathrm{ch}_{\rm UB}(v)$ on the length $|E(Q_v)|$ of a leaf path Q_v attached to v.

- For each edge $e = uu' \in E_{\rm C}$, a lower bound $\mathrm{bd}_{m,\rm LB}(e)$ and an upper bound $\mathrm{bd}_{m,\rm UB}(e)$ on the number of edges with bond-multiplicity $m \in [2, 3]$ in u, u'-path P_e, where we regard P_e, $e \in E_{(0/1)} \cup E_{(=1)}$ as single edge e.

We call a graph H^* that satisfies an interior-specification $\sigma_{\rm int}$ a $\sigma_{\rm int}$-*extension of* $G_{\rm C}$, where the bond-multiplicity of each edge has been determined.

Table 1 shows an example of an interior-specification $\sigma_{\rm int}$ to the seed graph $G_{\rm C}$ in Figure 4.

Table 1: Example 1 of an interior-specification $\sigma_{\rm int}$.

$\mathrm{n}^{\rm int}_{\rm LB} = 20$	$\mathrm{n}^{\rm int}_{\rm UB} = 28$

	a_1	a_2	a_3	a_4	a_5	a_6
$\ell_{\rm LB}(a_i)$	2	2	2	3	2	1
$\ell_{\rm UB}(a_i)$	3	4	3	5	4	4
$\mathrm{bl}_{\rm LB}(a_i)$	0	0	0	1	1	0
$\mathrm{bl}_{\rm UB}(a_i)$	1	1	0	2	1	0
$\mathrm{ch}_{\rm LB}(a_i)$	0	1	0	4	3	0
$\mathrm{ch}_{\rm UB}(a_i)$	3	3	1	6	5	2

	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}	u_{11}	u_{12}	u_{23}
$\mathrm{bl}_{\rm LB}(u_i)$	0	0	0	0	0	0	0	0	0	0	0	0	0
$\mathrm{bl}_{\rm UB}(u_i)$	1	1	1	1	1	0	0	0	0	0	0	0	0
$\mathrm{ch}_{\rm LB}(u_i)$	0	0	0	0	1	0	0	0	0	0	0	0	0
$\mathrm{ch}_{\rm UB}(u_i)$	1	0	0	0	3	0	1	1	0	1	2	4	1

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}
$\mathrm{bd}_{2,\rm LB}(a_i)$	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0
$\mathrm{bd}_{2,\rm UB}(a_i)$	1	1	0	2	2	0	0	0	0	0	0	1	0	0	0	0	0
$\mathrm{bd}_{3,\rm LB}(a_i)$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$\mathrm{bd}_{3,\rm UB}(a_i)$	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0

Figure 5 illustrates an example of an $\sigma_{\rm int}$-extension H^* of seed graph $G_{\rm C}$ in Figure 4 under the interior-specification $\sigma_{\rm int}$ in Table 1.

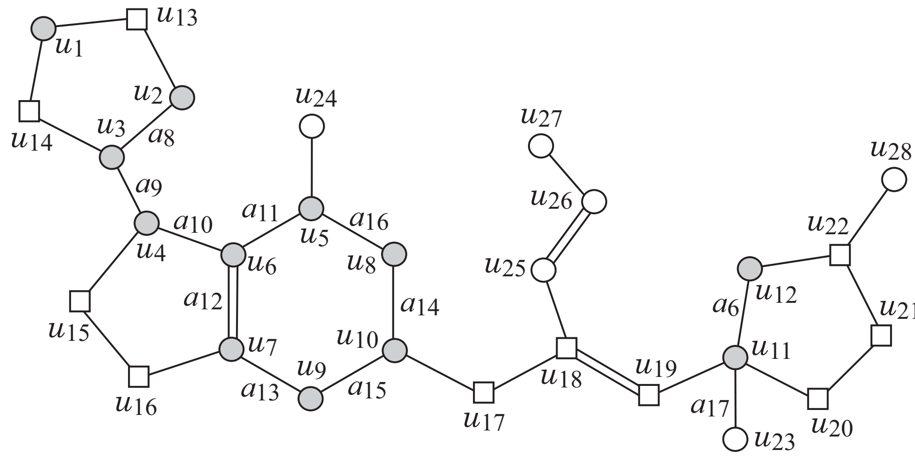


Figure 5: An illustration of a graph H^* that is obtained from the seed graph G_{C} in Figure 4 under the interior-specification σ_{int} in Table 1, where the vertices newly introduced by pure paths P_{a_i} (resp., by leaf paths Q_{v_i}) are $u_{13}, u_{14}, \ldots, u_{22}$ depicted with white squares (resp., $u_{23}, u_{24}, \ldots, u_{28}$ depicted with white circles).

Chemical-specification

Let H^* be a graph that serves as the interior H^{int} of a target chemical graph $\mathbb{C}$, where the bond-multiplicity of each edge in H^* has be determined. Finally we introduce a set of rules for constructing a target chemical graph $\mathbb{C}$ from H^* by choosing a chemical element $\mathtt{a} \in \Lambda$ and assigning a ρ-fringe-tree ψ to each interior-vertex $v \in V^{\mathrm{int}}$. We introduce the following rules for specifying the size of $\mathbb{C}$, a set of chemical rooted trees that are allowed to use as ρ-fringe-trees and lower and upper bounds on the frequency of a chemical element, a chemical symbol, and an edge-configuration, where we call the set of prescribed constants a *chemical specification* σ_{ce}:

- Lower and upper bounds $n_{\mathrm{LB}}, n^* \in \mathbb{Z}_+$ on the number of vertices, where $\mathrm{n}^{\mathrm{int}}_{\mathrm{LB}} \leq n_{\mathrm{LB}} \leq n^*$.

- Subsets $\mathcal{F}(v) \subseteq \mathcal{F}(D_\pi), v \in V_{\mathrm{C}}$ and $\mathcal{F}_E \subseteq \mathcal{F}(D_\pi)$ of chemical rooted trees ψ with $\mathrm{ht}(\langle\psi\rangle) \leq \rho$, where we require that every ρ-fringe-tree $\mathbb{C}[v]$ rooted at a vertex $v \in V_{\mathrm{C}}$ (resp., at an internal vertex v not in V_{C}) in $\mathbb{C}$ belongs to $\mathcal{F}(v)$ (resp., $\mathcal{F}_E$). Let $\mathcal{F}^* := \mathcal{F}_E \cup \bigcup_{v \in V_{\mathrm{C}}} \mathcal{F}(v)$ and Λ^{ex} denote the set of chemical elements assigned to non-root vertices over all chemical rooted trees in $\mathcal{F}^*$.

- A subset $\Lambda^{\mathrm{int}} \subseteq \Lambda^{\mathrm{int}}(D_\pi)$, where we require that every chemical element $\alpha(v)$ assigned to an interior-vertex v in $\mathbb{C}$ belongs to Λ^{int}. Let $\Lambda := \Lambda^{\mathrm{int}} \cup \Lambda^{\mathrm{ex}}$ and $\mathrm{na}_{\mathtt{a}}(\mathbb{C})$ (resp., $\mathrm{na}^{\mathrm{int}}_{\mathtt{a}}(\mathbb{C})$ and $\mathrm{na}^{\mathrm{ex}}_{\mathtt{a}}(\mathbb{C})$) denote the number of vertices (resp., interior-vertices and exterior-vertices) v such that $\alpha(v) = \mathtt{a}$ in $\mathbb{C}$.

- A set $\Lambda^{\mathrm{int}}_{\mathrm{dg}} \subseteq \Lambda \times [1, 4]$ of chemical symbols and a set $\Gamma^{\mathrm{int}} \subseteq \Gamma^{\mathrm{int}}(D_\pi)$ of edge-configurations (μ, μ', m) with $\mu \leq \mu'$, where we require that the edge-configuration $\mathrm{ec}(e)$ of an interior-edge e in $\mathbb{C}$ belongs to Γ^{int}. We do not distinguish (μ, μ', m) and (μ', μ, m).

- Define $\Gamma^{\mathrm{int}}_{\mathrm{ac}}$ to be the set of adjacency-configurations such that $\Gamma^{\mathrm{int}}_{\mathrm{ac}} := \{(\mathtt{a}, \mathtt{b}, m) \mid (\mathtt{a}d, \mathtt{b}d', m) \in \Gamma^{\mathrm{int}}\}$. Let $\mathrm{ac}^{\mathrm{int}}_{\nu}(\mathbb{C}), \nu \in \Gamma^{\mathrm{int}}_{\mathrm{ac}}$ denote the number of interior-edges e such that $\mathrm{ac}(e) = \nu$ in $\mathbb{C}$.

- Subsets $\Lambda^*(v) \subseteq \{\mathtt{a} \in \Lambda^{\mathrm{int}} \mid \mathrm{val}(\mathtt{a}) \geq 2\}$, $v \in V_{\mathrm{C}}$, we require that every chemical element $\alpha(v)$ assigned to a vertex $v \in V_{\mathrm{C}}$ in the seed graph belongs to $\Lambda^*(v)$.

- Lower and upper bound functions $\mathrm{na}_{\mathrm{LB}}, \mathrm{na}_{\mathrm{UB}} : \Lambda \to [1, n^*]$ and $\mathrm{na}^{\mathrm{int}}_{\mathrm{LB}}, \mathrm{na}^{\mathrm{int}}_{\mathrm{UB}} : \Lambda^{\mathrm{int}} \to [1, n^*]$ on the number of interior-vertices v such that $\alpha(v) = \mathtt{a}$ in $\mathbb{C}$.

- Lower and upper bound functions $\mathrm{ns}^{\mathrm{int}}_{\mathrm{LB}}, \mathrm{ns}^{\mathrm{int}}_{\mathrm{UB}} : \Lambda^{\mathrm{int}}_{\mathrm{dg}} \to [1, n^*]$ on the number of interior-vertices v such that $\mathrm{cs}(v) = \mu$ in $\mathbb{C}$.

- Lower and upper bound functions $\mathrm{ac}^{\mathrm{int}}_{\mathrm{LB}}, \mathrm{ac}^{\mathrm{int}}_{\mathrm{UB}} : \Gamma^{\mathrm{int}}_{\mathrm{ac}} \to \mathbb{Z}_+$ on the number of interior-edges e such that $\mathrm{ac}(e) = \nu$ in $\mathbb{C}$.

- Lower and upper bound functions $\mathrm{ec}^{\mathrm{int}}_{\mathrm{LB}}, \mathrm{ec}^{\mathrm{int}}_{\mathrm{UB}} : \Gamma^{\mathrm{int}} \to \mathbb{Z}_+$ on the number of interior-edges e such that $\mathrm{ec}(e) = \gamma$ in $\mathbb{C}$.

- Lower and upper bound functions $\mathrm{fc}_{\mathrm{LB}}, \mathrm{fc}_{\mathrm{UB}} : \mathcal{F}^* \to [0, n^*]$ on the number of interior-vertices v such that $\mathbb{C}[v]$ is r-isomorphic to $\psi \in \mathcal{F}^*$ in $\mathbb{C}$.

- Lower and upper bound functions $\mathrm{ac}^{\mathrm{lf}}_{\mathrm{LB}}, \mathrm{ac}^{\mathrm{lf}}_{\mathrm{UB}} : \Gamma^{\mathrm{lf}}_{\mathrm{ac}} \to [0, n^*]$ on the number of leaf-edges uv in ac_{C} with adjacency-configuration ν.

We call a chemical graph $\mathbb{C}$ that satisfies a chemical specification σ_{ce} a $(\sigma_{\mathrm{int}}, \sigma_{\mathrm{ce}})$-*extension of* G_{C}, and denote by $\mathcal{G}(G_{\mathrm{C}}, \sigma_{\mathrm{int}}, \sigma_{\mathrm{ce}})$ the set of all $(\sigma_{\mathrm{int}}, \sigma_{\mathrm{ce}})$-extensions of G_{C}.

Table 2 shows an example of a chemical-specification σ_{ce} to the seed graph G_{C} in Figure 4.

Figure 1 illustrates an example $\mathbb{C}$ of a $(\sigma_{\mathrm{int}}, \sigma_{\mathrm{ce}})$-extension of G_{C} obtained from the σ_{int}-extension H^* in Figure 5 under the chemical-specification σ_{ce} in Table 2. Note that $\mathrm{r}(\mathbb{C}) = \mathrm{r}(H^*) = \mathrm{r}(G_{\mathrm{C}}) - 1 = 4$ holds since the edge in $E_{(0/1)}$ is discarded in H^*.

3 Test Instances for Stages 4 and 5

We prepared the following instances (a)-(d) for conducting experiments of Stages 4 and 5 in Phase 2.

In Stages 4 and 5, we use five properties $\pi \in \{$Hc, Vd, OptR, IhcLiq, Vis$\}$ and define a set $\Lambda(\pi)$ of chemical elements as follows:
$\Lambda(\mathrm{Hc}) = \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}, \mathtt{S}_{(2)}, \mathtt{S}_{(6)}, \mathtt{Cl}\}$, $\Lambda(\mathrm{Vd}) = \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}, \mathtt{N}, \mathtt{Cl}, \mathtt{P}_{(3)}, \mathtt{P}_{(5)}\}$,
$\Lambda(\mathrm{OptR}) = \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}, \mathtt{S}_{(2)}, \mathtt{F}\}$, $\Lambda(\mathrm{IhcLiq}) = \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}, \mathtt{S}_{(2)}, \mathtt{S}_{(6)}, \mathtt{Cl}\}$ and
$\Lambda(\mathrm{Vis}) = \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{Si}\}$.

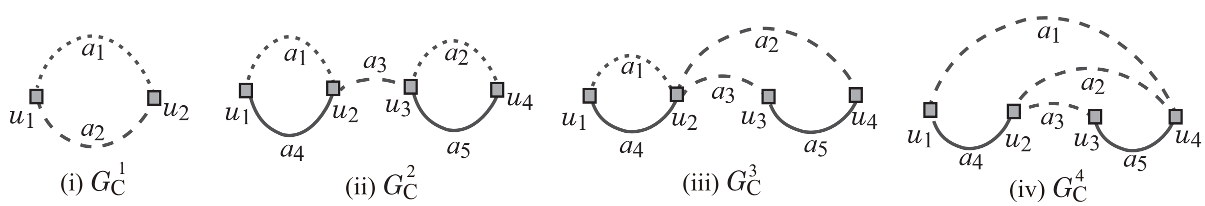


Figure 6: (i) Seed graph G^1_{C} for I^1_{b} and I_{d}; (ii) Seed graph G^2_{C} for I^2_{b}; (iii) Seed graph G^3_{C} for I^3_{b}; (iv) Seed graph G^4_{C} for I^4_{b}.

(a) $I_{\mathrm{a}} = (G_{\mathrm{C}}, \sigma_{\mathrm{int}}, \sigma_{\mathrm{ce}})$: The instance introduced in Appendix 2 to explain the target specification. For each property π, we replace $\Lambda = \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}, \mathtt{S}_{(2)}, \mathtt{S}_{(6)}, \mathtt{P}_{(5)}\}$ in Table 2 with $\Lambda(\pi) \cap \{\mathtt{S}_{(2)}, \mathtt{S}_{(6)}, \mathtt{P}_{(5)}\}$ and remove from the σ_{ce} all chemical symbols, edge-configurations and fringe-configurations that cannot be constructed from the replaced element set (i.e., those containing a chemical element in $\{\mathtt{S}_{(2)}, \mathtt{S}_{(6)}, \mathtt{P}_{(5)}\} \setminus \Lambda(\pi)$).

(b) $I^i_{\mathrm{b}} = (G^i_{\mathrm{C}}, \sigma^i_{\mathrm{int}}, \sigma^i_{\mathrm{ce}})$, $i = 1, 2, 3, 4$: An instance for inferring chemical graphs with rank at most 2. In the four instances I^i_{b}, $i = 1, 2, 3, 4$, the following specifications in $(\sigma_{\mathrm{int}}, \sigma_{\mathrm{ce}})$ are common.

Table 2: Example 2 of a chemical-specification $\sigma_{\rm ce}$.

$n_{\rm LB} = 30$, $n^* = 50$.
branch-parameter: $\rho = 2$
Each of sets $\mathcal{F}(v), v \in V_{\rm C}$ and $\mathcal{F}_E$ is set to be the set $\mathcal{F}$ of chemical rooted trees ψ with ${\rm ht}(\langle\psi\rangle) \leq \rho = 2$ in Figure 4(b).

$\Lambda = \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}, \mathtt{S}_{(2)}, \mathtt{S}_{(6)}, \mathtt{P} = \mathtt{P}_{(5)}\}$	$\Lambda^{\rm int}_{\rm dg} = \{\mathtt{C}2, \mathtt{C}3, \mathtt{C}4, \mathtt{N}2, \mathtt{N}3, \mathtt{O}2, \mathtt{S}_{(2)}2, \mathtt{S}_{(6)}3, \mathtt{P}4\}$

$\Gamma^{\rm int}_{\rm ac}$	$\nu_1 = (\mathtt{C}, \mathtt{C}, 1), \nu_2 = (\mathtt{C}, \mathtt{C}, 2), \nu_3 = (\mathtt{C}, \mathtt{N}, 1), \nu_4 = (\mathtt{C}, \mathtt{O}, 1), \nu_5 = (\mathtt{C}, \mathtt{S}_{(2)}, 1), \nu_6 = (\mathtt{C}, \mathtt{S}_{(6)}, 1), \nu_7 = (\mathtt{C}, \mathtt{P}, 1)$
$\Gamma^{\rm int}$	$\gamma_1 = (\mathtt{C}2, \mathtt{C}2, 1), \gamma_2 = (\mathtt{C}2, \mathtt{C}3, 1), \gamma_3 = (\mathtt{C}2, \mathtt{C}3, 2), \gamma_4 = (\mathtt{C}2, \mathtt{C}4, 1), \gamma_5 = (\mathtt{C}3, \mathtt{C}3, 1), \gamma_6 = (\mathtt{C}3, \mathtt{C}3, 2),$ $\gamma_7 = (\mathtt{C}3, \mathtt{C}4, 1), \gamma_8 = (\mathtt{C}2, \mathtt{N}2, 1), \gamma_9 = (\mathtt{C}3, \mathtt{N}2, 1), \gamma_{10} = (\mathtt{C}3, \mathtt{O}2, 1), \gamma_{11} = (\mathtt{C}2, \mathtt{C}2, 2), \gamma_{12} = (\mathtt{C}2, \mathtt{O}2, 1),$ $\gamma_{13} = (\mathtt{C}3, \mathtt{N}3, 1), \gamma_{14} = (\mathtt{C}4, \mathtt{S}_{(2)}2, 2), \gamma_{15} = (\mathtt{C}2, \mathtt{S}_{(6)}3, 1), \gamma_{16} = (\mathtt{C}3, \mathtt{S}_{(6)}3, 1), \gamma_{17} = (\mathtt{C}2, \mathtt{P}4, 2),$ $\gamma_{18} = (\mathtt{C}3, \mathtt{P}4, 1)$

$\Lambda^*(u_1) = \Lambda^*(u_8) = \{\mathtt{C}, \mathtt{N}\}$, $\Lambda^*(u_9) = \{\mathtt{C}, \mathtt{O}\}$, $\Lambda^*(u) = \{\mathtt{C}\}$, $u \in V_{\rm C} \setminus \{u_1, u_8, u_9\}$

	H	C	N	O	$\mathtt{S}_{(2)}$	$\mathtt{S}_{(6)}$	P
${\rm na}_{\rm LB}(\mathtt{a})$	40	27	1	1	0	0	0
${\rm na}_{\rm UB}(\mathtt{a})$	65	37	4	8	1	1	1

	C	N	O	$\mathtt{S}_{(2)}$	$\mathtt{S}_{(6)}$	P
${\rm na}^{\rm int}_{\rm LB}(\mathtt{a})$	9	1	0	0	0	0
${\rm na}^{\rm int}_{\rm UB}(\mathtt{a})$	23	4	5	1	1	1

	C2	C3	C4	N2	N3	O2	$\mathtt{S}_{(2)}2$	$\mathtt{S}_{(6)}3$	P4
${\rm ns}^{\rm int}_{\rm LB}(\mu)$	3	5	0	0	0	0	0	0	0
${\rm ns}^{\rm int}_{\rm UB}(\mu)$	8	15	2	2	3	5	1	1	1

	ν_1	ν_2	ν_3	ν_4	ν_5	ν_6	ν_7
${\rm ac}^{\rm int}_{\rm LB}(\nu)$	0	0	0	0	0	0	0
${\rm ac}^{\rm int}_{\rm UB}(\nu)$	30	10	10	10	1	1	1

	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7	γ_8	γ_9	γ_{10}	γ_{11}	γ_{12}	γ_{13}	γ_{14}	γ_{15}	γ_{16}	γ_{17}	γ_{18}
${\rm ec}^{\rm int}_{\rm LB}(\gamma)$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
${\rm ec}^{\rm int}_{\rm UB}(\gamma)$	4	15	4	4	10	5	4	4	6	4	4	4	2	2	2	2	2	2

	$\psi \in \{\psi_i \mid i = 1, 6, 11\}$	$\psi \in \mathcal{F}^* \setminus \{\psi_i \mid i = 1, 6, 11\}$
${\rm fc}_{\rm LB}(\psi)$	1	0
${\rm fc}_{\rm UB}(\psi)$	10	3

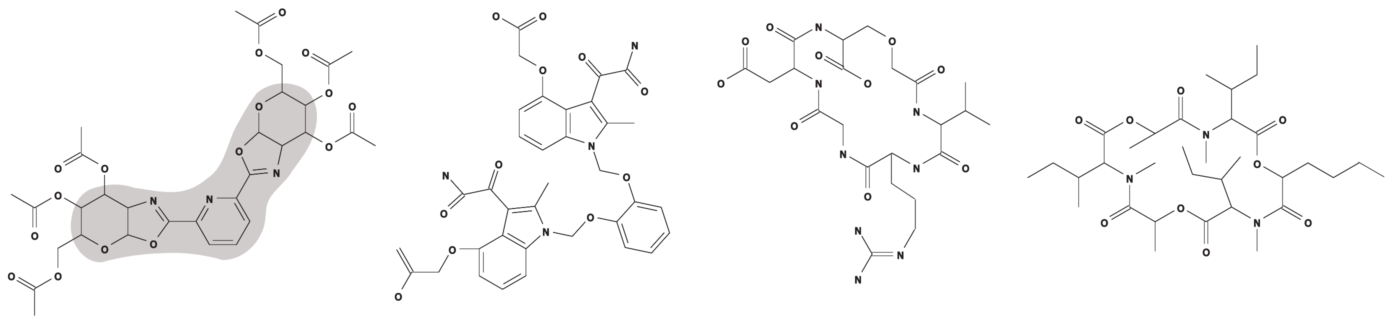
	$\nu \in \{(\mathtt{C}, \mathtt{C}, 1), (\mathtt{C}, \mathtt{C}, 2)\}$	$\nu \in \Gamma^{\rm lf}_{\rm ac} \setminus \{(\mathtt{C}, \mathtt{C}, 1), (\mathtt{C}, \mathtt{C}, 2)\}$
${\rm ac}^{\rm lf}_{\rm LB}(\nu)$	0	0
${\rm ac}^{\rm lf}_{\rm UB}(\nu)$	10	8

Set $\Lambda := \Lambda(\pi)$ for a given property $\pi \in \{$Hc, Vd, OptR, IhcLiq, Vis$\}$, set $\Lambda^{\rm int}_{\rm dg}$ to be the set of all possible symbols in $\Lambda \times [1, 4]$ that appear in the data set D_π and set $\Gamma^{\rm int}$ to be the set of all edge-configurations that appear in the data set D_π. Set $\Lambda^*(v) := \Lambda$, $v \in V_{\rm C}$.

