Adjustive Linear Regression and Its Application to the Inverse QSAR *

Jianshen Zhu[1], Kazuya Haraguchi[1], Hiroshi Nagamochi[1] and Tatsuya Akutsu[2]

[1] Department of Applied Mathematics and Physics, Kyoto University, Kyoto 606-8501, Japan
[2] Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji 611-0011, Japan

Abstract

In this paper, we propose a new machine learning method, called adjustive linear regression, which can be regarded as an ANN on an architecture with an input layer and an output layer of a single node, wherein an error function is minimized by choosing not only weights of the arcs but also an activation function at each node in the two layers simultaneously. Under some conditions, such a minimization can be formulated as a linear program (LP) and a prediction function with adjustive linear regression is obtained as an optimal solution to the LP. We apply the new machine learning method to a framework of inferring a chemical compound with a desired property (i.e., inverse QSAR). From the results of our computational experiments, we observe that a prediction function constructed by adjustive linear regression for some chemical properties drastically outperforms that by Lasso linear regression.

Keywords: Machine Learning, Linear Regression, Integer Programming, Linear Program, Cheminformatics, Materials Informatics, QSAR/QSPR, Molecular Design.

1 Introduction

In this paper, we design a new learning method, called "adjustive linear regression" in order to construct a function that predicts a chemical property of a given chemical compound. We start with the background and the recent results on the research.

Background Analysis of chemical compounds is one of the important applications of intelligent computing. Indeed, various machine learning methods have been applied to the prediction of chemical activities from their structural data, where such a problem is often referred to as *quantitative structure activity relationship* (QSAR) [1, 2]. Recently, neural networks and deep-learning technologies have extensively been applied to QSAR [3].

In addition to QSAR, extensive studies have been done on inverse quantitative structure activity relationship (inverse QSAR), which seeks for chemical structures having desired chemical activities under some constraints. Since it is difficult to directly handle chemical structures in both QSAR and inverse QSAR, chemical compounds are usually represented as vectors of real or integer numbers, which are often called *descriptors* in chemoinformatics and correspond to *feature vectors* in machine learning. One major approach in inverse QSAR is to infer feature vectors from given chemical activities and constraints and then reconstruct chemical structures from these feature

*Department of Applied Mathematics and Physics, Kyoto University, Technical Report, TR: 2021-002, September 3, 2021

vectors [4, 5, 6], where chemical structures are usually treated as undirected graphs. However, the reconstruction itself is a challenging task because the number of possible chemical graphs is huge. For example, chemical graphs with up to 30 atoms (vertices) `C`, `N`, `O`, and `S` may exceed 10^{60} [7]. Due to this difficulty, most existing methods for inverse QSAR do not guarantee optimal or exact solutions.

As a new approach, extensive studies have recently been done for inverse QSAR using *artificial neural networks* (ANNs), especially using graph convolutional networks [8]. For example, recurrent neural networks [10, 11], variational autoencoders [9], grammar variational autoencoders [12], generative adversarial networks [13], and invertible flow models [14, 15] have been applied. However, these methods do not yet guarantee optimal or exact solutions.

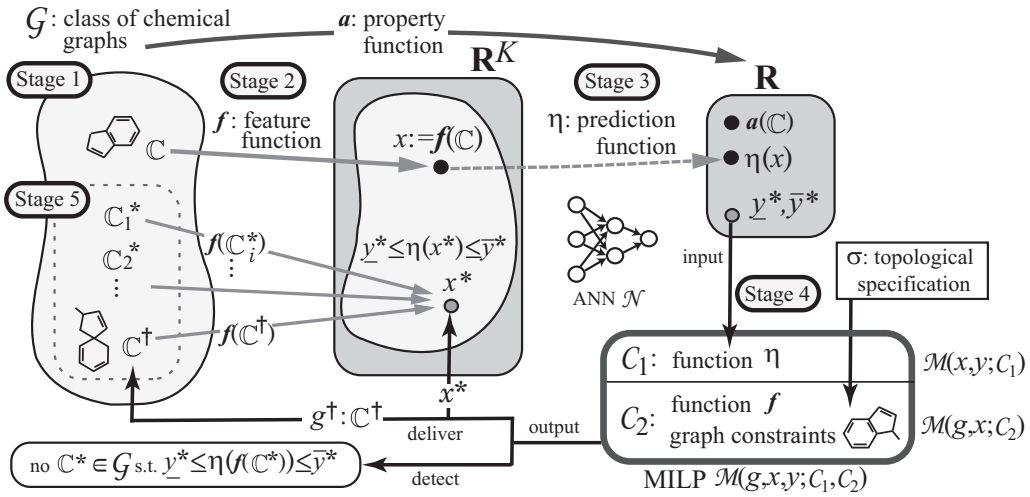


Figure 1: An illustration of a framework for inferring a set of chemical graphs $\mathbb{C}^*$.

Framework Akutsu and Nagamochi [16] proved that the computation process of a given ANN can be simulated with a mixed integer linear programming (MILP). Based on this, a novel framework for inferring chemical graphs has been developed and revised [17, 18], as illustrated in Figure 1. It constructs a prediction function in the first phase and infers a chemical graph in the second phase. The first phase of the framework consists of three stages. Stage 1 chooses a chemical property π and a class $\mathcal{G}$ of graphs, where a property function a is defined so that $a(\mathbb{C})$ is the value of π for a compound $\mathbb{C} \in \mathcal{G}$, and collects a data set D_π of chemical graphs in $\mathcal{G}$ such that $a(\mathbb{C})$ is available for every $\mathbb{C} \in D_\pi$. Stage 2 introduces a feature function $f : \mathcal{G} \to \mathbb{R}^K$ for a positive integer K. Stage 3 constructs a prediction function η with an ANN $\mathcal{N}$ that, given a vector $x \in \mathbb{R}^K$, returns a value $y = \eta(x) \in \mathbb{R}$ so that $\eta(f(\mathbb{C}))$ serves as a predicted value to the real value $a(\mathbb{C})$ of π for each $\mathbb{C} \in D_\pi$. Given two reals $\underline{y}^*$ and $\overline{y}^*$ as an interval for a target chemical value, the second phase infers chemical graphs $\mathbb{C}^*$ with $\underline{y}^* \le \eta(f(\mathbb{C}^*)) \le \overline{y}^*$ in the next two stages. After Stage 3, we have obtained a feature function f and a prediction function η. We can specify an additional constraint on the substructures of target chemical graphs, called a *topological specification* before we infer a target chemical graph. In Stage 4, the following two MILP formulations are prepared:

- MILP $\mathcal{M}(x, y; \mathcal{C}_1)$ with a set $\mathcal{C}_1$ of linear constraints on variables x and y (and some other auxiliary variables) simulates the process of computing $y := \eta(x)$ from a vector x; and

- MILP $\mathcal{M}(g, x; \mathcal{C}_2)$ with a set $\mathcal{C}_2$ of linear constraints on variable x and a variable vector g that represents a chemical graph $\mathbb{C}$ (and some other auxiliary variables) simulates the process of computing $x := f(\mathbb{C})$ from a chemical graph $\mathbb{C}$ and chooses a chemical graph $\mathbb{C}$ that satisfies the given topological specification σ.

Given an interval with boundaries $\underline{y}^*, \overline{y}^* \in \mathbb{R}$, Stage 4 solves the combined MILP $\mathcal{M}(g, x, y; \mathcal{C}_1, \mathcal{C}_2)$ to find a feature vector $x^* \in \mathbb{R}^K$ and a chemical graph $\mathbb{C}^\dagger$ with the specification σ such that $f(\mathbb{C}^\dagger) = x^*$ and $\underline{y}^* \leq \eta(x^*) \leq \overline{y}^*$ (where if the MILP instance is infeasible then this suggests that there does not exist such a desired chemical graph). Stage 5 generates other chemical graphs $\mathbb{C}^*$ such that $\underline{y}^* \leq \eta(f(\mathbb{C}^*)) \leq \overline{y}^*$ based on the output chemical graph $\mathbb{C}^\dagger$.

A modeling of chemical compounds together with an MILP formulation has been improved so that a chemical compound with any graph structure can be treated (see Shi et al. [18]). Not only ANNs but also other machine learning methods have been used to construct a prediction function η in Stage 3 recently. Tanaka et al. [19] (resp., Zhu et al. [20]) used a decision tree (resp., linear regression) to construct a prediction function η in Stage 3 in the framework and derived an MILP $\mathcal{M}(x, y; \mathcal{C}_1)$ that simulates the computation process of a decision tree (resp., linear regression).

