Supplementary Materials

An Inverse QSAR Method Based on a Two-layered Model and Integer Programming

Yu Shi[1], Jianshen Zhu[1], Naveed Ahmed Azam[1], Kazuya Haraguchi[1], Liang Zhao[2], Hiroshi Nagamochi[1], 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 Survivability (Shishu-Kan), Kyoto University, Kyoto 606-8306, Japan;
[3] Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji 611-0011, Japan

1 An MILP Formulation for Inferring a Target Chemical Graph in Stage 4

1.1 Constructing Target Chemical Graphs

This section describes how to construct a target chemical graph in Stages 4 and 5.

1.1.1 Formulating an MILP for a prediction function in Stage 4

In Stage 3, we construct a prediction function $\eta_{\mathcal{N}} : \mathbb{R}^K \to \mathbb{R}$. It is known that the computation process of $\eta_{\mathcal{N}}(x)$ from a vector $x^* \in \mathbb{R}^K$ can be formulated as an MILP with the following property.

Theorem 1. ([1, 2]) *Let $\mathcal{N}$ be an ANN with a piecewise-linear activation function for an input vector $x \in \mathbb{R}^K$, n_A denote the number of nodes in the architecture and n_B denote the total number of breakpoints over all activation functions. Then there is an MILP $\mathcal{M}(x, y; \mathcal{C}_1)$ that consists of variable vectors $x \in \mathbb{R}^K$, $y \in \mathbb{R}$, and an auxiliary variable vector $z \in \mathbb{R}^p$ for some integer $p = O(n_A + n_B)$ and a set $\mathcal{C}_1$ of $O(n_A + n_B)$ constraints on these variables such that: $\eta_{\mathcal{N}}(x^*) = y^*$ if and only if there is a vector (x^*, y^*) feasible to $\mathcal{M}(x, y; \mathcal{C}_1)$.*

Solving this MILP delivers a vector $x^* \in \mathbb{R}^K$ such that $\eta_{\mathcal{N}}(x^*) = y^*$ for a target value $\boldsymbol{y}^*$. However, the resulting vector x^* may not admit a chemical graph G^* such that $f(G^*) = x^*$. To ensure that such chemical graph always exists in Stage 4, we further introduce some more constraints for a set of new variables in the next section.

1.1.2 Formulating an MILP for a feature vector and a target specification in Stage 4

In this section, we show an outline of formulation of an MILP that represents the computation process of a feature function $f(G)$ from a chemical graph G and a construction of a target chemical graph $G \in \mathcal{G}(G_{\mathrm{C}}, \sigma_{\mathrm{int}}, \sigma_{\mathrm{ce}})$. Recall that the number of vertices in a target chemical graph is bounded by an upper bound n^* in a specification $(G_{\mathrm{C}}, \sigma_{\mathrm{int}}, \sigma_{\mathrm{ce}})$. However, if we introduce a set of $(n^*)^2$ variables for all pairs of n^* vertices to present all possible graphs for a target chemical graph, then the resulting MILP formulation is hard to solve for $n^* > 20$ due to a larger number of variables and constraints. To overcome this, a sparse representation of chemical graphs has been proposed in the previous applications of the framework for acyclic graphs [3] and ρ-lean graphs [4]. We also define a similar sparse representation to formulate an MILP for our two-layered model.

Scheme Graphs We first regard a given seed graph $G_{\rm C}$ as a digraph and then add some more vertices and edges to construct a digraph, called a *scheme graph* $\mathrm{SG} = (\mathcal{V}, \mathcal{E})$ so that any $(\sigma_{\rm int}, \sigma_{\rm ce})$-extension H of $G_{\rm C}$ can be chosen as a subgraph of SG.

For a given target specification $(G_{\rm C}, \sigma_{\rm int}, \sigma_{\rm ce})$, define integers that determine the size of a scheme graph SG as follows. $m_{\rm C} := |E_{\rm C}|$, $t_{\rm C} := |V_{\rm C}|$, $t_{\rm T} := \mathrm{n}^{\rm int}_{\rm UB} - |V_{\rm C}|$, and $t_{\rm F} := n^* - \mathrm{n}^{\rm int}_{\rm LB}$.

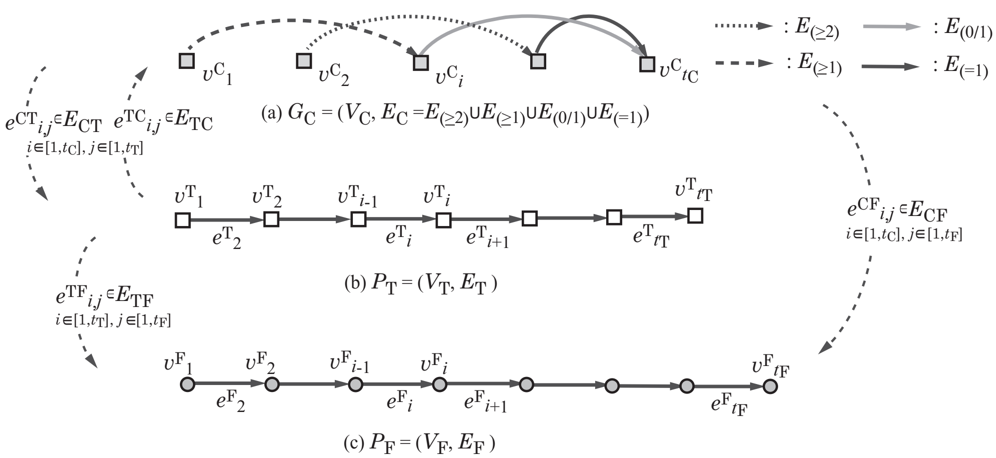

Figure 1: An illustration of a scheme graph SG: (a) A seed graph $G_{\rm C}$; (b) A path $P_{\rm T}$ of length $t_{\rm T} - 1$; (c) A path $P_{\rm F}$ of length $t_{\rm F} - 1$.

Formally the scheme graph $\mathrm{SG} = (\mathcal{V}, \mathcal{E})$ is defined with a vertex set $\mathcal{V} = V_{\rm C} \cup V_{\rm T} \cup V_{\rm F}$ and an edge set $\mathcal{E} = E_{\rm C} \cup E_{\rm T} \cup E_{\rm F} \cup E_{\rm CT} \cup E_{\rm TC} \cup E_{\rm CF} \cup E_{\rm TF}$ that consist of the following sets. See Figure 1 for an illustration of these sets.

Construction of a $\sigma_{\rm int}$-extension H^* of $G_{\rm C}$: Denote the vertex set $V_{\rm C}$ and the edge set $E_{\rm C}$ in the seed graph $G_{\rm C}$ by $V_{\rm C} = \{v^{\rm C}{}_i \mid i \in [1, t_{\rm C}]\}$ and $E_{\rm C} = \{a_i \mid i \in [1, m_{\rm C}]\}$, respectively, where $V_{\rm C}$ is always included in H^*. For including additional interior-vertices in H^*, introduce a path $P_{\rm T} = (V_{\rm T} = \{v^{\rm T}{}_1, v^{\rm T}{}_2, \ldots, v^{\rm T}{}_{t_{\rm T}}\}, E_{\rm T} = \{e^{\rm T}{}_2, e^{\rm T}{}_3, \ldots, e^{\rm T}{}_{t_{\rm T}}\})$ of length $t_{\rm T} - 1$ and a set $E_{\rm CT}$ (resp., $E_{\rm TC}$) of directed edges $e^{\rm CT}{}_{i,j} = (v^{\rm C}{}_i, v^{\rm T}{}_j)$ (resp., $e^{\rm TC}{}_{i,j} = (v^{\rm T}{}_j, v^{\rm C}{}_i)$) $i \in [1, t_{\rm C}]$, $j \in [1, t_{\rm T}]$. In H^*, an edge $a_k = (v^{\rm C}{}_i, v^{\rm C}{}_{i'}) \in E_{(\geq 2)} \cup E_{(\geq 1)}$ is allowed to be replaced with a pure path P_k from vertex $v^{\rm C}{}_i$ to vertex $v^{\rm C}{}_{i'}$ that visits a set of consecutive vertices $v^{\rm T}{}_j, v^{\rm T}{}_{j+1}, \ldots, v^{\rm T}{}_{j+p} \in V_{\rm T}$ and edge $e^{\rm TC}{}_{i,j} = (v^{\rm C}{}_i, v^{\rm T}{}_j) \in E_{\rm CT}$, then edges $e^{\rm T}{}_{j+1}, e^{\rm T}{}_{j+2}, \ldots, e^{\rm T}{}_{j+p} \in E_{\rm T}$ and finally edge $e^{\rm TC}{}_{i',j+p} = (v^{\rm T}{}_{j+p}, v^{\rm C}{}_{i'}) \in E_{\rm TC}$. The vertices in $V_{\rm T}$ selected in the path will be vertices in H^*.

Appending leaf paths with additional interior-edges in a $(\sigma_{\rm int}, \sigma_{\rm ce})$-extension H of $G_{\rm C}$: Introduce a path $P_{\rm F} = (V_{\rm F} = \{v^{\rm F}{}_1, v^{\rm F}{}_2, \ldots, v^{\rm F}{}_{t_{\rm F}}\}, E_{\rm F} = \{e^{\rm F}{}_2, e^{\rm F}{}_3, \ldots, e^{\rm F}{}_{t_{\rm F}}\})$ of length $t_{\rm F} - 1$, a set $E_{\rm CF}$ of directed edges $e^{\rm CF}{}_{i,j} = (v^{\rm C}{}_i, v^{\rm F}{}_j)$, $i \in [1, t_{\rm C}]$, $j \in [1, t_{\rm F}]$, and a set $E_{\rm TF}$ of directed edges $e^{\rm TF}{}_{i,j} = (v^{\rm T}{}_i, v^{\rm F}{}_j)$, $i \in [1, t_{\rm T}]$, $j \in [1, t_{\rm F}]$. In H, a leaf path Q with interior-edges that starts from a vertex $v^{\rm C}{}_i \in V_{\rm C}$ (resp., $v^{\rm T}{}_i \in V_{\rm T}$) visits a set of consecutive vertices $v^{\rm F}{}_j, v^{\rm F}{}_{j+1}, \ldots, v^{\rm F}{}_{j+p} \in V_{\rm F}$ and edge $e^{\rm CF}{}_{i,j} = (v^{\rm C}{}_i, v^{\rm F}{}_j) \in E_{\rm CF}$ (resp., $e^{\rm TF}{}_{i,j} = (v^{\rm T}{}_i, v^{\rm F}{}_j) \in E_{\rm TF}$) and edges $e^{\rm F}{}_{j+1}, e^{\rm F}{}_{j+2}, \ldots, e^{\rm F}{}_{j+p} \in E_{\rm F}$. In H, the edges and the vertices selected in the path Q are regarded as interior-edges and interior-vertices, respectively.

Construction of ρ-fringe-trees in a $(\sigma_{\rm int}, \sigma_{\rm ce})$-extension G of $G_{\rm C}$: In H, the root of a ρ-fringe-tree can be any vertex in $V_{\rm C} \cup V_{\rm T} \cup V_{\rm F}$. For each vertex $v = v^{\rm C}{}_i$ (resp., $v = v^{\rm T}{}_i$ or $v^{\rm F}{}_i$), we choose a chemical rooted tree T from the specified set $\mathcal{F}(v)$ (resp., $\mathcal{F}_E$).

Recall that the dimension K of a feature vector $x = f(G)$ used in constructing a prediction function $\eta_{\mathcal{N}}$ over a set of chemical graphs G is $K = 17 + |\Lambda^{\rm int}(D_\pi)| + |\Lambda^{\rm ex}(D_\pi)| + |\Gamma^{\rm int}(D_\pi)| + |\mathcal{F}(D_\pi)|$. For a target specification $(G_{\rm C}, \sigma_{\rm int}, \sigma_{\rm ce})$, let $\mathcal{F}^*$ denote the set of chemical rooted trees ψ in the sets $\mathcal{F}(v), v \in V_{\rm C}$ and $\mathcal{F}_E$ and $K^* := 17 + |\Lambda^{\rm int}(D_\pi)| + |\Lambda^{\rm ex}(D_\pi)| + |\Gamma^{\rm int}(D_\pi)| + |\mathcal{F}^*|$. Based on the scheme graph SG, we obtain the following MILP formulation $\mathcal{M}(x, g; \mathcal{C}_2)$.

Theorem 2. *Let $(G_{\rm C}, \sigma_{\rm int}, \sigma_{\rm ce})$ be a target specification and $\varphi^* = |\Lambda^{\rm int}(D_\pi)| + |\Lambda^{\rm ex}(D_\pi)| + |\Gamma^{\rm int}(D_\pi)| + |\mathcal{F}^*|$ for sets of chemical elements, edge-configurations and fringe-configurations in $\sigma_{\rm ce}$. Then there is an MILP $\mathcal{M}(x, g; \mathcal{C}_2)$ that consists of variable vectors $x \in \mathbb{R}^{K^*}$ and $g \in \mathbb{R}^q$ for an integer $q = O(\mathrm{n}^{\rm int}_{\rm UB}(|E_{\rm C}| + n^*) + (|E_{\rm C}| + |\mathcal{V}|)\varphi^*)$ and a set $\mathcal{C}_2$ of $O([\mathrm{n}^{\rm int}_{\rm UB}(|E_{\rm C}| + n^*) + |\mathcal{V}|]\varphi^*)$ constraints on x and g such that: (x^*, g^*) is feasible to $\mathcal{M}(x, g; \mathcal{C}_2)$ if and only if g^* forms a chemical graph $G \in \mathcal{G}(G_{\rm C}, \sigma_{\rm int}, \sigma_{\rm ce})$ such that $f(G) = x^*$.*

Note that our MILP requires only $O(n^*)$ variables and constraints when the branch-parameter ρ, integers $|E_{\rm C}|$, $\mathrm{n}^{\rm int}_{\rm UB}$ and φ^* are constant. We explain the basic idea of our MILP that satisfies Theorem 2. The MILP mainly consists of the following three types of constraints.

C1. Constraints for selecting an underlying graph H of a chemical graph $G \in \mathcal{G}(G_{\rm C}, \sigma_{\rm int}, \sigma_{\rm ce})$ as a subgraph of the scheme graph SG;

C2. Constraints for assigning chemical elements to interior-vertices and multiplicity to interior-edges to determine a chemical graph $G = (H, \alpha, \beta)$; and

C3. Constraints for computing descriptors in the feature vector $f(G)$ of the selected chemical graph G.

In the constraints of C1, more formally we prepare the following.

Variables:

- a binary variable $v^{\rm X}(i) \in \{0, 1\}$ for each vertex $v^{\rm X}{}_i \in V_{\rm X}$, $\mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$ so that $v^{\rm X}(i) = 1 \Leftrightarrow$ vertex $v^{\rm X}{}_i$ is used in a graph H selected from SG;
- a binary variable $e^{\rm X}(i) \in \{0, 1\}$ (resp., $e^{\rm C}(i) \in \{0, 1\}$) for each edge $e^{\rm X}{}_i \in E_{\rm T} \cup E_{\rm F}$ (resp., $e^{\rm C}{}_i = a_i \in E_{(\geq 2)} \cup E_{(\geq 1)} \cup E_{(0/1)}$) so that $e^{\rm X}(i) = 1 \Leftrightarrow$ edge $e^{\rm X}{}_i$ is used in a graph H selected from SG. To save the number of variables in our MILP formulation, we do not prepare a binary variable $e^{\rm X}(i, j) \in \{0, 1\}$ for any edge $e^{\rm X}{}_{i,j} \in E_{\rm CT} \cup E_{\rm TC} \cup E_{\rm CF} \cup E_{\rm TC}$, where we represent a choice of edges in these sets by a set of $O(n^*|E_{\rm C}|)$ variables (see Supplementary Materials for the details);
- binary variables $\delta^{\rm C}_{\rm fr}(i, \psi) \in \{0, 1\}, i \in [1, t_{\rm C}], \psi \in \mathcal{F}(v), v = v^{\rm C}{}_i \in V_{\rm C}$ and $\delta^{\rm T}_{\rm fr}(i, \psi) \in \{0, 1\}, i \in [1, t_{\rm T}], \delta^{\rm F}_{\rm fr}(i, \psi) \in \{0, 1\}, i \in [1, t_{\rm F}], \psi \in \mathcal{F}_E$, where $\delta^{\rm X}_{\rm fr}(i, \psi) = 1$ ($\mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$) if and only if the ρ-fringe-tree rooted at vertex $v^{\rm X}{}_i$ is r-isomorphic to ψ.