The lower bounds $\ell_{\rm LB}$, ${\rm bl}_{\rm LB}$, ${\rm ch}_{\rm LB}$, ${\rm bd}_{2,{\rm LB}}$, ${\rm bd}_{3,{\rm LB}}$, ${\rm na}_{\rm LB}$, ${\rm na}^{\rm int}_{\rm LB}$, ${\rm ns}^{\rm int}_{\rm LB}$, ${\rm ac}^{\rm int}_{\rm LB}$, ${\rm ec}^{\rm int}_{\rm LB}$ and ${\rm ac}^{\rm lf}_{\rm LB}$ are all set to be 0.

The upper bounds $\ell_{\rm UB}$, ${\rm bl}_{\rm UB}$, ${\rm ch}_{\rm UB}$, ${\rm bd}_{2,{\rm UB}}$, ${\rm bd}_{3,{\rm UB}}$, ${\rm na}_{\rm UB}$, ${\rm na}^{\rm int}_{\rm UB}$, ${\rm ns}^{\rm int}_{\rm UB}$, ${\rm ac}^{\rm int}_{\rm UB}$, ${\rm ec}^{\rm int}_{\rm UB}$ and ${\rm ac}^{\rm lf}_{\rm UB}$ are all set to be an upper bound n^* on $n(G^*)$.

For each property π, let $\mathcal{F}(D_\pi)$ denote the set of 2-fringe-trees in the compounds in D_π, and select a subset $\mathcal{F}^i_\pi \subseteq \mathcal{F}(D_\pi)$ with $|\mathcal{F}^i_\pi| = 45 - 5i$, $i \in [1, 5]$. For each instance $I^i_{\rm b}$, set $\mathcal{F}_E := \mathcal{F}(v) := \mathcal{F}^i_\pi$, $v \in V_{\rm C}$ and ${\rm fc}_{\rm LB}(\psi) := 0, {\rm fc}_{\rm UB}(\psi) := 10, \psi \in \mathcal{F}^i_\pi$.



(a) $\mathbb{C}_A$: CID 24822711 (b) $\mathbb{C}_B$: CID 59170444 (c) $\mathbb{C}_A$: CID 10076784 (d) $\mathbb{C}_B$: CID 44340250

Figure 7: An illustration of chemical compounds for instances $I_{\rm c}$ and $I_{\rm d}$: (a) $\mathbb{C}_A$: CID 24822711; (b) $\mathbb{C}_B$: CID 59170444; (c) $\mathbb{C}_A$: CID 10076784; (d) $\mathbb{C}_B$: CID 44340250, where hydrogens are omitted.

Instance $I_{\rm b}^1$ is given by the rank-1 seed graph $G_{\rm C}^1$ in Figure 6(i) and Instances $I_{\rm b}^i$, $i = 2, 3, 4$ are given by the rank-2 seed graph $G_{\rm C}^i$, $i = 2, 3, 4$ in Figure 6(ii)-(iv).

(i) For instance $I_{\rm b}^1$, select as a seed graph the monocyclic graph $G_{\rm C}^1 = (V_{\rm C}, E_{\rm C} = E_{(\geq 2)} \cup E_{(\geq 1)})$ in Figure 6(i), where $V_{\rm C} = \{u_1, u_2\}$, $E_{(\geq 2)} = \{a_1\}$ and $E_{(\geq 1)} = \{a_2\}$. Set $\mathrm{n}_{\rm LB}^{\rm int} := 5, \mathrm{n}_{\rm UB}^{\rm int} := 15, n_{\rm LB} := 35$ and $n^* := 38$. We include a linear constraint $\ell(a_1) \leq \ell(a_2)$ and $5 \leq \ell(a_1) + \ell(a_2) \leq 15$ as part of the side constraint.

(ii) For instance $I_{\rm b}^2$, select as a seed graph the graph $G_{\rm C}^2 = (V_{\rm C}, E_{\rm C} = E_{(\geq 2)} \cup E_{(\geq 1)} \cup E_{(=1)})$ in Figure 6(ii), where $V_{\rm C} = \{u_1, u_2, u_3, u_4\}$, $E_{(\geq 2)} = \{a_1, a_2\}$, $E_{(\geq 1)} = \{a_3\}$ and $E_{(=1)} = \{a_4, a_5\}$. Set $\mathrm{n}_{\rm LB}^{\rm int} := 25, \mathrm{n}_{\rm UB}^{\rm int} := 30, n_{\rm LB} := 45$ and $n^* := 50$. We include a linear constraint $\ell(a_1) \leq \ell(a_2)$ and $\ell(a_1) + \ell(a_2) + \ell(a_3) \leq 15$.

(iii) For instance $I_{\rm b}^3$, select as a seed graph the graph $G_{\rm C}^3 = (V_{\rm C}, E_{\rm C} = E_{(\geq 2)} \cup E_{(\geq 1)} \cup E_{(=1)})$ in Figure 6(iii), where $V_{\rm C} = \{u_1, u_2, u_3, u_4\}$, $E_{(\geq 2)} = \{a_1\}$, $E_{(\geq 1)} = \{a_2, a_3\}$ and $E_{(=1)} = \{a_4, a_5\}$. Set $\mathrm{n}_{\rm LB}^{\rm int} := 25, \mathrm{n}_{\rm UB}^{\rm int} := 30, n_{\rm LB} := 45$ and $n^* := 50$. We include linear constraints $\ell(a_1) \leq \ell(a_2) + \ell(a_3)$, $\ell(a_2) \leq \ell(a_3)$ and $\ell(a_1) + \ell(a_2) + \ell(a_3) \leq 15$.

(iv) For instance $I_{\rm b}^4$, select as a seed graph the graph $G_{\rm C}^4 = (V_{\rm C}, E_{\rm C} = E_{(\geq 2)} \cup E_{(\geq 1)} \cup E_{(=1)})$ in Figure 6(iv), where $V_{\rm C} = \{u_1, u_2, u_3, u_4\}$, $E_{(\geq 1)} = \{a_1, a_2, a_3\}$ and $E_{(=1)} = \{a_4, a_5\}$. Set $\mathrm{n}_{\rm LB}^{\rm int} := 25, \mathrm{n}_{\rm UB}^{\rm int} := 30, n_{\rm LB} := 45$ and $n^* := 50$. We include linear constraints $\ell(a_2) \leq \ell(a_1) + 1$, $\ell(a_2) \leq \ell(a_3) + 1$, $\ell(a_1) \leq \ell(a_3)$ and $\ell(a_1) + \ell(a_2) + \ell(a_3) \leq 15$.

We define instances in (c) and (d) in order to find chemical graphs that have an intermediate structure of given two chemical cyclic graphs $G_A = (H_A = (V_A, E_A), \alpha_A, \beta_A)$ and $G_B = (H_B = (V_B, E_B), \alpha_B, \beta_B)$. Let $\Lambda_A^{\rm int}$ and $\Lambda_{{\rm dg},A}^{\rm int}$ denote the sets of chemical elements and chemical symbols of the interior-vertices in G_A, $\Gamma_A^{\rm int}$ denote the sets of edge-configurations of the interior-edges in G_A, and $\mathcal{F}_A$ denote the set of 2-fringe-trees in G_A. Analogously define sets $\Lambda_B^{\rm int}$, $\Lambda_{{\rm dg},B}^{\rm int}$, $\Gamma_B^{\rm int}$ and $\mathcal{F}_B$ in G_B.

(c) $I_{\rm c} = (G_{\rm C}, \sigma_{\rm int}, \sigma_{\rm ce})$: An instance aimed to infer a chemical graph $G^\dagger$ such that the core of $G^\dagger$ is equal to the core of G_A and the frequency of each edge-configuration in the non-core of $G^\dagger$ is equal to that of G_B. We use chemical compounds CID 24822711 and CID 59170444 in Figure 7(a) and (b) for G_A and G_B, respectively.
Set a seed graph $G_{\rm C} = (V_{\rm C}, E_{\rm C} = E_{(=1)})$ to be the core of G_A.
Set $\Lambda := \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}\}$, and set $\Lambda_{\rm dg}^{\rm int}$ to be the set of all possible chemical symbols in $\Lambda \times [1, 4]$.

Set $\Gamma^{\rm int} := \Gamma^{\rm int}_A \cup \Gamma^{\rm int}_B$ and $\Lambda^*(v) := \{\alpha_A(v)\}$, $v \in V_{\rm C}$.
Set $\mathrm{n}^{\rm int}_{\rm LB} := \min\{\mathrm{n}^{\rm int}(G_A), \mathrm{n}^{\rm int}(G_B)\}$, $\mathrm{n}^{\rm int}_{\rm UB} := \max\{\mathrm{n}^{\rm int}(G_A), \mathrm{n}^{\rm int}(G_B)\}$,
$n_{\rm LB} := \min\{n(G_A), n(G_B)\} - 10$ and $n^* := \max\{n(G_A), n(G_B)\} + 5$.
Set lower bounds $\ell_{\rm LB}$, $\mathrm{bl}_{\rm LB}$, $\mathrm{ch}_{\rm LB}$, $\mathrm{bd}_{2,\rm LB}$, $\mathrm{bd}_{3,\rm LB}$, $\mathrm{na}_{\rm LB}$, $\mathrm{na}^{\rm int}_{\rm LB}$, $\mathrm{ns}^{\rm int}_{\rm LB}$, $\mathrm{ac}^{\rm int}_{\rm LB}$ and $\mathrm{ac}^{\rm lf}_{\rm LB}$ to be 0.
Set upper bounds $\ell_{\rm UB}$, $\mathrm{bl}_{\rm UB}$, $\mathrm{ch}_{\rm UB}$, $\mathrm{bd}_{2,\rm UB}$, $\mathrm{bd}_{3,\rm UB}$, $\mathrm{na}_{\rm UB}$, $\mathrm{na}^{\rm int}_{\rm UB}$, $\mathrm{ns}^{\rm int}_{\rm UB}$, $\mathrm{ac}^{\rm int}_{\rm UB}$ and $\mathrm{ac}^{\rm lf}_{\rm UB}$ to be n^*.
Set $\mathrm{ec}^{\rm int}_{\rm LB}(\gamma)$ to be the number of core-edges in G_A with $\gamma \in \Gamma^{\rm int}$ and $\mathrm{ec}^{\rm int}_{\rm UB}(\gamma)$ to be the number interior-edges in G_A and G_B with edge-configuration γ.
Let $\mathcal{F}^{(p)}_B, p \in [1, 2]$ denote the set of chemical rooted trees r-isomorphic p-fringe-trees in G_B;
Set $\mathcal{F}_E := \mathcal{F}(v) := \mathcal{F}^{(1)}_B \cup \mathcal{F}^{(2)}_B$, $v \in V_{\rm C}$ and $\mathrm{fc}_{\rm LB}(\psi) := 0, \mathrm{fc}_{\rm UB}(\psi) := 10, \psi \in \mathcal{F}^{(1)}_B \cup \mathcal{F}^{(2)}_B$.

(d) $I_{\rm d} = (G^1_{\rm C}, \sigma_{\rm int}, \sigma_{\rm ce})$: An instance aimed to infer a chemical monocyclic graph $G^\dagger$ such that the frequency vector of edge-configurations in $G^\dagger$ is a vector obtained by merging those of G_A and G_B. We use chemical monocyclic compounds CID 10076784 and CID 44340250 in Figure 7(c) and (d) for G_A and G_B, respectively. Set a seed graph to be the monocyclic seed graph $G^1_{\rm C} = (V_{\rm C}, E_{\rm C} = E_{(\geq 2)} \cup E_{(\geq 1)})$ with $V_{\rm C} = \{u_1, u_2\}$, $E_{(\geq 2)} = \{a_1\}$ and $E_{(\geq 1)} = \{a_2\}$ in Figure 6(i).
Set $\Lambda := \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}\}$, $\Lambda^{\rm int}_{\rm dg} := \Lambda^{\rm int}_{{\rm dg},A} \cup \Lambda^{\rm int}_{{\rm dg},B}$ and $\Gamma^{\rm int} := \Gamma^{\rm int}_A \cup \Gamma^{\rm int}_B$.
Set $\mathrm{n}^{\rm int}_{\rm LB} := \min\{\mathrm{n}^{\rm int}(G_A), \mathrm{n}^{\rm int}(G_B)\}$, $\mathrm{n}^{\rm int}_{\rm UB} := \max\{\mathrm{n}^{\rm int}(G_A), \mathrm{n}^{\rm int}(G_B)\}$,
$n_{\rm LB} := \min\{n(G_A), n(G_B)\}$ and $n^* := \max\{n(G_A), n(G_B)\}$.
Set lower bounds $\ell_{\rm LB}$, $\mathrm{bl}_{\rm LB}$, $\mathrm{ch}_{\rm LB}$, $\mathrm{bd}_{2,\rm LB}$, $\mathrm{bd}_{3,\rm LB}$, $\mathrm{na}_{\rm LB}$, $\mathrm{na}^{\rm int}_{\rm LB}$, $\mathrm{ns}^{\rm int}_{\rm LB}$, $\mathrm{ac}^{\rm int}_{\rm LB}$ and $\mathrm{ac}^{\rm lf}_{\rm LB}$ to be 0.
Set upper bounds $\ell_{\rm UB}$, $\mathrm{bl}_{\rm UB}$, $\mathrm{ch}_{\rm UB}$, $\mathrm{bd}_{2,\rm UB}$, $\mathrm{bd}_{3,\rm UB}$, $\mathrm{na}_{\rm UB}$, $\mathrm{na}^{\rm int}_{\rm UB}$, $\mathrm{ns}^{\rm int}_{\rm UB}$, $\mathrm{ac}^{\rm int}_{\rm UB}$ and $\mathrm{ac}^{\rm lf}_{\rm UB}$ to be n^*.
For each edge-configuration $\gamma \in \Gamma^{\rm int}$, let $x^*_A(\gamma^{\rm int})$ (resp., $x^*_B(\gamma^{\rm int})$) denote the number of interior-edges with γ in G_A (resp., G_B), $\gamma \in \Gamma^{\rm int}$ and set
$x^*_{\min}(\gamma) := \min\{x^*_A(\gamma), x^*_B(\gamma)\}$, $x^*_{\max}(\gamma) := \max\{x^*_A(\gamma), x^*_B(\gamma)\}$,
$\mathrm{ec}^{\rm int}_{\rm LB}(\gamma) := \lfloor (3/4)x^*_{\min}(\gamma) + (1/4)x^*_{\max}(\gamma) \rfloor$ and
$\mathrm{ec}^{\rm int}_{\rm UB}(\gamma) := \lceil (1/4)x^*_{\min}(\gamma) + (3/4)x^*_{\max}(\gamma) \rceil$.
Set $\mathcal{F}_E := \mathcal{F}(v) := \mathcal{F}_A \cup \mathcal{F}_B$, $v \in V_{\rm C}$ and $\mathrm{fc}_{\rm LB}(\psi) := 0, \mathrm{fc}_{\rm UB}(\psi) := 10, \psi \in \mathcal{F}_A \cup \mathcal{F}_B$.
We include a linear constraint $\ell(a_1) \leq \ell(a_2)$ and $5 \leq \ell(a_1) + \ell(a_2) \leq 15$ as part of the side constraint.

4 All Constraints in an MILP Formulation for Chemical Graphs

We define a standard encoding of a finite set A of elements to be a bijection $\sigma : A \to [1, |A|]$, where we denote by $[A]$ the set $[1, |A|]$ of integers and by $[\mathtt{e}]$ the encoded element $\sigma(\mathtt{e})$. Let ϵ denote *null*, a fictitious chemical element that does not belong to any set of chemical elements, chemical symbols, adjacency-configurations and edge-configurations in the following formulation. Given a finite set A, let A_ϵ denote the set $A \cup \{\epsilon\}$ and define a standard encoding of A_ϵ to be a bijection $\sigma : A \to [0, |A|]$ such that $\sigma(\epsilon) = 0$, where we denote by $[A_\epsilon]$ the set $[0, |A|]$ of integers and by $[\mathtt{e}]$ the encoded element $\sigma(\mathtt{e})$, where $[\epsilon] = 0$.

Let $\sigma = (G_{\rm C}, \sigma_{\rm int}, \sigma_{\rm ce})$ be a target specification, ρ denote the branch-parameter in the specification σ and $\mathbb{C}$ denote a chemical graph in $\mathcal{G}(G_{\rm C}, \sigma_{\rm int}, \sigma_{\rm ce})$.

4.1 Selecting a Cyclical-base

Recall that

$$E_{(=1)} = \{e \in E_{\rm C} \mid \ell_{\rm LB}(e) = \ell_{\rm UB}(e) = 1\}; \quad E_{(0/1)} = \{e \in E_{\rm C} \mid \ell_{\rm LB}(e) = 0, \ell_{\rm UB}(e) = 1\};$$
$$E_{(\geq 1)} = \{e \in E_{\rm C} \mid \ell_{\rm LB}(e) = 1, \ell_{\rm UB}(e) \geq 2\}; \quad E_{(\geq 2)} = \{e \in E_{\rm C} \mid \ell_{\rm LB}(e) \geq 2\};$$

- Every edge $a_i \in E_{(=1)}$ is included in $\langle \mathbb{C} \rangle$;

- Each edge $a_i \in E_{(0/1)}$ is included in $\langle \mathbb{C} \rangle$ if necessary;

- For each edge $a_i \in E_{(\geq 2)}$, edge a_i is not included in $\langle \mathbb{C} \rangle$ and instead a path

$$P_i = (v^{\mathrm{C}}{}_{\mathrm{tail}(i)}, v^{\mathrm{T}}{}_{j-1}, v^{\mathrm{T}}{}_{j}, \ldots, v^{\mathrm{T}}{}_{j+t}, v^{\mathrm{C}}{}_{\mathrm{head}(i)})$$

  of length at least 2 from vertex $v^{\mathrm{C}}{}_{\mathrm{tail}(i)}$ to vertex $v^{\mathrm{C}}{}_{\mathrm{head}(i)}$ visiting some vertices in V_{T} is constructed in $\langle \mathbb{C} \rangle$; and

- For each edge $a_i \in E_{(\geq 1)}$, either edge a_i is directly used in $\langle \mathbb{C} \rangle$ or the above path P_i of length at least 2 is constructed in $\langle \mathbb{C} \rangle$.

Let $t_{\mathrm{C}} \triangleq |V_{\mathrm{C}}|$ and denote V_{C} by $\{v^{\mathrm{C}}{}_i \mid i \in [1, t_{\mathrm{C}}]\}$. Regard the seed graph G_{C} as a digraph such that each edge a_i with end-vertices $v^{\mathrm{C}}{}_j$ and $v^{\mathrm{C}}{}_{j'}$ is directed from $v^{\mathrm{C}}{}_j$ to $v^{\mathrm{C}}{}_{j'}$ when $j < j'$. For each directed edge $a_i \in E_{\mathrm{C}}$, let head(i) and tail(i) denote the head and tail of $e^{\mathrm{C}}(i)$; i.e., $a_i = (v^{\mathrm{C}}{}_{\mathrm{tail}(i)}, v^{\mathrm{C}}{}_{\mathrm{head}(i)})$.

Define

$$k_{\mathrm{C}} \triangleq |E_{(\geq 2)} \cup E_{(\geq 1)}|, \quad \widetilde{k_{\mathrm{C}}} \triangleq |E_{(\geq 2)}|,$$

and denote $E_{\mathrm{C}} = \{a_i \mid i \in [1, m_{\mathrm{C}}]\}$, $E_{(\geq 2)} = \{a_k \mid k \in [1, \widetilde{k_{\mathrm{C}}}]\}$, $E_{(\geq 1)} = \{a_k \mid k \in [\widetilde{k_{\mathrm{C}}} + 1, k_{\mathrm{C}}]\}$, $E_{(0/1)} = \{a_i \mid i \in [k_{\mathrm{C}} + 1, k_{\mathrm{C}} + |E_{(0/1)}|]\}$ and $E_{(=1)} = \{a_i \mid i \in [k_{\mathrm{C}} + |E_{(0/1)}| + 1, m_{\mathrm{C}}]\}$. Let $I_{(=1)}$ denote the set of indices i of edges $a_i \in E_{(=1)}$. Similarly for $I_{(0/1)}$, $I_{(\geq 1)}$ and $I_{(\geq 2)}$.

To control the construction of such a path P_i for each edge $a_k \in E_{(\geq 2)} \cup E_{(\geq 1)}$, we regard the index $k \in [1, k_{\mathrm{C}}]$ of each edge $a_k \in E_{(\geq 2)} \cup E_{(\geq 1)}$ as the "color" of the edge. To introduce necessary linear constraints that can construct such a path P_k properly in our MILP, we assign the color k to the vertices $v^{\mathrm{T}}{}_{j-1}, v^{\mathrm{T}}{}_{j}, \ldots, v^{\mathrm{T}}{}_{j+t}$ in V_{T} when the above path P_k is used in $\langle \mathbb{C} \rangle$.