The novelty of the framework is based on the fact that a prediction process by linear regression or ANNs can be modeled as an MILP, to which we can find a mathematically exact solution by relying on the state-of-the-art solvers from Operations Research (OR). A sophisticated method has been studied by Shi et al. [18] in order to formulate a sparse MILP instance even for a complicated requirements in a topological specification. Currently an MILP instance in Stage 4 for inferring a chemical compound with 50 non-hydrogen atoms contains around 10,000 variables and 10,000 linear constraints, and can be solved in a few seconds.

Contribution In this paper, we apply a mathematical programming in OR to Stage 3 in order to design a new machine learning method for QSAR. Let us compare linear regression and ANNs. The former uses a hyperplane to explain a given data set and the latter can represent a more complex subspace than a hyperplane. Importantly a best hyperplane that minimizes an error function can be found exactly in the former whereas a local optimum solution to an error function is constructed by an iterative procedure in the latter and different local optimum solutions often appear depending on how we have tuned many parameters in ANNs. Linear regression can be regarded as an ANN on an architecture with an input layer and an output layer of a single node with a linear activation function. We consider an ANN on the same architecture such that each node in the input and out layers is equipped with a set Φ of activation functions. Given a data set, we consider a problem of minimizing an error function on the data set by choosing a weight of each arc, a bias of the output node and a best activation function for each node simultaneously. With some restriction on the set Φ of activation functions and the definition of an error function, we show that such an minimization problem can be formulated as a linear program, which is much easier than an MILP to solve exactly. We call this new method "adjustive linear regression" and implemented it in Stages 3 an 4 in the framework. We used the same MILP $\mathcal{M}(g, x; \mathcal{C}_2)$ formulation proposed by Zhu et al. [20] and omit the details in this paper. We compared adjustive linear regression with Lasso linear regression in constructing prediction functions for several chemical properties. From the results of our computational experiments, we observe that a prediction function constructed by adjustive linear regression for some chemical properties drastically outperform that by Lasso

linear regression.

The paper is organized as follows. Section 2 reviews the idea of prediction functions based on linear regression and ANNs and designs "adjustive linear regression," a new method for constructing a prediction function by solving a linear program to optimize a choice of weights/bias together with activation functions in an ANN with no hidden layers. Section 3.1 introduces some notions on graphs, a modeling of chemical compounds and a choice of descriptors. Section 4 reviews a method, called a *two-layered model* for representing the feature of a chemical graph in order to deal with an arbitrary graph in the framework. Section 5 reports the results on some computational experiments conducted for the framework of inferring chemical graphs by using our new method of adjustive linear regression. Section 6 makes some concluding remarks. Some technical details are given in Appendices: Appendix A for all descriptors in our feature function; Appendix B for a full description of a topological specification; and Appendix C for the detail of test instances used in our computational experiment for Stages 4 and 5.

2 Constructing Prediction Functions

Let $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{Z}$ and $\mathbb{Z}_+$ denote the sets of reals, non-negative reals, integers and non-negative integers, respectively. For two integers a and b, let $[a, b]$ denote the set of integers i with $a \leq i \leq b$. For a vector $x \in \mathbb{R}^p$, the j-th entry of x is denoted by $x(j), j \in [1, p]$.

2.1 Linear Prediction Functions

For an integer $K \geq 1$, define a feature space $\mathbb{R}^K$. Let $\mathcal{X} = \{x_1, x_2, \ldots, x_m\}$ be a set of feature vectors $x \in \mathbb{R}^K$ and let $a_i \in \mathbb{R}$ be a real assigned to a feature vector x_i. Let $A = \{a_i \mid i \in [1, m]\}$. A function $\eta : \mathbb{R}^K \to \mathbb{R}$ is called a *prediction function*. We wish to find a prediction function $\eta : \mathbb{R}^K \to \mathbb{R}$ based on a subset of $\{x_1, x_2, \ldots, x_m\}$ so that $\eta(x_i)$ is closed to the value a_i for many indices $i \in [1, m]$.

For a prediction function $\eta : \mathbb{R}^K \to \mathbb{R}$, define an error function

$$\mathrm{Err}(\eta; \mathcal{X}) \triangleq \sum_{i \in [1,m]} (a_i - \eta(x_i))^2,$$

and define the *coefficient of determination* $\mathrm{R}^2(\eta, \mathcal{X})$ to be

$$\mathrm{R}^2(\eta, \mathcal{X}) \triangleq 1 - \frac{\mathrm{Err}(\eta; \mathcal{X})}{\sum_{i \in [1,m]} (a_i - \widetilde{a})^2} \text{ for } \widetilde{a} = \frac{1}{m} \sum_{i \in [1,m]} a_i.$$

Many methods have been proposed in order to find a prediction function η that minimizes the error function $\mathrm{Err}(\eta_{w,b}; \mathcal{X})$ possibly without using all elements in $\mathcal{X}$.

For the feature space $\mathbb{R}^K$, a hyperplane is defined to be a pair (w, b) of a vector $w \in \mathbb{R}^K$ and a real $b \in \mathbb{R}$. A prediction function η is called *linear* if η is given by $\eta_{w,b}(x) = w \cdot x + b, x \in \mathbb{R}^K$ for a hyperplane (w, b). The linear regression is to find a hyperplane (w, b) that minimizes $\mathrm{Err}(\eta_{w,b}; \mathcal{X}) = \sum_{i \in [1,m]} (a_i - (w \cdot x_i + b))^2$.

In many cases, a feature vector f contains descriptors that do not play an essential role in constructing a good prediction function. When we solve the minimization problem, the entries $w(j)$ for some descriptors $j \in [1, K]$ in the resulting hyperplane (w, b) become zero, which means that these descriptors were not necessarily important for finding a prediction function $\eta_{w,b}$. It is proposed that solving the minimization with an additional penalty term to the error function often results in a more number of entries $w(j) = 0$, reducing a set of descriptors necessary for defining a prediction function $\eta_{w,b}$. For an error function with such a penalty term, a Ridge function $\frac{1}{2m}\mathrm{Err}(\eta_{w,b}; \mathcal{X}) + \lambda[\sum_{j\in[1,K]} w(j)^2 + b^2]$ [21, 22] and a Lasso function $\frac{1}{2m}\mathrm{Err}(\eta_{w,b}; \mathcal{X}) + \lambda[\sum_{j\in[1,K]} |w(j)| + |b|]$ [23] are known, where $\lambda \in \mathbb{R}$ is a given real number. As a hybridization of Ridge linear regression and Lasso linear regression, a linear regression that minimizes an error function defined to be $\frac{1}{2m}\mathrm{Err}(\eta_{w,b}; \mathcal{X}) + \lambda_2[\sum_{j\in[1,K]} w(j)^2 + b^2] + \lambda_1[\sum_{j\in[1,K]} |w(j)| + |b|]$ is called elastic net linear regression [24], where $\lambda_1, \lambda_2 \in \mathbb{R}$ are given real numbers.

Zhu et al. [20] used Lasso linear regression to construct a prediction function η in Stage 3 in the framework.

2.2 ANNs for Linear Prediction Functions

It is not difficult to see that a linear prediction function η with a hyperplane (w, b) can be represented by an ANN $\mathcal{N}$ with an input layer $L_{\text{in}} = \{u_1, u_2, \ldots, u_K\}$ of K input nodes and an output layer $L_{\text{out}} = \{v\}$ of a single output node v such that the weight of an arc (u_j, v) from an input node u_j to the output node v is given by $w(j), j \in [1, K]$; the bias at node v is given by b; and the activation function at node v is linear. See Figure 2(a) for an illustration of an ANN $\mathcal{N}$ that represents a linear prediction function η with a hyperplane (w, b). Given a vector $x \in \mathbb{R}^K$, the ANN $\mathcal{N}$ outputs $y := \sum_{j\in[1,K]} w(j)x(j) + b$.