Constraints:

- linear constraints so that each ρ-fringe-tree rooted at a vertex $v^{\rm X}{}_i$ in a graph H from SG is selected from the given set $\mathcal{F}(v^{\rm C}{}_i)$ for X=C (or $\mathcal{F}_E$ for $\mathrm{X} \in \{\mathrm{T}, \mathrm{F}\}$);
- linear constraints such that each edge $e^{\rm C}{}_i = a_i \in E_{(=1)}$ is always used as an edge in H and each edge $e^{\rm C}{}_i = a_i \in E_{(0/1)}$ is used as an edge in H if necessary;

- linear constraints such that for each edge $a_k = (v^{\mathrm{C}}{}_i, v^{\mathrm{C}}{}_{i'}) \in E_{(\geq 2)}$, vertex $v^{\mathrm{C}}{}_i \in V_{\mathrm{C}}$ is connected to vertex $v^{\mathrm{C}}{}_{i'} \in V_{\mathrm{C}}$ in H by a pure path P_k that passes through some vertices in V_{T} and edges $e^{\mathrm{CT}}{}_{i,j}, e^{\mathrm{T}}{}_{j+1}, e^{\mathrm{T}}{}_{j+2}, \ldots, e^{\mathrm{T}}{}_{j+p}, e^{\mathrm{TC}}{}_{i',j+p}$ for some integers j and p;
- linear constraints such that for each edge $a_k = (v^{\mathrm{C}}{}_i, v^{\mathrm{C}}{}_{i'}) \in E_{(\geq 1)}$, either the edge a_k is used as an edge in H or vertex $v^{\mathrm{C}}{}_i \in V_{\mathrm{C}}$ is connected to vertex $v^{\mathrm{C}}{}_{i'} \in V_{\mathrm{C}}$ in H by a pure path P_k as in the case of edges in $E_{(\geq 2)}$;
- linear constraints for selecting a leaf path Q_v rooted at a vertex $v = v^{\mathrm{C}}{}_i$ (resp., $v = v^{\mathrm{T}}{}_i$) with ρ-internal edges $e^{\mathrm{CF}}{}_{i,j}$ (resp., $e^{\mathrm{TF}}{}_{i,j}$), $e^{\mathrm{F}}{}_{j+1}, e^{\mathrm{F}}{}_{j+2}, \ldots, e^{\mathrm{F}}{}_{j+p}$ for some integers j and p.

In the constraints of C2, we prepare an integer variable $\alpha^{\mathrm{X}}(i)$ for each vertex $v^{\mathrm{X}}{}_i \in \mathcal{V}$, $\mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$ in the scheme graph that represents the chemical element $\alpha(v^{\mathrm{X}}{}_i) \in \Lambda$ if $v^{\mathrm{X}}{}_i$ is in a selected graph H (or $\alpha(v^{\mathrm{X}}{}_i) = 0$ otherwise); integer variables $\beta^{\mathrm{C}} : E_{\mathrm{C}} \to [0, 3]$, $\beta^{\mathrm{T}} : E_{\mathrm{T}} \to [0, 3]$ and $\beta^{\mathrm{F}} : E_{\mathrm{F}} \to [0, 3]$ that represent the bond-multiplicity of edges in $E_{\mathrm{C}} \cup E_{\mathrm{T}} \cup E_{\mathrm{F}}$; and integer variables $\beta^{+}, \beta^{-} : E_{(\geq 2)} \cup E_{(\geq 1)} \to [0, 3]$ and $\beta^{\mathrm{in}} : V_{\mathrm{C}} \cup V_{\mathrm{T}} \to [0, 3]$ that represent the bond-multiplicity of edges in $E_{\mathrm{CT}} \cup E_{\mathrm{TC}} \cup E_{\mathrm{CF}} \cup E_{\mathrm{TF}}$. This determines a chemical graph $G = (H, \alpha, \beta)$. Also we include constraints for a selected chemical graph G to satisfy the valence condition at each interior-vertex v with the edge-configurations $\mathrm{ec}(e)$ of the edges e incident to v and the chemical specification σ_{ce}.

In the constraints of C3, we introduce a variable for each descriptor and constraints with some more variables to compute the value of each descriptor in $f(G)$ for a selected chemical graph G.

The details of the MILP can be found in Section 3.

2 A Dynamic Programming Algorithm for Generating Isomers in Stage 5

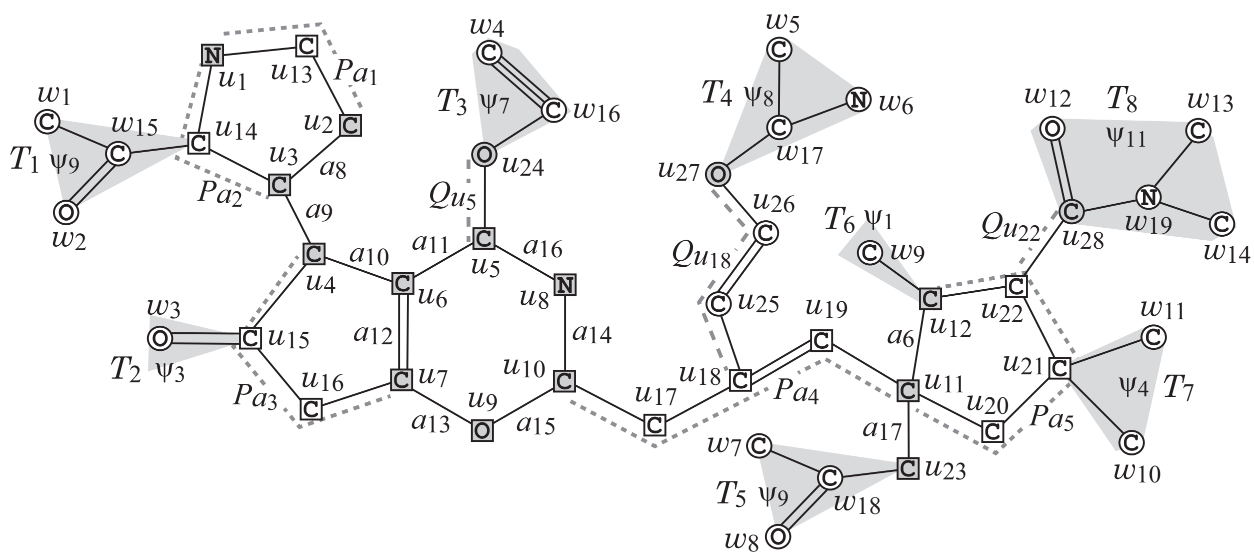

Figure 2: An illustration of a chemical graph G, where for $\rho = 2$, the exterior-vertices are $w_1, w_2, \ldots, w_{19}$ and the interior-vertices are $u_1, u_2, \ldots, u_{28}$.

This section briefly reviews the method [4] for Stage 5. Let $G^{\dagger}$ be a chemical graph that is a $(\sigma_{\mathrm{int}}, \sigma_{\mathrm{ce}})$-extension of a seed graph $G_{\mathrm{C}} = (V_{\mathrm{C}}, E_{\mathrm{C}})$, where we denote by $E_{(=0)}$ the set of the edges in $E_{(0/1)}$ that are not used in $G^\dagger$. We define a *base-graph* $G_B = (V_B, E_B)$ to be the seed graph $(V_\mathrm{C}, E_\mathrm{C} \setminus E_{(=0)})$ after removing the edges in $E_{(=0)}$. We call a chemical graph G^* a *chemical isomer* of $G^\dagger$ if $f(G^*) = f(G^\dagger)$ and G^* is also a $(\sigma_\mathrm{int}, \sigma_\mathrm{ce})$-extension of G_B.

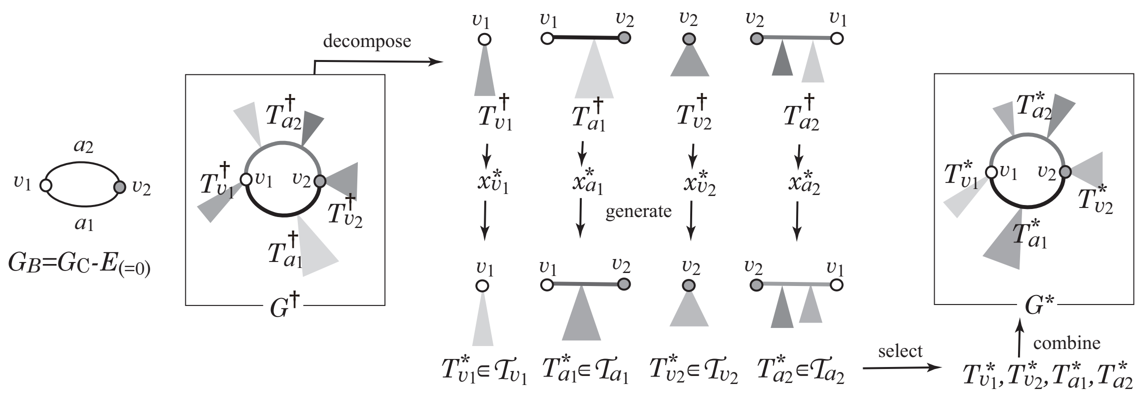

Figure 3: An illustration of generating a chemical isomer G^* of a chemical graph $G^\dagger$ with a base-graph $G_B = (V_B, E_B)$.

The method generates chemical isomers G^* of $G^\dagger$ in the following way, where Figure 3 illustrates the whole process in the case of $V_B = \{v_1, v_2\}$ and $E_B = \{a_1, a_2\}$.

1. We first decompose a given chemical graph $G^\dagger$ into a collection of chemical rooted or bi-rooted trees.

   - For each vertex $v \in V_B$, let $T^\dagger_v$ denote the chemical rooted tree rooted at v in G that is constructed with a leaf path Q_v and fringe-trees attached to Q_v. Possibly $T^\dagger_v$ consists of a single vertex v and we call such a tree *trivial*.
   - For each edge $a = uv \in E_{(\geq 2)} \cup E_{(\geq 1)}$, let $T^\dagger_a$ denote the chemical bi-rooted tree rooted at vertices u and v in G that consists of a pure u,v-path P_a, leaf paths rooted at internal vertices in P_a and fringe-trees attached to theses leaf paths. Possibly $T^\dagger_a$ consists of a single edge a and we call such a tree *trivial*.

   Figure 4 illustrates the non-trivial chemical trees $T^\dagger_\mathrm{t}, \mathrm{t} \in V^*_B \cup E^*_B$ of the $(\sigma_\mathrm{int}, \sigma_\mathrm{ce})$-extension $G^\dagger = G$ in Figure 2.

2. Let V^*_B (resp., E^*_B) denote the set of vertices $v \in V_B$ (resp., $a \in E_B$) such that $T^\dagger_v$ (resp., $T^\dagger_a$) is not trivial. For each vertex or edge $\mathrm{t} \in V^*_B \cup E^*_B$, compute the feature vector $x^*_\mathrm{t} = f(T^\dagger_\mathrm{t})$ and then generate a set $\mathcal{T}_\mathrm{t}$ of all (or a limited number of) chemical acyclic graphs T^*_t such that $f(T^*_\mathrm{t}) = x^*_\mathrm{t}$ and the structure of T^*_t satisfies the lower and upper bounds in the interior-specification σ_int by using the dynamic programming algorithm for chemical acyclic graphs [3].

3. For each combination of chemical trees $T^*_\mathrm{t} \in \mathcal{T}_\mathrm{t}, \mathrm{t} \in V^*_B \cup E^*_B$, a chemical graph G^* such that $f(G^*) = f(G^\dagger)$ is obtained from $G^\dagger$ by replacing each tree $T^\dagger_\mathrm{t}$ with a new tree T^*_t. The number of such combinations is $\prod_{\mathrm{t} \in V^*_B \cup E^*_B} |\mathcal{T}_\mathrm{t}|$, where we ignore a possible automorphism of the resulting graphs G^*.

The above method [4] can be used to generate chemical isomers in Stage 5 in our two-layered model by making a minor modification to the definition of a feature vector $f(G)$.

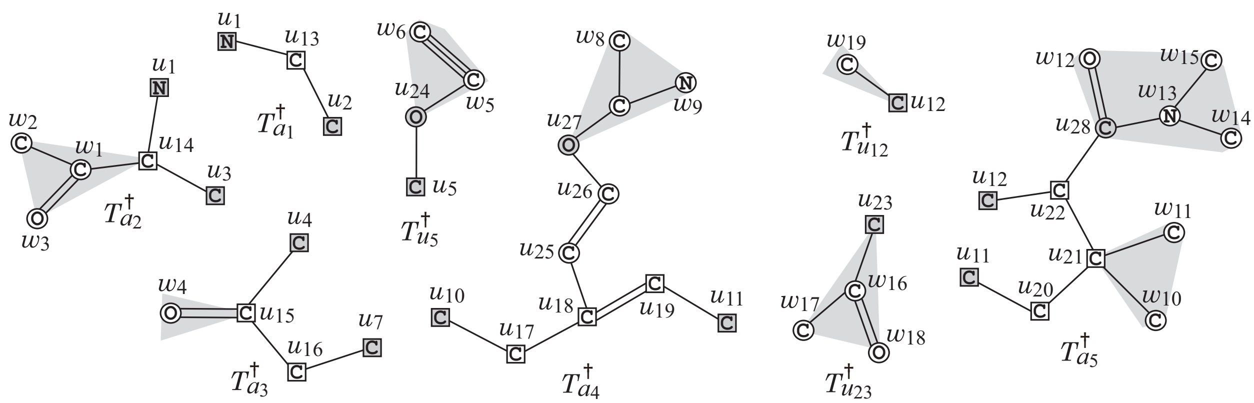

Figure 4: The non-trivial chemical rooted trees $T^\dagger_v$ for $v \in \{u_5, u_{12}, u_{23}\} = V^*_B$ and the non-trivial chemical bi-rooted trees $T^\dagger_a$ for $a \in \{a_1 = u_1u_2, a_2 = u_1u_3, a_3 = u_4u_7, a_4 = u_{10}u_{11}, a_5 = u_{11}u_{12}\} = E^*_B$ for the $(\sigma_{\mathrm{int}}, \sigma_{\mathrm{ce}})$-extension $G^\dagger = G$ in Figure 2, where the gray squares indicate the roots of these rooted and bi-rooted trees.

3 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 $[\epsilon] = 0$.

3.1 Selecting a Cyclical-base

Recall that

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

- Every edge $a_i \in E_{(=1)}$ is included in G;

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

- For each edge $a_i \in E_{(\geq 2)}$, edge a_i is not included in G 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 G; and

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

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., $a_i = (v^{\mathrm{C}}{}_{\mathrm{tail}(i)}, v^{\mathrm{C}}{}_{\mathrm{head}(i)})$.

Assume that $E_{\mathrm{C}} = \{a_i \mid i \in [1, m_{\mathrm{C}}]\}$, $E_{(\geq 2)} = \{a_k \mid k \in [1, p]\}$, $E_{(\geq 1)} = \{a_k \mid k \in [p+1, q]\}$, $E_{(0/1)} = \{a_i \mid i \in [q+1, t]\}$ and $E_{(=1)} = \{a_i \mid i \in [t+1, m_{\mathrm{C}}]\}$ for integers p, q and t. 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 $I_{(\geq 2)}$.

Define

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

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 G.