For each index $s \in [1, t_{\mathrm{C}}]$, let $I_{\mathrm{C}}(s)$ denote the set of edges $e \in E_{\mathrm{C}}$ incident to vertex $v^{\mathrm{C}}{}_s$, and $E^{+}_{(=1)}(s)$ (resp., $E^{-}_{(=1)}(s)$) denote the set of edges $a_i \in E_{(=1)}$ such that the tail (resp., head) of a_i is vertex $v^{\mathrm{C}}{}_s$. Similarly for $E^{+}_{(0/1)}(s)$, $E^{-}_{(0/1)}(s)$, $E^{+}_{(\geq 1)}(s)$, $E^{-}_{(\geq 1)}(s)$, $E^{+}_{(\geq 2)}(s)$ and $E^{-}_{(\geq 2)}(s)$. Let $I_{\mathrm{C}}(s)$ denote the set of indices i of edges $a_i \in I_{\mathrm{C}}(s)$. Similarly for $I^{+}_{(=1)}(s)$, $I^{-}_{(=1)}(s)$, $I^{+}_{(0/1)}(s)$, $I^{-}_{(0/1)}(s)$, $I^{+}_{(\geq 1)}(s)$, $I^{-}_{(\geq 1)}(s)$, $I^{+}_{(\geq 2)}(s)$ and $I^{-}_{(\geq 2)}(s)$. Note that $[1, k_{\mathrm{C}}] = I_{(\geq 2)} \cup I_{(\geq 1)}$ and $[\widetilde{k_{\mathrm{C}}} + 1, m_{\mathrm{C}}] = I_{(\geq 1)} \cup I_{(0/1)} \cup I_{(=1)}$.

constants:

- $t_{\mathrm{C}} = |V_{\mathrm{C}}|$, $\widetilde{k_{\mathrm{C}}} = |E_{(\geq 2)}|$, $k_{\mathrm{C}} = |E_{(\geq 2)} \cup E_{(\geq 1)}|$, $t_{\mathrm{T}} = \mathrm{n}^{\mathrm{int}}_{\mathrm{UB}} - |V_{\mathrm{C}}|$, $m_{\mathrm{C}} = |E_{\mathrm{C}}|$. Note that $a_i \in E_{\mathrm{C}} \setminus (E_{(\geq 2)} \cup E_{(\geq 1)})$ holds $i \in [k_{\mathrm{C}} + 1, m_{\mathrm{C}}]$;

- $\ell_{\mathrm{LB}}(k), \ell_{\mathrm{UB}}(k) \in [1, t_{\mathrm{T}}]$, $k \in [1, k_{\mathrm{C}}]$: lower and upper bounds on the length of path P_k;

- $r_{G_{\mathrm{C}}} \in [1, m_{\mathrm{C}}]$: the rank $\mathrm{r}(G_{\mathrm{C}})$ of seed graph G_{C};

variables:

- $e^{\mathrm{C}}(i) \in [0, 1]$, $i \in [1, m_{\mathrm{C}}]$: $e^{\mathrm{C}}(i)$ represents edge $a_i \in E_{\mathrm{C}}$, $i \in [1, m_{\mathrm{C}}]$ ($e^{\mathrm{C}}(i) = 1$, $i \in I_{(=1)}$; $e^{\mathrm{C}}(i) = 0$, $i \in I_{(\geq 2)}$) ($e^{\mathrm{C}}(i) = 1 \Leftrightarrow$ edge a_i is used in $\langle \mathbb{C} \rangle$);

- $v^{\mathrm{T}}(i) \in [0, 1]$, $i \in [1, t_{\mathrm{T}}]$: $v^{\mathrm{T}}(i) = 1 \Leftrightarrow$ vertex $v^{\mathrm{T}}{}_i$ is used in $\langle \mathbb{C} \rangle$;

- $e^{\mathrm{T}}(i) \in [0, 1]$, $i \in [1, t_{\mathrm{T}} + 1]$: $e^{\mathrm{T}}(i)$ represents edge $e^{\mathrm{T}}{}_i = (v^{\mathrm{T}}{}_{i-1}, v^{\mathrm{T}}{}_i) \in E_{\mathrm{T}}$, where $e^{\mathrm{T}}{}_1$ and $e^{\mathrm{T}}{}_{t_{\mathrm{T}}+1}$ are fictitious edges ($e^{\mathrm{T}}(i) = 1 \Leftrightarrow$ edge $e^{\mathrm{T}}{}_i$ is used in $\langle \mathbb{C} \rangle$);

- $\chi^{\mathrm{T}}(i) \in [0, k_{\mathrm{C}}]$, $i \in [1, t_{\mathrm{T}}]$: $\chi^{\mathrm{T}}(i)$ represents the color assigned to vertex $v^{\mathrm{T}}{}_i$ ($\chi^{\mathrm{T}}(i) = k > 0 \Leftrightarrow$ vertex $v^{\mathrm{T}}{}_i$ is assigned color k; $\chi^{\mathrm{T}}(i) = 0$ means that vertex $v^{\mathrm{T}}{}_i$ is not used in $\langle \mathbb{C} \rangle$);
- $\mathrm{clr}^{\mathrm{T}}(k) \in [\ell_{\mathrm{LB}}(k) - 1, \ell_{\mathrm{UB}}(k) - 1]$, $k \in [1, k_{\mathrm{C}}]$, $\mathrm{clr}^{\mathrm{T}}(0) \in [0, t_{\mathrm{T}}]$: the number of vertices $v^{\mathrm{T}}{}_i \in V_{\mathrm{T}}$ with color c;
- $\delta^{\mathrm{T}}_{\chi}(k) \in [0, 1]$, $k \in [0, k_{\mathrm{C}}]$: $\delta^{\mathrm{T}}_{\chi}(k) = 1 \Leftrightarrow \chi^{\mathrm{T}}(i) = k$ for some $i \in [1, t_{\mathrm{T}}]$;
- $\chi^{\mathrm{T}}(i, k) \in [0, 1]$, $i \in [1, t_{\mathrm{T}}]$, $k \in [0, k_{\mathrm{C}}]$ ($\chi^{\mathrm{T}}(i, k) = 1 \Leftrightarrow \chi^{\mathrm{T}}(i) = k$);
- $\widetilde{\deg}^{+}_{\mathrm{C}}(i) \in [0, 4]$, $i \in [1, t_{\mathrm{C}}]$: the out-degree of vertex $v^{\mathrm{C}}{}_i$ with the used edges e^{C} in E_{C};
- $\widetilde{\deg}^{-}_{\mathrm{C}}(i) \in [0, 4]$, $i \in [1, t_{\mathrm{C}}]$: the in-degree of vertex $v^{\mathrm{C}}{}_i$ with the used edges e^{C} in E_{C};
- rank: the rank $\mathrm{r}(\mathbb{C})$ of a target chemical graph $\mathbb{C}$;

constraints:

$$\mathrm{rank} = r_{G_{\mathrm{C}}} - \sum_{i \in I_{(0/1)}} (1 - e^{\mathrm{C}}(i)), \tag{1}$$

$$e^{\mathrm{C}}(i) = 1, \quad i \in I_{(=1)}, \tag{2}$$

$$e^{\mathrm{C}}(i) = 0, \quad \mathrm{clr}^{\mathrm{T}}(i) \geq 1, \quad i \in I_{(\geq 2)}, \tag{3}$$

$$e^{\mathrm{C}}(i) + \mathrm{clr}^{\mathrm{T}}(i) \geq 1, \qquad \mathrm{clr}^{\mathrm{T}}(i) \leq t_{\mathrm{T}} \cdot (1 - e^{\mathrm{C}}(i)), \quad i \in I_{(\geq 1)}, \tag{4}$$

$$\sum_{c \in I^{-}_{(\geq 1)}(i) \cup I^{-}_{(0/1)}(i) \cup I^{-}_{(=1)}(i)} e^{\mathrm{C}}(c) = \widetilde{\deg}^{-}_{\mathrm{C}}(i), \qquad \sum_{c \in I^{+}_{(\geq 1)}(i) \cup I^{+}_{(0/1)}(i) \cup I^{+}_{(=1)}(i)} e^{\mathrm{C}}(c) = \widetilde{\deg}^{+}_{\mathrm{C}}(i), \qquad i \in [1, t_{\mathrm{C}}], \tag{5}$$

$$\chi^{\mathrm{T}}(i, 0) = 1 - v^{\mathrm{T}}(i), \qquad \sum_{k \in [0, k_{\mathrm{C}}]} \chi^{\mathrm{T}}(i, k) = 1, \qquad \sum_{k \in [0, k_{\mathrm{C}}]} k \cdot \chi^{\mathrm{T}}(i, k) = \chi^{\mathrm{T}}(i), \qquad i \in [1, t_{\mathrm{T}}], \tag{6}$$

$$\sum_{i \in [1, t_{\mathrm{T}}]} \chi^{\mathrm{T}}(i, k) = \mathrm{clr}^{\mathrm{T}}(k), \quad t_{\mathrm{T}} \cdot \delta^{\mathrm{T}}_{\chi}(k) \geq \sum_{i \in [1, t_{\mathrm{T}}]} \chi^{\mathrm{T}}(i, k) \geq \delta^{\mathrm{T}}_{\chi}(k), \qquad k \in [0, k_{\mathrm{C}}], \tag{7}$$

$$\begin{aligned} v^{\mathrm{T}}(i-1) &\geq v^{\mathrm{T}}(i), \\ k_{\mathrm{C}} \cdot (v^{\mathrm{T}}(i-1) - e^{\mathrm{T}}(i)) \geq \chi^{\mathrm{T}}(i-1) - \chi^{\mathrm{T}}(i) &\geq v^{\mathrm{T}}(i-1) - e^{\mathrm{T}}(i), \end{aligned} \qquad i \in [2, t_{\mathrm{T}}]. \tag{8}$$

4.2 Constraints for Including Leaf Paths

Let $\widetilde{t_{\mathrm{C}}}$ denote the number of vertices $u \in V_{\mathrm{C}}$ such that $\mathrm{bl}_{\mathrm{UB}}(u) = 1$ and assume that $V_{\mathrm{C}} = \{u_1, u_2, \ldots, u_p\}$ so that

$$\mathrm{bl}_{\mathrm{UB}}(u_i) = 1,\ i \in [1, \widetilde{t_{\mathrm{C}}}] \text{ and } \mathrm{bl}_{\mathrm{UB}}(u_i) = 0,\ i \in [\widetilde{t_{\mathrm{C}}} + 1, t_{\mathrm{C}}].$$

Define the set of colors for the vertex set $\{u_i \mid i \in [1, \widetilde{t_{\mathrm{C}}}]\} \cup V_{\mathrm{T}}$ to be $[1, c_{\mathrm{F}}]$ with

$$c_{\mathrm{F}} \triangleq \widetilde{t_{\mathrm{C}}} + t_{\mathrm{T}} = |\{u_i \mid i \in [1, \widetilde{t_{\mathrm{C}}}]\} \cup V_{\mathrm{T}}|.$$

Let each vertex $v^{\mathrm{C}}{}_i$, $i \in [1, \widetilde{t_{\mathrm{C}}}]$ (resp., $v^{\mathrm{T}}{}_i \in V_{\mathrm{T}}$) correspond to a color $i \in [1, c_{\mathrm{F}}]$ (resp., $i + \widetilde{t_{\mathrm{C}}} \in [1, c_{\mathrm{F}}]$). When a path $P = (u, v^{\mathrm{F}}{}_j, v^{\mathrm{F}}{}_{j+1}, \ldots, v^{\mathrm{F}}{}_{j+t})$ from a vertex $u \in V_{\mathrm{C}} \cup V_{\mathrm{T}}$ is used in $\langle \mathbb{C} \rangle$, we assign the color $i \in [1, c_{\mathrm{F}}]$ of the vertex u to the vertices $v^{\mathrm{F}}{}_j, v^{\mathrm{F}}{}_{j+1}, \ldots, v^{\mathrm{F}}{}_{j+t} \in V_{\mathrm{F}}$.

constants:

- $c_{\rm F}$: the maximum number of different colors assigned to the vertices in $V_{\rm F}$;

- n^*: an upper bound on the number $n(\mathbb{C})$ of non-hydrogen atoms in $\mathbb{C}$;

- $\mathrm{n}^{\rm int}_{\rm LB}, \mathrm{n}^{\rm int}_{\rm UB} \in [2, n^*]$: lower and upper bounds on the number of interior-vertices in $\mathbb{C}$;

- $\mathrm{bl}_{\rm LB}(i) \in [0, 1]$, $i \in [1, \widetilde{t_{\rm C}}]$: a lower bound on the number of leaf ρ-branches in the leaf path rooted at a vertex $v^{\rm C}{}_i$;

- $\mathrm{bl}_{\rm LB}(k), \mathrm{bl}_{\rm UB}(k) \in [0, \ell_{\rm UB}(k) - 1]$, $k \in [1, k_{\rm C}] = I_{(\geq 2)} \cup I_{(\geq 1)}$: lower and upper bounds on the number of leaf ρ-branches in the trees rooted at internal vertices of a pure path P_k for an edge $a_k \in E_{(\geq 1)} \cup E_{(\geq 2)}$;

variables:

- $\mathrm{n}^{\rm int}_G \in [\mathrm{n}^{\rm int}_{\rm LB}, \mathrm{n}^{\rm int}_{\rm UB}]$: the number of interior-vertices in $\mathbb{C}$;

- $v^{\rm F}(i) \in [0, 1]$, $i \in [1, t_{\rm F}]$: $v^{\rm F}(i) = 1 \Leftrightarrow$ vertex $v^{\rm F}{}_i$ is used in $\mathbb{C}$;

- $e^{\rm F}(i) \in [0, 1]$, $i \in [1, t_{\rm F} + 1]$: $e^{\rm F}(i)$ represents edge $e^{\rm F}{}_i = v^{\rm F}{}_{i-1} v^{\rm F}{}_i$, where $e^{\rm F}{}_1$ and $e^{\rm F}{}_{t_{\rm F}+1}$ are fictitious edges ($e^{\rm F}(i) = 1 \Leftrightarrow$ edge $e^{\rm F}{}_i$ is used in $\mathbb{C}$);

- $\chi^{\rm F}(i) \in [0, c_{\rm F}]$, $i \in [1, t_{\rm F}]$: $\chi^{\rm F}(i)$ represents the color assigned to vertex $v^{\rm F}{}_i$ ($\chi^{\rm F}(i) = c \Leftrightarrow$ vertex $v^{\rm F}{}_i$ is assigned color c);

- $\mathrm{clr}^{\rm F}(c) \in [0, t_{\rm F}]$, $c \in [0, c_{\rm F}]$: the number of vertices $v^{\rm F}{}_i$ with color c;

- $\delta^{\rm F}_\chi(c) \in [\mathrm{bl}_{\rm LB}(c), 1]$, $c \in [1, \widetilde{t_{\rm C}}]$: $\delta^{\rm F}_\chi(c) = 1 \Leftrightarrow \chi^{\rm F}(i) = c$ for some $i \in [1, t_{\rm F}]$;

- $\delta^{\rm F}_\chi(c) \in [0, 1]$, $c \in [\widetilde{t_{\rm C}} + 1, c_{\rm F}]$: $\delta^{\rm F}_\chi(c) = 1 \Leftrightarrow \chi^{\rm F}(i) = c$ for some $i \in [1, t_{\rm F}]$;

- $\chi^{\rm F}(i, c) \in [0, 1]$, $i \in [1, t_{\rm F}]$, $c \in [0, c_{\rm F}]$: $\chi^{\rm F}(i, c) = 1 \Leftrightarrow \chi^{\rm F}(i) = c$;

- $\mathrm{bl}(k, i) \in [0, 1]$, $k \in [1, k_{\rm C}] = I_{(\geq 2)} \cup I_{(\geq 1)}$, $i \in [1, t_{\rm T}]$: $\mathrm{bl}(k, i) = 1 \Leftrightarrow$ path P_k contains vertex $v^{\rm T}{}_i$ as an internal vertex and the ρ-fringe-tree rooted at $v^{\rm T}{}_i$ contains a leaf ρ-branch;

constraints:

$$\chi^{\rm F}(i, 0) = 1 - v^{\rm F}(i), \qquad \sum_{c \in [0, c_{\rm F}]} \chi^{\rm F}(i, c) = 1, \qquad \sum_{c \in [0, c_{\rm F}]} c \cdot \chi^{\rm F}(i, c) = \chi^{\rm F}(i), \qquad i \in [1, t_{\rm F}], \tag{9}$$

$$\sum_{i \in [1, t_{\rm F}]} \chi^{\rm F}(i, c) = \mathrm{clr}^{\rm F}(c), \quad t_{\rm F} \cdot \delta^{\rm F}_\chi(c) \geq \sum_{i \in [1, t_{\rm F}]} \chi^{\rm F}(i, c) \geq \delta^{\rm F}_\chi(c), \qquad c \in [0, c_{\rm F}], \tag{10}$$

$$e^{\rm F}(1) = e^{\rm F}(t_{\rm F} + 1) = 0, \tag{11}$$

$$\begin{aligned} v^{\rm F}(i-1) &\geq v^{\rm F}(i), \\ c_{\rm F} \cdot (v^{\rm F}(i-1) - e^{\rm F}(i)) \geq \chi^{\rm F}(i-1) - \chi^{\rm F}(i) &\geq v^{\rm F}(i-1) - e^{\rm F}(i), \end{aligned} \qquad i \in [2, t_{\rm F}], \tag{12}$$

$$\mathrm{bl}(k, i) \geq \delta^{\rm F}_\chi(\widetilde{t_{\rm C}} + i) + \chi^{\rm T}(i, k) - 1, \qquad k \in [1, k_{\rm C}], i \in [1, t_{\rm T}], \tag{13}$$

$$\sum_{k\in[1,k_{\mathrm{C}}],i\in[1,t_{\mathrm{T}}]} \mathrm{bl}(k,i) \le \sum_{i\in[1,t_{\mathrm{T}}]} \delta^{\mathrm{F}}_{\chi}(\widetilde{t_{\mathrm{C}}}+i), \tag{14}$$

$$\mathrm{bl}_{\mathrm{LB}}(k) \le \sum_{i\in[1,t_{\mathrm{T}}]} \mathrm{bl}(k,i) \le \mathrm{bl}_{\mathrm{UB}}(k), \qquad k\in[1,k_{\mathrm{C}}], \tag{15}$$

$$t_{\mathrm{C}} + \sum_{i\in[1,t_{\mathrm{T}}]} v^{\mathrm{T}}(i) + \sum_{i\in[1,t_{\mathrm{F}}]} v^{\mathrm{F}}(i) = \mathrm{n}^{\mathrm{int}}_{G}. \tag{16}$$

4.3 Constraints for Including Fringe-trees

Recall that $\mathcal{F}(D_\pi)$ denotes the set of chemical rooted trees ψ r-isomorphic to a chemical rooted tree in $\mathcal{T}(\mathbb{C})$ over all chemical graphs $\mathbb{C}\in D_\pi$, where possibly a chemical rooted tree $\psi\in\mathcal{F}(D_\pi)$ consists of a single chemical element $\mathtt{a}\in\Lambda\setminus\{\mathtt{H}\}$.

To express the condition that the ρ-fringe-tree is chosen from a rooted tree C_i, T_i or F_i, we introduce the following set of variables and constraints.

constants:

- n_{LB}: a lower bound on the number $n(\mathbb{C})$ of non-hydrogen atoms in $\mathbb{C}$, where $n_{\mathrm{LB}}, n^* \ge \mathrm{n}^{\mathrm{int}}_{\mathrm{LB}}$;
- $\mathrm{ch}_{\mathrm{LB}}(i), \mathrm{ch}_{\mathrm{UB}}(i)\in[0,n^*]$, $i\in[1,t_{\mathrm{T}}]$: lower and upper bounds on $\mathrm{ht}(\langle T_i\rangle)$ of the tree T_i rooted at a vertex ${v^{\mathrm{C}}}_i$;
- $\mathrm{ch}_{\mathrm{LB}}(k), \mathrm{ch}_{\mathrm{UB}}(k)\in[0,n^*]$, $k\in[1,k_{\mathrm{C}}]=I_{(\ge 2)}\cup I_{(\ge 1)}$: lower and upper bounds on the maximum height $\mathrm{ht}(\langle T\rangle)$ of the tree $T\in\mathcal{F}(P_k)$ rooted at an internal vertex of a path P_k for an edge $a_k\in E_{(\ge 1)}\cup E_{(\ge 2)}$;
- Prepare a coding of the set $\mathcal{F}(D_\pi)$ and let $[\psi]$ denote the coded integer of an element ψ in $\mathcal{F}(D_\pi)$;
- Sets $\mathcal{F}(v)\subseteq\mathcal{F}(D_\pi), v\in V_{\mathrm{C}}$ and $\mathcal{F}_E\subseteq\mathcal{F}(D_\pi)$ of chemical rooted trees T with $\mathrm{ht}(T)\in[1,\rho]$;
- Define $\mathcal{F}^* := \bigcup_{v\in V_{\mathrm{C}}}\mathcal{F}(v)\cup\mathcal{F}_E$, $\mathcal{F}^{\mathrm{C}}_i := \mathcal{F}({v^{\mathrm{C}}}_i)$, $i\in[1,t_{\mathrm{C}}]$, $\mathcal{F}^{\mathrm{T}}_i := \mathcal{F}_E$, $i\in[1,t_{\mathrm{T}}]$ and $\mathcal{F}^{\mathrm{F}}_i := \mathcal{F}_E$, $i\in[1,t_{\mathrm{F}}]$;
- $\mathrm{fc}_{\mathrm{LB}}(\psi), \mathrm{fc}_{\mathrm{UB}}(\psi)\in[0,n^*], \psi\in\mathcal{F}^*$: lower and upper bound functions on the number of interior-vertices v such that $\mathbb{C}[v]$ is r-isomorphic to ψ in $\mathbb{C}$;
- $\mathcal{F}^{\mathrm{X}}_i[p], p\in[1,\rho], \mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F}\}$: the set of chemical rooted trees $T\in\mathcal{F}^{\mathrm{X}}_i$ with $\mathrm{ht}(\langle T\rangle)=p$;
- $n_{\overline{\mathtt{H}}}([\psi])\in[0,3^\rho], \psi\in\mathcal{F}^*$: the number $n(\langle\psi\rangle)$ of non-root hydrogen vertices in a chemical rooted tree ψ;
- $\mathrm{ht}_{\overline{\mathtt{H}}}([\psi])\in[0,\rho], \psi\in\mathcal{F}^*$: the height $\mathrm{ht}(\langle\psi\rangle)$ of the hydrogen-suppressed chemical rooted tree $\langle\psi\rangle$;
- $\mathrm{deg}^{\overline{\mathtt{H}}}_{\mathrm{r}}([\psi])\in[0,3], \psi\in\mathcal{F}^*$: the number $\mathrm{deg}_{\mathrm{r}}(\langle\psi\rangle)$ of non-hydrogen children of the root r of a chemical rooted tree ψ;
- $\mathrm{deg}^{\mathrm{hyd}}_{\mathrm{r}}([\psi])\in[0,3], \psi\in\mathcal{F}^*$: the number $\mathrm{deg}_{\mathrm{r}}(\psi)-\mathrm{deg}_{\mathrm{r}}(\langle\psi\rangle)$ of hydrogen children of the root r of a chemical rooted tree ψ;
- $\mathrm{v}_{\mathrm{ion}}(\psi)\in[-3,+3], \psi\in\mathcal{F}^*$: the ion-valence of the root in ψ;