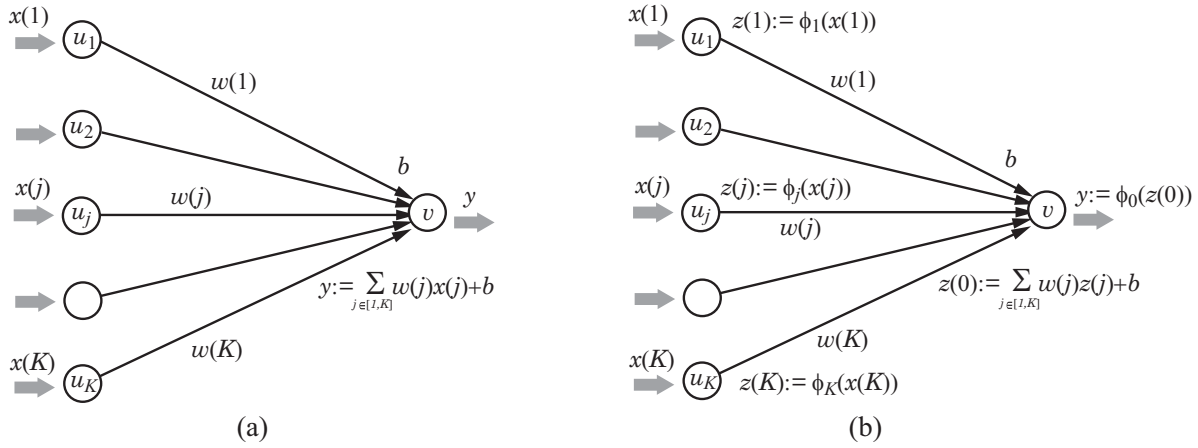


Figure 2: An illustration of the process in ANNs with no hidden layers: (a) An ANN $\mathcal{N}$ that represents a linear prediction function η with a hyperplane (w, b); (b) an ANN $\mathcal{N}_\phi$ with activation functions $\phi_j, j \in [0, K]$ at all nodes.

We consider an ANN $\mathcal{N}_\phi$ with the same architecture with the ANN $\mathcal{N}$ and introduce activation functions ϕ_j at nodes $u_j, j \in [1, K]$ and an activation function ϕ_0 at node v. Given a vector $x \in \mathbb{R}^K$,

the ANN $\mathcal{N}_\phi$ outputs $y := \phi_0(z(0))$ for $z(0) := \sum_{j\in[1,K]} w(j)z(j)+b$ and $z(j) := \phi_j(x(j)), j \in [1, K]$.

In a standard method of a prediction function $\eta_{\mathcal{N}_\phi}$ with the above ANN $\mathcal{N}_\phi$, we specify each activation function ϕ_j and determine weights w and a bias b by executing an iterative procedure that tries to minimize an error function between the real values a_i and the predicted values $\eta_{\mathcal{N}_\phi}(x_i)$.

2.3 Adjustive Linear Regression

In this paper, we design a new method of constructing a prediction function with the above ANN $\mathcal{N}_\phi$ so that (i) not only weights w and a bias b but also prediction functions ϕ_j are chosen so as to minimize an error function and (ii) the minimization problem is formulated as a linear programming problem.

We introduce a class Φ_j of functions for a choice of each activation function $\phi_j, j \in [0, K]$. When we choose a function $\phi_j \in \Phi_j$ for each $j \in [0, K]$ and a hyperplane (w, b), we define a prediction function $\eta_{\Psi,w,b}$ such that

$$\eta_{\Psi,w,b}(x) \triangleq \phi_0(\sum_{j\in[1,K]} w(j)(\phi_j(x(j))) - b)$$

for the set $\Psi = \{\phi_j \mid j \in [0, K]\}$ of the functions.

In this paper, we use a function $\xi(t) = ct+c't^2+c''(1-(t-1)^2), 0 \le t \le 1$ for a function $\phi_j, j \in [1, K]$ or the inverse ϕ_0^{-1} of a function ϕ_0, where c, c' and c'' are nonnegative constant constants with $c+c'+c'' = 1$ which will be determined for each $j \in [0, K]$ by our method. Note that, for a domain $0 \le t \le 1$, $\xi(t)$ is a monotone increasing function and admits an inverse function $\xi^{-1}(t)$, where $\xi^{-1}(t) = t/(c + 2c'')$ when $c' = c''$ and $\xi^{-1}(t) = (-c - 2c'' + \sqrt{(c + 2c'')^2 + 4(c' - c'')t})/(2c' - 2c'')$ when $c' \neq c''$.

We introduce a class Φ_j of functions in the following way.

1. Normalize the set $\{x_i(j) \mid x_i \in \mathcal{X}\}, j \in [1, K]$ and the set $\{a_i(j) \mid x_i \in \mathcal{X}\}$ so that the minimum and maximum in the set become 0 and 1.

2. For each index $j \in [0, K]$, define a class Φ_j of functions to be

$$\Phi_j \triangleq \{c_0(j)t + c_1(j)t^2 + c_2(j)(1 - (t - 1)^2), 0 \le t \le 1 \mid c_q(j) \ge 0, q \in [0, 2], \sum_{q\in[0,2]} c_q(j) = 1\}, j \in [1, K].$$

   Define

$$\widetilde{\Phi}_0 \triangleq \{c_0(0)t + c_1(0)t^2 + c_2(0)(1 - (t - 1)^2), 0 \le t \le 1 \mid c_q(0) \ge 0, q \in [0, 2], \sum_{q\in[0,2]} c_q(0) = 1\},$$

$$\Phi_0 \triangleq \{\xi^{-1}(t), 0 \le t \le 1 \mid \xi(t) \in \widetilde{\Phi}_0\}.$$

To use linear programming, we measure an error of a prediction function η over a data set $\mathcal{X}$ by the sum of the absolute errors:

$$\mathrm{SAE}(\eta; \mathcal{X}) \triangleq \sum_{x_i\in\mathcal{X}} |a_i - \eta(x_i)|.$$

Now our aim is to find a prediction function $\eta_{\Psi,w,b}$ that minimizes the sum of the absolute errors $\mathrm{SAE}(\eta_{\Psi,w,b};\mathcal{X})$ or equivalently

$$\sum_{i\in[1,m]} |\phi_0^{-1}(a_i) - (\sum_{j\in[1,K]} w(j)(\phi_j(x_i(j))) - b)| \tag{1}$$

over all functions $\phi_0 \in \widetilde{\Phi}_0, \phi_j \in \Phi_j, j \in [1,K]$ and hyperplanes (w,b).

To formulate this minimization problem as a linear programming problem, we predetermine the sign of $w(j)$ for each descriptor j in a hyperplane (w,b) that we will choose. Compute the correlation coefficient $\sigma(X_j,A)$ between $X_j = \{x_i(j) \mid i \in [1,m]\}$ and $A = \{a_i \mid i \in [1,m]\}$ and partition the set of descriptors into two sets $I^+ := \{j \in [1,K] \mid \sigma(X_j,A) \geq 0\}$ and $I^- := \{j \in [1,K] \mid \sigma(X_j,A) < 0\}$. We impose an additional constraint that $w(j) \geq 0, j \in I^+$ and $w(j) \leq 0, j \in I^-$. Then the objective function (1) is described as follows, where we rewrite each term $w(j), j \in I^+$ (resp., $-w(j), j \in I^-$) as $w'(j)$:

$$\begin{aligned}\sum_{i\in[1,m]} \Big| & c_0(0)a_i + c_1(0)a_i^2 + c_2(0)(1-(a_i-1)^2) \\ & - \sum_{j\in I^+}[w'(j)\big(c_0(j)x_i(j) + c_1(j)x_i(j)^2 + c_2(j)(1-(x_i(j)-1)^2)\big)] \\ & + \sum_{j\in I^-}[w'(j)\big(c_0(j)x_i(j) + c_1(j)x_i(j)^2 + c_2(j)(1-(x_i(j)-1)^2)\big)] - b\Big|.\end{aligned}$$

We minimize (1) over all nonnegative reals $c_q(j), q \in [0,2], j \in [1,K]$, nonnegative reals $w(j), j \in [1,K]$ and a real $b \in \mathbb{R}$ such that $\sum_{q\in[0,2]} c_q(j) = 1, j \in [1,K]$.

Before we derive a linear programming formulation to the above minimization problem, we include a penalty term for the weights $w(j), j \in [1,K]$ analogously with the Lasso linear regression. We consider the following problem which we call *adjustive linear regression*, where $w'(j)c_q(j), q \in [0,2]$ is rewritten as $w_q(j)$.