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;

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 G);
- $v^{\mathrm{T}}(i) \in [0, 1]$, $i \in [1, t_{\mathrm{T}}]$: $v^{\mathrm{T}}(i) = 1 \Leftrightarrow$ vertex $v^{\mathrm{T}}{}_i$ is used in G;
- $e^{\mathrm{T}}(i) \in [0, 1]$, $i \in [1, t_{\mathrm{T}}+1]$: $e^{\mathrm{T}}(i)$ 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 G);
- $\chi^{\mathrm{T}}(i) \in [0, k_{\mathrm{C}}]$, $i \in [1, t_{\mathrm{T}}]$: $\chi^{\mathrm{T}}(i)$ 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 G);
- $\mathrm{clr}^{\mathrm{T}}(k) \in [\ell_{\mathrm{LB}}(k) - 1, \ell_{\mathrm{UB}}(k) - 1]$, $k \in [1, k_{\mathrm{C}}]$, $\mathrm{clr}^{\mathrm{T}}(0) \in [0, t_{\mathrm{T}}]$: 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};

constraints:

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

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

$$e^{\mathrm{C}}(i) + \mathrm{clr}^{\mathrm{T}}(i) \geq 1, \qquad \mathrm{clr}^{\mathrm{T}}(i) \leq t_{\mathrm{T}} \cdot (1 - e^{\mathrm{C}}(i)), \quad i \in I_{(\geq 1)}, \tag{3}$$

$$\sum_{c \in I^{-}_{(\geq 1)}(i) \cup I^{-}_{(0/1)}(i) \cup I^{-}_{(=1)}(i)} e^{\mathrm{C}}(c) = \widetilde{\deg}^{-}_{\mathrm{C}}(i), \qquad \sum_{c \in I^{+}_{(\geq 1)}(i) \cup I^{+}_{(0/1)}(i) \cup I^{+}_{(=1)}(i)} e^{\mathrm{C}}(c) = \widetilde{\deg}^{+}_{\mathrm{C}}(i), \quad i \in [1, t_{\mathrm{C}}], \tag{4}$$

$$\chi^{\mathrm{T}}(i, 0) = 1 - v^{\mathrm{T}}(i), \qquad \sum_{k \in [0, k_{\mathrm{C}}]} \chi^{\mathrm{T}}(i, k) = 1, \qquad \sum_{k \in [0, k_{\mathrm{C}}]} k \cdot \chi^{\mathrm{T}}(i, k) = \chi^{\mathrm{T}}(i), \qquad i \in [1, t_{\mathrm{T}}], \tag{5}$$

$$\sum_{i \in [1, t_{\mathrm{T}}]} \chi^{\mathrm{T}}(i, k) = \mathrm{clr}^{\mathrm{T}}(k), \quad t_{\mathrm{T}} \cdot \delta^{\mathrm{T}}_{\chi}(k) \geq \sum_{i \in [1, t_{\mathrm{T}}]} \chi^{\mathrm{T}}(i, k) \geq \delta^{\mathrm{T}}_{\chi}(k), \qquad k \in [0, k_{\mathrm{C}}], \tag{6}$$

$$\begin{aligned} v^{\mathrm{T}}(i-1) &\geq v^{\mathrm{T}}(i), \\ k_{\mathrm{C}} \cdot (v^{\mathrm{T}}(i-1) - e^{\mathrm{T}}(i)) \geq \chi^{\mathrm{T}}(i-1) - \chi^{\mathrm{T}}(i) &\geq v^{\mathrm{T}}(i-1) - e^{\mathrm{T}}(i), \end{aligned} \qquad i \in [2, t_{\mathrm{T}}]. \tag{7}$$

3.2 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 G, 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_{F}: the maximum number of different colors assigned to the vertices in V_{F};
- $\mathrm{n}^{\mathrm{int}}_{\mathrm{LB}}, \mathrm{n}^{\mathrm{int}}_{\mathrm{UB}} \in [2, n^*]$: lower and upper bounds on the number of interior-vertices in G;
- $\mathrm{bl}_{\mathrm{LB}}(i) \in [0, 1]$, $i \in [1, \widetilde{t_{\mathrm{C}}}]$: a lower bound on the number of leaf ρ-branches in the leaf path rooted at a vertex ${v^{\mathrm{C}}}_i$;
- $\mathrm{bl}_{\mathrm{LB}}(k), \mathrm{bl}_{\mathrm{UB}}(k) \in [0, \ell_{\mathrm{UB}}(k) - 1]$, $k \in [1, k_{\mathrm{C}}] = I_{(\geq 2)} \cup I_{(\geq 1)}$: 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:

- $\mathrm{n}_G^{\mathrm{int}} \in [\mathrm{n}_{\mathrm{LB}}^{\mathrm{int}}, \mathrm{n}_{\mathrm{UB}}^{\mathrm{int}}]$: the number of interior-vertices in G;
- $v^{\mathrm{F}}(i) \in [0,1]$, $i \in [1, t_{\mathrm{F}}]$: $v^{\mathrm{F}}(i) = 1 \Leftrightarrow$ vertex ${v^{\mathrm{F}}}_i$ is used in G;
- $e^{\mathrm{F}}(i) \in [0,1]$, $i \in [1, t_{\mathrm{F}}+1]$: $e^{\mathrm{F}}(i)$ represents edge ${e^{\mathrm{F}}}_i = {v^{\mathrm{F}}}_{i-1}{v^{\mathrm{F}}}_i$, where ${e^{\mathrm{F}}}_1$ and ${e^{\mathrm{F}}}_{t_{\mathrm{F}}+1}$ are fictitious edges ($e^{\mathrm{F}}(i) = 1 \Leftrightarrow$ edge ${e^{\mathrm{F}}}_i$ is used in G);
- $\chi^{\mathrm{F}}(i) \in [0, c_{\mathrm{F}}]$, $i \in [1, t_{\mathrm{F}}]$: $\chi^{\mathrm{F}}(i)$ represents the color assigned to vertex ${v^{\mathrm{F}}}_i$ ($\chi^{\mathrm{F}}(i) = c \Leftrightarrow$ vertex ${v^{\mathrm{F}}}_i$ is assigned color c);
- $\mathrm{clr}^{\mathrm{F}}(c) \in [0, t_{\mathrm{F}}]$, $c \in [0, c_{\mathrm{F}}]$: the number of vertices ${v^{\mathrm{F}}}_i$ with color c;
- $\delta_\chi^{\mathrm{F}}(c) \in [\mathrm{bl}_{\mathrm{LB}}(c), 1]$, $c \in [1, \widetilde{t_{\mathrm{C}}}]$: $\delta_\chi^{\mathrm{F}}(c) = 1 \Leftrightarrow \chi^{\mathrm{F}}(i) = c$ for some $i \in [1, t_{\mathrm{F}}]$;
- $\delta_\chi^{\mathrm{F}}(c) \in [0, 1]$, $c \in [\widetilde{t_{\mathrm{C}}}+1, c_{\mathrm{F}}]$: $\delta_\chi^{\mathrm{F}}(c) = 1 \Leftrightarrow \chi^{\mathrm{F}}(i) = c$ for some $i \in [1, t_{\mathrm{F}}]$;
- $\chi^{\mathrm{F}}(i,c) \in [0,1]$, $i \in [1, t_{\mathrm{F}}]$, $c \in [0, c_{\mathrm{F}}]$: $\chi^{\mathrm{F}}(i,c) = 1 \Leftrightarrow \chi^{\mathrm{F}}(i) = c$;
- $\mathrm{bl}(k,i) \in [0,1]$, $k \in [1, k_{\mathrm{C}}] = I_{(\geq 2)} \cup I_{(\geq 1)}$, $i \in [1, t_{\mathrm{T}}]$: $\mathrm{bl}(k,i) = 1 \Leftrightarrow$ path P_k contains vertex ${v^{\mathrm{T}}}_i$ as an internal vertex and the ρ-fringe-tree rooted at ${v^{\mathrm{T}}}_i$ contains a leaf ρ-branch;

constraints:

$$\chi^{\mathrm{F}}(i,0) = 1 - v^{\mathrm{F}}(i), \quad \sum_{c \in [0, c_{\mathrm{F}}]} \chi^{\mathrm{F}}(i,c) = 1, \quad \sum_{c \in [0, c_{\mathrm{F}}]} c \cdot \chi^{\mathrm{F}}(i,c) = \chi^{\mathrm{F}}(i), \qquad i \in [1, t_{\mathrm{F}}], \tag{8}$$

$$\sum_{i \in [1, t_{\mathrm{F}}]} \chi^{\mathrm{F}}(i,c) = \mathrm{clr}^{\mathrm{F}}(c), \quad t_{\mathrm{F}} \cdot \delta_\chi^{\mathrm{F}}(c) \geq \sum_{i \in [1, t_{\mathrm{F}}]} \chi^{\mathrm{F}}(i,c) \geq \delta_\chi^{\mathrm{F}}(c), \qquad c \in [0, c_{\mathrm{F}}], \tag{9}$$

$$e^{\mathrm{F}}(1) = e^{\mathrm{F}}(t_{\mathrm{F}}+1) = 0, \tag{10}$$

$$\begin{aligned} v^{\mathrm{F}}(i-1) &\geq v^{\mathrm{F}}(i), \\ c_{\mathrm{F}} \cdot (v^{\mathrm{F}}(i-1) - e^{\mathrm{F}}(i)) \geq \chi^{\mathrm{F}}(i-1) - \chi^{\mathrm{F}}(i) &\geq v^{\mathrm{F}}(i-1) - e^{\mathrm{F}}(i), \end{aligned} \qquad i \in [2, t_{\mathrm{F}}], \tag{11}$$

$$\mathrm{bl}(k,i) \geq \delta_\chi^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i) + \chi^{\mathrm{T}}(i,k) - 1, \qquad k \in [1, k_{\mathrm{C}}], i \in [1, t_{\mathrm{T}}], \tag{12}$$

$$\sum_{k \in [1, k_{\mathrm{C}}], i \in [1, t_{\mathrm{T}}]} \mathrm{bl}(k,i) \leq \sum_{i \in [1, t_{\mathrm{T}}]} \delta_\chi^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i), \tag{13}$$

$$\mathrm{bl}_{\mathrm{LB}}(k) \leq \sum_{i \in [1, t_{\mathrm{T}}]} \mathrm{bl}(k,i) \leq \mathrm{bl}_{\mathrm{UB}}(k), \qquad k \in [1, k_{\mathrm{C}}], \tag{14}$$

$$t_{\mathrm{C}} + \sum_{i \in [1, t_{\mathrm{T}}]} v^{\mathrm{T}}(i) + \sum_{i \in [1, t_{\mathrm{F}}]} v^{\mathrm{F}}(i) = \mathrm{n}_G^{\mathrm{int}}. \tag{15}$$

3.3 Constraints for Including Fringe-trees

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:

- $n_{\rm LB}, n^*$: lower and upper bounds on $n(G)$, where $n_{\rm LB}, n^* \geq {\rm n}^{\rm int}_{\rm LB}$;
- ${\rm ch}_{\rm LB}(i), {\rm ch}_{\rm UB}(i) \in [0, n^*]$, $i \in [1, t_{\rm T}]$: lower and upper bounds on ${\rm ht}(T_i)$ of the tree T_i rooted at a vertex $v^{\rm C}{}_i$;
- ${\rm ch}_{\rm LB}(k), {\rm ch}_{\rm UB}(k) \in [0, n^*]$, $k \in [1, k_{\rm C}] = I_{(\geq 2)} \cup I_{(\geq 1)}$: lower and upper bounds on the maximum height ${\rm ht}(T)$ of the tree $T \in \mathcal{F}(P_k)$ rooted at an internal vertex of a path P_k for an edge $a_k \in E_{(\geq 1)} \cup E_{(\geq 2)}$;
- Let $\mathcal{F}_\Lambda$ denote the set of chemical rooted trees $\psi = (\{v\}, \emptyset)$ with ${\rm ht}(\psi) = 0$ and $\alpha(v) = \mathtt{a}$ for each chemical element $\mathtt{a} \in \Lambda$;
- 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 $\mathcal{F}(v) \subseteq \mathcal{F}(D_\pi), v \in V_{\rm C}$ and $\mathcal{F}_E \subseteq \mathcal{F}(D_\pi)$ of chemical rooted trees T with ${\rm ht}(T) \in [1, \rho]$;
- Define $\mathcal{F}^* := \bigcup_{v \in V_{\rm C}} \mathcal{F}(v) \cup \mathcal{F}_E$, $\mathcal{F}^{\rm C}_i := \mathcal{F}(v^{\rm C}{}_i)$, $i \in [1, t_{\rm C}]$, $\mathcal{F}^{\rm T}_i := \mathcal{F}_E$, $i \in [1, t_{\rm T}]$ and $\mathcal{F}^{\rm F}_i := \mathcal{F}_E$, $i \in [1, t_{\rm F}]$;
- $\mathcal{F}^{\rm X}_i[p], p \in [1, \rho], {\rm X} \in \{{\rm C}, {\rm T}, {\rm F}\}$: the set of chemical rooted trees $T \in \mathcal{F}^{\rm X}_i$ with ${\rm ht}(T) = p$;
- $n([\psi]) \in [0, 3^\rho], \psi \in \mathcal{F}^*$: the number of non-root vertices in a chemical rooted tree ψ;
- ${\rm ht}([\psi]) \in [0, \rho], \psi \in \mathcal{F}^*$: the height of a chemical rooted tree ψ;
- ${\rm deg}_{\rm r}([\psi]) \in [0, 4], \psi \in \mathcal{F}^*$: the number of children of the root r of a chemical rooted tree ψ;

variables:

- $n_G \in [n_{\rm LB}, n^*]$: $n(G)$;
- $v^{\rm X}(i) \in [0, 1], i \in [1, t_{\rm X}]$, ${\rm X} \in \{{\rm T}, {\rm F}\}$: $v^{\rm X}(i) = 1 \Leftrightarrow$ vertex $v^{\rm X}{}_i$ is used in G;
- $h^{\rm X}(i) \in [0, \rho]$, $i \in [1, t_{\rm X}]$, ${\rm X} \in \{{\rm C}, {\rm T}, {\rm F}\}$: the height of the ρ-fringe-tree rooted at vertex $v^{\rm X}{}_i$ in G;
- $\delta^{\rm X}_{\rm fr}(i, [\psi]) \in [0, 1], i \in [1, t_{\rm X}], \psi \in \mathcal{F}^{\rm X}_i \cup \mathcal{F}_\Lambda, {\rm X} \in \{{\rm T}, {\rm F}\}$: $\delta^{\rm X}_{\rm fr}(i, [\psi]) = 1 \Leftrightarrow \psi$ is the ρ-fringe-tree at vertex $v^{\rm X}{}_i$, where $\psi \in \mathcal{F}_\Lambda$ means that the height of the ρ-fringe-tree is 0;
- ${\rm deg}^{\rm ex}_{\rm X}(i) \in [0, 3], i \in [1, t_{\rm X}], {\rm X} \in \{{\rm C}, {\rm T}, {\rm F}\}$: the number of children of the root of the ρ-fringe-tree rooted at vertex $v^{\rm X}{}_i$ in G;
- $\sigma(k, i) \in [0, 1]$, $k \in [1, k_{\rm C}] = I_{(\geq 2)} \cup I_{(\geq 1)}, i \in [1, t_{\rm T}]$: $\sigma(k, i) = 1 \Leftrightarrow$ the ρ-fringe-tree T_v rooted at vertex $v = v^{\rm T}{}_i$ with color k has the largest height among such trees;

constraints:

$$\sum_{\psi \in \mathcal{F}^{\rm C}_i \cup \mathcal{F}_\Lambda} \delta^{\rm C}_{\rm fr}(i, [\psi]) = 1, \quad \sum_{\psi \in \mathcal{F}^{\rm C}_i \cup \mathcal{F}_\Lambda} {\rm deg}_{\rm r}([\psi]) \cdot \delta^{\rm C}_{\rm fr}(i, [\psi]) = {\rm deg}^{\rm ex}_{\rm C}(i), \qquad i \in [1, t_{\rm C}],$$

$$\sum_{\psi \in \mathcal{F}^{\rm X}_i \cup \mathcal{F}_\Lambda} \delta^{\rm X}_{\rm fr}(i, [\psi]) = v^{\rm X}(i), \quad \sum_{\psi \in \mathcal{F}^{\rm X}_i \cup \mathcal{F}_\Lambda} {\rm deg}_{\rm r}([\psi]) \cdot \delta^{\rm X}_{\rm fr}(i, [\psi]) = {\rm deg}^{\rm ex}_{\rm X}(i), \quad i \in [1, t_{\rm X}], {\rm X} \in \{{\rm T}, {\rm F}\}, \qquad (16)$$

$$\sum_{\psi \in \mathcal{F}_i^{\mathrm{F}}[\rho]} \delta_{\mathrm{fr}}^{\mathrm{F}}(i,[\psi]) \geq v^{\mathrm{F}}(i) - e^{\mathrm{F}}(i+1), \qquad i \in [1, t_{\mathrm{F}}] \ (e^{\mathrm{F}}(t_{\mathrm{F}}+1)=0), \tag{17}$$

$$\sum_{\psi \in \mathcal{F}_i^{\mathrm{X}}} \mathrm{ht}([\psi]) \cdot \delta_{\mathrm{fr}}^{\mathrm{X}}(i,[\psi]) = h^{\mathrm{X}}(i), \qquad i \in [1, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{C},\mathrm{T},\mathrm{F}\}, \tag{18}$$

$$\sum_{\substack{\psi \in \mathcal{F}_i^{\mathrm{X}} \\ i \in [1,t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{C},\mathrm{T},\mathrm{F}\}}} n([\psi]) \cdot \delta_{\mathrm{fr}}^{\mathrm{X}}(i,[\psi]) + \sum_{i \in [1,t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{T},\mathrm{F}\}} v^{\mathrm{X}}(i) + t_{\mathrm{C}} = n_G, \tag{19}$$

$$\begin{gathered} h^{\mathrm{C}}(i) \geq \mathrm{ch}_{\mathrm{LB}}(i) - n^* \delta_{\chi}^{\mathrm{F}}(i), \quad \mathrm{clr}^{\mathrm{F}}(i) + \rho \geq \mathrm{ch}_{\mathrm{LB}}(i), \\ h^{\mathrm{C}}(i) \leq \mathrm{ch}_{\mathrm{UB}}(i), \quad \mathrm{clr}^{\mathrm{F}}(i) + \rho \leq \mathrm{ch}_{\mathrm{UB}}(i) + n^*(1 - \delta_{\chi}^{\mathrm{F}}(i)), \ i \in [1, \widetilde{t_{\mathrm{C}}}], \end{gathered} \tag{20}$$

$$\mathrm{ch}_{\mathrm{LB}}(i) \leq h^{\mathrm{C}}(i) \leq \mathrm{ch}_{\mathrm{UB}}(i), \quad i \in [\widetilde{t_{\mathrm{C}}}+1, t_{\mathrm{C}}], \tag{21}$$

$$\begin{aligned} h^{\mathrm{T}}(i) &\leq \mathrm{ch}_{\mathrm{UB}}(k) + n^*(\delta_{\chi}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i) + 1 - \chi^{\mathrm{T}}(i,k)), \\ \mathrm{clr}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i) + \rho &\leq \mathrm{ch}_{\mathrm{UB}}(k) + n^*(2 - \delta_{\chi}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i) - \chi^{\mathrm{T}}(i,k)), \\ & k \in [1, k_{\mathrm{C}}], i \in [1, t_{\mathrm{T}}], \end{aligned} \tag{22}$$

$$\sum_{i \in [1,t_{\mathrm{T}}]} \sigma(k,i) = \delta_{\chi}^{\mathrm{T}}(k), \qquad k \in [1, k_{\mathrm{C}}], \tag{23}$$

$$\begin{aligned} \chi^{\mathrm{T}}(i,k) &\geq \sigma(k,i), \\ h^{\mathrm{T}}(i) &\geq \mathrm{ch}_{\mathrm{LB}}(k) - n^*(\delta_{\chi}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i) + 1 - \sigma(k,i)), \\ \mathrm{clr}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i) + \rho &\geq \mathrm{ch}_{\mathrm{LB}}(k) - n^*(2 - \delta_{\chi}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}}+i) - \sigma(k,i)), \qquad k \in [1, k_{\mathrm{C}}], i \in [1, t_{\mathrm{T}}]. \end{aligned} \tag{24}$$