- $\mathrm{ac}^{\mathrm{lf}}_{\nu}(\psi), \nu \in \Gamma^{\mathrm{lf}}_{\mathrm{ac}}$: the frequency of leaf-edges with adjacency-configuration ν in ψ;
- $\mathrm{ac}^{\mathrm{lf}}_{\mathrm{LB}}, \mathrm{ac}^{\mathrm{lf}}_{\mathrm{UB}} : \Gamma^{\mathrm{lf}}_{\mathrm{ac}} \to [0, n^*]$: lower and upper bound functions on the number of leaf-edges uv in ac_{C} with adjacency-configuration ν;

variables:

- $n_G \in [n_{\mathrm{LB}}, n^*]$: the number $n(\mathbb{C})$ of non-hydrogen atoms in $\mathbb{C}$;
- $v^{\mathrm{X}}(i) \in [0,1], i \in [1, t_{\mathrm{X}}]$, $\mathrm{X} \in \{\mathrm{T}, \mathrm{F}\}$: $v^{\mathrm{X}}(i) = 1 \Leftrightarrow$ vertex $v^{\mathrm{X}}{}_i$ is used in $\mathbb{C}$;
- $\delta^{\mathrm{X}}_{\mathrm{fr}}(i, [\psi]) \in [0,1], i \in [1, t_{\mathrm{X}}], \psi \in \mathcal{F}^{\mathrm{X}}_i, \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: $\delta^{\mathrm{X}}_{\mathrm{fr}}(i, [\psi]) = 1 \Leftrightarrow \psi$ is the ρ-fringe-tree rooted at vertex $v^{\mathrm{X}}{}_i$ in $\mathbb{C}$;
- $\mathrm{fc}([\psi]) \in [\mathrm{fc}_{\mathrm{LB}}(\psi), \mathrm{fc}_{\mathrm{UB}}(\psi)], \psi \in \mathcal{F}^*$: the number of interior-vertices v such that $\mathbb{C}[v]$ is r-isomorphic to ψ in $\mathbb{C}$;
- $\mathrm{ac}^{\mathrm{lf}}([\nu]) \in [\mathrm{ac}^{\mathrm{lf}}_{\mathrm{LB}}(\nu), \mathrm{ac}^{\mathrm{lf}}_{\mathrm{UB}}(\nu)], \nu \in \Gamma^{\mathrm{lf}}_{\mathrm{ac}}$: the number of leaf-edge with adjacency-configuration ν in $\mathbb{C}$;
- $\deg^{\mathrm{ex}}_{\mathrm{X}}(i) \in [0,3], i \in [1, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: the number of non-hydrogen children of the root of the ρ-fringe-tree rooted at vertex $v^{\mathrm{X}}{}_i$ in $\mathbb{C}$;
- $\mathrm{hyddeg}^{\mathrm{X}}(i) \in [0,4]$, $i \in [1, t_{\mathrm{X}}]$, $\mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: the number of hydrogen atoms adjacent to vertex $v^{\mathrm{X}}{}_i$ (i.e., $\mathrm{hyddeg}(v^{\mathrm{X}}{}_i)$) in $\mathbb{C} = (H, \alpha, \beta)$;
- $\mathrm{eledeg}_{\mathrm{X}}(i) \in [-3, +3]$, $i \in [1, t_{\mathrm{X}}]$, $\mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: the ion-valence $\mathrm{v}_{\mathrm{ion}}(\psi)$ of vertex $v^{\mathrm{X}}{}_i$ (i.e., $\mathrm{eledeg}_{\mathrm{X}}(i) = \mathrm{v}_{\mathrm{ion}}(\psi)$ for the ρ-fringe-tree ψ rooted at $v^{\mathrm{X}}{}_i$) in $\mathbb{C} = (H, \alpha, \beta)$;
- $h^{\mathrm{X}}(i) \in [0, \rho]$, $i \in [1, t_{\mathrm{X}}]$, $\mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: the height $\mathrm{ht}(\langle T \rangle)$ of the hydrogen-suppressed chemical rooted tree $\langle T \rangle$ of the ρ-fringe-tree T rooted at vertex $v^{\mathrm{X}}{}_i$ in $\mathbb{C}$;
- $\sigma(k,i) \in [0,1]$, $k \in [1, k_{\mathrm{C}}] = I_{(\geq 2)} \cup I_{(\geq 1)}, i \in [1, t_{\mathrm{T}}]$: $\sigma(k,i) = 1 \Leftrightarrow$ the ρ-fringe-tree T_v rooted at vertex $v = v^{\mathrm{T}}{}_i$ with color k has the largest height $\mathrm{ht}(\langle \mathcal{T}_v \rangle)$ among such trees $T_v, v \in V_{\mathrm{T}}$;

constraints:

$$\sum_{\psi \in \mathcal{F}^{\mathrm{C}}_i} \delta^{\mathrm{C}}_{\mathrm{fr}}(i, [\psi]) = 1, \qquad i \in [1, t_{\mathrm{C}}],$$
$$\sum_{\psi \in \mathcal{F}^{\mathrm{X}}_i} \delta^{\mathrm{X}}_{\mathrm{fr}}(i, [\psi]) = v^{\mathrm{X}}(i), \qquad i \in [1, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{T}, \mathrm{F}\}, \tag{17}$$

$$\sum_{\psi \in \mathcal{F}^{\mathrm{X}}_i} \deg^{\overline{\mathrm{H}}}_{\mathrm{r}}([\psi]) \cdot \delta^{\mathrm{X}}_{\mathrm{fr}}(i, [\psi]) = \deg^{\mathrm{ex}}_{\mathrm{X}}(i),$$
$$\sum_{\psi \in \mathcal{F}^{\mathrm{X}}_i} \deg^{\mathrm{hyd}}_{\mathrm{r}}([\psi]) \cdot \delta^{\mathrm{X}}_{\mathrm{fr}}(i, [\psi]) = \mathrm{hyddeg}^{\mathrm{X}}(i),$$
$$\sum_{\psi \in \mathcal{F}^{\mathrm{X}}_i} \mathrm{v}_{\mathrm{ion}}([\psi]) \cdot \delta^{\mathrm{X}}_{\mathrm{fr}}(i, [\psi]) = \mathrm{eledeg}_{\mathrm{X}}(i), \qquad i \in [1, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}, \tag{18}$$

$$\sum_{\psi \in \mathcal{F}^{\mathrm{F}}_i[\rho]} \delta^{\mathrm{F}}_{\mathrm{fr}}(i, [\psi]) \geq v^{\mathrm{F}}(i) - e^{\mathrm{F}}(i+1), \qquad i \in [1, t_{\mathrm{F}}] \ (e^{\mathrm{F}}(t_{\mathrm{F}}+1) = 0), \tag{19}$$

$$\sum_{\psi\in\mathcal{F}_i^{\mathrm{X}}} \mathrm{ht}_{\overline{\mathtt{H}}}([\psi])\cdot\delta_{\mathrm{fr}}^{\mathrm{X}}(i,[\psi]) = h^{\mathrm{X}}(i), \qquad i\in[1,t_{\mathrm{X}}], \mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F}\}, \tag{20}$$

$$\sum_{\substack{\psi\in\mathcal{F}_i^{\mathrm{X}}\\ i\in[1,t_{\mathrm{X}}],\mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F}\}}} n_{\overline{\mathtt{H}}}([\psi])\cdot\delta_{\mathrm{fr}}^{\mathrm{X}}(i,[\psi]) + \sum_{i\in[1,t_{\mathrm{X}}],\mathrm{X}\in\{\mathrm{T},\mathrm{F}\}} v^{\mathrm{X}}(i) + t_{\mathrm{C}} = n_G, \tag{21}$$

$$\sum_{i\in[1,t_{\mathrm{X}}],\mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F}\}} \delta_{\mathrm{fr}}^{\mathrm{X}}(i,[\psi]) = \mathrm{fc}([\psi]), \qquad \psi\in\mathcal{F}^*, \tag{22}$$

$$\sum_{\psi\in\mathcal{F}_i^{\mathrm{X}},i\in[1,t_{\mathrm{X}}],\mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F}\}} \mathrm{ac}_{\nu}^{\mathrm{lf}}(\psi)\cdot\delta_{\mathrm{fr}}^{\mathrm{X}}(i,[\psi]) = \mathrm{ac}^{\mathrm{lf}}([\nu]), \qquad \nu\in\Gamma_{\mathrm{ac}}^{\mathrm{lf}}, \tag{23}$$

$$\begin{aligned} h^{\mathrm{C}}(i) \ge \mathrm{ch}_{\mathrm{LB}}(i) - n^*\cdot\delta_{\chi}^{\mathrm{F}}(i), \quad \mathrm{clr}^{\mathrm{F}}(i)+\rho \ge \mathrm{ch}_{\mathrm{LB}}(i),\\ h^{\mathrm{C}}(i) \le \mathrm{ch}_{\mathrm{UB}}(i), \quad \mathrm{clr}^{\mathrm{F}}(i)+\rho \le \mathrm{ch}_{\mathrm{UB}}(i) + n^*\cdot(1-\delta_{\chi}^{\mathrm{F}}(i)), \end{aligned} \qquad i\in[1,\widetilde{t_{\mathrm{C}}}], \tag{24}$$

$$\mathrm{ch}_{\mathrm{LB}}(i) \le h^{\mathrm{C}}(i) \le \mathrm{ch}_{\mathrm{UB}}(i), \qquad i\in[\widetilde{t_{\mathrm{C}}}+1, t_{\mathrm{C}}], \tag{25}$$

$$\begin{aligned} h^{\mathrm{T}}(i) \le \mathrm{ch}_{\mathrm{UB}}(k) + n^*\cdot(\delta_{\chi}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i) + 1 - \chi^{\mathrm{T}}(i,k)),\\ \mathrm{clr}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i)+\rho \le \mathrm{ch}_{\mathrm{UB}}(k) + n^*\cdot(2-\delta_{\chi}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i) - \chi^{\mathrm{T}}(i,k)), \end{aligned} \qquad k\in[1,k_{\mathrm{C}}], i\in[1,t_{\mathrm{T}}], \tag{26}$$

$$\sum_{i\in[1,t_{\mathrm{T}}]} \sigma(k,i) = \delta_{\chi}^{\mathrm{T}}(k), \qquad k\in[1,k_{\mathrm{C}}], \tag{27}$$

$$\begin{aligned} \chi^{\mathrm{T}}(i,k) \ge \sigma(k,i),\\ h^{\mathrm{T}}(i) \ge \mathrm{ch}_{\mathrm{LB}}(k) - n^*\cdot(\delta_{\chi}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i) + 1 - \sigma(k,i)),\\ \mathrm{clr}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i)+\rho \ge \mathrm{ch}_{\mathrm{LB}}(k) - n^*\cdot(2-\delta_{\chi}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i) - \sigma(k,i)), \end{aligned} \qquad k\in[1,k_{\mathrm{C}}], i\in[1,t_{\mathrm{T}}]. \tag{28}$$

4.4 Descriptor for the Number of Specified Degree

We include constraints to compute descriptors for degrees in $\mathbb{C}$.

variables:

- $\deg^{\mathrm{X}}(i)\in[0,4]$, $i\in[1,t_{\mathrm{X}}]$, $\mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F}\}$: the number of non-hydrogen atoms adjacent to vertex $v={v^{\mathrm{X}}}_i$ (i.e., $\deg_{\langle\mathbb{C}\rangle}(v)=\deg_H(v)-\mathrm{hyddeg}_{\mathbb{C}}(v)$) in $\mathbb{C}=(H,\alpha,\beta)$;
- $\deg_{\mathrm{CT}}(i)\in[0,4]$, $i\in[1,t_{\mathrm{C}}]$: the number of edges from vertex ${v^{\mathrm{C}}}_i$ to vertices ${v^{\mathrm{T}}}_j$, $j\in[1,t_{\mathrm{T}}]$;

- $\mathrm{deg_{TC}}(i) \in [0,4]$, $i \in [1, t_{\mathrm{C}}]$: the number of edges from vertices $v^{\mathrm{T}}{}_{j}$, $j \in [1, t_{\mathrm{T}}]$ to vertex $v^{\mathrm{C}}{}_{i}$;
- $\delta_{\mathrm{dg}}^{\mathrm{C}}(i,d) \in [0,1]$, $i \in [1, t_{\mathrm{C}}]$, $d \in [1,4]$, $\delta_{\mathrm{dg}}^{\mathrm{X}}(i,d) \in [0,1]$, $i \in [1, t_{\mathrm{X}}]$, $d \in [0,4]$, $\mathrm{X} \in \{\mathrm{T},\mathrm{F}\}$: $\delta_{\mathrm{dg}}^{\mathrm{X}}(i,d) = 1 \Leftrightarrow \mathrm{deg}^{\mathrm{X}}(i) + \mathrm{hyddeg}^{\mathrm{X}}(i) = d$;
- $\mathrm{dg}(d) \in [\mathrm{dg_{LB}}(d), \mathrm{dg_{UB}}(d)]$, $d \in [1,4]$: the number of interior-vertices v with $\mathrm{deg}_H(v^{\mathrm{X}}{}_i) = d$ in $\mathbb{C} = (H, \alpha, \beta)$;
- $\mathrm{deg}_{\mathrm{C}}^{\mathrm{int}}(i) \in [1,4]$, $i \in [1, t_{\mathrm{C}}]$, $\mathrm{deg}_{\mathrm{X}}^{\mathrm{int}}(i) \in [0,4]$, $i \in [1, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{T},\mathrm{F}\}$: the interior-degree $\mathrm{deg}_{H^{\mathrm{int}}}(v^{\mathrm{X}}{}_i)$ in the interior $H^{\mathrm{int}} = (V^{\mathrm{int}}(\mathbb{C}), E^{\mathrm{int}}(\mathbb{C}))$ of $\mathbb{C}$; i.e., the number of interior-edges incident to vertex $v^{\mathrm{X}}{}_i$;
- $\delta_{\mathrm{dg,C}}^{\mathrm{int}}(i,d) \in [0,1]$, $i \in [1, t_{\mathrm{C}}]$, $d \in [1,4]$, $\delta_{\mathrm{dg,X}}^{\mathrm{int}}(i,d) \in [0,1]$, $i \in [1, t_{\mathrm{X}}]$, $d \in [0,4]$, $\mathrm{X} \in \{\mathrm{T},\mathrm{F}\}$: $\delta_{\mathrm{dg,X}}^{\mathrm{int}}(i,d) = 1 \Leftrightarrow \mathrm{deg}_{\mathrm{X}}^{\mathrm{int}}(i) = d$;
- $\mathrm{dg}^{\mathrm{int}}(d) \in [\mathrm{dg_{LB}}(d), \mathrm{dg_{UB}}(d)]$, $d \in [1,4]$: the number of interior-vertices v with the interior-degree $\mathrm{deg}_{H^{\mathrm{int}}}(v) = d$ in the interior $H^{\mathrm{int}} = (V^{\mathrm{int}}(\mathbb{C}), E^{\mathrm{int}}(\mathbb{C}))$ of $\mathbb{C} = (H, \alpha, \beta)$.

constraints:

$$\sum_{k \in I^{+}_{(\geq 2)}(i) \cup I^{+}_{(\geq 1)}(i)} \delta_{\chi}^{\mathrm{T}}(k) = \mathrm{deg_{CT}}(i), \qquad \sum_{k \in I^{-}_{(\geq 2)}(i) \cup I^{-}_{(\geq 1)}(i)} \delta_{\chi}^{\mathrm{T}}(k) = \mathrm{deg_{TC}}(i), \qquad i \in [1, t_{\mathrm{C}}], \tag{29}$$

$$\widetilde{\mathrm{deg}}_{\mathrm{C}}^{-}(i) + \widetilde{\mathrm{deg}}_{\mathrm{C}}^{+}(i) + \mathrm{deg_{CT}}(i) + \mathrm{deg_{TC}}(i) + \delta_{\chi}^{\mathrm{F}}(i) = \mathrm{deg}_{\mathrm{C}}^{\mathrm{int}}(i), \qquad i \in [1, \widetilde{t_{\mathrm{C}}}], \tag{30}$$

$$\widetilde{\mathrm{deg}}_{\mathrm{C}}^{-}(i) + \widetilde{\mathrm{deg}}_{\mathrm{C}}^{+}(i) + \mathrm{deg_{CT}}(i) + \mathrm{deg_{TC}}(i) = \mathrm{deg}_{\mathrm{C}}^{\mathrm{int}}(i), \qquad i \in [\widetilde{t_{\mathrm{C}}} + 1, t_{\mathrm{C}}], \tag{31}$$

$$\mathrm{deg}_{\mathrm{C}}^{\mathrm{int}}(i) + \mathrm{deg}_{\mathrm{C}}^{\mathrm{ex}}(i) = \mathrm{deg}^{\mathrm{C}}(i), \qquad i \in [1, t_{\mathrm{C}}], \tag{32}$$

$$\sum_{\psi \in \mathcal{F}_i^{\mathrm{C}}[\rho]} \delta_{\mathrm{fr}}^{\mathrm{C}}(i, [\psi]) \geq 2 - \mathrm{deg}_{\mathrm{C}}^{\mathrm{int}}(i) \qquad i \in [1, t_{\mathrm{C}}], \tag{33}$$

$$\begin{aligned} 2v^{\mathrm{T}}(i) + \delta_{\chi}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}} + i) &= \mathrm{deg}_{\mathrm{T}}^{\mathrm{int}}(i), \\ \mathrm{deg}_{\mathrm{T}}^{\mathrm{int}}(i) + \mathrm{deg}_{\mathrm{T}}^{\mathrm{ex}}(i) &= \mathrm{deg}^{\mathrm{T}}(i), \qquad i \in [1, t_{\mathrm{T}}] \ (e^{\mathrm{T}}(1) = e^{\mathrm{T}}(t_{\mathrm{T}} + 1) = 0), \end{aligned} \tag{34}$$

$$\begin{aligned} v^{\mathrm{F}}(i) + e^{\mathrm{F}}(i+1) &= \mathrm{deg}_{\mathrm{F}}^{\mathrm{int}}(i), \\ \mathrm{deg}_{\mathrm{F}}^{\mathrm{int}}(i) + \mathrm{deg}_{\mathrm{F}}^{\mathrm{ex}}(i) &= \mathrm{deg}^{\mathrm{F}}(i), \qquad i \in [1, t_{\mathrm{F}}] \ (e^{\mathrm{F}}(1) = e^{\mathrm{F}}(t_{\mathrm{F}} + 1) = 0), \end{aligned} \tag{35}$$

$$\begin{aligned} &\sum_{d \in [0,4]} \delta_{\mathrm{dg}}^{\mathrm{X}}(i,d) = 1, \quad \sum_{d \in [1,4]} d \cdot \delta_{\mathrm{dg}}^{\mathrm{X}}(i,d) = \mathrm{deg}^{\mathrm{X}}(i) + \mathrm{hyddeg}^{\mathrm{X}}(i), \\ &\sum_{d \in [0,4]} \delta_{\mathrm{dg,X}}^{\mathrm{int}}(i,d) = 1, \quad \sum_{d \in [1,4]} d \cdot \delta_{\mathrm{dg,X}}^{\mathrm{int}}(i,d) = \mathrm{deg}_{\mathrm{X}}^{\mathrm{int}}(i), \qquad i \in [1, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{T},\mathrm{C},\mathrm{F}\}, \end{aligned} \tag{36}$$

$$\begin{aligned} \sum_{i \in [1, t_{\mathrm{C}}]} \delta_{\mathrm{dg}}^{\mathrm{C}}(i,d) + \sum_{i \in [1, t_{\mathrm{T}}]} \delta_{\mathrm{dg}}^{\mathrm{T}}(i,d) + \sum_{i \in [1, t_{\mathrm{F}}]} \delta_{\mathrm{dg}}^{\mathrm{F}}(i,d) &= \mathrm{dg}(d), \\ \sum_{i \in [1, t_{\mathrm{C}}]} \delta_{\mathrm{dg,C}}^{\mathrm{int}}(i,d) + \sum_{i \in [1, t_{\mathrm{T}}]} \delta_{\mathrm{dg,T}}^{\mathrm{int}}(i,d) + \sum_{i \in [1, t_{\mathrm{F}}]} \delta_{\mathrm{dg,F}}^{\mathrm{int}}(i,d) &= \mathrm{dg}^{\mathrm{int}}(d), \qquad d \in [1,4]. \end{aligned} \tag{37}$$

4.5 Assigning Multiplicity

We prepare an integer variable $\beta(e)$ for each edge e in the scheme graph SG to denote the bond-multiplicity of e in a selected graph H and include necessary constraints for the variables to satisfy in H.

constants:

- $\beta_{\mathrm{r}}([\psi])$: the sum $\beta_\psi(r)$ of bond-multiplicities of edges incident to the root r of a chemical rooted tree $\psi \in \mathcal{F}^*$;

variables:

- $\beta^{\mathrm{X}}(i) \in [0,3]$, $i \in [2, t_{\mathrm{X}}]$, $\mathrm{X} \in \{\mathrm{T}, \mathrm{F}\}$: the bond-multiplicity of edge $e^{\mathrm{X}}{}_i$ in $\mathbb{C}$;
- $\beta^{\mathrm{C}}(i) \in [0,3]$, $i \in [\widetilde{k_{\mathrm{C}}}+1, m_{\mathrm{C}}] = I_{(\geq 1)} \cup I_{(0/1)} \cup I_{(=1)}$: the bond-multiplicity of edge $a_i \in E_{(\geq 1)} \cup E_{(0/1)} \cup E_{(=1)}$ in $\mathbb{C}$;
- $\beta^{\mathrm{CT}}(k), \beta^{\mathrm{TC}}(k) \in [0,3]$, $k \in [1, k_{\mathrm{C}}] = I_{(\geq 2)} \cup I_{(\geq 1)}$: the bond-multiplicity of the first (resp., last) edge of the pure path P_k in $\mathbb{C}$;
- $\beta^{*\mathrm{F}}(c) \in [0,3], c \in [1, c_{\mathrm{F}} = \widetilde{t_{\mathrm{C}}} + t_{\mathrm{T}}]$: the bond-multiplicity of the first edge of the leaf path Q_c rooted at vertex $v^{\mathrm{C}}{}_c, c \leq \widetilde{t_{\mathrm{C}}}$ or $v^{\mathrm{T}}{}_{c-\widetilde{t_{\mathrm{C}}}}, c > \widetilde{t_{\mathrm{C}}}$ in $\mathbb{C}$;
- $\beta^{\mathrm{X}}_{\mathrm{ex}}(i) \in [0,4], i \in [1, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: the sum $\beta_{\mathbb{C}[v]}(v)$ of bond-multiplicities of edges in the ρ-fringe-tree $\mathbb{C}[v]$ rooted at interior-vertex $v = v^{\mathrm{X}}{}_i$;
- $\delta^{\mathrm{X}}_{\beta}(i,m) \in [0,1]$, $i \in [2, t_{\mathrm{X}}]$, $m \in [0,3]$, $\mathrm{X} \in \{\mathrm{T}, \mathrm{F}\}$: $\delta^{\mathrm{X}}_{\beta}(i,m) = 1 \Leftrightarrow \beta^{\mathrm{X}}(i) = m$;
- $\delta^{\mathrm{C}}_{\beta}(i,m) \in [0,1]$, $i \in [\widetilde{k_{\mathrm{C}}}, m_{\mathrm{C}}] = I_{(\geq 1)} \cup I_{(0/1)} \cup I_{(=1)}$, $m \in [0,3]$: $\delta^{\mathrm{C}}_{\beta}(i,m) = 1 \Leftrightarrow \beta^{\mathrm{C}}(i) = m$;
- $\delta^{\mathrm{CT}}_{\beta}(k,m), \delta^{\mathrm{TC}}_{\beta}(k,m) \in [0,1]$, $k \in [1, k_{\mathrm{C}}] = I_{(\geq 2)} \cup I_{(\geq 1)}$, $m \in [0,3]$: $\delta^{\mathrm{CT}}_{\beta}(k,m) = 1$ (resp., $\delta^{\mathrm{TC}}_{\beta}(k,m) = 1$) $\Leftrightarrow$ $\beta^{\mathrm{CT}}(k) = m$ (resp., $\beta^{\mathrm{TC}}(k) = m$);
- $\delta^{*\mathrm{F}}_{\beta}(c,m) \in [0,1]$, $c \in [1, c_{\mathrm{F}}]$, $m \in [0,3], \mathrm{X} \in \{\mathrm{C}, \mathrm{T}\}$: $\delta^{*\mathrm{F}}_{\beta}(c,m) = 1 \Leftrightarrow \beta^{*\mathrm{F}}(c) = m$;
- $\mathrm{bd}^{\mathrm{int}}(m) \in [0, 2\mathrm{n}^{\mathrm{int}}_{\mathrm{UB}}]$, $m \in [1,3]$: the number of interior-edges with bond-multiplicity m in $\mathbb{C}$;
- $\mathrm{bd}_{\mathrm{X}}(m) \in [0, 2\mathrm{n}^{\mathrm{int}}_{\mathrm{UB}}], \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{CT}, \mathrm{TC}\}$, $\mathrm{bd}_{\mathrm{X}}(m) \in [0, 2\mathrm{n}^{\mathrm{int}}_{\mathrm{UB}}], \mathrm{X} \in \{\mathrm{F}, \mathrm{CF}, \mathrm{TF}\}$, $m \in [1,3]$: the number of interior-edges $e \in E_{\mathrm{X}}$ with bond-multiplicity m in $\mathbb{C}$;

constraints:

$$e^{\mathrm{C}}(i) \leq \beta^{\mathrm{C}}(i) \leq 3e^{\mathrm{C}}(i), i \in [\widetilde{k_{\mathrm{C}}}+1, m_{\mathrm{C}}] = I_{(\geq 1)} \cup I_{(0/1)} \cup I_{(=1)}, \tag{38}$$

$$e^{\mathrm{X}}(i) \leq \beta^{\mathrm{X}}(i) \leq 3e^{\mathrm{X}}(i), \qquad i \in [2, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{T}, \mathrm{F}\}, \tag{39}$$

$$\delta^{\mathrm{T}}_{\chi}(k) \leq \beta^{\mathrm{CT}}(k) \leq 3\delta^{\mathrm{T}}_{\chi}(k), \quad \delta^{\mathrm{T}}_{\chi}(k) \leq \beta^{\mathrm{TC}}(k) \leq 3\delta^{\mathrm{T}}_{\chi}(k), \qquad k \in [1, k_{\mathrm{C}}], \tag{40}$$

$$\delta^{\mathrm{F}}_{\chi}(c) \leq \beta^{\mathrm{XF}}(c) \leq 3\delta^{\mathrm{F}}_{\chi}(c), \qquad c \in [1, c_{\mathrm{F}}], \tag{41}$$

$$\sum_{m \in [0,3]} \delta^{\mathrm{X}}_{\beta}(i,m) = 1, \quad \sum_{m \in [0,3]} m \cdot \delta^{\mathrm{X}}_{\beta}(i,m) = \beta^{\mathrm{X}}(i), \qquad i \in [2, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{T}, \mathrm{F}\}, \tag{42}$$

$$\sum_{m\in[0,3]} \delta_\beta^{\mathrm{C}}(i,m)=1, \quad \sum_{m\in[0,3]} m\cdot\delta_\beta^{\mathrm{C}}(i,m)=\beta^{\mathrm{C}}(i), \qquad i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}], \tag{43}$$

$$\begin{aligned}
&\sum_{m\in[0,3]} \delta_\beta^{\mathrm{CT}}(k,m)=1, \quad \sum_{m\in[0,3]} m\cdot\delta_\beta^{\mathrm{CT}}(k,m)=\beta^{\mathrm{CT}}(k), && k\in[1,k_{\mathrm{C}}],\\
&\sum_{m\in[0,3]} \delta_\beta^{\mathrm{TC}}(k,m)=1, \quad \sum_{m\in[0,3]} m\cdot\delta_\beta^{\mathrm{TC}}(k,m)=\beta^{\mathrm{TC}}(k), && k\in[1,k_{\mathrm{C}}],\\
&\sum_{m\in[0,3]} \delta_\beta^{*\mathrm{F}}(c,m)=1, \quad \sum_{m\in[0,3]} m\cdot\delta_\beta^{*\mathrm{F}}(c,m)=\beta^{*\mathrm{F}}(c), && c\in[1,c_{\mathrm{F}}],
\end{aligned} \tag{44}$$

$$\sum_{\psi\in\mathcal{F}_i^{\mathrm{X}}} \beta_{\mathrm{r}}([\psi])\cdot\delta_{\mathrm{fr}}^{\mathrm{X}}(i,[\psi])=\beta_{\mathrm{ex}}^{\mathrm{X}}(i), \qquad i\in[1,t_{\mathrm{X}}],\mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F}\}, \tag{45}$$

$$\begin{aligned}
\sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]} \delta_\beta^{\mathrm{C}}(i,m)=\mathrm{bd}_{\mathrm{C}}(m), \quad \sum_{i\in[2,t_{\mathrm{T}}]} \delta_\beta^{\mathrm{T}}(i,m)=\mathrm{bd}_{\mathrm{T}}(m),\\
\sum_{k\in[1,k_{\mathrm{C}}]} \delta_\beta^{\mathrm{CT}}(k,m)=\mathrm{bd}_{\mathrm{CT}}(m), \quad \sum_{k\in[1,k_{\mathrm{C}}]} \delta_\beta^{\mathrm{TC}}(k,m)=\mathrm{bd}_{\mathrm{TC}}(m),\\
\sum_{i\in[2,t_{\mathrm{F}}]} \delta_\beta^{\mathrm{F}}(i,m)=\mathrm{bd}_{\mathrm{F}}(m), \quad \sum_{c\in[1,\widetilde{t_{\mathrm{C}}}]} \delta_\beta^{*\mathrm{F}}(c,m)=\mathrm{bd}_{\mathrm{CF}}(m),\\
\sum_{c\in[\widetilde{t_{\mathrm{C}}}+1,c_{\mathrm{F}}]} \delta_\beta^{*\mathrm{F}}(c,m)=\mathrm{bd}_{\mathrm{TF}}(m),\\
\mathrm{bd}_{\mathrm{C}}(m)+\mathrm{bd}_{\mathrm{T}}(m)+\mathrm{bd}_{\mathrm{F}}(m)+\mathrm{bd}_{\mathrm{CT}}(m)+\mathrm{bd}_{\mathrm{TC}}(m)+\mathrm{bd}_{\mathrm{TF}}(m)+\mathrm{bd}_{\mathrm{CF}}(m)=\mathrm{bd}^{\mathrm{int}}(m),\\
m\in[1,3].
\end{aligned} \tag{46}$$

4.6 Assigning Chemical Elements and Valence Condition

We include constraints so that each vertex v in a selected graph H satisfies the valence condition; i.e., $\beta_{\mathbb{C}}(v) = \mathrm{val}(\alpha(v)) + \mathrm{eledeg}_{\mathbb{C}}(v)$, where $\mathrm{eledeg}_{\mathbb{C}}(v) = \mathrm{v_{ion}}(\psi)$ for the ρ-fringe-tree $\mathbb{C}[v]$ r-isomorphic to ψ. With these constraints, a chemical graph $\mathbb{C} = (H, \alpha, \beta)$ on a selected subgraph H will be constructed.

constants:

- Subsets $\Lambda^{\mathrm{int}} \subseteq \Lambda \setminus \{\mathtt{H}\}, \Lambda^{\mathrm{ex}} \subseteq \Lambda$ of chemical elements, where we denote by $[\mathtt{e}]$ (resp., $[\mathtt{e}]^{\mathrm{int}}$ and $[\mathtt{e}]^{\mathrm{ex}}$) of a standard encoding of an element $\mathtt{e}$ in the set Λ (resp., $\Lambda_\epsilon^{\mathrm{int}}$ and $\Lambda_\epsilon^{\mathrm{ex}}$);
- A valence function: $\mathrm{val} : \Lambda \to [1, 6]$;
- A function $\mathrm{mass}^* : \Lambda \to \mathbb{Z}$ (we let $\mathrm{mass}(\mathtt{a})$ denote the observed mass of a chemical element $\mathtt{a} \in \Lambda$, and define $\mathrm{mass}^*(\mathtt{a}) \triangleq \lfloor 10 \cdot \mathrm{mass}(\mathtt{a}) \rfloor$);
- Subsets $\Lambda^*(i) \subseteq \Lambda^{\mathrm{int}}$, $i \in [1, t_{\mathrm{C}}]$;
- $\mathrm{na_{LB}}(\mathtt{a}), \mathrm{na_{UB}}(\mathtt{a}) \in [0, n^*]$, $\mathtt{a} \in \Lambda$: lower and upper bounds on the number of vertices v with $\alpha(v) = \mathtt{a}$;
- $\mathrm{na_{LB}^{int}}(\mathtt{a}), \mathrm{na_{UB}^{int}}(\mathtt{a}) \in [0, n^*]$, $\mathtt{a} \in \Lambda^{\mathrm{int}}$: lower and upper bounds on the number of interior-vertices v with $\alpha(v) = \mathtt{a}$;

- $\alpha_{\rm r}([\psi]) \in [\Lambda^{\rm ex}], \in \mathcal{F}^*$: the chemical element $\alpha(r)$ of the root r of ψ;

- $\mathrm{na}^{\rm ex}_{\tt a}([\psi]) \in [0, n^*]$, ${\tt a} \in \Lambda^{\rm ex}, \psi \in \mathcal{F}^*$: the frequency of chemical element ${\tt a}$ in the set of non-rooted vertices in ψ, where possibly ${\tt a} = {\tt H}$;

- M: an upper bound for the average $\overline{\rm ms}(\mathbb{C})$ of mass* over all atoms in $\mathbb{C}$;

variables:

- $\beta^{\rm CT}(i), \beta^{\rm TC}(i) \in [0, 3], i \in [1, t_{\rm T}]$: the bond-multiplicity of edge $e^{\rm CT}{}_{j,i}$ (resp., $e^{\rm TC}{}_{j,i}$) if one exists;

- $\beta^{\rm CF}(i), \beta^{\rm TF}(i) \in [0, 3], i \in [1, t_{\rm F}]$: the bond-multiplicity of $e^{\rm CF}{}_{j,i}$ (resp., $e^{\rm TF}{}_{j,i}$) if one exists;

- $\alpha^{\rm X}(i) \in [\Lambda^{\rm int}_\epsilon], \delta^{\rm X}_\alpha(i, [{\tt a}]^{\rm int}) \in [0, 1], {\tt a} \in \Lambda^{\rm int}_\epsilon, i \in [1, t_{\rm X}], {\rm X} \in \{{\rm C}, {\rm T}, {\rm F}\}$: $\alpha^{\rm X}(i) = [{\tt a}]^{\rm int} \geq 1$ (resp., $\alpha^{\rm X}(i) = 0$) $\Leftrightarrow$ $\delta^{\rm X}_\alpha(i, [{\tt a}]^{\rm int}) = 1$ (resp., $\delta^{\rm X}_\alpha(i, 0) = 0$) $\Leftrightarrow \alpha(v^{\rm X}{}_i) = {\tt a} \in \Lambda$ (resp., vertex $v^{\rm X}{}_i$ is not used in $\mathbb{C}$);

- $\delta^{\rm X}_\alpha(i, [{\tt a}]^{\rm int}) \in [0, 1], i \in [1, t_{\rm X}], {\tt a} \in \Lambda^{\rm int}, {\rm X} \in \{{\rm C}, {\rm T}, {\rm F}\}$: $\delta^{\rm X}_\alpha(i, [{\tt a}]^{\rm t}) = 1 \Leftrightarrow \alpha(v^{\rm X}{}_i) = {\tt a}$;

- ${\rm Mass} \in \mathbb{Z}_+$: $\sum_{v \in V(H)} {\rm mass}^*(\alpha(v))$;

- $\overline{\rm ms} \in \mathbb{R}_+$: $\sum_{v \in V(H)} {\rm mass}^*(\alpha(v))/|V(H)|$;

- $\delta_{\rm atm}(i) \in [0, 1], i \in [n_{\rm LB} + {\rm na}_{\rm LB}({\tt H}), n^* + {\rm na}_{\rm UB}({\tt H})]$: $\delta_{\rm atm}(i) = 1 \Leftrightarrow |V(H)| = i$;

- ${\rm na}([{\tt a}]) \in [{\rm na}_{\rm LB}({\tt a}), {\rm na}_{\rm UB}({\tt a})]$, ${\tt a} \in \Lambda$: the number of vertices $v \in V(H)$ with $\alpha(v) = {\tt a}$, where possibly ${\tt a} = {\tt H}$;

- ${\rm na}^{\rm int}([{\tt a}]^{\rm int}) \in [{\rm na}^{\rm int}_{\rm LB}({\tt a}), {\rm na}^{\rm int}_{\rm UB}({\tt a})]$, ${\tt a} \in \Lambda, {\rm X} \in \{{\rm C}, {\rm T}, {\rm F}\}$: the number of interior-vertices $v \in V(\mathbb{C})$ with $\alpha(v) = {\tt a}$;

- ${\rm na}^{\rm ex}_{\rm X}([{\tt a}]^{\rm ex}), {\rm na}^{\rm ex}([{\tt a}]^{\rm ex}) \in [0, {\rm na}_{\rm UB}({\tt a})]$, ${\tt a} \in \Lambda$, ${\rm X} \in \{{\rm C}, {\rm T}, {\rm F}\}$: the number of exterior-vertices rooted at vertices $v \in V_{\rm X}$ and the number of exterior-vertices v such that $\alpha(v) = {\tt a}$;

constraints:

$$\begin{aligned}\beta^{\rm CT}(k) - 3(e^{\rm T}(i) - \chi^{\rm T}(i, k) + 1) \leq \beta^{\rm CT}(i) \leq \beta^{\rm CT}(k) + 3(e^{\rm T}(i) - \chi^{\rm T}(i, k) + 1), i \in [1, t_{\rm T}],\\ \beta^{\rm TC}(k) - 3(e^{\rm T}(i+1) - \chi^{\rm T}(i, k) + 1) \leq \beta^{\rm TC}(i) \leq \beta^{\rm TC}(k) + 3(e^{\rm T}(i+1) - \chi^{\rm T}(i, k) + 1), i \in [1, t_{\rm T}],\\ k \in [1, k_{\rm C}],\end{aligned} \tag{47}$$

$$\begin{aligned}\beta^{*{\rm F}}(c) - 3(e^{\rm F}(i) - \chi^{\rm F}(i, c) + 1) \leq \beta^{\rm CF}(i) \leq \beta^{*{\rm F}}(c) + 3(e^{\rm F}(i) - \chi^{\rm F}(i, c) + 1), i \in [1, t_{\rm F}], \quad c \in [1, \widetilde{t_{\rm C}}],\\ \beta^{*{\rm F}}(c) - 3(e^{\rm F}(i) - \chi^{\rm F}(i, c) + 1) \leq \beta^{\rm TF}(i) \leq \beta^{*{\rm F}}(c) + 3(e^{\rm F}(i) - \chi^{\rm F}(i, c) + 1), i \in [1, t_{\rm F}], \quad c \in [\widetilde{t_{\rm C}} + 1, c_{\rm F}],\end{aligned} \tag{48}$$

$$\begin{aligned}\sum_{{\tt a} \in \Lambda^{\rm int}} \delta^{\rm C}_\alpha(i, [{\tt a}]^{\rm int}) = 1, \quad \sum_{{\tt a} \in \Lambda^{\rm int}} [{\tt a}]^{\rm int} \cdot \delta^{\rm X}_\alpha(i, [{\tt a}]^{\rm int}) = \alpha^{\rm C}(i), \quad & i \in [1, t_{\rm C}],\\ \sum_{{\tt a} \in \Lambda^{\rm int}} \delta^{\rm X}_\alpha(i, [{\tt a}]^{\rm int}) = v^{\rm X}(i), \quad \sum_{{\tt a} \in \Lambda^{\rm int}} [{\tt a}]^{\rm int} \cdot \delta^{\rm X}_\alpha(i, [{\tt a}]^{\rm int}) = \alpha^{\rm X}(i), \quad & i \in [1, t_{\rm X}], {\rm X} \in \{{\rm T}, {\rm F}\},\end{aligned} \tag{49}$$

$$\sum_{\psi\in\mathcal{F}_i^{\mathrm{X}}} \alpha_{\mathrm{r}}([\psi])\cdot\delta_{\mathrm{fr}}^{\mathrm{X}}(i,[\psi]) = \alpha^{\mathrm{X}}(i), \qquad i\in[1,t_{\mathrm{X}}], \mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F}\}, \tag{50}$$

$$\begin{aligned}\sum_{j\in I_{\mathrm{C}}(i)}\beta^{\mathrm{C}}(j) + \sum_{k\in I^+_{(\geq 2)}(i)\cup I^+_{(\geq 1)}(i)}\beta^{\mathrm{CT}}(k) + \sum_{k\in I^-_{(\geq 2)}(i)\cup I^-_{(\geq 1)}(i)}\beta^{\mathrm{TC}}(k)&\\ +\beta^{*\mathrm{F}}(i) + \beta^{\mathrm{C}}_{\mathrm{ex}}(i) - \mathrm{eledeg}_{\mathrm{C}}(i) = \sum_{\mathtt{a}\in\Lambda^{\mathrm{int}}}\mathrm{val}(\mathtt{a})\delta^{\mathrm{C}}_{\alpha}(i,[\mathtt{a}]^{\mathrm{int}}),& \qquad i\in[1,\widetilde{t_{\mathrm{C}}}],\end{aligned} \tag{51}$$

$$\begin{aligned}\sum_{j\in I_{\mathrm{C}}(i)}\beta^{\mathrm{C}}(j) + \sum_{k\in I^+_{(\geq 2)}(i)\cup I^+_{(\geq 1)}(i)}\beta^{\mathrm{CT}}(k) + \sum_{k\in I^-_{(\geq 2)}(i)\cup I^-_{(\geq 1)}(i)}\beta^{\mathrm{TC}}(k)&\\ +\beta^{\mathrm{C}}_{\mathrm{ex}}(i) - \mathrm{eledeg}_{\mathrm{C}}(i) = \sum_{\mathtt{a}\in\Lambda^{\mathrm{int}}}\mathrm{val}(\mathtt{a})\delta^{\mathrm{C}}_{\alpha}(i,[\mathtt{a}]^{\mathrm{int}}),& \qquad i\in[\widetilde{t_{\mathrm{C}}}+1,t_{\mathrm{C}}],\end{aligned} \tag{52}$$

$$\begin{aligned}\beta^{\mathrm{T}}(i)+\beta^{\mathrm{T}}(i{+}1)+\beta^{\mathrm{T}}_{\mathrm{ex}}(i)+\beta^{\mathrm{CT}}(i)+\beta^{\mathrm{TC}}(i)&\\ +\beta^{*\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i)-\mathrm{eledeg}_{\mathrm{T}}(i) = \sum_{\mathtt{a}\in\Lambda^{\mathrm{int}}}\mathrm{val}(\mathtt{a})\delta^{\mathrm{T}}_{\alpha}(i,[\mathtt{a}]^{\mathrm{int}}),&\\ i\in[1,t_{\mathrm{T}}]\ (\beta^{\mathrm{T}}(1)=\beta^{\mathrm{T}}(t_{\mathrm{T}}+1)=0),&\end{aligned} \tag{53}$$

$$\begin{aligned}\beta^{\mathrm{F}}(i)+\beta^{\mathrm{F}}(i{+}1)+\beta^{\mathrm{CF}}(i)+\beta^{\mathrm{TF}}(i)&\\ +\beta^{\mathrm{F}}_{\mathrm{ex}}(i)-\mathrm{eledeg}_{\mathrm{F}}(i) = \sum_{\mathtt{a}\in\Lambda^{\mathrm{int}}}\mathrm{val}(\mathtt{a})\delta^{\mathrm{F}}_{\alpha}(i,[\mathtt{a}]^{\mathrm{int}}),&\\ i\in[1,t_{\mathrm{F}}]\ (\beta^{\mathrm{F}}(1)=\beta^{\mathrm{F}}(t_{\mathrm{F}}+1)=0),&\end{aligned} \tag{54}$$

$$\sum_{i\in[1,t_{\mathrm{X}}]}\delta^{\mathrm{X}}_{\alpha}(i,[\mathtt{a}]^{\mathrm{int}}) = \mathrm{na}_{\mathrm{X}}([\mathtt{a}]^{\mathrm{int}}), \qquad \mathtt{a}\in\Lambda^{\mathrm{int}}, \mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F}\}, \tag{55}$$