Adjustive Linear Regression$(\mathcal{X},\lambda)$

$$\begin{aligned}\text{Minimize: } & \frac{1}{2m}\sum_{i\in[1,m]} \Big| c_0(0)a_i + c_1(0)a_i^2 + c_2(0)(1-(a_i-1)^2) \\ & - \sum_{j\in I^+}[w_0(j)x_i(j) + w_1(j)x_i(j)^2 + w_2(j)(1-(x_i(j)-1)^2)] \\ & + \sum_{j\in I^-}[w_0(j)x_i(j) + w_1(j)x_i(j)^2 + w_2(j)(1-(x_i(j)-1)^2)] - b\Big| \\ & + \lambda\sum_{j\in[1,K]} w_0(j) + \lambda|b| \end{aligned} \tag{2}$$

subject to

$$c_0(0) + c_1(0) + c_2(0) = 1.$$

When the set $S = \{a_i \mid i \in [1,m]\}$ (resp., $S = \{x_i(j) \mid i \in [1,m]\}$ for an index $j \in [1,m]$) is binary (i.e., $S = \{0,1\}$), we always set $c_1(0) = c_2(0) = 0$ (resp., $c_1(j) = c_2(j) = 0$).

We observe that adjustive linear regression is an extension of the Lasso linear regression except that the error function is the sum of absolute errors in the former and the sum of square errors in the latter.

We solve the above minimization problem (2) to construct a prediction function $\eta_{\Psi,w,b}$. Let $c_q^*(0), q \in [0,2]$, $w_q^*(j), q \in [0,2], j \in [1,K]$ and b^* denote the values of variables $c_q(0), q \in [0,2]$, $w_q(j), q \in [0,2], j \in [1,K]$ and b in an optimal solution, respectively. Let K' denote the number of descriptors $j \in [1,K]$ with $w_0^*(j) > 0$ and $I_{K'}$ denote the set of $j \in [1,K]$ with $w_0^*(j) > 0$. Then we set

$w^*(j) := 0$ for $j \in [1,K]$ with $w_0^*(j) = 0$,
$w^*(j) := w_0^*(j)/(w_0^*(j) + w_1^*(j) + w_2^*(j))$ for $j \in I^+ \cap I_{K'}$,
$w^*(j) := -w_0^*(j)/(w_0^*(j) + w_1^*(j) + w_2^*(j))$ for $j \in I^- \cap I_{K'}$,
$c_q^*(j) := w_q^*(j)/w^*(j), q \in [0,2]$ for $j \in I_{K'}$ and
$w^* := (w_0^*(1), w_0^*(2), \ldots, w_0^*(K)) \in \mathbb{R}^K$.

For a set Ψ^* of selected functions $\phi_j(t) = c_0^*(j)t + c_1^*(j)t^2 + c_2^*(j)(1-(t-1)^2), j \in I_{K'}$ with and $\phi_0(t)$ with $\phi_0^{-1}(t) = c_0^*(0)t + c_1^*(0)t^2 + c_2^*(0)(1-(t-1)^2)$ and a hyperplane (w^*, b^*), we construct a prediction function η_{Ψ^*,w^*,b^*}.

We propose the following scheme of executing adjustive linear regression for constructing a prediction function and evaluating the performance.

1. Given a data set $\mathcal{X} = \{x_i \in \mathbb{R}^K \mid i \in [1,m]\}$ of normalized feature vectors and a set $A = \{a_i \in \mathbb{R} \mid i \in [1,m]\}$ of normalized observed values, we choose a real $\lambda > 0$ possibly from a set of candidates for $\lambda > 0$ so that the performance of a prediction function η_{Ψ^*,w^*,b^*} obtained from an optimal solution (Ψ^*, w^*, b^*) to the adjustive linear regression $(\mathcal{X}, \lambda)$ attains a criterion, where we may use cross-validation and the test coefficient of determination to know the performance. Note that the reals $c_q^*(j), q \in [1,2], j \in I_{K'}$ indicate the non-linearity of a function ϕ_j and the resulting prediction function η_{Ψ^*,w^*,b^*} may easily cause overfitting when the bounds $c_q^*(j)$ are unnecessarily large. To bound reals $c_q^*(j), q \in [1,2]$, we also penalize weights $w_q(j), q \in [1,2], j \in I_{K'}$ with the same real λ.

2. With the real λ determined in 1, we evaluate the performance of a prediction function obtained with adjustive linear regression based on cross-validation. We divide the entire set $\mathcal{X}$ into five subsets $\mathcal{X}^{(k)}, k \in [1,5]$. For each $k \in [1,5]$, we use the set $\mathcal{X} \setminus \mathcal{X}^{(k)}$ as a training data to construct a prediction function $\eta_{\Psi,w,b}$ with adjustive linear regression $(\mathcal{X} \setminus \mathcal{X}^{(k)}, \lambda)$ and compute the coefficient of determination $\mathrm{R}^2(\eta_{\Psi,w,b}; \mathcal{X}^{(k)})$.

2.4 An LP formulation for Adjustive Linear Regression

We formulate a linear programming problem to the adjustive linear regression $(\mathcal{X}, \lambda)$.

LP$(\mathcal{X}, \lambda)$:
constants:
- A set $\mathcal{X} = \{x_i \in \mathbb{R}^K \mid i \in [1,m]\}$ of feature vectors and a set $A = \{a_i \in \mathbb{R} \mid i \in [1,m]\}$ of observed values. Assume that each of the sets $X_j = \{x_i(j) \mid i \in [1,m]\}$, $j \in [1,K]$ and A is standardized;
- A positive real $\lambda \in \mathbb{R}$: a coefficient for the penalty term;

variables:
- Nonnegative reals $c_q(0) \in \mathbb{R}, q \in [0,2]$;

- Nonnegative vectors $w_q \in \mathbb{R}^K, q \in [0,2]$ and a real $b \in \mathbb{R}$;
- Nonnegative real $\overline{b} \in \mathbb{R}$;
- Nonnegative reals $\Delta_i \geq 0, i \in [1,m]$;

constraints:

$$c_0(0) + c_1(0) + c_2(0) = 1, \tag{3}$$

$$\begin{aligned} &\Delta_i \geq c_0(0)a_i + c_1(0)a_i^2 + c_2(0)(1-(a_i-1)^2) \\ &- \sum_{j \in I^+} [w_0(j)x_i(j) + w_1(j)x_i(j)^2 + w_2(j)(1-(x_i(j)-1)^2)] \\ &+ \sum_{j \in I^-} [w_0(j)x_i(j) + w_1(j)x_i(j)^2 + w_2(j)(1-(x_i(j)-1)^2)] - b \geq -\Delta_i, \qquad i \in [1,m], \end{aligned} \tag{4}$$

$$\overline{b} \geq b \geq -\overline{b}, \tag{5}$$

objective function:

$$\text{Minimize} \quad \frac{1}{2m} \sum_{i \in [1,m]} \Delta_i + \lambda \sum_{q \in [0,2], j \in [1,K]} w_q(j) + \lambda \overline{b}.$$

We see that the numbers of variables and constraints in the linear program $\mathrm{LP}(\mathcal{X}, \lambda)$ are both $O(m+K)$.

Let $w_q^*(j), q \in [0,2], j \in [1,K]$ and b^* denote the values of variables $w_q(j), q \in [0,2], j \in [1,K]$ and b in an optimal solution to linear program $\mathrm{LP}(\mathcal{X}, \lambda)$, respectively. Let K' denote the number of descriptors $j \in [1,K]$ with $w_0^*(j) > 0$ and $I_{K'}$ denote the set of $j \in [1,K]$ with $w_0^*(j) > 0$. Then we obtain an optimal solution to the adjustive linear regression (2) by setting $w^*(j) := w_0^*(j)/(w_0^*(j)+w_1^*(j)+w_2^*(j)), j \in I^+ \cap I_{K'}$, $w^*(j) := -w_0^*(j)/(w_0^*(j)+w_1^*(j)+w_2^*(j)), j \in I^- \cap I_{K'}$, and $c_q^*(j) := w_q^*(j)/w^*(j), q \in [1,2], j \in I_{K'}$.

3 Modeling of Chemical Compounds

This section introduces some notions and terminologies on graphs and reviews the modeling of chemical compounds due to Zhu et al. [20].