3.4 Descriptor for the Number of Specified Degree

We include constraints to compute descriptors $\mathrm{dg}_d^{\mathrm{int}}(G)$, $d \in [1,4]$.

variables:

- $\deg^{\mathrm{X}}(i) \in [0,4]$, $i \in [1, t_{\mathrm{X}}]$, $\mathrm{X} \in \{\mathrm{C},\mathrm{T},\mathrm{F}\}$: the degree $\deg_G(v^{\mathrm{X}}{}_i)$ of vertex $v^{\mathrm{X}}{}_i$ in G;
- $\deg_{\mathrm{CT}}(i) \in [0,4]$, $i \in [1, t_{\mathrm{C}}]$: the number of edges from vertex $v^{\mathrm{C}}{}_i$ to vertices $v^{\mathrm{T}}{}_j$, $j \in [1, t_{\mathrm{T}}]$;
- $\deg_{\mathrm{TC}}(i) \in [0,4]$, $i \in [1, t_{\mathrm{C}}]$: the number of edges from vertices $v^{\mathrm{T}}{}_j$, $j \in [1, t_{\mathrm{T}}]$ to vertex $v^{\mathrm{C}}{}_i$;

- $\delta^{\rm C}_{\rm dg}(i,d) \in [0,1]$, $i \in [1,t_{\rm C}]$, $d \in [1,4]$, $\delta^{\rm X}_{\rm dg}(i,d) \in [0,1]$, $i \in [1,t_{\rm X}]$, $d \in [0,4]$, ${\rm X} \in \{{\rm T},{\rm F}\}$: $\delta^{\rm X}_{\rm dg}(i,d) = 1 \Leftrightarrow {\rm deg}^{\rm X}(i) = d$;
- ${\rm dg}(d) \in [{\rm dg}_{\rm LB}(d), {\rm dg}_{\rm UB}(d)]$, $d \in [1,4]$: the number of interior-vertices v with ${\rm deg}_G(v) = d$;
- ${\rm deg}^{\rm int}_{\rm C}(i) \in [1,4]$, $i \in [1,t_{\rm C}]$, ${\rm deg}^{\rm int}_{\rm X}(i) \in [0,4]$, $i \in [1,t_{\rm X}], {\rm X} \in \{{\rm T},{\rm F}\}$: the interior-degree ${\rm deg}_{(V^{\rm int},E^{\rm int})}(v^{\rm X}{}_i)$; i.e., the number of interior-edges incident to vertex $v^{\rm X}{}_i$;
- $\delta^{\rm int}_{\rm dg,C}(i,d) \in [0,1]$, $i \in [1,t_{\rm C}]$, $d \in [1,4]$, $\delta^{\rm int}_{\rm dg,X}(i,d) \in [0,1]$, $i \in [1,t_{\rm X}]$, $d \in [0,4]$, ${\rm X} \in \{{\rm T},{\rm F}\}$: $\delta^{\rm int}_{\rm dg,X}(i,d) = 1 \Leftrightarrow {\rm deg}^{\rm int}_{\rm X}(i) = d$;
- ${\rm dg}^{\rm int}(d) \in [{\rm dg}_{\rm LB}(d), {\rm dg}_{\rm UB}(d)]$, $d \in [1,4]$: the number of interior-vertices v with the interior-degree ${\rm deg}_{(V^{\rm int},E^{\rm int})}(v) = d$;

constraints:

$$\sum_{k \in I^+_{(\ge 2)}(i) \cup I^+_{(\ge 1)}(i)} \delta^{\rm T}_\chi(k) = {\rm deg}_{\rm CT}(i), \quad \sum_{k \in I^-_{(\ge 2)}(i) \cup I^-_{(\ge 1)}(i)} \delta^{\rm T}_\chi(k) = {\rm deg}_{\rm TC}(i), \qquad i \in [1,t_{\rm C}], \tag{25}$$

$$\widetilde{{\rm deg}}^-_{\rm C}(i) + \widetilde{{\rm deg}}^+_{\rm C}(i) + {\rm deg}_{\rm CT}(i) + {\rm deg}_{\rm TC}(i) + \delta^{\rm F}_\chi(i) = {\rm deg}^{\rm int}_{\rm C}(i), \qquad i \in [1,\widetilde{t_{\rm C}}], \tag{26}$$

$$\widetilde{{\rm deg}}^-_{\rm C}(i) + \widetilde{{\rm deg}}^+_{\rm C}(i) + {\rm deg}_{\rm CT}(i) + {\rm deg}_{\rm TC}(i) = {\rm deg}^{\rm int}_{\rm C}(i), \qquad i \in [\widetilde{t_{\rm C}}+1, t_{\rm C}], \tag{27}$$

$${\rm deg}^{\rm int}_{\rm C}(i) + {\rm deg}^{\rm ex}_{\rm C}(i) = {\rm deg}^{\rm C}(i), \qquad i \in [1,t_{\rm C}], \tag{28}$$

$$\sum_{\psi \in \mathcal{F}^{\rm C}_i[\rho]} \delta^{\rm C}_{\rm fr}(i,[\psi]) \ge 2 - {\rm deg}^{\rm int}_{\rm C}(i) \qquad i \in [1,t_{\rm C}], \tag{29}$$

$$\begin{aligned} 2v^{\rm T}(i) + \delta^{\rm F}_\chi(\widetilde{t_{\rm C}}+i) &= {\rm deg}^{\rm int}_{\rm T}(i), \\ {\rm deg}^{\rm int}_{\rm T}(i) + {\rm deg}^{\rm ex}_{\rm T}(i) &= {\rm deg}^{\rm T}(i), \qquad i \in [1,t_{\rm T}]\ (e^{\rm T}(1) = e^{\rm T}(t_{\rm T}+1) = 0), \end{aligned} \tag{30}$$

$$\begin{aligned} v^{\rm F}(i) + e^{\rm F}(i+1) &= {\rm deg}^{\rm int}_{\rm F}(i), \\ {\rm deg}^{\rm int}_{\rm F}(i) + {\rm deg}^{\rm ex}_{\rm F}(i) &= {\rm deg}^{\rm F}(i), \qquad i \in [1,t_{\rm F}]\ (e^{\rm F}(1) = e^{\rm F}(t_{\rm F}+1) = 0), \end{aligned} \tag{31}$$

$$\begin{aligned} &\sum_{d \in [0,4]} \delta^{\rm X}_{\rm dg}(i,d) = 1, \quad \sum_{d \in [1,4]} d \cdot \delta^{\rm X}_{\rm dg}(i,d) = {\rm deg}^{\rm X}(i), \\ &\sum_{d \in [0,4]} \delta^{\rm int}_{\rm dg,X}(i,d) = 1, \quad \sum_{d \in [1,4]} d \cdot \delta^{\rm int}_{\rm dg,X}(i,d) = {\rm deg}^{\rm int}_{\rm X}(i), \qquad i \in [1,t_{\rm X}], {\rm X} \in \{{\rm T},{\rm C},{\rm F}\}, \end{aligned} \tag{32}$$

$$\begin{aligned} &\sum_{i \in [1,t_{\rm C}]} \delta^{\rm C}_{\rm dg}(i,d) + \sum_{i \in [1,t_{\rm T}]} \delta^{\rm T}_{\rm dg}(i,d) + \sum_{i \in [1,t_{\rm F}]} \delta^{\rm F}_{\rm dg}(i,d) = {\rm dg}(d), \\ &\sum_{i \in [1,t_{\rm C}]} \delta^{\rm int}_{\rm dg,C}(i,d) + \sum_{i \in [1,t_{\rm T}]} \delta^{\rm int}_{\rm dg,T}(i,d) + \sum_{i \in [1,t_{\rm F}]} \delta^{\rm int}_{\rm dg,F}(i,d) = {\rm dg}^{\rm int}(d), \qquad d \in [1,4]. \end{aligned} \tag{33}$$

3.5 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 G and include necessary constraints for the variables to satisfy in G.

constants:

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

constraints:

$$e^{\mathrm{C}}(i) \leq \beta^{\mathrm{C}}(i) \leq 3e^{\mathrm{C}}(i), i \in [\widetilde{k_{\mathrm{C}}}+1, m_{\mathrm{C}}] = I_{(\geq 1)} \cup I_{(0/1)} \cup I_{(=1)}, \tag{34}$$

$$e^{\mathrm{X}}(i) \leq \beta^{\mathrm{X}}(i) \leq 3e^{\mathrm{X}}(i), \qquad i \in [2, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{T}, \mathrm{F}\}, \tag{35}$$

$$\delta^{\mathrm{T}}_{\chi}(k) \leq \beta^+(k) \leq 3\delta^{\mathrm{T}}_{\chi}(k), \quad \delta^{\mathrm{T}}_{\chi}(k) \leq \beta^-(k) \leq 3\delta^{\mathrm{T}}_{\chi}(k), \qquad k \in [1, k_{\mathrm{C}}], \tag{36}$$

$$\delta^{\mathrm{F}}_{\chi}(c) \leq \beta^{\mathrm{in}}(c) \leq 3\delta^{\mathrm{F}}_{\chi}(c), \qquad c \in [1, c_{\mathrm{F}}], \tag{37}$$

$$\sum_{m\in[0,3]}\delta_\beta^{\mathrm{X}}(i,m)=1,\quad \sum_{m\in[0,3]} m\cdot\delta_\beta^{\mathrm{X}}(i,m)=\beta^{\mathrm{X}}(i),\qquad i\in[2,t_{\mathrm{X}}],\mathrm{X}\in\{\mathrm{T},\mathrm{F}\},\tag{38}$$

$$\sum_{m\in[0,3]}\delta_\beta^{\mathrm{C}}(i,m)=1,\quad \sum_{m\in[0,3]} m\cdot\delta_\beta^{\mathrm{C}}(i,m)=\beta^{\mathrm{C}}(i),\qquad i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}],\tag{39}$$

$$\begin{aligned}
&\sum_{m\in[0,3]}\delta_\beta^{+}(k,m)=1,\quad \sum_{m\in[0,3]} m\cdot\delta_\beta^{+}(k,m)=\beta^{+}(k), && k\in[1,k_{\mathrm{C}}],\\
&\sum_{m\in[0,3]}\delta_\beta^{-}(k,m)=1,\quad \sum_{m\in[0,3]} m\cdot\delta_\beta^{-}(k,m)=\beta^{-}(k), && k\in[1,k_{\mathrm{C}}],\\
&\sum_{m\in[0,3]}\delta_\beta^{\mathrm{in}}(c,m)=1,\quad \sum_{m\in[0,3]} m\cdot\delta_\beta^{\mathrm{in}}(c,m)=\beta^{\mathrm{in}}(c), && c\in[1,c_{\mathrm{F}}],
\end{aligned}\tag{40}$$

$$\sum_{\psi\in\mathcal{F}_i^{\mathrm{X}}}\beta_{\mathrm{r}}([\psi])\cdot\delta_{\mathrm{fr}}^{\mathrm{X}}(i,[\psi])=\beta_{\mathrm{ex}}^{\mathrm{X}}(i),\qquad i\in[1,t_{\mathrm{X}}],\mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F}\},\tag{41}$$

$$\begin{aligned}
\sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]}\delta_\beta^{\mathrm{C}}(i,m)=\mathrm{bd}_{\mathrm{C}}(m),\quad &\sum_{i\in[2,t_{\mathrm{T}}]}\delta_\beta^{\mathrm{T}}(i,m)=\mathrm{bd}_{\mathrm{T}}(m),\\
\sum_{k\in[1,k_{\mathrm{C}}]}\delta_\beta^{+}(k,m)=\mathrm{bd}_{\mathrm{CT}}(m),\quad &\sum_{k\in[1,k_{\mathrm{C}}]}\delta_\beta^{-}(k,m)=\mathrm{bd}_{\mathrm{TC}}(m),\\
\sum_{i\in[2,t_{\mathrm{F}}]}\delta_\beta^{\mathrm{F}}(i,m)=\mathrm{bd}_{\mathrm{F}}(m),\quad &\sum_{c\in[1,\widetilde{t_{\mathrm{C}}}]}\delta_\beta^{\mathrm{in}}(c,m)=\mathrm{bd}_{\mathrm{CF}}(m),\\
&\sum_{c\in[\widetilde{t_{\mathrm{C}}}+1,c_{\mathrm{F}}]}\delta_\beta^{\mathrm{in}}(c,m)=\mathrm{bd}_{\mathrm{TF}}(m),\\
\mathrm{bd}_{\mathrm{C}}(m)+\mathrm{bd}_{\mathrm{T}}(m)+\mathrm{bd}_{\mathrm{F}}(m)+\mathrm{bd}_{\mathrm{CT}}(m)+\mathrm{bd}_{\mathrm{TC}}(m)&+\mathrm{bd}_{\mathrm{TF}}(m)+\mathrm{bd}_{\mathrm{CF}}(m)=\mathrm{bd}^{\mathrm{int}}(m),\\
&m\in[1,3].
\end{aligned}\tag{42}$$

3.6 Assigning Chemical Elements and Valence Condition

We include constraints so that each vertex u in a selected graph H satisfies the valence condition; i.e., $\sum_{uv\in E(H)}\beta(uv)\le\mathrm{val}(\alpha(u))$. With these constraints, a chemical graph $G=(H,\alpha,\beta)$ on a selected subgraph H will be constructed.

constants:

- Subsets $\Lambda^{\mathrm{int}},\Lambda^{\mathrm{ex}}\subseteq\Lambda$ 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: $\mathrm{val}:\Lambda\to[1,4]$;
- 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 $\Lambda^*(i) \subseteq \Lambda^{\mathrm{int}}$, $i \in [1, t_{\mathrm{C}}]$;
- $\mathrm{na}_{\mathrm{LB}}(\mathtt{a}), \mathrm{na}_{\mathrm{UB}}(\mathtt{a}) \in [0, n^*]$, $\mathtt{a} \in \Lambda$: lower and upper bounds on the number of vertices v with $\alpha(v) = \mathtt{a}$;
- $\mathrm{na}^{\mathrm{int}}_{\mathrm{LB}}(\mathtt{a}), \mathrm{na}^{\mathrm{int}}_{\mathrm{UB}}(\mathtt{a}) \in [0, n^*]$, $\mathtt{a} \in \Lambda^{\mathrm{int}}$: lower and upper bounds on the number of interior-vertices v with $\alpha(v) = \mathtt{a}$;
- $\alpha_{\mathrm{r}}([\psi]) \in [\Lambda^{\mathrm{ex}}], \in \mathcal{F}^* \cup \mathcal{F}_\Lambda$: the chemical element $\alpha(r)$ of the root r of ψ;
- $\mathrm{na}^{\mathrm{ex}}_{\mathtt{a}}([\psi]) \in [0, n^*]$, $\mathtt{a} \in \Lambda^{\mathrm{ex}}, \psi \in \mathcal{F}^*$: the frequency of chemical element $\mathtt{a}$ in the set of non-rooted vertices in ψ;
- $n_{\mathtt{H}}([\psi], d) \in [0, 3^\rho], \psi \in \mathcal{F}^* \cup \mathcal{F}_\Lambda, d \in [0, 3]$: the number of non-root vertices with $\deg_{\mathrm{hyd}}(v) = d$ in ψ.