$$\sum_{\psi\in\mathcal{F}_i^{\mathrm{X}}, i\in[1,t_{\mathrm{X}}]}\mathrm{na}^{\mathrm{ex}}_{\mathtt{a}}([\psi])\cdot\delta^{\mathrm{X}}_{\mathrm{fr}}(i,[\psi]) = \mathrm{na}^{\mathrm{ex}}_{\mathrm{X}}([\mathtt{a}]^{\mathrm{ex}}), \qquad \mathtt{a}\in\Lambda^{\mathrm{ex}}, \mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F}\}, \tag{56}$$

$$\begin{aligned}\mathrm{na}_{\mathrm{C}}([\mathtt{a}]^{\mathrm{int}})+\mathrm{na}_{\mathrm{T}}([\mathtt{a}]^{\mathrm{int}})+\mathrm{na}_{\mathrm{F}}([\mathtt{a}]^{\mathrm{int}}) = \mathrm{na}^{\mathrm{int}}([\mathtt{a}]^{\mathrm{int}}),&\qquad \mathtt{a}\in\Lambda^{\mathrm{int}},\\ \sum_{\mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F}\}}\mathrm{na}^{\mathrm{ex}}_{\mathrm{X}}([\mathtt{a}]^{\mathrm{ex}}) = \mathrm{na}^{\mathrm{ex}}([\mathtt{a}]^{\mathrm{ex}}),&\qquad \mathtt{a}\in\Lambda^{\mathrm{ex}},\\ \mathrm{na}^{\mathrm{int}}([\mathtt{a}]^{\mathrm{int}})+\mathrm{na}^{\mathrm{ex}}([\mathtt{a}]^{\mathrm{ex}}) = \mathrm{na}([\mathtt{a}]),&\qquad \mathtt{a}\in\Lambda^{\mathrm{int}}\cap\Lambda^{\mathrm{ex}},\\ \mathrm{na}^{\mathrm{int}}([\mathtt{a}]^{\mathrm{int}}) = \mathrm{na}([\mathtt{a}]),&\qquad \mathtt{a}\in\Lambda^{\mathrm{int}}\setminus\Lambda^{\mathrm{ex}},\\ \mathrm{na}^{\mathrm{ex}}([\mathtt{a}]^{\mathrm{ex}}) = \mathrm{na}([\mathtt{a}]),&\qquad \mathtt{a}\in\Lambda^{\mathrm{ex}}\setminus\Lambda^{\mathrm{int}},\end{aligned} \tag{57}$$

$$\sum_{\mathtt{a}\in\Lambda^*(i)}\delta^{\mathrm{C}}_{\alpha}(i,[\mathtt{a}]^{\mathrm{int}}) = 1, \qquad i\in[1,t_{\mathrm{C}}], \tag{58}$$

$$\sum_{\mathtt{a}\in\Lambda} \mathrm{mass}^*(\mathtt{a}) \cdot \mathrm{na}([\mathtt{a}]) = \mathrm{Mass}, \tag{59}$$

$$\sum_{i\in[n_{\mathrm{LB}}+\mathrm{na}_{\mathrm{LB}}(\mathtt{H}),n^*+\mathrm{na}_{\mathrm{UB}}(\mathtt{H})]} \delta_{\mathrm{atm}}(i) = 1, \tag{60}$$

$$\sum_{i\in[n_{\mathrm{LB}}+\mathrm{na}_{\mathrm{LB}}(\mathtt{H}),n^*+\mathrm{na}_{\mathrm{UB}}(\mathtt{H})]} i\cdot\delta_{\mathrm{atm}}(i) = n_G + \mathrm{na}^{\mathrm{ex}}([\mathtt{H}]^{\mathrm{ex}}), \tag{61}$$

$$\mathrm{Mass}/i - \mathrm{M}\cdot(1-\delta_{\mathrm{atm}}(i)) \le \overline{\mathrm{ms}} \le \mathrm{Mass}/i + \mathrm{M}\cdot(1-\delta_{\mathrm{atm}}(i)), \quad i\in[n_{\mathrm{LB}}+\mathrm{na}_{\mathrm{LB}}(\mathtt{H}), n^*+\mathrm{na}_{\mathrm{UB}}(\mathtt{H})]. \tag{62}$$

4.7 Constraints for Bounds on the Number of Bonds

We include constraints for specification of lower and upper bounds $\mathrm{bd}_{\mathrm{LB}}$ and $\mathrm{bd}_{\mathrm{UB}}$.

constants:

- $\mathrm{bd}_{m,\mathrm{LB}}(i), \mathrm{bd}_{m,\mathrm{UB}}(i) \in [0, \mathrm{n}^{\mathrm{int}}_{\mathrm{UB}}]$, $i\in[1,m_{\mathrm{C}}]$, $m\in[2,3]$: lower and upper bounds on the number of edges $e\in E(P_i)$ with bond-multiplicity $\beta(e)=m$ in the pure path P_i for edge $e_i\in E_{\mathrm{C}}$;

variables :

- $\mathrm{bd}_{\mathrm{T}}(k,i,m)\in[0,1]$, $k\in[1,k_{\mathrm{C}}]$, $i\in[2,t_{\mathrm{T}}]$, $m\in[2,3]$: $\mathrm{bd}_{\mathrm{T}}(k,i,m)=1 \Leftrightarrow$ the pure path P_k for edge $e_k\in E_{\mathrm{C}}$ contains edge ${e^{\mathrm{T}}}_i$ with $\beta({e^{\mathrm{T}}}_i)=m$;

constraints:

$$\mathrm{bd}_{m,\mathrm{LB}}(i) \le \delta^{\mathrm{C}}_{\beta}(i,m) \le \mathrm{bd}_{m,\mathrm{UB}}(i), i\in I_{(=1)}\cup I_{(0/1)}, m\in[2,3], \tag{63}$$

$$\mathrm{bd}_{\mathrm{T}}(k,i,m) \ge \delta^{\mathrm{T}}_{\beta}(i,m) + \chi^{\mathrm{T}}(i,k) - 1, \quad k\in[1,k_{\mathrm{C}}], i\in[2,t_{\mathrm{T}}], m\in[2,3], \tag{64}$$

$$\sum_{j\in[2,t_{\mathrm{T}}]} \delta^{\mathrm{T}}_{\beta}(j,m) \ge \sum_{k\in[1,k_{\mathrm{C}}],i\in[2,t_{\mathrm{T}}]} \mathrm{bd}_{\mathrm{T}}(k,i,m), \quad m\in[2,3], \tag{65}$$

$$\mathrm{bd}_{m,\mathrm{LB}}(k) \le \sum_{i\in[2,t_{\mathrm{T}}]} \mathrm{bd}_{\mathrm{T}}(k,i,m) + \delta^{\mathrm{CT}}_{\beta}(k,m) + \delta^{\mathrm{TC}}_{\beta}(k,m) \le \mathrm{bd}_{m,\mathrm{UB}}(k),$$
$$k\in[1,k_{\mathrm{C}}], m\in[2,3]. \tag{66}$$

4.8 Descriptor for the Number of Adjacency-configurations

We call a tuple $(\mathtt{a},\mathtt{b},m)\in(\Lambda\setminus\{\mathtt{H}\})\times(\Lambda\setminus\{\mathtt{H}\})\times[1,3]$ an *adjacency-configuration*. The adjacency-configuration of an edge-configuration $(\mu=\mathtt{a}d, \mu'=\mathtt{b}d', m)$ is defined to be $(\mathtt{a},\mathtt{b},m)$. We include constraints to compute the frequency of each adjacency-configuration in an inferred chemical graph $\mathbb{C}$.

constants:

- A set Γ^{int} of edge-configurations $\gamma=(\mu,\mu',m)$ with $\mu\le\mu'$;

- Let $\overline{\gamma}$ of an edge-configuration $\gamma = (\mu, \mu', m)$ denote the edge-configuration (μ', μ, m);

- Let $\Gamma^{\rm int}_{<} = \{(\mu, \mu', m) \in \Gamma^{\rm int} \mid \mu < \mu'\}$, $\Gamma^{\rm int}_{=} = \{(\mu, \mu', m) \in \Gamma^{\rm int} \mid \mu = \mu'\}$ and $\Gamma^{\rm int}_{>} = \{\overline{\gamma} \mid \gamma \in \Gamma^{\rm int}_{<}\}$;

- Let $\Gamma^{\rm int}_{{\rm ac},<}$, $\Gamma^{\rm int}_{{\rm ac},=}$ and $\Gamma^{\rm int}_{{\rm ac},>}$ denote the sets of the adjacency-configurations of edge-configurations in the sets $\Gamma^{\rm int}_{<}$, $\Gamma^{\rm int}_{=}$ and $\Gamma^{\rm int}_{>}$, respectively;

- Let $\overline{\nu}$ of an adjacency-configuration $\nu = (\mathtt{a}, \mathtt{b}, m)$ denote the adjacency-configuration $(\mathtt{b}, \mathtt{a}, m)$;

- Prepare a coding of the set $\Gamma^{\rm int}_{\rm ac} \cup \Gamma^{\rm int}_{{\rm ac},>}$ and let $[\nu]^{\rm int}$ denote the coded integer of an element ν in $\Gamma^{\rm int}_{\rm ac} \cup \Gamma^{\rm int}_{{\rm ac},>}$;

- Choose subsets $\widetilde{\Gamma}^{\rm C}_{\rm ac}, \widetilde{\Gamma}^{\rm T}_{\rm ac}, \widetilde{\Gamma}^{\rm CT}_{\rm ac}, \widetilde{\Gamma}^{\rm TC}_{\rm ac}, \widetilde{\Gamma}^{\rm F}_{\rm ac}, \widetilde{\Gamma}^{\rm CF}_{\rm ac}, \widetilde{\Gamma}^{\rm TF}_{\rm ac} \subseteq \Gamma^{\rm int}_{\rm ac} \cup \Gamma^{\rm int}_{{\rm ac},>}$; To compute the frequency of adjacency-configurations exactly, set $\widetilde{\Gamma}^{\rm C}_{\rm ac} := \widetilde{\Gamma}^{\rm T}_{\rm ac} := \widetilde{\Gamma}^{\rm CT}_{\rm ac} := \widetilde{\Gamma}^{\rm TC}_{\rm ac} := \widetilde{\Gamma}^{\rm F}_{\rm ac} := \widetilde{\Gamma}^{\rm CF}_{\rm ac} := \widetilde{\Gamma}^{\rm TF}_{\rm ac} := \Gamma^{\rm int}_{\rm ac} \cup \Gamma^{\rm int}_{{\rm ac},>}$;

- $\mathrm{ac}^{\rm int}_{\rm LB}(\nu), \mathrm{ac}^{\rm int}_{\rm UB}(\nu) \in [0, 2\mathrm{n}^{\rm int}_{\rm UB}], \nu = (\mathtt{a}, \mathtt{b}, m) \in \Gamma^{\rm int}_{\rm ac}$: lower and upper bounds on the number of interior-edges $e = uv$ with $\alpha(u) = \mathtt{a}$, $\alpha(v) = \mathtt{b}$ and $\beta(e) = m$;

variables:

- $\mathrm{ac}^{\rm int}([\nu]^{\rm int}) \in [\mathrm{ac}^{\rm int}_{\rm LB}(\nu), \mathrm{ac}^{\rm int}_{\rm UB}(\nu)], \nu \in \Gamma^{\rm int}_{\rm ac}$: the number of interior-edges with adjacency-configuration ν;

- $\mathrm{ac}_{\rm C}([\nu]^{\rm int}) \in [0, m_{\rm C}], \nu \in \widetilde{\Gamma}^{\rm C}_{\rm ac}$, $\mathrm{ac}_{\rm T}([\nu]^{\rm int}) \in [0, t_{\rm T}], \nu \in \widetilde{\Gamma}^{\rm T}_{\rm ac}$, $\mathrm{ac}_{\rm F}([\nu]^{\rm int}) \in [0, t_{\rm F}], \nu \in \widetilde{\Gamma}^{\rm F}_{\rm ac}$: the number of edges $e^{\rm C} \in E_{\rm C}$ (resp., edges $e^{\rm T} \in E_{\rm T}$ and edges $e^{\rm F} \in E_{\rm F}$) with adjacency-configuration ν;

- $\mathrm{ac}_{\rm CT}([\nu]^{\rm int}) \in [0, \min\{k_{\rm C}, t_{\rm T}\}], \nu \in \widetilde{\Gamma}^{\rm CT}_{\rm ac}$, $\mathrm{ac}_{\rm TC}([\nu]^{\rm int}) \in [0, \min\{k_{\rm C}, t_{\rm T}\}], \nu \in \widetilde{\Gamma}^{\rm CT}_{\rm ac}$, $\mathrm{ac}_{\rm CF}([\nu]^{\rm int}) \in [0, \widetilde{t_{\rm C}}], \nu \in \widetilde{\Gamma}^{\rm CF}_{\rm ac}$, $\mathrm{ac}_{\rm TF}([\nu]^{\rm int}) \in [0, t_{\rm T}], \nu \in \widetilde{\Gamma}^{\rm TF}_{\rm ac}$: the number of edges $e^{\rm CT} \in E_{\rm CT}$ (resp., edges $e^{\rm TC} \in E_{\rm TC}$ and edges $e^{\rm CF} \in E_{\rm CF}$ and $e^{\rm TF} \in E_{\rm TF}$) with adjacency-configuration ν;

- $\delta^{\rm C}_{\rm ac}(i, [\nu]^{\rm int}) \in [0, 1], i \in [\widetilde{k_{\rm C}} + 1, m_{\rm C}] = I_{(\geq 1)} \cup I_{(0/1)} \cup I_{(=1)}, \nu \in \widetilde{\Gamma}^{\rm C}_{\rm ac}$, $\delta^{\rm T}_{\rm ac}(i, [\nu]^{\rm int}) \in [0, 1], i \in [2, t_{\rm T}], \nu \in \widetilde{\Gamma}^{\rm T}_{\rm ac}$, $\delta^{\rm F}_{\rm ac}(i, [\nu]^{\rm int}) \in [0, 1], i \in [2, t_{\rm F}], \nu \in \widetilde{\Gamma}^{\rm F}_{\rm ac}$: $\delta^{\rm X}_{\rm ac}(i, [\nu]^{\rm int}) = 1 \Leftrightarrow$ edge ${e^{\rm X}}_i$ has adjacency-configuration ν;

- $\delta^{\rm CT}_{\rm ac}(k, [\nu]^{\rm int}), \delta^{\rm TC}_{\rm ac}(k, [\nu]^{\rm int}) \in [0, 1], k \in [1, k_{\rm C}] = I_{(\geq 2)} \cup I_{(\geq 1)}, \nu \in \widetilde{\Gamma}^{\rm CT}_{\rm ac}$: $\delta^{\rm CT}_{\rm ac}(k, [\nu]^{\rm int}) = 1$ (resp., $\delta^{\rm TC}_{\rm ac}(k, [\nu]^{\rm int}) = 1$) $\Leftrightarrow$ edge ${e^{\rm CT}}_{{\rm tail}(k),j}$ (resp., ${e^{\rm TC}}_{{\rm head}(k),j}$) for some $j \in [1, t_{\rm T}]$ has adjacency-configuration ν;

- $\delta^{\rm CF}_{\rm ac}(c, [\nu]^{\rm int}) \in [0, 1], c \in [1, \widetilde{t_{\rm C}}], \nu \in \widetilde{\Gamma}^{\rm CF}_{\rm ac}$: $\delta^{\rm CF}_{\rm ac}(c, [\nu]^{\rm int}) = 1 \Leftrightarrow$ edge ${e^{\rm CF}}_{c,i}$ for some $i \in [1, t_{\rm F}]$ has adjacency-configuration ν;

- $\delta^{\rm TF}_{\rm ac}(i, [\nu]^{\rm int}) \in [0, 1], i \in [1, t_{\rm T}], \nu \in \widetilde{\Gamma}^{\rm TF}_{\rm ac}$: $\delta^{\rm TF}_{\rm ac}(i, [\nu]^{\rm int}) = 1 \Leftrightarrow$ edge ${e^{\rm TF}}_{i,j}$ for some $j \in [1, t_{\rm F}]$ has adjacency-configuration ν;

- $\alpha^{\rm CT}(k), \alpha^{\rm TC}(k) \in [0, |\Lambda^{\rm int}|], k \in [1, k_{\rm C}]$: $\alpha(v)$ of the edge $({v^{\rm C}}_{{\rm tail}(k)}, v) \in E_{\rm CT}$ (resp., $(v, {v^{\rm C}}_{{\rm head}(k)}) \in E_{\rm TC}$) if any;

- $\alpha^{\rm CF}(c) \in [0, |\Lambda^{\rm int}|], c \in [1, \widetilde{t_{\rm C}}]$: $\alpha(v)$ of the edge $({v^{\rm C}}_c, v) \in E_{\rm CF}$ if any;

- $\alpha^{\rm TF}(i) \in [0, |\Lambda^{\rm int}|], i \in [1, t_{\rm T}]$: $\alpha(v)$ of the edge $({v^{\rm T}}_i, v) \in E_{\rm TF}$ if any;

- $\Delta^{\rm C+}_{\rm ac}(i), \Delta^{\rm C-}_{\rm ac}(i), \in [0, |\Lambda^{\rm int}|], i \in [\widetilde{k_{\rm C}} + 1, m_{\rm C}]$, $\Delta^{\rm T+}_{\rm ac}(i), \Delta^{\rm T-}_{\rm ac}(i) \in [0, |\Lambda^{\rm int}|], i \in [2, t_{\rm T}]$, $\Delta^{\rm F+}_{\rm ac}(i), \Delta^{\rm F-}_{\rm ac}(i) \in [0, |\Lambda^{\rm int}|], i \in [2, t_{\rm F}]$: $\Delta^{\rm X+}_{\rm ac}(i) = \Delta^{\rm X-}_{\rm ac}(i) = 0$ (resp., $\Delta^{\rm X+}_{\rm ac}(i) = \alpha(u)$ and $\Delta^{\rm X-}_{\rm ac}(i) = \alpha(v)$) $\Leftrightarrow$ edge ${e^{\rm X}}_i = (u, v) \in E_{\rm X}$ is used in $\mathbb{C}$ (resp., ${e^{\rm X}}_i \not\in E(G)$);

- $\Delta^{\rm CT+}_{\rm ac}(k), \Delta^{\rm CT-}_{\rm ac}(k) \in [0, |\Lambda^{\rm int}|], k \in [1, k_{\rm C}] = I_{(\geq 2)} \cup I_{(\geq 1)}$: $\Delta^{\rm CT+}_{\rm ac}(k) = \Delta^{\rm CT-}_{\rm ac}(k) = 0$ (resp., $\Delta^{\rm CT+}_{\rm ac}(k) = \alpha(u)$ and $\Delta^{\rm CT-}_{\rm ac}(k) = \alpha(v)$) $\Leftrightarrow$ edge ${e^{\rm CT}}_{{\rm tail}(k),j} = (u, v) \in E_{\rm CT}$ for some $j \in [1, t_{\rm T}]$ is used in $\mathbb{C}$ (resp., otherwise);

- $\Delta_{\mathrm{ac}}^{\mathrm{TC}+}(k), \Delta_{\mathrm{ac}}^{\mathrm{TC}-}(k) \in [0, |\Lambda^{\mathrm{int}}|], k \in [1, k_{\mathrm{C}}] = I_{(\geq 2)} \cup I_{(\geq 1)}$: Analogous with $\Delta_{\mathrm{ac}}^{\mathrm{CT}+}(k)$ and $\Delta_{\mathrm{ac}}^{\mathrm{CT}-}(k)$;
- $\Delta_{\mathrm{ac}}^{\mathrm{CF}+}(c) \in [0, |\Lambda^{\mathrm{int}}|], \Delta_{\mathrm{ac}}^{\mathrm{CF}-}(c) \in [0, |\Lambda^{\mathrm{int}}|], c \in [1, \widetilde{t_{\mathrm{C}}}]$: $\Delta_{\mathrm{ac}}^{\mathrm{CF}+}(c) = \Delta_{\mathrm{ac}}^{\mathrm{CF}-}(c) = 0$ (resp., $\Delta_{\mathrm{ac}}^{\mathrm{CF}+}(c) = \alpha(u)$ and $\Delta_{\mathrm{ac}}^{\mathrm{CF}-}(c) = \alpha(v)) \Leftrightarrow$ edge ${e^{\mathrm{CF}}}_{c,i} = (u, v) \in E_{\mathrm{CF}}$ for some $i \in [1, t_{\mathrm{F}}]$ is used in $\mathbb{C}$ (resp., otherwise);
- $\Delta_{\mathrm{ac}}^{\mathrm{TF}+}(i) \in [0, |\Lambda^{\mathrm{int}}|], \Delta_{\mathrm{ac}}^{\mathrm{TF}-}(i) \in [0, |\Lambda^{\mathrm{int}}|], i \in [1, t_{\mathrm{T}}]$: Analogous with $\Delta_{\mathrm{ac}}^{\mathrm{CF}+}(c)$ and $\Delta_{\mathrm{ac}}^{\mathrm{CF}-}(c)$;

constraints:

$$
\begin{aligned}
\mathrm{ac}_{\mathrm{C}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma_{\mathrm{ac}}^{\mathrm{int}} \setminus \widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{C}}, \\
\mathrm{ac}_{\mathrm{T}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma_{\mathrm{ac}}^{\mathrm{int}} \setminus \widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{T}}, \\
\mathrm{ac}_{\mathrm{F}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma_{\mathrm{ac}}^{\mathrm{int}} \setminus \widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{F}}, \\
\mathrm{ac}_{\mathrm{CT}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma_{\mathrm{ac}}^{\mathrm{int}} \setminus \widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{CT}}, \\
\mathrm{ac}_{\mathrm{TC}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma_{\mathrm{ac}}^{\mathrm{int}} \setminus \widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{TC}}, \\
\mathrm{ac}_{\mathrm{CF}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma_{\mathrm{ac}}^{\mathrm{int}} \setminus \widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{CF}}, \\
\mathrm{ac}_{\mathrm{TF}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma_{\mathrm{ac}}^{\mathrm{int}} \setminus \widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{TF}},
\end{aligned} \tag{67}
$$