Graph Given a graph G, let $V(G)$ and $E(G)$ denote the sets of vertices and edges, respectively. For a subset $V' \subseteq V(G)$ (resp., $E' \subseteq E(G)$) of a graph G, let $G - V'$ (resp., $G - E'$) denote the graph obtained from G by removing the vertices in V' (resp., the edges in E'), where we remove all edges incident to a vertex in V' in $G - V'$. An edge subset $E' \subseteq E(G)$ in a connected graph G is called *separating* (resp., *non-separating*) if $G - E'$ remains connected (resp., $G - E'$ becomes disconnected). The *rank* $\mathrm{r}(G)$ of a graph G is defined to be the minimum $|F|$ of an edge subset

$F \subseteq E(G)$ such that $G - F$ contains no cycle, where $\mathrm{r}(G) = |E(G)| - |V(G)| + 1$. Observe that $\mathrm{r}(G - E') = \mathrm{r}(G) - |E'|$ holds for any non-separating edge subset $E' \subseteq E(G)$. An edge $e = u_1u_2 \in E(G)$ in a connected graph G is called a *bridge* if $\{e\}$ is separating, i.e., $G - e$ consists of two connected graphs G_i containing vertex u_i, $i = 1, 2$. For a connected cyclic graph G, an edge e is called a *core-edge* if it is in a cycle of G or is a bridge $e = u_1u_2$ such that each of the connected graphs G_i, $i = 1, 2$ of $G - e$ contains a cycle. A vertex incident to a core-edge is called a *core-vertex* of G. A path with two end-vertices u and v is called a *u, v-path*.

A vertex designated in a graph G is called a *root*. In this paper, we designate at most two vertices as roots, and denote by $\mathrm{Rt}(G)$ the set of roots of G. We call a graph G *rooted* (resp., *bi-rooted*) if $|\mathrm{Rt}(G)| = 1$ (resp., $|\mathrm{Rt}(G)| = 2$), where we call G *unrooted* if $\mathrm{Rt}(G) = \emptyset$.

For a graph G possibly with roots a *leaf-vertex* is defined to be a non-root vertex $v \in V(G) \setminus \mathrm{Rt}(G)$ with degree 1, call the edge uv incident to a leaf vertex v a *leaf-edge*, and denote by $V_{\mathrm{leaf}}(G)$ and $E_{\mathrm{leaf}}(G)$ the sets of leaf-vertices and leaf-edges in G, respectively. For a graph or a rooted graph G, we define graphs $G_i, i \in \mathbb{Z}_+$ obtained from G by removing the set of leaf-vertices i times so that

$$G_0 := G; \quad G_{i+1} := G_i - V_{\mathrm{leaf}}(G_i),$$

where we call a vertex $v \in V_{\mathrm{leaf}}(G_k)$ a *leaf k-branch* and we say that a vertex $v \in V_{\mathrm{leaf}}(G_k)$ has *height* $\mathrm{ht}(v) = k$ in G. The *height* $\mathrm{ht}(T)$ of a rooted tree T is defined to be the maximum of $\mathrm{ht}(v)$ of a vertex $v \in V(T)$. For an integer $k \geq 0$, we call a rooted tree T *k-lean* if T has at most one leaf k-branch. For an unrooted cyclic graph G, we regard that the set of non-core-edges in G induces a collection $\mathcal{T}$ of trees each of which is rooted at a core-vertex, where we call G *k-lean* if each of the rooted trees in $\mathcal{T}$ is k-lean.

3.1 Chemical Graphs

To represent a chemical compound, we introduce a set of chemical elements such as $\mathtt{H}$ (hydrogen), $\mathtt{C}$ (carbon), $\mathtt{O}$ (oxygen), $\mathtt{N}$ (nitrogen) and so on. To distinguish a chemical element $\mathtt{a}$ with multiple valences such as $\mathtt{S}$ (sulfur), we denote a chemical element $\mathtt{a}$ with a valence i by $\mathtt{a}_{(i)}$, where we do not use such a suffix (i) for a chemical element $\mathtt{a}$ with a unique valence. Let Λ be a set of chemical elements $\mathtt{a}_{(i)}$. For example, $\Lambda = \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{N}, \mathtt{P}, \mathtt{S}_{(2)}, \mathtt{S}_{(4)}, \mathtt{S}_{(6)}\}$. Let $\mathrm{val} : \Lambda \to [1, 6]$ be a valence function. For example, $\mathrm{val}(\mathtt{H}) = 1$, $\mathrm{val}(\mathtt{C}) = 4$, $\mathrm{val}(\mathtt{O}) = 2$, $\mathrm{val}(\mathtt{P}) = 5$, $\mathrm{val}(\mathtt{S}_{(2)}) = 2$, $\mathrm{val}(\mathtt{S}_{(4)}) = 4$ and $\mathrm{val}(\mathtt{S}_{(6)}) = 6$. For each chemical element $\mathtt{a} \in \Lambda$, let $\mathrm{mass}(\mathtt{a})$ denote the mass of $\mathtt{a}$.

A chemical compound is represented by a *chemical graph* defined to be a tuple $\mathbb{C} = (H, \alpha, \beta)$ of a simple, connected undirected graph H and functions $\alpha : V(H) \to \Lambda$ and $\beta : E(H) \to [1, 3]$. The set of atoms and the set of bonds in the compound are represented by the vertex set $V(H)$ and the edge set $E(H)$, respectively. The chemical element assigned to a vertex $v \in V(H)$ is represented by $\alpha(v)$ and the bond-multiplicity between two adjacent vertices $u, v \in V(H)$ is represented by $\beta(e)$ of the edge $e = uv \in E(H)$. We say that two tuples $(H_i, \alpha_i, \beta_i), i = 1, 2$ are *isomorphic* if they admit an isomorphism ϕ, i.e., a bijection $\phi : V(H_1) \to V(H_2)$ such that $uv \in E(H_1), \alpha_1(u) = \mathtt{a}, \alpha_1(v) = \mathtt{b}, \beta_1(uv) = m \leftrightarrow \phi(u)\phi(v) \in E(H_2), \alpha_2(\phi(u)) = \mathtt{a}, \alpha_2(\phi(v)) = \mathtt{b}, \beta_2(\phi(u)\phi(v)) = m$. When H_i is rooted at a vertex $r_i, i = 1, 2$, $(H_i, \alpha_i, \beta_i), i = 1, 2$ are *rooted-isomorphic* (r-isomorphic) if they admit an isomorphism ϕ such that $\phi(r_1) = r_2$.

For a notational convenience, we use a function $\beta_{\mathbb{C}} : V(H) \to [0, 12]$ for a chemical graph $\mathbb{C} = (H, \alpha, \beta)$ such that $\beta_{\mathbb{C}}(u)$ means the sum of bond-multiplicities of edges incident to a vertex u; i.e.,

$$\beta_{\mathbb{C}}(u) \triangleq \sum_{uv \in E(H)} \beta(uv) \text{ for each vertex } u \in V(H).$$

For each vertex $u \in V(H)$, define the *electron-degree* $\mathrm{eledeg}_{\mathbb{C}}(u)$ to be

$$\mathrm{eledeg}_{\mathbb{C}}(u) \triangleq \beta_{\mathbb{C}}(u) - \mathrm{val}(\alpha(u)).$$

For each vertex $u \in V(H)$, let $\deg_{\mathbb{C}}(u)$ denote the number of vertices adjacent to the vertex u in $\mathbb{C}$.

For a chemical graph $\mathbb{C} = (H, \alpha, \beta)$, let $V_{\mathtt{a}}(\mathbb{C})$, $\mathtt{a} \in \Lambda$ denote the set vertices $v \in V(H)$ such that $\alpha(v) = \mathtt{a}$ in $\mathbb{C}$ and define the *hydrogen-suppressed chemical graph* $\langle \mathbb{C} \rangle$ to be the graph obtained from H by removing all the vertices $v \in V_{\mathtt{H}}(\mathbb{C})$.