variables:

- $\beta^{\mathrm{CT}}(i), \beta^{\mathrm{TC}}(i) \in [0, 3], i \in [1, t_{\mathrm{T}}]$: the bond-multiplicity of edge ${e^{\mathrm{CT}}}_{j,i}$ (resp., ${e^{\mathrm{TC}}}_{j,i}$) if one exists;
- $\beta^{\mathrm{CF}}(i), \beta^{\mathrm{TF}}(i) \in [0, 3], i \in [1, t_{\mathrm{F}}]$: the bond-multiplicity of ${e^{\mathrm{CF}}}_{j,i}$ (resp., ${e^{\mathrm{TF}}}_{j,i}$) if one exists;
- $\alpha^{\mathrm{X}}(i) \in [\Lambda^{\mathrm{int}}_\epsilon], \delta^{\mathrm{X}}_\alpha(i, [\mathtt{a}]^{\mathrm{int}}) \in [0, 1], \mathtt{a} \in \Lambda^{\mathrm{int}}_\epsilon, i \in [1, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: $\alpha^{\mathrm{X}}(i) = [\mathtt{a}]^{\mathrm{int}} \geq 1$ (resp., $\alpha^{\mathrm{X}}(i) = 0$) $\Leftrightarrow \delta^{\mathrm{X}}_\alpha(i, [\mathtt{a}]^{\mathrm{int}}) = 1$ (resp., $\delta^{\mathrm{X}}_\alpha(i, 0) = 0$) $\Leftrightarrow \alpha({v^{\mathrm{X}}}_i) = \mathtt{a} \in \Lambda$ (resp., vertex ${v^{\mathrm{X}}}_i$ is not used in G);
- $\delta^{\mathrm{X}}_\alpha(i, [\mathtt{a}]^{\mathrm{int}}) \in [0, 1], i \in [1, t_{\mathrm{X}}], \mathtt{a} \in \Lambda^{\mathrm{int}}, \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: $\delta^{\mathrm{X}}_\alpha(i, [\mathtt{a}]^{\mathrm{t}}) = 1 \Leftrightarrow \alpha({v^{\mathrm{X}}}_i) = \mathtt{a}$;
- $\mathrm{Mass} \in \mathbb{Z}_+$: $\sum_{v \in V(H)} \mathrm{mass}^*(\alpha(v))$;
- $\mathrm{na}([\mathtt{a}]) \in [\mathrm{na}_{\mathrm{LB}}(\mathtt{a}), \mathrm{na}_{\mathrm{UB}}(\mathtt{a})]$, $\mathtt{a} \in \Lambda$: the number of vertices $v \in V(H)$ with $\alpha(v) = \mathtt{a}$;
- $\mathrm{na}^{\mathrm{int}}([\mathtt{a}]^{\mathrm{int}}) \in [\mathrm{na}^{\mathrm{int}}_{\mathrm{LB}}(\mathtt{a}), \mathrm{na}^{\mathrm{int}}_{\mathrm{UB}}(\mathtt{a})]$, $\mathtt{a} \in \Lambda, \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: the number of interior-vertices $v \in V(G)$ with $\alpha(v) = \mathtt{a}$;
- $\mathrm{na}^{\mathrm{ex}}_{\mathrm{X}}([\mathtt{a}]^{\mathrm{ex}}), \mathrm{na}^{\mathrm{ex}}([\mathtt{a}]^{\mathrm{ex}}) \in [0, \mathrm{na}_{\mathrm{UB}}(\mathtt{a})]$, $\mathtt{a} \in \Lambda$, $\mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: the number of exterior-vertices rooted at vertices $v \in V_{\mathrm{X}}$ and the number of exterior-vertices v such that $\alpha(v) = \mathtt{a}$;
- $\delta^{\mathrm{X}}_{\mathrm{hyd}}(i, d) \in [0, 1], d \in [0, 3], \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: $\delta^{\mathrm{X}}_{\mathrm{hyd}}(i, d) \Leftrightarrow \deg_{\mathrm{hyd}}({v^{\mathrm{X}}}_i) = d$;
- $\mathrm{hydg}(d), d \in [0, 3]$: the number of vertices v with $\deg_{\mathrm{hyd}}({v^{\mathrm{X}}}_i) = d$;

constraints:

$$\begin{aligned}
\beta^+(k) - 3(e^{\mathrm{T}}(i) - \chi^{\mathrm{T}}(i, k) + 1) \leq \beta^{\mathrm{CT}}(i) \leq \beta^+(k) + 3(e^{\mathrm{T}}(i) - \chi^{\mathrm{T}}(i, k) + 1), i \in [1, t_{\mathrm{T}}],\\
\beta^-(k) - 3(e^{\mathrm{T}}(i+1) - \chi^{\mathrm{T}}(i, k) + 1) \leq \beta^{\mathrm{TC}}(i) \leq \beta^-(k) + 3(e^{\mathrm{T}}(i+1) - \chi^{\mathrm{T}}(i, k) + 1), i \in [1, t_{\mathrm{T}}],\\
k \in [1, k_{\mathrm{C}}],
\end{aligned} \tag{43}$$

$$\begin{aligned}
&\beta^{\mathrm{in}}(c) - 3(e^{\mathrm{F}}(i) - \chi^{\mathrm{F}}(i, c) + 1) \leq \beta^{\mathrm{CF}}(i) \leq \beta^{\mathrm{in}}(c) + 3(e^{\mathrm{F}}(i) - \chi^{\mathrm{F}}(i, c) + 1), i \in [1, t_{\mathrm{F}}], \quad c \in [1, \widetilde{t_{\mathrm{C}}}],\\
&\beta^{\mathrm{in}}(c) - 3(e^{\mathrm{F}}(i) - \chi^{\mathrm{F}}(i, c) + 1) \leq \beta^{\mathrm{TF}}(i) \leq \beta^{\mathrm{in}}(c) + 3(e^{\mathrm{F}}(i) - \chi^{\mathrm{F}}(i, c) + 1), i \in [1, t_{\mathrm{F}}], \quad c \in [\widetilde{t_{\mathrm{C}}} + 1, c_{\mathrm{F}}],
\end{aligned} \tag{44}$$

$$\sum_{\mathtt{a}\in\Lambda^{\rm int}} \delta_\alpha^{\rm C}(i,[\mathtt{a}]^{\rm int}) = 1, \quad \sum_{\mathtt{a}\in\Lambda^{\rm int}} [\mathtt{a}]^{\rm int}\cdot\delta_\alpha^{\rm X}(i,[\mathtt{a}]^{\rm int}) = \alpha^{\rm C}(i), \qquad i\in[1,t_{\rm C}],$$
$$\sum_{\mathtt{a}\in\Lambda^{\rm int}} \delta_\alpha^{\rm X}(i,[\mathtt{a}]^{\rm int}) = v^{\rm X}(i), \quad \sum_{\mathtt{a}\in\Lambda^{\rm int}} [\mathtt{a}]^{\rm int}\cdot\delta_\alpha^{\rm X}(i,[\mathtt{a}]^{\rm int}) = \alpha^{\rm X}(i), \qquad i\in[1,t_{\rm X}], {\rm X}\in\{{\rm T},{\rm F}\}, \tag{45}$$

$$\sum_{\psi\in\mathcal{F}_i^{\rm X}\cup\mathcal{F}_\Lambda} \alpha_{\rm r}([\psi])\cdot\delta_{\rm fr}^{\rm X}(i,[\psi]) = \alpha^{\rm X}(i), \qquad i\in[1,t_{\rm X}], {\rm X}\in\{{\rm C},{\rm T},{\rm F}\}, \tag{46}$$

$$\sum_{j\in I_{\rm C}(i)}\beta^{\rm C}(j) + \sum_{k\in I^+_{(\geq 2)}(i)\cup I^+_{(\geq 1)}(i)}\beta^+(k) + \sum_{k\in I^-_{(\geq 2)}(i)\cup I^-_{(\geq 1)}(i)}\beta^-(k)$$
$$+\beta^{\rm in}(i) + \beta^{\rm C}_{\rm ex}(i) + \sum_{d\in[0,3]} d\cdot\delta^{\rm C}_{\rm hyd}(i,d) = \sum_{\mathtt{a}\in\Lambda^{\rm int}} {\rm val}(\mathtt{a})\delta_\alpha^{\rm C}(i,[\mathtt{a}]^{\rm int}), \qquad i\in[1,\widetilde{t_{\rm C}}], \tag{47}$$

$$\sum_{j\in I_{\rm C}(i)}\beta^{\rm C}(j) + \sum_{k\in I^+_{(\geq 2)}(i)\cup I^+_{(\geq 1)}(i)}\beta^+(k) + \sum_{k\in I^-_{(\geq 2)}(i)\cup I^-_{(\geq 1)}(i)}\beta^-(k)$$
$$+\beta^{\rm C}_{\rm ex}(i) + \sum_{d\in[0,3]} d\cdot\delta^{\rm C}_{\rm hyd}(i,d) = \sum_{\mathtt{a}\in\Lambda^{\rm int}} {\rm val}(\mathtt{a})\delta_\alpha^{\rm C}(i,[\mathtt{a}]^{\rm int}), \qquad i\in[\widetilde{t_{\rm C}}+1,t_{\rm C}], \tag{48}$$

$$\beta^{\rm T}(i) + \beta^{\rm T}(i{+}1) + \beta^{\rm T}_{\rm ex}(i) + \beta^{\rm CT}(i) + \beta^{\rm TC}(i)$$
$$+\beta^{\rm in}(\widetilde{t_{\rm C}}+i) + \sum_{d\in[0,3]} d\cdot\delta^{\rm T}_{\rm hyd}(i,d) = \sum_{\mathtt{a}\in\Lambda^{\rm int}} {\rm val}(\mathtt{a})\delta_\alpha^{\rm T}(i,[\mathtt{a}]^{\rm int}),$$
$$i\in[1,t_{\rm T}]\ (\beta^{\rm T}(1)=\beta^{\rm T}(t_{\rm T}+1)=0), \tag{49}$$

$$\beta^{\rm F}(i) + \beta^{\rm F}(i{+}1) + \beta^{\rm CF}(i) + \beta^{\rm TF}(i)$$
$$+\beta^{\rm F}_{\rm ex}(i) + \sum_{d\in[0,3]} d\cdot\delta^{\rm F}_{\rm hyd}(i,d) = \sum_{\mathtt{a}\in\Lambda^{\rm int}} {\rm val}(\mathtt{a})\delta_\alpha^{\rm F}(i,[\mathtt{a}]^{\rm int}),$$
$$i\in[1,t_{\rm F}]\ (\beta^{\rm F}(1)=\beta^{\rm F}(t_{\rm F}+1)=0), \tag{50}$$

$$\sum_{i\in[1,t_{\rm X}]}\delta_\alpha^{\rm X}(i,[\mathtt{a}]^{\rm int}) = {\rm na}_{\rm X}([\mathtt{a}]^{\rm int}), \qquad \mathtt{a}\in\Lambda^{\rm int}, {\rm X}\in\{{\rm C},{\rm T},{\rm F}\}, \tag{51}$$

$$\sum_{\psi\in\mathcal{F}_i^{\rm X}} {\rm na}^{\rm ex}_{\mathtt{a}}([\psi])\cdot\delta_{\rm fr}^{\rm X}(i,[\psi]) = {\rm na}^{\rm ex}_{\rm X}([\mathtt{a}]^{\rm ex}), \qquad \mathtt{a}\in\Lambda^{\rm ex}, {\rm X}\in\{{\rm C},{\rm T},{\rm F}\}, \tag{52}$$

$$
\begin{aligned}
\mathrm{na}_{\mathrm{C}}([\mathtt{a}]^{\mathrm{int}}) + \mathrm{na}_{\mathrm{T}}([\mathtt{a}]^{\mathrm{int}}) + \mathrm{na}_{\mathrm{F}}([\mathtt{a}]^{\mathrm{int}}) &= \mathrm{na}^{\mathrm{int}}([\mathtt{a}]^{\mathrm{int}}), & \mathtt{a} \in \Lambda^{\mathrm{int}}, \\
\sum_{\mathrm{X}\in\{\mathrm{C,T,F}\}} \mathrm{na}^{\mathrm{ex}}_{\mathrm{X}}([\mathtt{a}]^{\mathrm{ex}}) &= \mathrm{na}^{\mathrm{ex}}([\mathtt{a}]^{\mathrm{ex}}), & \mathtt{a} \in \Lambda^{\mathrm{ex}}, \\
\mathrm{na}^{\mathrm{int}}([\mathtt{a}]^{\mathrm{int}}) + \mathrm{na}^{\mathrm{ex}}([\mathtt{a}]^{\mathrm{ex}}) &= \mathrm{na}([\mathtt{a}]), & \mathtt{a} \in \Lambda^{\mathrm{int}} \cap \Lambda^{\mathrm{ex}}, \\
\mathrm{na}^{\mathrm{int}}([\mathtt{a}]^{\mathrm{int}}) &= \mathrm{na}([\mathtt{a}]), & \mathtt{a} \in \Lambda^{\mathrm{int}} \setminus \Lambda^{\mathrm{ex}}, \\
\mathrm{na}^{\mathrm{ex}}([\mathtt{a}]^{\mathrm{ex}}) &= \mathrm{na}([\mathtt{a}]), & \mathtt{a} \in \Lambda^{\mathrm{ex}} \setminus \Lambda^{\mathrm{int}}, \quad (53)
\end{aligned}
$$

$$
\sum_{\mathtt{a}\in\Lambda} \mathrm{mass}^*(\mathtt{a}) \cdot \mathrm{na}([\mathtt{a}]) = \mathrm{Mass}, \quad (54)
$$

$$
\begin{aligned}
\sum_{d\in[0,3]} \delta^{\mathrm{C}}_{\mathrm{hyd}}(i,d) &= 1, i \in [1, t_{\mathrm{C}}], \\
\sum_{d\in[0,3]} \delta^{\mathrm{X}}_{\mathrm{hyd}}(i,d) &= v^{\mathrm{X}}(i), i \in [1, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{T,F}\}, \quad (55)
\end{aligned}
$$

$$
\sum_{i\in[1,t_{\mathrm{X}}],\mathrm{X}\in\{\mathrm{C,T,F}\}} \delta^{\mathrm{X}}_{\mathrm{hyd}}(i,d) + \sum_{\psi\in\mathcal{F}^{\mathrm{X}}_i, i\in[1,t_{\mathrm{X}}],\mathrm{X}\in\{\mathrm{C,T,F}\}} n_{\mathtt{H}}([\psi],d) \cdot \delta^{\mathrm{X}}_{\mathrm{fr}}(i,[\psi]) = \mathrm{hydg}(d), d \in [0,3], \quad (56)
$$

$$
\sum_{\mathtt{a}\in\Lambda^*(i)} \delta^{\mathrm{C}}_{\alpha}(i, [\mathtt{a}]^{\mathrm{int}}) = 1, \qquad i \in [1, t_{\mathrm{C}}]. \quad (57)
$$

3.7 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:

$$
\mathrm{bd}_{m,\mathrm{LB}}(i) \le \delta^{\mathrm{C}}_{\beta}(i,m) \le \mathrm{bd}_{m,\mathrm{UB}}(i), i \in I_{(=1)} \cup I_{(0/1)}, m \in [2,3], \quad (58)
$$

$$
\mathrm{bd}_{\mathrm{T}}(k,i,m) \ge \delta^{\mathrm{T}}_{\beta}(i,m) + \chi^{\mathrm{T}}(i,k) - 1, \quad k \in [1, k_{\mathrm{C}}], i \in [2, t_{\mathrm{T}}], m \in [2,3], \quad (59)
$$

$$\sum_{j\in[2,t_{\mathrm{T}}]} \delta_\beta^{\mathrm{T}}(j,m) \geq \sum_{k\in[1,k_{\mathrm{C}}],i\in[2,t_{\mathrm{T}}]} \mathrm{bd}_{\mathrm{T}}(k,i,m), \quad m\in[2,3], \tag{60}$$

$$\mathrm{bd}_{m,\mathrm{LB}}(k) \leq \sum_{i\in[2,t_{\mathrm{T}}]} \mathrm{bd}_{\mathrm{T}}(k,i,m) + \delta_\beta^{+}(k,m) + \delta_\beta^{-}(k,m) \leq \mathrm{bd}_{m,\mathrm{UB}}(k),$$
$$k\in[1,k_{\mathrm{C}}], m\in[2,3]. \tag{61}$$