$$
\begin{aligned}
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{C}}([\nu]^{\mathrm{int}}) &= \sum_{i\in[\widetilde{k_{\mathrm{C}}}+1, m_{\mathrm{C}}]} \delta_{\beta}^{\mathrm{C}}(i, m), & m &\in [1, 3], \\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{T}}([\nu]^{\mathrm{int}}) &= \sum_{i\in[2, t_{\mathrm{T}}]} \delta_{\beta}^{\mathrm{T}}(i, m), & m &\in [1, 3], \\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{F}}([\nu]^{\mathrm{int}}) &= \sum_{i\in[2, t_{\mathrm{F}}]} \delta_{\beta}^{\mathrm{F}}(i, m), & m &\in [1, 3], \\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{CT}}([\nu]^{\mathrm{int}}) &= \sum_{k\in[1, k_{\mathrm{C}}]} \delta_{\beta}^{\mathrm{CT}}(k, m), & m &\in [1, 3], \\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{TC}}([\nu]^{\mathrm{int}}) &= \sum_{k\in[1, k_{\mathrm{C}}]} \delta_{\beta}^{\mathrm{TC}}(k, m), & m &\in [1, 3], \\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{CF}}([\nu]^{\mathrm{int}}) &= \sum_{c\in[1, \widetilde{t_{\mathrm{C}}}]} \delta_{\beta}^{*\mathrm{F}}(c, m), & m &\in [1, 3], \\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{TF}}([\nu]^{\mathrm{int}}) &= \sum_{c\in[\widetilde{t_{\mathrm{C}}}+1, c_{\mathrm{F}}]} \delta_{\beta}^{*\mathrm{F}}(c, m), & m &\in [1, 3],
\end{aligned} \tag{68}
$$

$$
\begin{aligned}
\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ac}}} m\cdot\delta^{\mathrm{C}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{C}}(i), \\
\Delta^{\mathrm{C}+}_{\mathrm{ac}}(i) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ac}}} [\mathtt{a}]^{\mathrm{int}}\delta^{\mathrm{C}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{tail}(i)), \\
\Delta^{\mathrm{C}-}_{\mathrm{ac}}(i) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ac}}} [\mathtt{b}]^{\mathrm{int}}\delta^{\mathrm{C}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{head}(i)), \\
\Delta^{\mathrm{C}+}_{\mathrm{ac}}(i) + \Delta^{\mathrm{C}-}_{\mathrm{ac}}(i) &\le 2|\Lambda^{\mathrm{int}}|(1-e^{\mathrm{C}}(i)), && i\in[\widetilde{k_{\mathrm{C}}}+1, m_{\mathrm{C}}], \\
\sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]} \delta^{\mathrm{C}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{C}}([\nu]^{\mathrm{int}}), && \nu\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ac}}, \qquad (69)
\end{aligned}
$$

$$
\begin{aligned}
\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ac}}} m\cdot\delta^{\mathrm{T}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{T}}(i), \\
\Delta^{\mathrm{T}+}_{\mathrm{ac}}(i) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ac}}} [\mathtt{a}]^{\mathrm{int}}\delta^{\mathrm{T}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{T}}(i-1), \\
\Delta^{\mathrm{T}-}_{\mathrm{ac}}(i) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ac}}} [\mathtt{b}]^{\mathrm{int}}\delta^{\mathrm{T}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{T}}(i), \\
\Delta^{\mathrm{T}+}_{\mathrm{ac}}(i) + \Delta^{\mathrm{T}-}_{\mathrm{ac}}(i) &\le 2|\Lambda^{\mathrm{int}}|(1-e^{\mathrm{T}}(i)), && i\in[2,t_{\mathrm{T}}], \\
\sum_{i\in[2,t_{\mathrm{T}}]} \delta^{\mathrm{T}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{T}}([\nu]^{\mathrm{int}}), && \nu\in\widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ac}}, \qquad (70)
\end{aligned}
$$

$$
\begin{aligned}
\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}} m\cdot\delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{F}}(i), \\
\Delta^{\mathrm{F}+}_{\mathrm{ac}}(i) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}} [\mathtt{a}]^{\mathrm{int}}\delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{F}}(i-1), \\
\Delta^{\mathrm{F}-}_{\mathrm{ac}}(i) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}} [\mathtt{b}]^{\mathrm{int}}\delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{F}}(i), \\
\Delta^{\mathrm{F}+}_{\mathrm{ac}}(i) + \Delta^{\mathrm{F}-}_{\mathrm{ac}}(i) &\le 2|\Lambda^{\mathrm{ex}}|(1-e^{\mathrm{F}}(i)), && i\in[2,t_{\mathrm{F}}], \\
\sum_{i\in[2,t_{\mathrm{F}}]} \delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{F}}([\nu]^{\mathrm{int}}), && \nu\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}, \qquad (71)
\end{aligned}
$$

$$
\begin{aligned}
\alpha^{\mathrm{T}}(i) + |\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i)) &\ge \alpha^{\mathrm{CT}}(k), \\
\alpha^{\mathrm{CT}}(k) &\ge \alpha^{\mathrm{T}}(i) - |\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i)), && i\in[1,t_{\mathrm{T}}], \\
\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}} m\cdot\delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{CT}}(k), \\
\Delta^{\mathrm{CT}+}_{\mathrm{ac}}(k) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}} [\mathtt{a}]^{\mathrm{int}}\delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{tail}(k)), \\
\Delta^{\mathrm{CT}-}_{\mathrm{ac}}(k) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}} [\mathtt{b}]^{\mathrm{int}}\delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{CT}}(k), \\
\Delta^{\mathrm{CT}+}_{\mathrm{ac}}(k) + \Delta^{\mathrm{CT}-}_{\mathrm{ac}}(k) &\le 2|\Lambda^{\mathrm{int}}|(1-\delta^{\mathrm{T}}_{\chi}(k)), && k\in[1,k_{\mathrm{C}}], \\
\sum_{k\in[1,k_{\mathrm{C}}]} \delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{CT}}([\nu]^{\mathrm{int}}), && \nu\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}, \qquad (72)
\end{aligned}
$$

$$\begin{aligned}
\alpha^{\mathrm{T}}(i) + |\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i+1)) &\geq \alpha^{\mathrm{TC}}(k), \\
\alpha^{\mathrm{TC}}(k) &\geq \alpha^{\mathrm{T}}(i) - |\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i+1)), && i \in [1, t_{\mathrm{T}}], \\
\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}} m \cdot \delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{TC}}(k), \\
\Delta^{\mathrm{TC}+}_{\mathrm{ac}}(k) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}} [\mathtt{a}]^{\mathrm{int}}\delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{TC}}(k), \\
\Delta^{\mathrm{TC}-}_{\mathrm{ac}}(k) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}} [\mathtt{b}]^{\mathrm{int}}\delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{head}(k)), \\
\Delta^{\mathrm{TC}+}_{\mathrm{ac}}(k) + \Delta^{\mathrm{TC}-}_{\mathrm{ac}}(k) &\leq 2|\Lambda^{\mathrm{int}}|(1-\delta^{\mathrm{T}}_{\chi}(k)), && k \in [1, k_{\mathrm{C}}], \\
\sum_{k\in[1,k_{\mathrm{C}}]} \delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{TC}}([\nu]^{\mathrm{int}}), && \nu \in \widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}, \qquad (73)
\end{aligned}$$

$$\begin{aligned}
\alpha^{\mathrm{F}}(i) + |\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{F}}(i,c)+e^{\mathrm{F}}(i)) &\geq \alpha^{\mathrm{CF}}(c), \\
\alpha^{\mathrm{CF}}(c) &\geq \alpha^{\mathrm{F}}(i) - |\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{F}}(i,c)+e^{\mathrm{F}}(i)), && i \in [1, t_{\mathrm{F}}], \\
\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ac}}} m \cdot \delta^{\mathrm{CF}}_{\mathrm{ac}}(c,[\nu]^{\mathrm{int}}) &= \beta^{*\mathrm{F}}(c), \\
\Delta^{\mathrm{CF}+}_{\mathrm{ac}}(c) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ac}}} [\mathtt{a}]^{\mathrm{int}}\delta^{\mathrm{CF}}_{\mathrm{ac}}(c,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{head}(c)), \\
\Delta^{\mathrm{CF}-}_{\mathrm{ac}}(c) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ac}}} [\mathtt{b}]^{\mathrm{int}}\delta^{\mathrm{CF}}_{\mathrm{ac}}(c,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{CF}}(c), \\
\Delta^{\mathrm{CF}+}_{\mathrm{ac}}(c) + \Delta^{\mathrm{CF}-}_{\mathrm{ac}}(c) &\leq 2\max\{|\Lambda^{\mathrm{int}}|, |\Lambda^{\mathrm{int}}|\}(1-\delta^{\mathrm{F}}_{\chi}(c)), && c \in [1, \widetilde{t_{\mathrm{C}}}], \\
\sum_{c\in[1,\widetilde{t_{\mathrm{C}}}]} \delta^{\mathrm{CF}}_{\mathrm{ac}}(c,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{CF}}([\nu]^{\mathrm{int}}), && \nu \in \widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ac}}, \qquad (74)
\end{aligned}$$

$$\begin{aligned}
\alpha^{\mathrm{F}}(j) + |\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{F}}(j,i+\widetilde{t_{\mathrm{C}}})+e^{\mathrm{F}}(j)) &\geq \alpha^{\mathrm{TF}}(i), \\
\alpha^{\mathrm{TF}}(i) &\geq \alpha^{\mathrm{F}}(j) - |\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{F}}(j,i+\widetilde{t_{\mathrm{C}}})+e^{\mathrm{F}}(j)), && j \in [1, t_{\mathrm{F}}], \\
\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}} m \cdot \delta^{\mathrm{TF}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \beta^{*\mathrm{F}}(i+\widetilde{t_{\mathrm{C}}}), \\
\Delta^{\mathrm{TF}+}_{\mathrm{ac}}(i) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}} [\mathtt{a}]^{\mathrm{int}}\delta^{\mathrm{TF}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{T}}(i), \\
\Delta^{\mathrm{TF}-}_{\mathrm{ac}}(i) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}} [\mathtt{b}]^{\mathrm{int}}\delta^{\mathrm{TF}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{TF}}(i), \\
\Delta^{\mathrm{TF}+}_{\mathrm{ac}}(i) + \Delta^{\mathrm{TF}-}_{\mathrm{ac}}(i) &\leq 2\max\{|\Lambda^{\mathrm{int}}|, |\Lambda^{\mathrm{int}}|\}(1-\delta^{\mathrm{F}}_{\chi}(i+\widetilde{t_{\mathrm{C}}})), && i \in [1, t_{\mathrm{T}}], \\
\sum_{i\in[1,t_{\mathrm{T}}]} \delta^{\mathrm{TF}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{TF}}([\nu]^{\mathrm{int}}), && \nu \in \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}, \qquad (75)
\end{aligned}$$

$$\begin{aligned}
\sum_{\mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F},\mathrm{CT},\mathrm{TC},\mathrm{CF},\mathrm{TF}\}} (\mathrm{ac}_{\mathrm{X}}([\nu]^{\mathrm{int}}) + \mathrm{ac}_{\mathrm{X}}([\overline{\nu}]^{\mathrm{int}})) &= \mathrm{ac}^{\mathrm{int}}([\nu]^{\mathrm{int}}), && \nu \in \Gamma^{\mathrm{int}}_{\mathrm{ac},<}, \\
\sum_{\mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F},\mathrm{CT},\mathrm{TC},\mathrm{CF},\mathrm{TF}\}} \mathrm{ac}_{\mathrm{X}}([\nu]^{\mathrm{int}}) &= \mathrm{ac}^{\mathrm{int}}([\nu]^{\mathrm{int}}), && \nu \in \Gamma^{\mathrm{int}}_{\mathrm{ac},=}. \qquad (76)
\end{aligned}$$

4.9 Descriptor for the Number of Chemical Symbols

We include constraints for computing the frequency of each chemical symbol in $\Lambda_{\rm dg}$. Let $\mathrm{cs}(v)$ denote the chemical symbol of an interior-vertex v in a chemical graph $\mathbb{C}$ to be inferred; i.e., $\mathrm{cs}(v) = \mu = \mathtt{a}d \in \Lambda_{\rm dg}$ such that $\alpha(v) = \mathtt{a}$ and $\deg_{\langle \mathbb{C}\rangle}(v) = \deg_H(v) - \deg_{\mathbb{C}}^{\rm hyd}(v) = d$ in $\mathbb{C} = (H, \alpha, \beta)$.

constants:

- A set $\Lambda^{\rm int}_{\rm dg}$ of chemical symbols;
- Prepare a coding of each of the two sets $\Lambda^{\rm int}_{\rm dg}$ and let $[\mu]^{\rm int}$ denote the coded integer of an element $\mu \in \Lambda^{\rm int}_{\rm dg}$;
- Choose subsets $\widetilde{\Lambda}^{\rm C}_{\rm dg}, \widetilde{\Lambda}^{\rm T}_{\rm dg}, \widetilde{\Lambda}^{\rm F}_{\rm dg} \subseteq \Lambda^{\rm int}_{\rm dg}$: To compute the frequency of chemical symbols exactly, set $\widetilde{\Lambda}^{\rm C}_{\rm dg} := \widetilde{\Lambda}^{\rm T}_{\rm dg} := \widetilde{\Lambda}^{\rm F}_{\rm dg} := \Lambda^{\rm int}_{\rm dg}$;

variables:

- $\mathrm{ns}^{\rm int}([\mu]^{\rm int}) \in [0, \mathrm{n}^{\rm int}_{\rm UB}]$, $\mu \in \Lambda^{\rm int}_{\rm dg}$: the number of interior-vertices v with $\mathrm{cs}(v) = \mu$;
- $\delta^{\rm X}_{\rm ns}(i, [\mu]^{\rm int}) \in [0, 1]$, $i \in [1, t_{\rm X}], \mu \in \Lambda^{\rm int}_{\rm dg}$, $\mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$;

constraints:

$$\sum_{\mu \in \widetilde{\Lambda}^{\rm X}_{\rm dg} \cup \{\epsilon\}} \delta^{\rm X}_{\rm ns}(i, [\mu]^{\rm int}) = 1, \quad \sum_{\mu = \mathtt{a}d \in \widetilde{\Lambda}^{\rm X}_{\rm dg}} [\mathtt{a}]^{\rm int} \cdot \delta^{\rm X}_{\rm ns}(i, [\mu]^{\rm int}) = \alpha^{\rm X}(i),$$
$$\sum_{\mu = \mathtt{a}d \in \widetilde{\Lambda}^{\rm X}_{\rm dg}} d \cdot \delta^{\rm X}_{\rm ns}(i, [\mu]^{\rm int}) = \deg^{\rm X}(i),$$
$$i \in [1, t_{\rm X}], \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}, \tag{77}$$

$$\sum_{i \in [1, t_{\rm C}]} \delta^{\rm C}_{\rm ns}(i, [\mu]^{\rm int}) + \sum_{i \in [1, t_{\rm T}]} \delta^{\rm T}_{\rm ns}(i, [\mu]^{\rm int}) + \sum_{i \in [1, t_{\rm F}]} \delta^{\rm F}_{\rm ns}(i, [\mu]^{\rm int}) = \mathrm{ns}^{\rm int}([\mu]^{\rm int}), \qquad \mu \in \Lambda^{\rm int}_{\rm dg}. \tag{78}$$

4.10 Descriptor for the Number of Edge-configurations

We include constraints to compute the frequency of each edge-configuration in an inferred chemical graph $\mathbb{C}$.

constants:

- A set $\Gamma^{\rm int}$ of edge-configurations $\gamma = (\mu, \mu', m)$ with $\mu \leq \mu'$;
- Let $\Gamma^{\rm int}_{<} = \{(\mu, \mu', m) \in \Gamma^{\rm int} \mid \mu < \mu'\}$, $\Gamma^{\rm int}_{=} = \{(\mu, \mu', m) \in \Gamma^{\rm int} \mid \mu = \mu'\}$ and $\Gamma^{\rm int}_{>} = \{(\mu', \mu, m) \mid (\mu, \mu', m) \in \Gamma^{\rm int}_{<}\}$;
- Prepare a coding of the set $\Gamma^{\rm int} \cup \Gamma^{\rm int}_{>}$ and let $[\gamma]^{\rm int}$ denote the coded integer of an element γ in $\Gamma^{\rm int} \cup \Gamma^{\rm int}_{>}$;
- Choose subsets $\widetilde{\Gamma}^{\rm C}_{\rm ec}, \widetilde{\Gamma}^{\rm T}_{\rm ec}, \widetilde{\Gamma}^{\rm CT}_{\rm ec}, \widetilde{\Gamma}^{\rm TC}_{\rm ec}, \widetilde{\Gamma}^{\rm F}_{\rm ec}, \widetilde{\Gamma}^{\rm CF}_{\rm ec}, \widetilde{\Gamma}^{\rm TF}_{\rm ec} \subseteq \Gamma^{\rm int} \cup \Gamma^{\rm int}_{>}$; To compute the frequency of edge-configurations exactly, set $\widetilde{\Gamma}^{\rm C}_{\rm ec} := \widetilde{\Gamma}^{\rm T}_{\rm ec} := \widetilde{\Gamma}^{\rm CT}_{\rm ec} := \widetilde{\Gamma}^{\rm TC}_{\rm ec} := \widetilde{\Gamma}^{\rm F}_{\rm ec} := \widetilde{\Gamma}^{\rm CF}_{\rm ec} := \widetilde{\Gamma}^{\rm TF}_{\rm ec} := \Gamma^{\rm int} \cup \Gamma^{\rm int}_{>}$;
- $\mathrm{ec}^{\rm int}_{\rm LB}(\gamma), \mathrm{ec}^{\rm int}_{\rm UB}(\gamma) \in [0, 2\mathrm{n}^{\rm int}_{\rm UB}], \gamma = (\mu, \mu', m) \in \Gamma^{\rm int}$: lower and upper bounds on the number of interior-edges $e = uv$ with $\mathrm{cs}(u) = \mu$, $\mathrm{cs}(v) = \mu'$ and $\beta(e) = m$;

variables:

- $\mathrm{ec}^{\mathrm{int}}([\gamma]^{\mathrm{int}}) \in [\mathrm{ec}^{\mathrm{int}}_{\mathrm{LB}}(\gamma), \mathrm{ec}^{\mathrm{int}}_{\mathrm{UB}}(\gamma)], \gamma \in \Gamma^{\mathrm{int}}$: the number of interior-edges with edge-configuration γ;

- $\mathrm{ec}_{\mathrm{C}}([\gamma]^{\mathrm{int}}) \in [0, m_{\mathrm{C}}], \gamma \in \widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ec}}$, $\mathrm{ec}_{\mathrm{T}}([\gamma]^{\mathrm{int}}) \in [0, t_{\mathrm{T}}], \gamma \in \widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ec}}$, $\mathrm{ec}_{\mathrm{F}}([\gamma]^{\mathrm{int}}) \in [0, t_{\mathrm{F}}], \gamma \in \widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ec}}$: the number of edges $e^{\mathrm{C}} \in E_{\mathrm{C}}$ (resp., edges $e^{\mathrm{T}} \in E_{\mathrm{T}}$ and edges $e^{\mathrm{F}} \in E_{\mathrm{F}}$) with edge-configuration γ;

- $\mathrm{ec}_{\mathrm{CT}}([\gamma]^{\mathrm{int}}) \in [0, \min\{k_{\mathrm{C}}, t_{\mathrm{T}}\}], \gamma \in \widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ec}}$, $\mathrm{ec}_{\mathrm{TC}}([\gamma]^{\mathrm{int}}) \in [0, \min\{k_{\mathrm{C}}, t_{\mathrm{T}}\}], \gamma \in \widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ec}}$, $\mathrm{ec}_{\mathrm{CF}}([\gamma]^{\mathrm{int}}) \in [0, \widetilde{t_{\mathrm{C}}}], \gamma \in \widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}}$, $\mathrm{ec}_{\mathrm{TF}}([\gamma]^{\mathrm{int}}) \in [0, t_{\mathrm{T}}], \gamma \in \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}}$: the number of edges $e^{\mathrm{CT}} \in E_{\mathrm{CT}}$ (resp., edges $e^{\mathrm{TC}} \in E_{\mathrm{TC}}$ and edges $e^{\mathrm{CF}} \in E_{\mathrm{CF}}$ and $e^{\mathrm{TF}} \in E_{\mathrm{TF}}$) with edge-configuration γ;

- $\delta^{\mathrm{C}}_{\mathrm{ec}}(i, [\gamma]^{\mathrm{int}}) \in [0, 1], i \in [\widetilde{k_{\mathrm{C}}} + 1, m_{\mathrm{C}}] = I_{(\geq 1)} \cup I_{(0/1)} \cup I_{(=1)}, \gamma \in \widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ec}}$, $\delta^{\mathrm{T}}_{\mathrm{ec}}(i, [\gamma]^{\mathrm{int}}) \in [0, 1], i \in [2, t_{\mathrm{T}}], \gamma \in \widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ec}}$, $\delta^{\mathrm{F}}_{\mathrm{ec}}(i, [\gamma]^{\mathrm{int}}) \in [0, 1], i \in [2, t_{\mathrm{F}}], \gamma \in \widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ec}}$: $\delta^{\mathrm{X}}_{\mathrm{ec}}(i, [\gamma]^{\mathrm{t}}) = 1 \Leftrightarrow$ edge ${e^{\mathrm{X}}}_i$ has edge-configuration γ;

- $\delta^{\mathrm{CT}}_{\mathrm{ec,C}}(k, [\gamma]^{\mathrm{int}}), \delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k, [\gamma]^{\mathrm{int}}) \in [0, 1], k \in [1, k_{\mathrm{C}}] = I_{(\geq 2)} \cup I_{(\geq 1)}, \gamma \in \widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ec}}$: $\delta^{\mathrm{CT}}_{\mathrm{ec,C}}(k, [\gamma]^{\mathrm{int}}) = 1$ (resp., $\delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k, [\gamma]^{\mathrm{int}}) = 1$) $\Leftrightarrow$ edge ${e^{\mathrm{CT}}}_{\mathrm{tail}(k),j}$ (resp., ${e^{\mathrm{TC}}}_{\mathrm{head}(k),j}$) for some $j \in [1, t_{\mathrm{T}}]$ has edge-configuration γ;

- $\delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c, [\gamma]^{\mathrm{int}}) \in [0, 1], c \in [1, \widetilde{t_{\mathrm{C}}}], \gamma \in \widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}}$: $\delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c, [\gamma]^{\mathrm{int}}) = 1 \Leftrightarrow$ edge ${e^{\mathrm{CF}}}_{c,i}$ for some $i \in [1, t_{\mathrm{F}}]$ has edge-configuration γ;

- $\delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i, [\gamma]^{\mathrm{int}}) \in [0, 1], i \in [1, t_{\mathrm{T}}], \gamma \in \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}}$: $\delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i, [\gamma]^{\mathrm{int}}) = 1 \Leftrightarrow$ edge ${e^{\mathrm{TF}}}_{i,j}$ for some $j \in [1, t_{\mathrm{F}}]$ has edge-configuration γ;