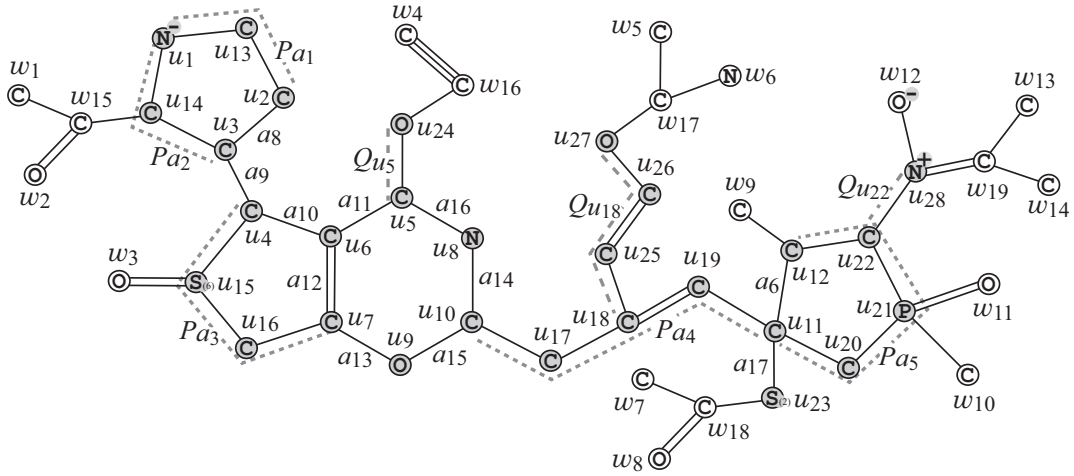


Figure 3: An illustration of a hydrogen-suppressed chemical graph $\langle \mathbb{C} \rangle$ obtained from a chemical graph $\mathbb{C}$ with $\mathrm{r}(\mathbb{C}) = 4$ by removing all the hydrogens, where for $\rho = 2$, $V^{\mathrm{ex}}(\mathbb{C}) = \{w_i \mid i \in [1, 19]\}$ and $V^{\mathrm{int}}(\mathbb{C}) = \{u_i \mid i \in [1, 28]\}$.

4 Two-layered Model

This section reviews the two-layered model introduced by Zhu et al. [20].

Let $\mathbb{C} = (H, \alpha, \beta)$ be a chemical graph and $\rho \geq 1$ be an integer, which we call a *branch-parameter*.

A *two-layered model* of $\mathbb{C}$ is a partition of the hydrogen-suppressed chemical graph $\langle \mathbb{C} \rangle$ into an "interior" and an "exterior" in the following way. We call a vertex $v \in V(\langle \mathbb{C} \rangle)$ (resp., an edge $e \in E(\langle \mathbb{C} \rangle)$) of $\mathbb{C}$ an *exterior-vertex* (resp., *exterior-edge*) if $\mathrm{ht}(v) < \rho$ (resp., e is incident to an exterior-vertex) and denote the sets of exterior-vertices and exterior-edges by $V^{\mathrm{ex}}(\mathbb{C})$ and $E^{\mathrm{ex}}(\mathbb{C})$, respectively and denote $V^{\mathrm{int}}(\mathbb{C}) = V(\langle \mathbb{C} \rangle) \setminus V^{\mathrm{ex}}(\mathbb{C})$ and $E^{\mathrm{int}}(\mathbb{C}) = E(\langle \mathbb{C} \rangle) \setminus E^{\mathrm{ex}}(\mathbb{C})$, respectively. We call a vertex in $V^{\mathrm{int}}(\mathbb{C})$ (resp., an edge in $E^{\mathrm{int}}(\mathbb{C})$) an *interior-vertex* (resp., *interior-edge*).

The set $E^{\mathrm{ex}}(\mathbb{C})$ of exterior-edges forms a collection of connected graphs each of which is regarded as a rooted tree T rooted at the vertex $v \in V(T)$ with the maximum $\mathrm{ht}(v)$. Let $\mathcal{T}^{\mathrm{ex}}(\langle\mathbb{C}\rangle)$ denote the set of these chemical rooted trees in $\langle\mathbb{C}\rangle$. The *interior* $\mathbb{C}^{\mathrm{int}}$ of $\mathbb{C}$ is defined to be the subgraph $(V^{\mathrm{int}}(\mathbb{C}), E^{\mathrm{int}}(\mathbb{C}))$ of $\langle\mathbb{C}\rangle$.

Figure 3 illustrates an example of a hydrogen-suppressed chemical graph $\langle\mathbb{C}\rangle$. For a branch-parameter $\rho = 2$, the interior of the chemical graph $\langle\mathbb{C}\rangle$ in Figure 3 is obtained by removing the set of vertices with degree 1 $\rho = 2$ times; i.e., first remove the set $V_1 = \{w_1, w_2, \ldots, w_{14}\}$ of vertices of degree 1 in $\langle\mathbb{C}\rangle$ and then remove the set $V_2 = \{w_{15}, w_{16}, \ldots, w_{19}\}$ of vertices of degree 1 in $\langle\mathbb{C}\rangle - V_1$, where the removed vertices become the exterior-vertices of $\langle\mathbb{C}\rangle$.

For each interior-vertex $u \in V^{\mathrm{int}}(\mathbb{C})$, let $T_u \in \mathcal{T}^{\mathrm{ex}}(\langle\mathbb{C}\rangle)$ denote the chemical tree rooted at u (where possibly T_u consists of vertex u) and define the *ρ-fringe-tree* $\mathbb{C}[u]$ to be the chemical rooted tree obtained from T_u by putting back the hydrogens originally attached with T_u in $\mathbb{C}$. Let $\mathcal{T}(\mathbb{C})$ denote the set of ρ-fringe-trees $\mathbb{C}[u], u \in V^{\mathrm{int}}(\mathbb{C})$. Figure 4 illustrates the set $\mathcal{T}(\mathbb{C}) = \{\mathbb{C}[u_i] \mid i \in [1, 28]\}$ of the 2-fringe-trees of the example $\mathbb{C}$ with $\langle\mathbb{C}\rangle$ in Figure 3.

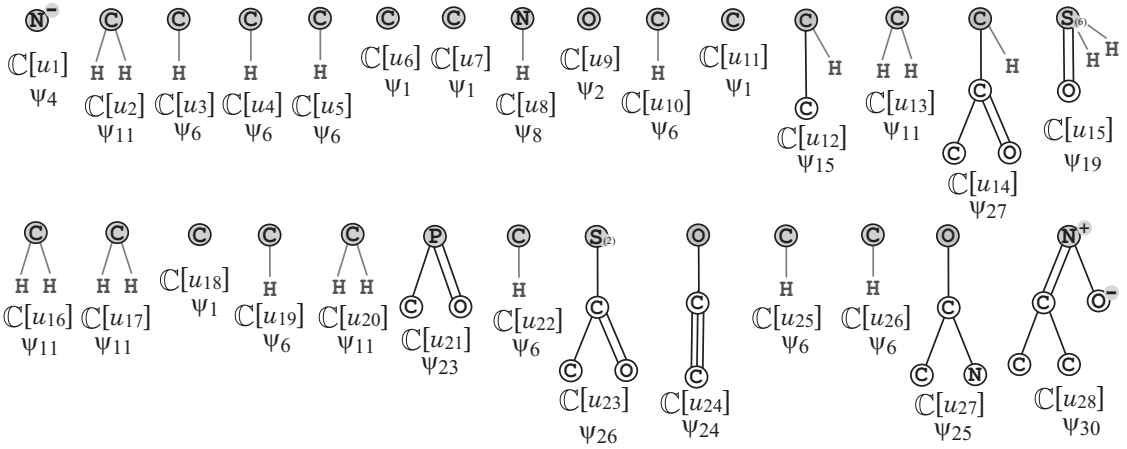


Figure 4: The set $\mathcal{T}(\mathbb{C})$ of 2-fringe-trees $\mathbb{C}[u_i], i \in [1, 28]$ of the example $\mathbb{C}$ with $\langle\mathbb{C}\rangle$ in Figure 3, where the root of each tree is depicted with a gray circle and the hydrogens attached to non-root vertices are omitted in the figure.

Feature Function The feature of an interior-edge $e = uv \in E^{\mathrm{int}}(\mathbb{C})$ such that $\alpha(u) = \mathtt{a}$, $\deg_{\langle\mathbb{C}\rangle}(u) = d$, $\alpha(v) = \mathtt{b}$, $\deg_{\langle\mathbb{C}\rangle}(v) = d'$ and $\beta(e) = m$ is represented by a tuple $(\mathtt{a}d, \mathtt{b}d', m)$, which is called the *edge-configuration* of the edge e, where we call the tuple $(\mathtt{a}, \mathtt{b}, m)$ the *adjacency-configuration* of the edge e.

For an integer K, a feature vector $f(\mathbb{C})$ of a chemical graph $\mathbb{C}$ is defined by a *feature function* f that consists of K descriptors. We call $\mathbb{R}^K$ *the feature space*.