3.8 Descriptor for the Number of Adjacency-configurations

We call a tuple $(\mathtt{a},\mathtt{b},m)\in\Lambda\times\Lambda\times[1,3]$ 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 G.

constants:

- A set Γ^{int} of edge-configurations $\gamma=(\mu,\xi,m)$ with $\mu\leq\xi$;
- Let $\overline{\gamma}$ of an edge-configuration $\gamma=(\mu,\xi,m)$ denote the edge-configuration (ξ,μ,m);
- Let $\Gamma^{\mathrm{int}}_{<}=\{(\mu,\xi,m)\in\Gamma^{\mathrm{int}}\mid\mu<\xi\}$, $\Gamma^{\mathrm{int}}_{=}=\{(\mu,\xi,m)\in\Gamma^{\mathrm{int}}\mid\mu=\xi\}$ and $\Gamma^{\mathrm{int}}_{>}=\{\overline{\gamma}\mid\gamma\in\Gamma^{\mathrm{int}}_{<}\}$;
- Let $\Gamma^{\mathrm{int}}_{\mathrm{ac},<}$, $\Gamma^{\mathrm{int}}_{\mathrm{ac},=}$ and $\Gamma^{\mathrm{int}}_{\mathrm{ac},>}$ denote the sets of the adjacency-configurations of edge-configurations in the sets $\Gamma^{\mathrm{int}}_{<}$, $\Gamma^{\mathrm{int}}_{=}$ and $\Gamma^{\mathrm{int}}_{>}$, 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 $\Gamma^{\mathrm{int}}_{\mathrm{ac}}\cup\Gamma^{\mathrm{int}}_{\mathrm{ac},>}$ and let $[\nu]^{\mathrm{int}}$ denote the coded integer of an element ν in $\Gamma^{\mathrm{int}}_{\mathrm{ac}}\cup\Gamma^{\mathrm{int}}_{\mathrm{ac},>}$;
- Choose subsets $\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ac}},\widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ac}},\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}},\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}},\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}},\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ac}},\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}\subseteq\Gamma^{\mathrm{int}}_{\mathrm{ac}}\cup\Gamma^{\mathrm{int}}_{\mathrm{ac},>}$; To compute the frequency of adjacency-configurations exactly, set $\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ac}}:=\widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ac}}:=\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}:=\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}:=\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}:=\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ac}}:=\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}:=\Gamma^{\mathrm{int}}_{\mathrm{ac}}\cup\Gamma^{\mathrm{int}}_{\mathrm{ac},>}$;
- $\mathrm{ac}^{\mathrm{int}}_{\mathrm{LB}}(\nu),\mathrm{ac}^{\mathrm{int}}_{\mathrm{UB}}(\nu)\in[0,2\mathrm{n}^{\mathrm{int}}_{\mathrm{UB}}],\nu=(\mathtt{a},\mathtt{b},m)\in\Gamma^{\mathrm{int}}_{\mathrm{ac}}$: 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:

- $\mathrm{ac}^{\mathrm{int}}([\nu]^{\mathrm{int}})\in[\mathrm{ac}^{\mathrm{int}}_{\mathrm{LB}}(\nu),\mathrm{ac}^{\mathrm{int}}_{\mathrm{UB}}(\nu)],\nu\in\Gamma^{\mathrm{int}}_{\mathrm{ac}}$: the number of interior-edges with adjacency-configuration ν;
- $\mathrm{ac}_{\mathrm{C}}([\nu]^{\mathrm{int}})\in[0,m_{\mathrm{C}}],\nu\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ac}}$, $\mathrm{ac}_{\mathrm{T}}([\nu]^{\mathrm{int}})\in[0,t_{\mathrm{T}}],\nu\in\widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ac}}$, $\mathrm{ac}_{\mathrm{F}}([\nu]^{\mathrm{int}})\in[0,t_{\mathrm{F}}],\nu\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}$: 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 adjacency-configuration ν;
- $\mathrm{ac}_{\mathrm{CT}}([\nu]^{\mathrm{int}})\in[0,\min\{k_{\mathrm{C}},t_{\mathrm{T}}\}],\nu\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}$, $\mathrm{ac}_{\mathrm{TC}}([\nu]^{\mathrm{int}})\in[0,\min\{k_{\mathrm{C}},t_{\mathrm{T}}\}],\nu\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}$, $\mathrm{ac}_{\mathrm{CF}}([\nu]^{\mathrm{int}})\in[0,\widetilde{t_{\mathrm{C}}}],\nu\in\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ac}}$, $\mathrm{ac}_{\mathrm{TF}}([\nu]^{\mathrm{int}})\in[0,t_{\mathrm{T}}],\nu\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}$: 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 adjacency-configuration ν;
- $\delta^{\mathrm{C}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}})\in[0,1],i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]=I_{(\geq1)}\cup I_{(0/1)}\cup I_{(=1)},\nu\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ac}}$, $\delta^{\mathrm{T}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}})\in[0,1],i\in[2,t_{\mathrm{T}}],\nu\in\widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ac}}$, $\delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}})\in[0,1],i\in[2,t_{\mathrm{F}}],\nu\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}$: $\delta^{\mathrm{X}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}})=1\Leftrightarrow$ edge ${e^{\mathrm{X}}}_i$ 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 G (resp., ${e^{\rm X}}_i\notin E(G)$);

- $\Delta^{\rm CT+}_{\rm ac}(k), \Delta^{\rm CT-}_{\rm ac}(k)\in[0,|\Lambda^{\rm int}|], k\in[1,k_{\rm C}]=I_{(\geq 2)}\cup I_{(\geq 1)}$: $\Delta^{\rm CT+}_{\rm ac}(k)=\Delta^{\rm CT-}_{\rm ac}(k)=0$ (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 G (resp., otherwise);

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

- $\Delta^{\rm CF+}_{\rm ac}(c)\in[0,|\Lambda^{\rm int}|], \Delta^{\rm CF-}_{\rm ac}(c)\in[0,|\Lambda^{\rm int}|], c\in[1,\widetilde{t_{\rm C}}]$: $\Delta^{\rm CF+}_{\rm ac}(c)=\Delta^{\rm CF-}_{\rm ac}(c)=0$ (resp., $\Delta^{\rm CF+}_{\rm ac}(c)=\alpha(u)$ and $\Delta^{\rm CF-}_{\rm ac}(c)=\alpha(v)$) $\Leftrightarrow$ edge ${e^{\rm CF}}_{c,i}=(u,v)\in E_{\rm CF}$ for some $i\in[1,t_{\rm F}]$ is used in G (resp., otherwise);

- $\Delta^{\rm TF+}_{\rm ac}(i)\in[0,|\Lambda^{\rm int}|], \Delta^{\rm TF-}_{\rm ac}(i)\in[0,|\Lambda^{\rm int}|], i\in[1,t_{\rm T}]$: Analogous with $\Delta^{\rm CF+}_{\rm ac}(c)$ and $\Delta^{\rm CF-}_{\rm ac}(c)$;

constraints:

$$
\begin{aligned}
{\rm ac}_{\rm C}([\nu]^{\rm int})&=0, && \nu\in\Gamma^{\rm int}_{\rm ac}\setminus\widetilde{\Gamma}^{\rm C}_{\rm ac},\\
{\rm ac}_{\rm T}([\nu]^{\rm int})&=0, && \nu\in\Gamma^{\rm int}_{\rm ac}\setminus\widetilde{\Gamma}^{\rm T}_{\rm ac},\\
{\rm ac}_{\rm F}([\nu]^{\rm int})&=0, && \nu\in\Gamma^{\rm int}_{\rm ac}\setminus\widetilde{\Gamma}^{\rm F}_{\rm ac},\\
{\rm ac}_{\rm CT}([\nu]^{\rm int})&=0, && \nu\in\Gamma^{\rm int}_{\rm ac}\setminus\widetilde{\Gamma}^{\rm CT}_{\rm ac},\\
{\rm ac}_{\rm TC}([\nu]^{\rm int})&=0, && \nu\in\Gamma^{\rm int}_{\rm ac}\setminus\widetilde{\Gamma}^{\rm TC}_{\rm ac},\\
{\rm ac}_{\rm CF}([\nu]^{\rm int})&=0, && \nu\in\Gamma^{\rm int}_{\rm ac}\setminus\widetilde{\Gamma}^{\rm CF}_{\rm ac},\\
{\rm ac}_{\rm TF}([\nu]^{\rm int})&=0, && \nu\in\Gamma^{\rm int}_{\rm ac}\setminus\widetilde{\Gamma}^{\rm TF}_{\rm ac},
\end{aligned}
\tag{62}
$$

$$\begin{aligned}
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{C}}([\nu]^{\mathrm{int}}) &= \sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]} \delta_{\beta}^{\mathrm{C}}(i,m), && m\in[1,3],\\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{T}}([\nu]^{\mathrm{int}}) &= \sum_{i\in[2,t_{\mathrm{T}}]} \delta_{\beta}^{\mathrm{T}}(i,m), && m\in[1,3],\\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{F}}([\nu]^{\mathrm{int}}) &= \sum_{i\in[2,t_{\mathrm{F}}]} \delta_{\beta}^{\mathrm{F}}(i,m), && m\in[1,3],\\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{CT}}([\nu]^{\mathrm{int}}) &= \sum_{k\in[1,k_{\mathrm{C}}]} \delta_{\beta}^{+}(k,m), && m\in[1,3],\\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{TC}}([\nu]^{\mathrm{int}}) &= \sum_{k\in[1,k_{\mathrm{C}}]} \delta_{\beta}^{-}(k,m), && m\in[1,3],\\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{CF}}([\nu]^{\mathrm{int}}) &= \sum_{c\in[1,\widetilde{t_{\mathrm{C}}}]} \delta_{\beta}^{\mathrm{in}}(c,m), && m\in[1,3],\\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{TF}}([\nu]^{\mathrm{int}}) &= \sum_{c\in[\widetilde{t_{\mathrm{C}}}+1,c_{\mathrm{F}}]} \delta_{\beta}^{\mathrm{in}}(c,m), && m\in[1,3],
\end{aligned} \tag{63}$$

$$\begin{aligned}
\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{C}}} m\cdot\delta_{\mathrm{ac}}^{\mathrm{C}}(i,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{C}}(i), &&\\
\Delta_{\mathrm{ac}}^{\mathrm{C}+}(i) + \sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{C}}} [\mathsf{a}]^{\mathrm{int}}\delta_{\mathrm{ac}}^{\mathrm{C}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{tail}(i)), &&\\
\Delta_{\mathrm{ac}}^{\mathrm{C}-}(i) + \sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{C}}} [\mathsf{b}]^{\mathrm{int}}\delta_{\mathrm{ac}}^{\mathrm{C}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{head}(i)), &&\\
\Delta_{\mathrm{ac}}^{\mathrm{C}+}(i) + \Delta_{\mathrm{ac}}^{\mathrm{C}-}(i) &\le 2|\Lambda^{\mathrm{int}}|(1-e^{\mathrm{C}}(i)), && i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}],\\
\sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]} \delta_{\mathrm{ac}}^{\mathrm{C}}(i,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{C}}([\nu]^{\mathrm{int}}), && \nu\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{C}},
\end{aligned} \tag{64}$$

$$\begin{aligned}
\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{T}}} m\cdot\delta_{\mathrm{ac}}^{\mathrm{T}}(i,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{T}}(i), &&\\
\Delta_{\mathrm{ac}}^{\mathrm{T}+}(i) + \sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{T}}} [\mathsf{a}]^{\mathrm{int}}\delta_{\mathrm{ac}}^{\mathrm{T}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{T}}(i-1), &&\\
\Delta_{\mathrm{ac}}^{\mathrm{T}-}(i) + \sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{T}}} [\mathsf{b}]^{\mathrm{int}}\delta_{\mathrm{ac}}^{\mathrm{T}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{T}}(i), &&\\
\Delta_{\mathrm{ac}}^{\mathrm{T}+}(i) + \Delta_{\mathrm{ac}}^{\mathrm{T}-}(i) &\le 2|\Lambda^{\mathrm{int}}|(1-e^{\mathrm{T}}(i)), && i\in[2,t_{\mathrm{T}}],\\
\sum_{i\in[2,t_{\mathrm{T}}]} \delta_{\mathrm{ac}}^{\mathrm{T}}(i,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{T}}([\nu]^{\mathrm{int}}), && \nu\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{T}},
\end{aligned} \tag{65}$$

$$
\begin{aligned}
\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}} m\cdot\delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{F}}(i), \\
\Delta^{\mathrm{F}+}_{\mathrm{ac}}(i) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}} [\mathtt{a}]^{\mathrm{int}}\delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{F}}(i-1), \\
\Delta^{\mathrm{F}-}_{\mathrm{ac}}(i) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}} [\mathtt{b}]^{\mathrm{int}}\delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{F}}(i), \\
\Delta^{\mathrm{F}+}_{\mathrm{ac}}(i) + \Delta^{\mathrm{F}-}_{\mathrm{ac}}(i) &\le 2|\Lambda^{\mathrm{ex}}|(1-e^{\mathrm{F}}(i)), & i\in[2,t_{\mathrm{F}}], \\
\sum_{i\in[2,t_{\mathrm{F}}]} \delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{F}}([\nu]^{\mathrm{int}}), & \nu\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}},
\end{aligned} \tag{66}
$$

$$
\begin{aligned}
\alpha^{\mathrm{T}}(i) + |\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i)) &\ge \alpha^{\mathrm{CT}}(k), \\
\alpha^{\mathrm{CT}}(k) &\ge \alpha^{\mathrm{T}}(i) - |\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i)), & i\in[1,t_{\mathrm{T}}], \\
\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}} m\cdot\delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \beta^{+}(k), \\
\Delta^{\mathrm{CT}+}_{\mathrm{ac}}(k) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}} [\mathtt{a}]^{\mathrm{int}}\delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{tail}(k)), \\
\Delta^{\mathrm{CT}-}_{\mathrm{ac}}(k) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}} [\mathtt{b}]^{\mathrm{int}}\delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{CT}}(k), \\
\Delta^{\mathrm{CT}+}_{\mathrm{ac}}(k) + \Delta^{\mathrm{CT}-}_{\mathrm{ac}}(k) &\le 2|\Lambda^{\mathrm{int}}|(1-\delta^{\mathrm{T}}_{\chi}(k)), & k\in[1,k_{\mathrm{C}}], \\
\sum_{k\in[1,k_{\mathrm{C}}]} \delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{CT}}([\nu]^{\mathrm{int}}), & \nu\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}},
\end{aligned} \tag{67}
$$

$$
\begin{aligned}
\alpha^{\mathrm{T}}(i) + |\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i+1)) &\ge \alpha^{\mathrm{TC}}(k), \\
\alpha^{\mathrm{TC}}(k) &\ge \alpha^{\mathrm{T}}(i) - |\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i+1)), & i\in[1,t_{\mathrm{T}}], \\
\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}} m\cdot\delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \beta^{-}(k), \\
\Delta^{\mathrm{TC}+}_{\mathrm{ac}}(k) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}} [\mathtt{a}]^{\mathrm{int}}\delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{TC}}(k), \\
\Delta^{\mathrm{TC}-}_{\mathrm{ac}}(k) + \sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}} [\mathtt{b}]^{\mathrm{int}}\delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{head}(k)), \\
\Delta^{\mathrm{TC}+}_{\mathrm{ac}}(k) + \Delta^{\mathrm{TC}-}_{\mathrm{ac}}(k) &\le 2|\Lambda^{\mathrm{int}}|(1-\delta^{\mathrm{T}}_{\chi}(k)), & k\in[1,k_{\mathrm{C}}], \\
\sum_{k\in[1,k_{\mathrm{C}}]} \delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{TC}}([\nu]^{\mathrm{int}}), & \nu\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}},
\end{aligned} \tag{68}
$$

$$\begin{aligned}
\alpha^{\mathrm{F}}(i)+|\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{F}}(i,c)+e^{\mathrm{F}}(i)) &\geq \alpha^{\mathrm{CF}}(c), & \\
\alpha^{\mathrm{CF}}(c) &\geq \alpha^{\mathrm{F}}(i)-|\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{F}}(i,c)+e^{\mathrm{F}}(i)), & i\in[1,t_{\mathrm{F}}],\\
\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{CF}}} m\cdot\delta_{\mathrm{ac}}^{\mathrm{CF}}(c,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{in}}(c), & \\
\Delta_{\mathrm{ac}}^{\mathrm{CF}+}(c)+\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{CF}}}[\mathtt{a}]^{\mathrm{int}}\delta_{\mathrm{ac}}^{\mathrm{CF}}(c,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{head}(c)), & \\
\Delta_{\mathrm{ac}}^{\mathrm{CF}-}(c)+\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{CF}}}[\mathtt{b}]^{\mathrm{int}}\delta_{\mathrm{ac}}^{\mathrm{CF}}(c,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{CF}}(c), & \\
\Delta_{\mathrm{ac}}^{\mathrm{CF}+}(c)+\Delta_{\mathrm{ac}}^{\mathrm{CF}-}(c) &\leq 2\max\{|\Lambda^{\mathrm{int}}|,|\Lambda^{\mathrm{int}}|\}(1-\delta_{\chi}^{\mathrm{F}}(c)), & c\in[1,\widetilde{t_{\mathrm{C}}}],\\
\sum_{c\in[1,\widetilde{t_{\mathrm{C}}}]}\delta_{\mathrm{ac}}^{\mathrm{CF}}(c,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{CF}}([\nu]^{\mathrm{int}}), & \nu\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{CF}},
\end{aligned} \tag{69}$$