- $\deg^{\mathrm{CT}}_{\mathrm{T}}(k), \deg^{\mathrm{TC}}_{\mathrm{T}}(k) \in [0, 4], k \in [1, k_{\mathrm{C}}]$: $\deg_{\langle \mathbb{C} \rangle}(v)$ of an end-vertex $v \in V_{\mathrm{T}}$ of the edge $({v^{\mathrm{C}}}_{\mathrm{tail}(k)}, v) \in E_{\mathrm{CT}}$ (resp., $(v, {v^{\mathrm{C}}}_{\mathrm{head}(k)}) \in E_{\mathrm{TC}}$) if any;

- $\deg^{\mathrm{CF}}_{\mathrm{F}}(c) \in [0, 4], c \in [1, \widetilde{t_{\mathrm{C}}}]$: $\deg_{\langle \mathbb{C} \rangle}(v)$ of an end-vertex $v \in V_{\mathrm{F}}$ of the edge $({v^{\mathrm{C}}}_c, v) \in E_{\mathrm{CF}}$ if any;

- $\deg^{\mathrm{TF}}_{\mathrm{F}}(i) \in [0, 4], i \in [1, t_{\mathrm{T}}]$: $\deg_{\langle \mathbb{C} \rangle}(v)$ of an end-vertex $v \in V_{\mathrm{F}}$ of the edge $({v^{\mathrm{T}}}_i, v) \in E_{\mathrm{TF}}$ if any;

- $\Delta^{\mathrm{C}+}_{\mathrm{ec}}(i), \Delta^{\mathrm{C}-}_{\mathrm{ec}}(i), \in [0, 4], i \in [\widetilde{k_{\mathrm{C}}} + 1, m_{\mathrm{C}}]$, $\Delta^{\mathrm{T}+}_{\mathrm{ec}}(i), \Delta^{\mathrm{T}-}_{\mathrm{ec}}(i) \in [0, 4], i \in [2, t_{\mathrm{T}}]$, $\Delta^{\mathrm{F}+}_{\mathrm{ec}}(i), \Delta^{\mathrm{F}-}_{\mathrm{ec}}(i) \in [0, 4], i \in [2, t_{\mathrm{F}}]$: $\Delta^{\mathrm{X}+}_{\mathrm{ec}}(i) = \Delta^{\mathrm{X}-}_{\mathrm{ec}}(i) = 0$ (resp., $\Delta^{\mathrm{X}+}_{\mathrm{ec}}(i) = \deg_{\langle \mathbb{C} \rangle}(u)$ and $\Delta^{\mathrm{X}-}_{\mathrm{ec}}(i) = \deg_{\langle \mathbb{C} \rangle}(v)$) $\Leftrightarrow$ edge ${e^{\mathrm{X}}}_i = (u, v) \in E_{\mathrm{X}}$ is used in $\langle \mathbb{C} \rangle$ (resp., ${e^{\mathrm{X}}}_i \notin E(\langle \mathbb{C} \rangle)$);

- $\Delta^{\mathrm{CT}+}_{\mathrm{ec}}(k), \Delta^{\mathrm{CT}-}_{\mathrm{ec}}(k) \in [0, 4], k \in [1, k_{\mathrm{C}}] = I_{(\geq 2)} \cup I_{(\geq 1)}$: $\Delta^{\mathrm{CT}+}_{\mathrm{ec}}(k) = \Delta^{\mathrm{CT}-}_{\mathrm{ec}}(k) = 0$ (resp., $\Delta^{\mathrm{CT}+}_{\mathrm{ec}}(k) = \deg_{\langle \mathbb{C} \rangle}(u)$ and $\Delta^{\mathrm{CT}-}_{\mathrm{ec}}(k) = \deg_{\langle \mathbb{C} \rangle}(v)$) $\Leftrightarrow$ edge ${e^{\mathrm{CT}}}_{\mathrm{tail}(k),j} = (u, v) \in E_{\mathrm{CT}}$ for some $j \in [1, t_{\mathrm{T}}]$ is used in $\langle \mathbb{C} \rangle$ (resp., otherwise);

- $\Delta^{\mathrm{TC}+}_{\mathrm{ec}}(k), \Delta^{\mathrm{TC}-}_{\mathrm{ec}}(k) \in [0, 4], k \in [1, k_{\mathrm{C}}] = I_{(\geq 2)} \cup I_{(\geq 1)}$: Analogous with $\Delta^{\mathrm{CT}+}_{\mathrm{ec}}(k)$ and $\Delta^{\mathrm{CT}-}_{\mathrm{ec}}(k)$;

- $\Delta^{\mathrm{CF}+}_{\mathrm{ac}}(c), \Delta^{\mathrm{CF}-}_{\mathrm{ec}}(c) \in [0, 4], c \in [1, \widetilde{t_{\mathrm{C}}}]$: $\Delta^{\mathrm{CF}+}_{\mathrm{ec}}(c) = \Delta^{\mathrm{CF}-}_{\mathrm{ec}}(c) = 0$ (resp., $\Delta^{\mathrm{CF}+}_{\mathrm{ec}}(c) = \deg_{\langle \mathbb{C} \rangle}(u)$ and $\Delta^{\mathrm{CF}-}_{\mathrm{ec}}(c) = \deg_{\langle \mathbb{C} \rangle}(v)$) $\Leftrightarrow$ edge ${e^{\mathrm{CF}}}_{c,j} = (u, v) \in E_{\mathrm{CF}}$ for some $j \in [1, t_{\mathrm{F}}]$ is used in $\langle \mathbb{C} \rangle$ (resp., otherwise);

- $\Delta^{\mathrm{TF}+}_{\mathrm{ec}}(i), \Delta^{\mathrm{TF}-}_{\mathrm{ec}}(i) \in [0, 4], i \in [1, t_{\mathrm{T}}]$: Analogous with $\Delta^{\mathrm{CF}+}_{\mathrm{ec}}(c)$ and $\Delta^{\mathrm{CF}-}_{\mathrm{ec}}(c)$;

constraints:

$$
\begin{aligned}
\mathrm{ec_C}([\gamma]^{\mathrm{int}}) &= 0, & \gamma &\in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ec}},\\
\mathrm{ec_T}([\gamma]^{\mathrm{int}}) &= 0, & \gamma &\in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ec}},\\
\mathrm{ec_F}([\gamma]^{\mathrm{int}}) &= 0, & \gamma &\in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ec}},\\
\mathrm{ec_{CT}}([\gamma]^{\mathrm{int}}) &= 0, & \gamma &\in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ec}},\\
\mathrm{ec_{TC}}([\gamma]^{\mathrm{int}}) &= 0, & \gamma &\in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}},\\
\mathrm{ec_{CF}}([\gamma]^{\mathrm{int}}) &= 0, & \gamma &\in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}},\\
\mathrm{ec_{TF}}([\gamma]^{\mathrm{int}}) &= 0, & \gamma &\in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}},
\end{aligned} \tag{79}
$$

$$
\begin{aligned}
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec_C}([\gamma]^{\mathrm{int}}) &= \sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]} \delta^{\mathrm{C}}_{\beta}(i,m), & m &\in [1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec_T}([\gamma]^{\mathrm{int}}) &= \sum_{i\in[2,t_{\mathrm{T}}]} \delta^{\mathrm{T}}_{\beta}(i,m), & m &\in [1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec_F}([\gamma]^{\mathrm{int}}) &= \sum_{i\in[2,t_{\mathrm{F}}]} \delta^{\mathrm{F}}_{\beta}(i,m), & m &\in [1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec_{CT}}([\gamma]^{\mathrm{int}}) &= \sum_{k\in[1,k_{\mathrm{C}}]} \delta^{\mathrm{CT}}_{\beta}(k,m), & m &\in [1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec_{TC}}([\gamma]^{\mathrm{int}}) &= \sum_{k\in[1,k_{\mathrm{C}}]} \delta^{\mathrm{TC}}_{\beta}(k,m), & m &\in [1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec_{CF}}([\gamma]^{\mathrm{int}}) &= \sum_{c\in[1,\widetilde{t_{\mathrm{C}}}]} \delta^{*\mathrm{F}}_{\beta}(c,m), & m &\in [1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec_{TF}}([\gamma]^{\mathrm{int}}) &= \sum_{c\in[\widetilde{t_{\mathrm{C}}}+1,c_{\mathrm{F}}]} \delta^{*\mathrm{F}}_{\beta}(c,m), & m &\in [1,3],
\end{aligned} \tag{80}
$$

$$
\begin{aligned}
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ec}}} [(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}} \cdot \delta^{\mathrm{C}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}}) &= \sum_{\nu\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ac}}} [\nu]^{\mathrm{int}} \cdot \delta^{\mathrm{C}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}),\\
\Delta^{\mathrm{C}+}_{\mathrm{ec}}(i) + \sum_{\gamma=(\mathtt{a}d,\mu',m)\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ec}}} d\cdot\delta^{\mathrm{C}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}}) &= \mathrm{deg}^{\mathrm{C}}(\mathrm{tail}(i)),\\
\Delta^{\mathrm{C}-}_{\mathrm{ec}}(i) + \sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ec}}} d\cdot\delta^{\mathrm{C}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}}) &= \mathrm{deg}^{\mathrm{C}}(\mathrm{head}(i)),\\
\Delta^{\mathrm{C}+}_{\mathrm{ec}}(i) + \Delta^{\mathrm{C}-}_{\mathrm{ec}}(i) &\le 8(1-e^{\mathrm{C}}(i)), \qquad i \in [\widetilde{k_{\mathrm{C}}}+1, m_{\mathrm{C}}],\\
\sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]} \delta^{\mathrm{C}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}}) &= \mathrm{ec_C}([\gamma]^{\mathrm{int}}), \qquad \gamma \in \widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ec}},
\end{aligned} \tag{81}
$$

$$
\begin{aligned}
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\rm T}_{\rm ec}} [(\mathtt{a},\mathtt{b},m)]^{\rm int}\cdot\delta^{\rm T}_{\rm ec}(i,[\gamma]^{\rm int}) &= \sum_{\nu\in\widetilde{\Gamma}^{\rm T}_{\rm ac}}[\nu]^{\rm int}\cdot\delta^{\rm T}_{\rm ac}(i,[\nu]^{\rm int}),\\
\Delta^{\rm T+}_{\rm ec}(i)+\sum_{\gamma=(\mathtt{a}d,\mu',m)\in\widetilde{\Gamma}^{\rm T}_{\rm ec}} d\cdot\delta^{\rm T}_{\rm ec}(i,[\gamma]^{\rm int}) &= \deg^{\rm T}(i-1),\\
\Delta^{\rm T-}_{\rm ec}(i)+\sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\rm T}_{\rm ec}} d\cdot\delta^{\rm T}_{\rm ec}(i,[\gamma]^{\rm int}) &= \deg^{\rm T}(i),\\
\Delta^{\rm T+}_{\rm ec}(i)+\Delta^{\rm T-}_{\rm ec}(i) &\leq 8(1-e^{\rm T}(i)), && i\in[2,t_{\rm T}],\\
\sum_{i\in[2,t_{\rm T}]}\delta^{\rm T}_{\rm ec}(i,[\gamma]^{\rm int}) &= {\rm ec}_{\rm T}([\gamma]^{\rm int}), && \gamma\in\widetilde{\Gamma}^{\rm T}_{\rm ec},
\end{aligned}
\tag{82}
$$

$$
\begin{aligned}
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\rm F}_{\rm ec}} [(\mathtt{a},\mathtt{b},m)]^{\rm int}\cdot\delta^{\rm F}_{\rm ec}(i,[\gamma]^{\rm int}) &= \sum_{\nu\in\widetilde{\Gamma}^{\rm F}_{\rm ac}}[\nu]^{\rm int}\cdot\delta^{\rm F}_{\rm ac}(i,[\nu]^{\rm int}),\\
\Delta^{\rm F+}_{\rm ec}(i)+\sum_{\gamma=(\mathtt{a}d,\mu',m)\in\widetilde{\Gamma}^{\rm F}_{\rm ec}} d\cdot\delta^{\rm F}_{\rm ec}(i,[\gamma]^{\rm int}) &= \deg^{\rm F}(i-1),\\
\Delta^{\rm F-}_{\rm ec}(i)+\sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\rm F}_{\rm ec}} d\cdot\delta^{\rm F}_{\rm ec}(i,[\gamma]^{\rm int}) &= \deg^{\rm F}(i,0),\\
\Delta^{\rm F+}_{\rm ec}(i)+\Delta^{\rm F-}_{\rm ec}(i) &\leq 8(1-e^{\rm F}(i)), && i\in[2,t_{\rm F}],\\
\sum_{i\in[2,t_{\rm F}]}\delta^{\rm F}_{\rm ec}(i,[\gamma]^{\rm int}) &= {\rm ec}_{\rm F}([\gamma]^{\rm int}), && \gamma\in\widetilde{\Gamma}^{\rm F}_{\rm ec},
\end{aligned}
\tag{83}
$$

$$
\begin{aligned}
\deg^{\rm T}(i)+4(1-\chi^{\rm T}(i,k)+e^{\rm T}(i)) &\geq \deg^{\rm CT}_{\rm T}(k),\\
\deg^{\rm CT}_{\rm T}(k) &\geq \deg^{\rm T}(i)-4(1-\chi^{\rm T}(i,k)+e^{\rm T}(i)), && i\in[1,t_{\rm T}],\\
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\rm CT}_{\rm ec}} [(\mathtt{a},\mathtt{b},m)]^{\rm int}\cdot\delta^{\rm CT}_{\rm ec,C}(k,[\gamma]^{\rm int}) &= \sum_{\nu\in\widetilde{\Gamma}^{\rm CT}_{\rm ac}}[\nu]^{\rm int}\cdot\delta^{\rm CT}_{\rm ac}(k,[\nu]^{\rm int}),\\
\Delta^{\rm CT+}_{\rm ec}(k)+\sum_{\gamma=(\mathtt{a}d,\mu',m)\in\widetilde{\Gamma}^{\rm CT}_{\rm ec}} d\cdot\delta^{\rm CT}_{\rm ec,C}(k,[\gamma]^{\rm int}) &= \deg^{\rm C}({\rm tail}(k)),\\
\Delta^{\rm CT-}_{\rm ec}(k)+\sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\rm CT}_{\rm ec}} d\cdot\delta^{\rm CT}_{\rm ec,C}(k,[\gamma]^{\rm int}) &= \deg^{\rm CT}_{\rm T}(k),\\
\Delta^{\rm CT+}_{\rm ec}(k)+\Delta^{\rm CT-}_{\rm ec}(k) &\leq 8(1-\delta^{\rm T}_{\chi}(k)), && k\in[1,k_{\rm C}],\\
\sum_{k\in[1,k_{\rm C}]}\delta^{\rm CT}_{\rm ec,C}(k,[\gamma]^{\rm int}) &= {\rm ec}_{\rm CT}([\gamma]^{\rm int}), && \gamma\in\widetilde{\Gamma}^{\rm CT}_{\rm ec},
\end{aligned}
\tag{84}
$$

$$
\begin{aligned}
\deg^{\mathrm{T}}(i) + 4(1 - \chi^{\mathrm{T}}(i,k) + e^{\mathrm{T}}(i+1)) &\geq \deg^{\mathrm{TC}}_{\mathrm{T}}(k), \\
\deg^{\mathrm{TC}}_{\mathrm{T}}(k) &\geq \deg^{\mathrm{T}}(i) - 4(1 - \chi^{\mathrm{T}}(i,k) + e^{\mathrm{T}}(i+1)), && i \in [1, t_{\mathrm{T}}], \\
\sum_{\gamma=(\mathtt{a}d, \mathtt{b}d', m) \in \widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}}} [(\mathtt{a}, \mathtt{b}, m)]^{\mathrm{int}} \cdot \delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k, [\gamma]^{\mathrm{int}}) &= \sum_{\nu \in \widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}} [\nu]^{\mathrm{int}} \cdot \delta^{\mathrm{TC}}_{\mathrm{ac}}(k, [\nu]^{\mathrm{int}}), \\
\Delta^{\mathrm{TC}+}_{\mathrm{ec}}(k) + \sum_{\gamma=(\mathtt{a}d, \mu', m) \in \widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k, [\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{TC}}_{\mathrm{T}}(k), \\
\Delta^{\mathrm{TC}-}_{\mathrm{ec}}(k) + \sum_{\gamma=(\mu, \mathtt{b}d, m) \in \widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k, [\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{C}}(\mathrm{head}(k)), \\
\Delta^{\mathrm{TC}+}_{\mathrm{ec}}(k) + \Delta^{\mathrm{TC}-}_{\mathrm{ec}}(k) &\leq 8(1 - \delta^{\mathrm{T}}_{\chi}(k)), && k \in [1, k_{\mathrm{C}}], \\
\sum_{k \in [1, k_{\mathrm{C}}]} \delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k, [\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{TC}}([\gamma]^{\mathrm{int}}), && \gamma \in \widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}}, \qquad (85)
\end{aligned}
$$

$$
\begin{aligned}
\deg^{\mathrm{F}}(i) + 4(1 - \chi^{\mathrm{F}}(i,c) + e^{\mathrm{F}}(i)) &\geq \deg^{\mathrm{CF}}_{\mathrm{F}}(c), \\
\deg^{\mathrm{CF}}_{\mathrm{F}}(c) &\geq \deg^{\mathrm{F}}(i) - 4(1 - \chi^{\mathrm{F}}(i,c) + e^{\mathrm{F}}(i)), && i \in [1, t_{\mathrm{F}}], \\
\sum_{\gamma=(\mathtt{a}d, \mathtt{b}d', m) \in \widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}}} [(\mathtt{a}, \mathtt{b}, m)]^{\mathrm{int}} \cdot \delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c, [\gamma]^{\mathrm{int}}) &= \sum_{\nu \in \widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ac}}} [\nu]^{\mathrm{int}} \cdot \delta^{\mathrm{CF}}_{\mathrm{ac}}(c, [\nu]^{\mathrm{int}}), \\
\Delta^{\mathrm{CF}+}_{\mathrm{ec}}(c) + \sum_{\gamma=(\mathtt{a}d, \mu', m) \in \widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c, [\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{C}}(c), \\
\Delta^{\mathrm{CF}-}_{\mathrm{ec}}(c) + \sum_{\gamma=(\mu, \mathtt{b}d, m) \in \widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c, [\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{CF}}_{\mathrm{F}}(c), \\
\Delta^{\mathrm{CF}+}_{\mathrm{ec}}(c) + \Delta^{\mathrm{CF}-}_{\mathrm{ec}}(c) &\leq 8(1 - \delta^{\mathrm{F}}_{\chi}(c)), && c \in [1, \widetilde{t_{\mathrm{C}}}], \\
\sum_{c \in [1, \widetilde{t_{\mathrm{C}}}]} \delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c, [\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{CF}}([\gamma]^{\mathrm{int}}), && \gamma \in \widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}}, \qquad (86)
\end{aligned}
$$

$$
\begin{aligned}
\deg^{\mathrm{F}}(j) + 4(1 - \chi^{\mathrm{F}}(j, i + \widetilde{t_{\mathrm{C}}}) + e^{\mathrm{F}}(j)) &\geq \deg^{\mathrm{TF}}_{\mathrm{F}}(i), \\
\deg^{\mathrm{TF}}_{\mathrm{F}}(i) &\geq \deg^{\mathrm{F}}(j) - 4(1 - \chi^{\mathrm{F}}(j, i + \widetilde{t_{\mathrm{C}}}) + e^{\mathrm{F}}(j)), && j \in [1, t_{\mathrm{F}}], \\
\sum_{\gamma=(\mathtt{a}d, \mathtt{b}d', m) \in \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}}} [(\mathtt{a}, \mathtt{b}, m)]^{\mathrm{int}} \cdot \delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i, [\gamma]^{\mathrm{int}}) &= \sum_{\nu \in \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}} [\nu]^{\mathrm{int}} \cdot \delta^{\mathrm{TF}}_{\mathrm{ac}}(i, [\nu]^{\mathrm{int}}), \\
\Delta^{\mathrm{TF}+}_{\mathrm{ec}}(i) + \sum_{\gamma=(\mathtt{a}d, \mu', m) \in \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i, [\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{T}}(i), \\
\Delta^{\mathrm{TF}-}_{\mathrm{ec}}(i) + \sum_{\gamma=(\mu, \mathtt{b}d, m) \in \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i, [\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{TF}}_{\mathrm{F}}(i), \\
\Delta^{\mathrm{TF}+}_{\mathrm{ec}}(i) + \Delta^{\mathrm{TF}-}_{\mathrm{ec}}(i) &\leq 8(1 - \delta^{\mathrm{F}}_{\chi}(i + \widetilde{t_{\mathrm{C}}})), && i \in [1, t_{\mathrm{T}}], \\
\sum_{i \in [1, t_{\mathrm{T}}]} \delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i, [\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{TF}}([\gamma]^{\mathrm{int}}), && \gamma \in \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}}, \qquad (87)
\end{aligned}
$$

$$
\begin{aligned}
\sum_{\mathrm{X} \in \{\mathrm{C,T,F,CT,TC,CF,TF}\}} (\mathrm{ec}_{\mathrm{X}}([\gamma]^{\mathrm{int}}) + \mathrm{ec}_{\mathrm{X}}([\overline{\gamma}]^{\mathrm{int}})) &= \mathrm{ec}^{\mathrm{int}}([\gamma]^{\mathrm{int}}), && \gamma \in \Gamma^{\mathrm{int}}_{<}, \\
\sum_{\mathrm{X} \in \{\mathrm{C,T,F,CT,TC,CF,TF}\}} \mathrm{ec}_{\mathrm{X}}([\gamma]^{\mathrm{int}}) &= \mathrm{ec}^{\mathrm{int}}([\gamma]^{\mathrm{int}}), && \gamma \in \Gamma^{\mathrm{int}}_{=}. \qquad (88)
\end{aligned}
$$

4.11 Constraints for Normalization of Feature Vectors

By introducing a tolerance $\varepsilon > 0$ in the conversion between integers and reals, we include the following constraints for normalizing of a feature vector $f(\mathbb{C}) = (x_1, x_2, \ldots, x_K)$:

$$\frac{(1-\varepsilon)(x_i - \min(\mathrm{dcp}_i; D_\pi))}{\max(\mathrm{dcp}_i; D_\pi) - \min(\mathrm{dcp}_i; D_\pi)} \leq \widehat{x}_i \leq \frac{(1+\varepsilon)(x_i - \min(\mathrm{dcp}_i; D_\pi))}{\max(\mathrm{dcp}_i; D_\pi) - \min(\mathrm{dcp}_i; D_\pi)}, \ i \in [1, K]. \tag{89}$$

An example of a tolerance is $\varepsilon = 1 \times 10^{-5}$.