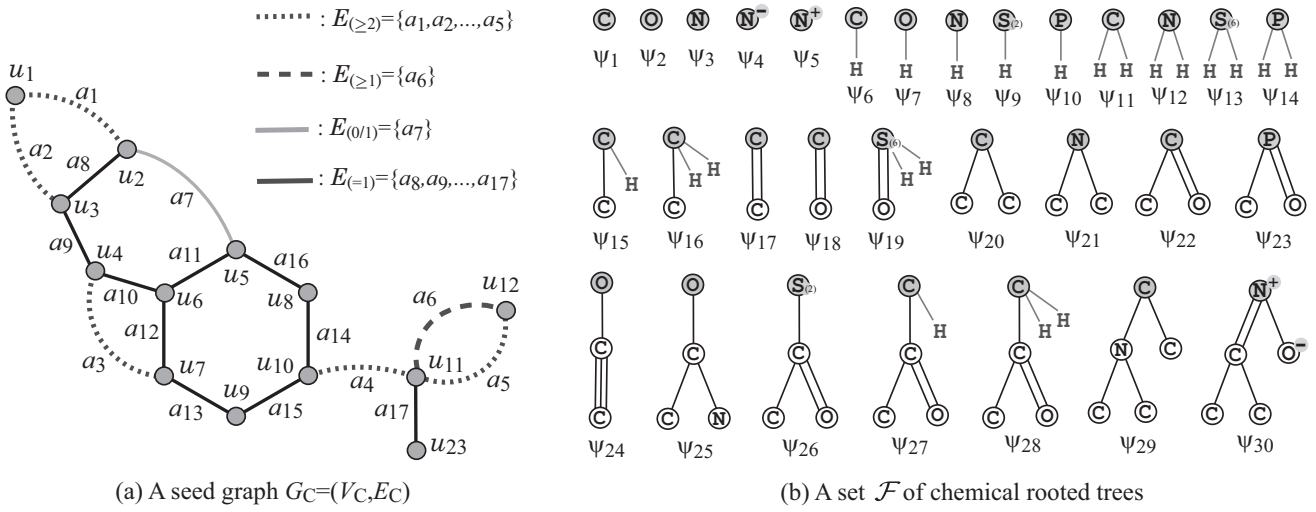
Tanaka et al. [19] defined a feature vector $f(\mathbb{C}) \in \mathbb{R}^K$ to be a combination of the frequency of edge-configurations of the interior-edges and the frequency of chemical rooted trees among the set of chemical rooted trees $\mathbb{C}[u]$ over all interior-vertices u.

Topological Specification A topological specification is described as a set of the following rules proposed by Shi et al. [18] and modified by Tanaka et al. [19]:

(i) a *seed graph* $G_{\rm C}$ as an abstract form of a target chemical graph $\mathbb{C}$;

(ii) a set $\mathcal{F}$ of chemical rooted trees as candidates for a tree $\mathbb{C}[u]$ rooted at each interior-vertex u in $\mathbb{C}$; and

(iii) lower and upper bounds on the number of components in a target chemical graph such as chemical elements, double/triple bonds and the interior-vertices in $\mathbb{C}$.



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

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

Figure 5(a) and (b) illustrate examples of a seed graph $G_{\rm C}$ and a set $\mathcal{F}$ of chemical rooted trees, respectively. Given a seed graph $G_{\rm C}$, the interior of a target chemical graph $\mathbb{C}$ is constructed from $G_{\rm C}$ by replacing some edges $a = uv$ with paths P_a between the end-vertices u and v and by attaching new paths Q_v to some vertices v. For example, a chemical graph $\mathbb{C}$ with $\langle\mathbb{C}\rangle$ in Figure 3 is constructed from the seed graph $G_{\rm C}$ in Figure 5(a) as follows.

- First replace five edges $a_1 = u_1u_2, a_2 = u_1u_3, a_3 = u_4u_7, a_4 = u_{10}u_{11}$ and $a_5 = u_{11}u_{12}$ in $G_{\rm C}$ with new paths $P_{a_1} = (u_1, u_{13}, u_2)$, $P_{a_2} = (u_1, u_{14}, u_3)$, $P_{a_3} = (u_4, u_{15}, u_{16}, u_7)$, $P_{a_4} = (u_{10}, u_{17}, u_{18}, u_{19}, u_{11})$ and $P_{a_5} = (u_{11}, u_{20}, u_{21}, u_{22}, u_{12})$, respectively to obtain a subgraph G_1 of $\langle\mathbb{C}\rangle$.
- Next attach to this graph G_1 three new paths $Q_{u_5} = (u_5, u_{24})$, $Q_{u_{18}} = (u_{18}, u_{25}, u_{26}, u_{27})$ and $Q_{u_{22}} = (u_{22}, u_{28})$ to obtain the interior of $\langle\mathbb{C}\rangle$ in Figure 3.
- Finally attach to the interior 28 trees selected from the set $\mathcal{F}$ and assign chemical elements and bond-multiplicities in the interior to obtain a chemical graph $\mathbb{C}$ with $\langle\mathbb{C}\rangle$ in Figure 3. In Figure 4, $\psi_1 \in \mathcal{F}$ is selected for $\mathbb{C}[u_i]$, $i \in \{6, 7, 11\}$. Similarly ψ_2 for $\mathbb{C}[u_9]$, ψ_4 for $\mathbb{C}[u_1]$, ψ_6 for $\mathbb{C}[u_i]$, $i \in \{3, 4, 5, 10, 19, 22, 25, 26\}$, ψ_8 for $\mathbb{C}[u_8]$, ψ_{11} for $\mathbb{C}[u_i]$, $i \in \{2, 13, 16, 17, 20\}$, ψ_{15} for $\mathbb{C}[u_{12}]$, ψ_{19} for $\mathbb{C}[u_{15}]$, ψ_{23} for $\mathbb{C}[u_{21}]$, ψ_{24} for $\mathbb{C}[u_{24}]$, ψ_{25} for $\mathbb{C}[u_{27}]$, ψ_{26} for $\mathbb{C}[u_{23}]$, ψ_{27} for $\mathbb{C}[u_{14}]$

and ψ_{30} for $\mathbb{C}[u_{28}]$.

Our definition of a topological specification is analogous with the one by Tanaka et al. [19] except for a necessary modification due to the introduction of multiple valences of chemical elements, cations and anions (see Appendix B for a full description of topological specification).

5 Results

We implemented our method of Stages 1 to 5 for inferring chemical graphs under a given topological specification and conducted experiments to evaluate the computational efficiency. We executed the experiments on a PC with Processor: Core i7-9700 (3.0GHz; 4.7 GHz at the maximum) and Memory: 16 GB RAM DDR4. We used scikit-learn version 0.23.2 with Python 3.8.5 for executing linear regression with Lasso function or constructing an ANN. To solve an LP in Stage 3 or an MILP in Stage 4, we used `CPLEX` version 12.10.

Results on Phase 1. We implemented Stages 1, 2 and 3 in Phase 1 as follows.

We have conducted experiments of adjustive linear regression and for 37 chemical properties of monomers (resp., ten chemical properties of polymers) using the feature function [20] (resp., [25]) and we found that the test coefficient of determination R^2 of ALR exceeds 0.6 for the following 28 properties of monomers:
isotropic polarizability (Alpha); boiling point (Bp); critical pressure (Cp); critical temperature (Ct); heat capacity at 298.15K (Cv); dissociation constants (Dc); electron density on the most positive atom (EDPA); flash point (Fp); energy difference between the highest and lowest unoccupied molecular orbitals (Gap); heat of atomization (Ha); heat of combustion (Hc); heat of formation (Hf); energy of highest occupied molecular orbital (Homo); heat of vaporization (Hv); isobaric heat capacities in liquid phase (IhcL); isobaric heat capacities in solid phase (IhcS); KvI-ats retention index (KvI), octanol/water partition coefficient (Kow); lipophilicity (Lp); energy of lowest unoccupied molecular orbital (Lumo); melting point (Mp); optical rotation (OptR); refractive index (Rf); solubility (Sl); surface tension (SfT); internal energy at 0K (U0); viscosity (Vis); and vapor density (Vd) and that the test coefficient of determination R^2 of ALR exceeds 0.8 for the following eight properties of polymers:
experimental amorphous density (AmD); characteristic ratio (ChaR); dielectric constant(DeC); heat capacity liquid (HcL); heat capacity solid (HcS); mol volume (MlV); refractive index (RfId); and glass transition (Tg), where we include the result of property permittivity (Prm) for a comparison with Lasso linear regression and ANN.