$$\begin{aligned}
\alpha^{\mathrm{F}}(j)+|\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{F}}(j,i+\widetilde{t_{\mathrm{C}}})+e^{\mathrm{F}}(j)) &\geq \alpha^{\mathrm{TF}}(i), & \\
\alpha^{\mathrm{TF}}(i) &\geq \alpha^{\mathrm{F}}(j)-|\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{F}}(j,i+\widetilde{t_{\mathrm{C}}})+e^{\mathrm{F}}(j)), & j\in[1,t_{\mathrm{F}}],\\
\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{TF}}} m\cdot\delta_{\mathrm{ac}}^{\mathrm{TF}}(i,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{in}}(i+\widetilde{t_{\mathrm{C}}}), & \\
\Delta_{\mathrm{ac}}^{\mathrm{TF}+}(i)+\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{TF}}}[\mathtt{a}]^{\mathrm{int}}\delta_{\mathrm{ac}}^{\mathrm{TF}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{T}}(i), & \\
\Delta_{\mathrm{ac}}^{\mathrm{TF}-}(i)+\sum_{\nu=(\mathtt{a},\mathtt{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{TF}}}[\mathtt{b}]^{\mathrm{int}}\delta_{\mathrm{ac}}^{\mathrm{TF}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{TF}}(i), & \\
\Delta_{\mathrm{ac}}^{\mathrm{TF}+}(i)+\Delta_{\mathrm{ac}}^{\mathrm{TF}-}(i) &\leq 2\max\{|\Lambda^{\mathrm{int}}|,|\Lambda^{\mathrm{int}}|\}(1-\delta_{\chi}^{\mathrm{F}}(i+\widetilde{t_{\mathrm{C}}})), & i\in[1,t_{\mathrm{T}}],\\
\sum_{i\in[1,t_{\mathrm{T}}]}\delta_{\mathrm{ac}}^{\mathrm{TF}}(i,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{TF}}([\nu]^{\mathrm{int}}), & \nu\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{TF}},
\end{aligned} \tag{70}$$

$$\begin{aligned}
\sum_{\mathrm{X}\in\{\mathrm{C,T,F,CT,TC,CF,TF}\}}(\mathrm{ac}_{\mathrm{X}}([\nu]^{\mathrm{int}})+\mathrm{ac}_{\mathrm{X}}([\overline{\nu}]^{\mathrm{int}})) &= \mathrm{ac}^{\mathrm{int}}([\nu]^{\mathrm{int}}), & \nu\in\Gamma_{\mathrm{ac},<}^{\mathrm{int}},\\
\sum_{\mathrm{X}\in\{\mathrm{C,T,F,CT,TC,CF,TF}\}}\mathrm{ac}_{\mathrm{X}}([\nu]^{\mathrm{int}}) &= \mathrm{ac}^{\mathrm{int}}([\nu]^{\mathrm{int}}), & \nu\in\Gamma_{\mathrm{ac},=}^{\mathrm{int}}.
\end{aligned} \tag{71}$$

3.9 Descriptor for the Number of Chemical Symbols

We include constraints for computing the frequency of each chemical symbol in Λ_{dg}. Let $\mathrm{cs}(v)$ denote the chemical symbol of a vertex v in a chemical graph G to be inferred; i.e., $\mathrm{cs}(v)=\mu=\mathtt{a}d\in\Lambda_{\mathrm{dg}}$ such that $\alpha(v)=\mathtt{a}$ and $\deg_G(v)=d$.

constants:

- A set $\Lambda_{\mathrm{dg}}^{\mathrm{int}}$ of chemical symbols;
- Prepare a coding of each of the two sets $\Lambda_{\mathrm{dg}}^{\mathrm{int}}$ and let $[\mu]^{\mathrm{int}}$ denote the coded integer of an element $\mu\in\Lambda_{\mathrm{dg}}^{\mathrm{int}}$;

- 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:

$$\sum_{\mu \in \widetilde{\Lambda}^{\rm X}_{\rm dg} \cup \{\epsilon\}} \delta^{\rm X}_{\rm ns}(i, [\mu]^{\rm int}) = 1, \qquad \sum_{\mu = \mathtt{a}d \in \widetilde{\Lambda}^{\rm X}_{\rm dg}} [\mathtt{a}]^{\rm int} \cdot \delta^{\rm X}_{\rm ns}(i, [\mu]^{\rm int}) = \alpha^{\rm X}(i),$$
$$\sum_{\mu = \mathtt{a}d \in \widetilde{\Lambda}^{\rm X}_{\rm dg}} d \cdot \delta^{\rm X}_{\rm ns}(i, [\mu]^{\rm int}) = \mathrm{deg}^{\rm X}(i),$$
$$i \in [1, t_{\rm X}], \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}, \tag{72}$$

$$\sum_{i \in [1, t_{\rm C}]} \delta^{\rm C}_{\rm ns}(i, [\mu]^{\rm int}) + \sum_{i \in [1, t_{\rm T}]} \delta^{\rm T}_{\rm ns}(i, [\mu]^{\rm int}) + \sum_{i \in [1, t_{\rm F}]} \delta^{\rm F}_{\rm ns}(i, [\mu]^{\rm int}) = \mathrm{ns}^{\rm int}([\mu]^{\rm int}), \qquad \mu \in \Lambda^{\rm int}_{\rm dg}. \tag{73}$$

3.10 Descriptor for the Number of Edge-configurations

We include constraints to compute the frequency of each edge-configuration in an inferred chemical graph G.

constants:

- A set $\Gamma^{\rm int}$ of edge-configurations $\gamma = (\mu, \xi, m)$ with $\mu \le \xi$;
- Let $\Gamma^{\rm int}_{<} = \{(\mu, \xi, m) \in \Gamma^{\rm int} \mid \mu < \xi\}$, $\Gamma^{\rm int}_{=} = \{(\mu, \xi, m) \in \Gamma^{\rm int} \mid \mu = \xi\}$ and $\Gamma^{\rm int}_{>} = \{(\xi, \mu, m) \mid (\mu, \xi, m) \in \Gamma^{\rm int}_{<}\}$;
- 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 $\widetilde{\Gamma}^{\rm C}_{\rm ec} := \widetilde{\Gamma}^{\rm T}_{\rm ec} := \widetilde{\Gamma}^{\rm CT}_{\rm ec} := \widetilde{\Gamma}^{\rm TC}_{\rm ec} := \widetilde{\Gamma}^{\rm F}_{\rm ec} := \widetilde{\Gamma}^{\rm CF}_{\rm ec} := \widetilde{\Gamma}^{\rm TF}_{\rm ec} := \Gamma^{\rm int} \cup \Gamma^{\rm int}_{>}$;
- $\mathrm{ec}^{\rm int}_{\rm LB}(\gamma), \mathrm{ec}^{\rm int}_{\rm UB}(\gamma) \in [0, 2\mathrm{n}^{\rm int}_{\rm UB}], \gamma = (\mu, \xi, m) \in \Gamma^{\rm int}$: lower and upper bounds on the number of interior-edges $e = uv$ with $\mathrm{cs}(u) = \mu$, $\mathrm{cs}(v) = \xi$ and $\beta(e) = m$;

variables:

- $\mathrm{ec}^{\rm int}([\gamma]^{\rm int}) \in [\mathrm{ec}^{\rm int}_{\rm LB}(\gamma), \mathrm{ec}^{\rm int}_{\rm UB}(\gamma)], \gamma \in \Gamma^{\rm int}$: the number of interior-edges with edge-configuration γ;
- $\mathrm{ec}_{\rm C}([\gamma]^{\rm int}) \in [0, m_{\rm C}], \gamma \in \widetilde{\Gamma}^{\rm C}_{\rm ec}$, $\mathrm{ec}_{\rm T}([\gamma]^{\rm int}) \in [0, t_{\rm T}], \gamma \in \widetilde{\Gamma}^{\rm T}_{\rm ec}$, $\mathrm{ec}_{\rm F}([\gamma]^{\rm int}) \in [0, t_{\rm F}], \gamma \in \widetilde{\Gamma}^{\rm F}_{\rm ec}$: 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 edge-configuration γ;
- $\mathrm{ec}_{\rm CT}([\gamma]^{\rm int}) \in [0, \min\{k_{\rm C}, t_{\rm T}\}], \gamma \in \widetilde{\Gamma}^{\rm CT}_{\rm ec}$, $\mathrm{ec}_{\rm TC}([\gamma]^{\rm int}) \in [0, \min\{k_{\rm C}, t_{\rm T}\}], \gamma \in \widetilde{\Gamma}^{\rm CT}_{\rm ec}$, $\mathrm{ec}_{\rm CF}([\gamma]^{\rm int}) \in [0, \widetilde{t_{\rm C}}], \gamma \in \widetilde{\Gamma}^{\rm CF}_{\rm ec}$, $\mathrm{ec}_{\rm TF}([\gamma]^{\rm int}) \in [0, t_{\rm T}], \gamma \in \widetilde{\Gamma}^{\rm TF}_{\rm ec}$: 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 edge-configuration γ;

- $\delta_{\rm ec}^{\rm C}(i,[\gamma]^{\rm int}) \in [0,1], i \in [\widetilde{k_{\rm C}}+1, m_{\rm C}] = I_{(\geq 1)} \cup I_{(0/1)} \cup I_{(=1)}, \gamma \in \widetilde{\Gamma}_{\rm ec}^{\rm C}$, $\delta_{\rm ec}^{\rm T}(i,[\gamma]^{\rm int}) \in [0,1], i \in [2,t_{\rm T}], \gamma \in \widetilde{\Gamma}_{\rm ec}^{\rm T}$, $\delta_{\rm ec}^{\rm F}(i,[\gamma]^{\rm int}) \in [0,1], i \in [2,t_{\rm F}], \gamma \in \widetilde{\Gamma}_{\rm ec}^{\rm F}$: $\delta_{\rm ec}^{\rm X}(i,[\gamma]^{\rm t}) = 1 \Leftrightarrow$ edge ${e^{\rm X}}_i$ has edge-configuration γ;
- $\delta_{\rm ec,C}^{\rm CT}(k,[\gamma]^{\rm int}), \delta_{\rm ec,C}^{\rm TC}(k,[\gamma]^{\rm int}) \in [0,1], k \in [1,k_{\rm C}] = I_{(\geq 2)} \cup I_{(\geq 1)}, \gamma \in \widetilde{\Gamma}_{\rm ec}^{\rm CT}$: $\delta_{\rm ec,C}^{\rm CT}(k,[\gamma]^{\rm int}) = 1$ (resp., $\delta_{\rm ec,C}^{\rm TC}(k,[\gamma]^{\rm int}) = 1$) $\Leftrightarrow$ 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 edge-configuration γ;
- $\delta_{\rm ec,C}^{\rm CF}(c,[\gamma]^{\rm int}) \in [0,1], c \in [1,\widetilde{t_{\rm C}}], \gamma \in \widetilde{\Gamma}_{\rm ec}^{\rm CF}$: $\delta_{\rm ec,C}^{\rm CF}(c,[\gamma]^{\rm int}) = 1 \Leftrightarrow$ edge ${e^{\rm CF}}_{c,i}$ for some $i \in [1,t_{\rm F}]$ has edge-configuration γ;
- $\delta_{\rm ec,T}^{\rm TF}(i,[\gamma]^{\rm int}) \in [0,1], i \in [1,t_{\rm T}], \gamma \in \widetilde{\Gamma}_{\rm ec}^{\rm TF}$: $\delta_{\rm ec,T}^{\rm TF}(i,[\gamma]^{\rm int}) = 1 \Leftrightarrow$ edge ${e^{\rm TF}}_{i,j}$ for some $j \in [1,t_{\rm F}]$ has edge-configuration γ;
- $\deg_{\rm T}^{\rm CT}(k), \deg_{\rm T}^{\rm TC}(k) \in [0,4], k \in [1,k_{\rm C}]$: $\deg_G(v)$ of an end-vertex $v \in V_{\rm T}$ 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;
- $\deg_{\rm F}^{\rm CF}(c) \in [0,4], c \in [1,\widetilde{t_{\rm C}}]$: $\deg_G(v)$ of an end-vertex $v \in V_{\rm F}$ of the edge $({v^{\rm C}}_c, v) \in E_{\rm CF}$ if any;
- $\deg_{\rm F}^{\rm TF}(i) \in [0,4], i \in [1,t_{\rm T}]$: $\deg_G(v)$ of an end-vertex $v \in V_{\rm F}$ of the edge $({v^{\rm T}}_i, v) \in E_{\rm TF}$ if any;
- $\Delta_{\rm ec}^{\rm C+}(i), \Delta_{\rm ec}^{\rm C-}(i), \in [0,4], i \in [\widetilde{k_{\rm C}}+1, m_{\rm C}]$, $\Delta_{\rm ec}^{\rm T+}(i), \Delta_{\rm ec}^{\rm T-}(i) \in [0,4], i \in [2,t_{\rm T}]$, $\Delta_{\rm ec}^{\rm F+}(i), \Delta_{\rm ec}^{\rm F-}(i) \in [0,4], i \in [2,t_{\rm F}]$: $\Delta_{\rm ec}^{\rm X+}(i) = \Delta_{\rm ec}^{\rm X-}(i) = 0$ (resp., $\Delta_{\rm ec}^{\rm X+}(i) = \deg_G(u)$ and $\Delta_{\rm ec}^{\rm X-}(i) = \deg_G(v)$) $\Leftrightarrow$ edge ${e^{\rm X}}_i = (u,v) \in E_{\rm X}$ is used in G (resp., ${e^{\rm X}}_i \notin E(G)$);
- $\Delta_{\rm ec}^{\rm CT+}(k), \Delta_{\rm ec}^{\rm CT-}(k) \in [0,4], k \in [1,k_{\rm C}] = I_{(\geq 2)} \cup I_{(\geq 1)}$: $\Delta_{\rm ec}^{\rm CT+}(k) = \Delta_{\rm ec}^{\rm CT-}(k) = 0$ (resp., $\Delta_{\rm ec}^{\rm CT+}(k) = \deg_G(u)$ and $\Delta_{\rm ec}^{\rm CT-}(k) = \deg_G(v)$) $\Leftrightarrow$ 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 G (resp., otherwise);
- $\Delta_{\rm ec}^{\rm TC+}(k), \Delta_{\rm ec}^{\rm TC-}(k) \in [0,4], k \in [1,k_{\rm C}] = I_{(\geq 2)} \cup I_{(\geq 1)}$: Analogous with $\Delta_{\rm ec}^{\rm CT+}(k)$ and $\Delta_{\rm ec}^{\rm CT-}(k)$;
- $\Delta_{\rm ac}^{\rm CF+}(c), \Delta_{\rm ec}^{\rm CF-}(c) \in [0,4], c \in [1,\widetilde{t_{\rm C}}]$: $\Delta_{\rm ec}^{\rm CF+}(c) = \Delta_{\rm ec}^{\rm CF-}(c) = 0$ (resp., $\Delta_{\rm ec}^{\rm CF+}(c) = \deg_G(u)$ and $\Delta_{\rm ec}^{\rm CF-}(c) = \deg_G(v)$) $\Leftrightarrow$ edge ${e^{\rm CF}}_{c,j} = (u,v) \in E_{\rm CF}$ for some $j \in [1,t_{\rm F}]$ is used in G (resp., otherwise);
- $\Delta_{\rm ec}^{\rm TF+}(i), \Delta_{\rm ec}^{\rm TF-}(i) \in [0,4], i \in [1,t_{\rm T}]$: Analogous with $\Delta_{\rm ec}^{\rm CF+}(c)$ and $\Delta_{\rm ec}^{\rm CF-}(c)$;

constraints:

$$\begin{aligned}
{\rm ec}_{\rm C}([\gamma]^{\rm int}) &= 0, & \gamma &\in \Gamma^{\rm int} \setminus \widetilde{\Gamma}_{\rm ec}^{\rm C}, \\
{\rm ec}_{\rm T}([\gamma]^{\rm int}) &= 0, & \gamma &\in \Gamma^{\rm int} \setminus \widetilde{\Gamma}_{\rm ec}^{\rm T}, \\
{\rm ec}_{\rm F}([\gamma]^{\rm int}) &= 0, & \gamma &\in \Gamma^{\rm int} \setminus \widetilde{\Gamma}_{\rm ec}^{\rm F}, \\
{\rm ec}_{\rm CT}([\gamma]^{\rm int}) &= 0, & \gamma &\in \Gamma^{\rm int} \setminus \widetilde{\Gamma}_{\rm ec}^{\rm CT}, \\
{\rm ec}_{\rm TC}([\gamma]^{\rm int}) &= 0, & \gamma &\in \Gamma^{\rm int} \setminus \widetilde{\Gamma}_{\rm ec}^{\rm TC}, \\
{\rm ec}_{\rm CF}([\gamma]^{\rm int}) &= 0, & \gamma &\in \Gamma^{\rm int} \setminus \widetilde{\Gamma}_{\rm ec}^{\rm CF}, \\
{\rm ec}_{\rm TF}([\gamma]^{\rm int}) &= 0, & \gamma &\in \Gamma^{\rm int} \setminus \widetilde{\Gamma}_{\rm ec}^{\rm TF},
\end{aligned} \tag{74}$$

$$
\begin{aligned}
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{C}}([\gamma]^{\mathrm{int}}) &= \sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]} \delta_{\beta}^{\mathrm{C}}(i,m), && m\in[1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{T}}([\gamma]^{\mathrm{int}}) &= \sum_{i\in[2,t_{\mathrm{T}}]} \delta_{\beta}^{\mathrm{T}}(i,m), && m\in[1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{F}}([\gamma]^{\mathrm{int}}) &= \sum_{i\in[2,t_{\mathrm{F}}]} \delta_{\beta}^{\mathrm{F}}(i,m), && m\in[1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{CT}}([\gamma]^{\mathrm{int}}) &= \sum_{k\in[1,k_{\mathrm{C}}]} \delta_{\beta}^{+}(k,m), && m\in[1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{TC}}([\gamma]^{\mathrm{int}}) &= \sum_{k\in[1,k_{\mathrm{C}}]} \delta_{\beta}^{-}(k,m), && m\in[1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{CF}}([\gamma]^{\mathrm{int}}) &= \sum_{c\in[1,\widetilde{t_{\mathrm{C}}}]} \delta_{\beta}^{\mathrm{in}}(c,m), && m\in[1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{TF}}([\gamma]^{\mathrm{int}}) &= \sum_{c\in[\widetilde{t_{\mathrm{C}}}+1,c_{\mathrm{F}}]} \delta_{\beta}^{\mathrm{in}}(c,m), && m\in[1,3],
\end{aligned}
\tag{75}
$$

$$
\begin{aligned}
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}_{\mathrm{ec}}^{\mathrm{C}}} [(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}}\cdot\delta_{\mathrm{ec}}^{\mathrm{C}}(i,[\gamma]^{\mathrm{int}}) &= \sum_{\nu\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{C}}} [\nu]^{\mathrm{int}}\cdot\delta_{\mathrm{ac}}^{\mathrm{C}}(i,[\nu]^{\mathrm{int}}),\\
\Delta_{\mathrm{ec}}^{\mathrm{C}+}(i) + \sum_{\gamma=(\mathtt{a}d,\xi,m)\in\widetilde{\Gamma}_{\mathrm{ec}}^{\mathrm{C}}} d\cdot\delta_{\mathrm{ec}}^{\mathrm{C}}(i,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{C}}(\mathrm{tail}(i)),\\
\Delta_{\mathrm{ec}}^{\mathrm{C}-}(i) + \sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}_{\mathrm{ec}}^{\mathrm{C}}} d\cdot\delta_{\mathrm{ec}}^{\mathrm{C}}(i,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{C}}(\mathrm{head}(i)),\\
\Delta_{\mathrm{ec}}^{\mathrm{C}+}(i) + \Delta_{\mathrm{ec}}^{\mathrm{C}-}(i) &\le 8(1-e^{\mathrm{C}}(i)), && i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}],\\
\sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]} \delta_{\mathrm{ec}}^{\mathrm{C}}(i,[\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{C}}([\gamma]^{\mathrm{int}}), && \gamma\in\widetilde{\Gamma}_{\mathrm{ec}}^{\mathrm{C}},
\end{aligned}
\tag{76}
$$

$$
\begin{aligned}
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}_{\mathrm{ec}}^{\mathrm{T}}} [(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}}\cdot\delta_{\mathrm{ec}}^{\mathrm{T}}(i,[\gamma]^{\mathrm{int}}) &= \sum_{\nu\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{T}}} [\nu]^{\mathrm{int}}\cdot\delta_{\mathrm{ac}}^{\mathrm{T}}(i,[\nu]^{\mathrm{int}}),\\
\Delta_{\mathrm{ec}}^{\mathrm{T}+}(i) + \sum_{\gamma=(\mathtt{a}d,\xi,m)\in\widetilde{\Gamma}_{\mathrm{ec}}^{\mathrm{T}}} d\cdot\delta_{\mathrm{ec}}^{\mathrm{T}}(i,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{T}}(i-1),\\
\Delta_{\mathrm{ec}}^{\mathrm{T}-}(i) + \sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}_{\mathrm{ec}}^{\mathrm{T}}} d\cdot\delta_{\mathrm{ec}}^{\mathrm{T}}(i,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{T}}(i),\\
\Delta_{\mathrm{ec}}^{\mathrm{T}+}(i) + \Delta_{\mathrm{ec}}^{\mathrm{T}-}(i) &\le 8(1-e^{\mathrm{T}}(i)), && i\in[2,t_{\mathrm{T}}],\\
\sum_{i\in[2,t_{\mathrm{T}}]} \delta_{\mathrm{ec}}^{\mathrm{T}}(i,[\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{T}}([\gamma]^{\mathrm{int}}), && \gamma\in\widetilde{\Gamma}_{\mathrm{ec}}^{\mathrm{T}},
\end{aligned}
\tag{77}
$$

$$\begin{aligned}
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ec}}}[(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}}\cdot\delta^{\mathrm{F}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}}) &= \sum_{\nu\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}}[\nu]^{\mathrm{int}}\cdot\delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}), \\
\Delta^{\mathrm{F}+}_{\mathrm{ec}}(i)+\sum_{\gamma=(\mathtt{a}d,\xi,m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ec}}}d\cdot\delta^{\mathrm{F}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{F}}(i-1), \\
\Delta^{\mathrm{F}-}_{\mathrm{ec}}(i)+\sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ec}}}d\cdot\delta^{\mathrm{F}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{F}}(i), \\
\Delta^{\mathrm{F}+}_{\mathrm{ec}}(i)+\Delta^{\mathrm{F}-}_{\mathrm{ec}}(i) &\le 8(1-e^{\mathrm{F}}(i)), && i\in[2,t_{\mathrm{F}}], \\
\sum_{i\in[2,t_{\mathrm{F}}]}\delta^{\mathrm{F}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{F}}([\gamma]^{\mathrm{int}}), && \gamma\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ec}}, \qquad (78)
\end{aligned}$$

$$\begin{aligned}
\deg^{\mathrm{T}}(i)+4(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i)) &\ge \deg^{\mathrm{CT}}_{\mathrm{T}}(k), \\
\deg^{\mathrm{CT}}_{\mathrm{T}}(k) &\ge \deg^{\mathrm{T}}(i)-4(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i)), && i\in[1,t_{\mathrm{T}}], \\
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ec}}}[(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}}\cdot\delta^{\mathrm{CT}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}}) &= \sum_{\nu\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}}[\nu]^{\mathrm{int}}\cdot\delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}), \\
\Delta^{\mathrm{CT}+}_{\mathrm{ec}}(k)+\sum_{\gamma=(\mathtt{a}d,\xi,m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ec}}}d\cdot\delta^{\mathrm{CT}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{C}}(\mathrm{tail}(k)), \\
\Delta^{\mathrm{CT}-}_{\mathrm{ec}}(k)+\sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ec}}}d\cdot\delta^{\mathrm{CT}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{CT}}_{\mathrm{T}}(k), \\
\Delta^{\mathrm{CT}+}_{\mathrm{ec}}(k)+\Delta^{\mathrm{CT}-}_{\mathrm{ec}}(k) &\le 8(1-\delta^{\mathrm{T}}_{\chi}(k)), && k\in[1,k_{\mathrm{C}}], \\
\sum_{k\in[1,k_{\mathrm{C}}]}\delta^{\mathrm{CT}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{CT}}([\gamma]^{\mathrm{int}}), && \gamma\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ec}}, \qquad (79)
\end{aligned}$$

$$\begin{aligned}
\deg^{\mathrm{T}}(i)+4(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i+1)) &\ge \deg^{\mathrm{TC}}_{\mathrm{T}}(k), \\
\deg^{\mathrm{TC}}_{\mathrm{T}}(k) &\ge \deg^{\mathrm{T}}(i)-4(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i+1)), && i\in[1,t_{\mathrm{T}}], \\
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}}}[(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}}\cdot\delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}}) &= \sum_{\nu\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}}[\nu]^{\mathrm{int}}\cdot\delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}), \\
\Delta^{\mathrm{TC}+}_{\mathrm{ec}}(k)+\sum_{\gamma=(\mathtt{a}d,\xi,m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}}}d\cdot\delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{TC}}_{\mathrm{T}}(k), \\
\Delta^{\mathrm{TC}-}_{\mathrm{ec}}(k)+\sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}}}d\cdot\delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{C}}(\mathrm{head}(k)), \\
\Delta^{\mathrm{TC}+}_{\mathrm{ec}}(k)+\Delta^{\mathrm{TC}-}_{\mathrm{ec}}(k) &\le 8(1-\delta^{\mathrm{T}}_{\chi}(k)), && k\in[1,k_{\mathrm{C}}], \\
\sum_{k\in[1,k_{\mathrm{C}}]}\delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{TC}}([\gamma]^{\mathrm{int}}), && \gamma\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}}, \qquad (80)
\end{aligned}$$

$$\begin{aligned}
\deg^{\mathrm{F}}(i) + 4(1 - \chi^{\mathrm{F}}(i,c) + e^{\mathrm{F}}(i)) &\geq \deg^{\mathrm{CF}}_{\mathrm{F}}(c), \\
\deg^{\mathrm{CF}}_{\mathrm{F}}(c) &\geq \deg^{\mathrm{F}}(i) - 4(1 - \chi^{\mathrm{F}}(i,c) + e^{\mathrm{F}}(i)), && i \in [1, t_{\mathrm{F}}], \\
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}}} [(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}} \cdot \delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c,[\gamma]^{\mathrm{int}}) &= \sum_{\nu\in\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ac}}} [\nu]^{\mathrm{int}} \cdot \delta^{\mathrm{CF}}_{\mathrm{ac}}(c,[\nu]^{\mathrm{int}}), \\
\Delta^{\mathrm{CF}+}_{\mathrm{ec}}(c) + \sum_{\gamma=(\mathtt{a}d,\xi,m)\in\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{C}}(c), \\
\Delta^{\mathrm{CF}-}_{\mathrm{ec}}(c) + \sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{CF}}_{\mathrm{F}}(c), \\
\Delta^{\mathrm{CF}+}_{\mathrm{ec}}(c) + \Delta^{\mathrm{CF}-}_{\mathrm{ec}}(c) &\leq 8(1 - \delta^{\mathrm{F}}_{\chi}(c)), && c \in [1, \widetilde{t_{\mathrm{C}}}], \\
\sum_{c\in[1,\widetilde{t_{\mathrm{C}}}]} \delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c,[\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{CF}}([\gamma]^{\mathrm{int}}), && \gamma \in \widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}},
\end{aligned} \tag{81}$$

$$\begin{aligned}
\deg^{\mathrm{F}}(j) + 4(1 - \chi^{\mathrm{F}}(j,i+\widetilde{t_{\mathrm{C}}}) + e^{\mathrm{F}}(j)) &\geq \deg^{\mathrm{TF}}_{\mathrm{F}}(i), \\
\deg^{\mathrm{TF}}_{\mathrm{F}}(i) &\geq \deg^{\mathrm{F}}(j) - 4(1 - \chi^{\mathrm{F}}(j,i+\widetilde{t_{\mathrm{C}}}) + e^{\mathrm{F}}(j)), && j \in [1, t_{\mathrm{F}}], \\
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}}} [(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}} \cdot \delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i,[\gamma]^{\mathrm{int}}) &= \sum_{\nu\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}} [\nu]^{\mathrm{int}} \cdot \delta^{\mathrm{TF}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}), \\
\Delta^{\mathrm{TF}+}_{\mathrm{ec}}(i) + \sum_{\gamma=(\mathtt{a}d,\xi,m)\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{T}}(i), \\
\Delta^{\mathrm{TF}-}_{\mathrm{ec}}(i) + \sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{TF}}_{\mathrm{F}}(i), \\
\Delta^{\mathrm{TF}+}_{\mathrm{ec}}(i) + \Delta^{\mathrm{TF}-}_{\mathrm{ec}}(i) &\leq 8(1 - \delta^{\mathrm{F}}_{\chi}(i+\widetilde{t_{\mathrm{C}}})), && i \in [1, t_{\mathrm{T}}], \\
\sum_{i\in[1,t_{\mathrm{T}}]} \delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i,[\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{TF}}([\gamma]^{\mathrm{int}}), && \gamma \in \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}},
\end{aligned} \tag{82}$$

$$\begin{aligned}
\sum_{\mathrm{X}\in\{\mathrm{C,T,F,CT,TC,CF,TF}\}} (\mathrm{ec}_{\mathrm{X}}([\gamma]^{\mathrm{int}}) + \mathrm{ec}_{\mathrm{X}}([\overline{\gamma}]^{\mathrm{int}})) &= \mathrm{ec}^{\mathrm{int}}([\gamma]^{\mathrm{int}}), && \gamma \in \Gamma^{\mathrm{int}}_{<}, \\
\sum_{\mathrm{X}\in\{\mathrm{C,T,F,CT,TC,CF,TF}\}} \mathrm{ec}_{\mathrm{X}}([\gamma]^{\mathrm{int}}) &= \mathrm{ec}^{\mathrm{int}}([\gamma]^{\mathrm{int}}), && \gamma \in \Gamma^{\mathrm{int}}_{=}.
\end{aligned} \tag{83}$$

3.11 Descriptor for the Number of of Fringe-configurations

We include constraints to compute the frequency of each fringe-configuration in an inferred chemical graph G.

variables:

$\mathrm{fc}([\psi]) \in [0, t_{\mathrm{C}} + t_{\mathrm{T}} + t_{\mathrm{F}}], \psi \in \mathcal{F}^*$: the frequency of a chemical rooted tree ψ in the set of ρ-fringe-trees in G;

constraints:

$$\sum_{i \in [1,t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{C,T,F}\}} \delta_{\mathrm{fr}}^{\mathrm{X}}(i, [\psi]) = \mathrm{fc}([\psi]), \qquad \psi \in \mathcal{F}^*. \tag{84}$$

3.12 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 $f(G) = (x_1, x_2, \ldots, x_K)$:

$$\frac{(1-\varepsilon)(x_i - \min(\mathrm{dcp}_i; D_\pi))}{\max(\mathrm{dcp}_i; D_\pi) - \min(\mathrm{dcp}_i; D_\pi)} \leq \widehat{x}_i \leq \frac{(1+\varepsilon)(x_i - \min(\mathrm{dcp}_i; D_\pi))}{\max(\mathrm{dcp}_i; D_\pi) - \min(\mathrm{dcp}_i; D_\pi)}, \ i \in [1, K]. \tag{85}$$

An example of a tolerance is $\varepsilon = 0.01$.

References

[1] R. Ito, N. A. Azam, C. Wang, A. Shurbevski, H. Nagamochi, T. Akutsu, A novel method for the inverse QSAR/QSPR to monocyclic chemical compounds based on artificial neural networks and integer programming, *BIOCOMP2020*, Las Vegas, Nevada, USA, 27-30 July 2020.

[2] F. Zhang, J. Zhu, R. Chiewvanichakorn, A. Shurbevski, H. Nagamochi, T. Akutsu, A new integer linear programming formulation to the inverse QSAR/QSPR for acyclic chemical compounds using skeleton trees, *The 33rd International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, September 22-25, 2020 Kitakyushu, Japan, Springer LNCS 12144, pp. 433–444.

[3] N. A. Azam, J. Zhu, Y. Sun, Y. Shi, A. Shurbevski, L. Zhao, H. Nagamochi, T. Akutsu, A novel method for inference of acyclic chemical compounds with bounded branch-height based on artificial neural networks and integer programming, arXiv:2009.09646

[4] T. Akutsu, H. Nagamochi, A novel method for inference of chemical compounds with prescribed topological substructures based on integer programming, arXiv: 2010.09203, 2020.