We used data sets are provided by HSDB from PubChem [27] for Cp, Ct, Dc, Fp, Hc, Hv, Kow, OptR, Rf and Vd M. Jalali-Heravi and M. Fatemi [28] for EDPA and KvI, Roy and Saha [29] for Bp, Ha, Hf and Mp, Ramakrishnan et al. [30] for Alpha, Cv, Lumo and U0, Goussardet al. [31] for SfT, Goussard et al. [32] for Vis, R. Naef [33] for IhcL and IhcS, Xiao [34] for Lp and Delaney [35] for Sl. Properties Alpha, Cv, Homo, Lumo and U0 share a common original data set D^* with more than 130,000 compounds, and we used a set D_π of 1,000 compounds randomly selected from D^* as a common data set of these four properties π in this experiment.

We used data sets of polymers provided by Bicerano [36], where we did not include any polymer

whose chemical formula could not be found by its name in the book. For property ChaR (resp., RfId), we remove the following polymer as an outlier from the original data set:
ethyleneTerephthalate, oxy(2-methyl-6-phenyl-1_4-phenylene) and N-vinylCarbazole (resp., 2-decyl-1_4-butadiene).

Stage 1. We set a graph class $\mathcal{G}$ to be the set of all chemical graphs with any graph structure, and set a branch-parameter ρ to be 2.

For each of the properties, we first select a set Λ of chemical elements and then collect a data set D_π on chemical graphs over the set Λ of chemical elements. To construct the data set D_π, we eliminated chemical compounds that do not satisfy one of the following: the graph is connected, the number of carbon atoms is at least four, and the number of non-hydrogen neighbors of each atom is at most 4.

Table 1 shows the size and range of data sets that we prepared for each chemical property in Stage 1, where we denote the following:

- Λ: the set of elements used in the data set D_π; Λ is one of the following 12 sets: $\Lambda_1 = \{\mathtt{H}, \mathtt{C}, \mathtt{O}\}$; $\Lambda_2 = \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{N}\}$; $\Lambda_3 = \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{S}\}$; $\Lambda_4 = \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{Si}_{(4)}\}$; $\Lambda_5 = \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{N}, \mathtt{S}_{(2)}, \mathtt{F}\}$; $\Lambda_6 = \{\mathtt{H}, \mathtt{C}_{(2)}, \mathtt{C}_{(3)}, \mathtt{C}_{(4)}, \mathtt{O}, \mathtt{N}_{(2)}, \mathtt{N}_{(3)}\}$; $\Lambda_7 = \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{N}, \mathtt{Cl}, \mathtt{Pb}\}$; $\Lambda_8 = \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{N}, \mathtt{S}_{(2)}, \mathtt{S}_{(6)}, \mathtt{Cl}\}$; $\Lambda_9 = \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{N}, \mathtt{S}_{(2)}, \mathtt{S}_{(4)}, \mathtt{S}_{(6)}, \mathtt{Cl}\}$; $\Lambda_{10} = \{\mathtt{H}, \mathtt{C}_{(2)}, \mathtt{C}_{(3)}, \mathtt{C}_{(4)}, \mathtt{C}_{(5)}, \mathtt{O}, \mathtt{N}_{(1)}, \mathtt{N}_{(2)}, \mathtt{N}_{(3)}, \mathtt{F}\}$;
  $\Lambda_{11} = \{\mathtt{H}, \mathtt{C}_{(2)}, \mathtt{C}_{(3)}, \mathtt{C}_{(4)}, \mathtt{O}, \mathtt{N}_{(2)}, \mathtt{N}_{(3)}, \mathtt{S}_{(2)}, \mathtt{S}_{(4)}, \mathtt{S}_{(6)}, \mathtt{Cl}\}$; $\Lambda_{12} = \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{N}, \mathtt{P}_{(2)}, \mathtt{P}_{(5)}, \mathtt{Cl}\}$;
  $\Lambda_{13} = \{\mathtt{H}, \mathtt{C}, \mathtt{O}_{(1)}, \mathtt{O}_{(2)}, \mathtt{N}\}$; $\Lambda_{14} = \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{N}, \mathtt{Cl}\}$; $\Lambda_{15} = \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{N}, \mathtt{Cl}, \mathtt{S}_{(2)}\}$; and
  $\Lambda_{16} = \{\mathtt{H}, \mathtt{C}, \mathtt{O}_{(1)}, \mathtt{O}_{(2)}, \mathtt{N}, \mathtt{Cl}, \mathtt{Si}_{(4)}, \mathtt{F}\}$, where $\mathtt{a}_{(i)}$ for a chemical element $\mathtt{a}$ and an integer $i \geq 1$ means that a chemical element $\mathtt{a}$ with valence i.
- $|D_\pi|$: the size of data set D_π over Λ for the property π;
- $\underline{n}$, $\overline{n}$: the minimum and maximum values of the number $n(\mathbb{C})$ of non-hydrogen atoms in compounds $\mathbb{C}$ in D_π;
- $\underline{a}$, $\overline{a}$: the minimum and maximum values of $a(\mathbb{C})$ for π over compounds $\mathbb{C}$ in D_π;
- $|\Gamma|$: the number of different edge-configurations of interior-edges over the compounds in D_π;
- $|\mathcal{F}|$: the number of non-isomorphic chemical rooted trees in the set of all 2-fringe-trees in the compounds in D_π; and
- K: the number of descriptors in a feature vector $f(\mathbb{C})$.

Stage 2. We used the feature function defined in our chemical model without suppressing hydrogen (see Appendix A for the detail). We standardize the range of each descriptor and the range $\{t \in \mathbb{R} \mid \underline{a} \leq t \leq \overline{a}\}$ of property values $a(\mathbb{C}), \mathbb{C} \in D_\pi$.

Stage 3. For each chemical property π, we select a penalty value λ_π for a constant λ in $\mathrm{ALR}(\mathcal{X}, \lambda)$ by conducting linear regression as a preliminary experiment.

We conducted an experiment in Stage 3 to evaluate the performance of the prediction function based on cross-validation. For a property π, an execution of a *cross-validation* consists of five trials of constructing a prediction function as follows. First partition the data set D_π into five subsets $D^{(k)}$, $k \in [1, 5]$ randomly. For each $k \in [1, 5]$, the k-th trial constructs a prediction function $\eta^{(k)}$ by conducting a linear regression with the penalty term λ_π using the set $D_\pi \setminus D^{(k)}$ as a training data set. For each property, we executed ten cross-validations and we show the median of test $\mathrm{R}^2(\eta^{(k)}, D^{(k)}), k \in [1, 5]$ over all ten cross-validations. Recall that a subset of descriptors is selected

in linear regression with ALR and let K' denote the average number of descriptors selected by ALR over all 50 trials.

Table 1: Results in Phase 1 for monomers.

π	Λ	$\lvert D_\pi \rvert$	$\underline{n},\ \overline{n}$	$\underline{a},\ \overline{a}$	$\lvert\Gamma\rvert$	$\lvert\mathcal{F}\rvert$	K	λ_π	K'	L-time	ALR	LLR	ANN
Alpha	Λ_{10}	977	4, 9	50.9, 99.6	59	190	297	6.0e−4	88.4	3.00	0.953	0.961	0.888
Bp	Λ_2	370	4, 67	-11.7, 470.0	22	130	184	5.3e−6	97.3	1.42	0.816	0.599	0.765
Bp	Λ_8	444	4, 67	-11.7, 470.0	26	163	230	2.8e−6	112.6	2.02	0.832	0.663	0.720
Cp	Λ_2	125	4, 63	4.7×10^{-6}, 5.52	8	75	112	5.9e−3	17.4	0.15	0.650	0.445	0.694
Cp	Λ_7	131	4, 63	4.7×10^{-6}, 5.52	8	79	119	8.3e−3	12.3	0.12	0.690	0.555	0.727
Ct	Λ_2	125	4, 63	56.1, 3607.5	8	76	113	3.1e−3	37.5	0.24	0.900	0.037	0.357
Ct	Λ_7	132	4, 63	56.1, 3607.5	8	81	121	2.6e−3	43.0	0.28	0.895	0.048	0.356
Cv	Λ_{10}	977	4, 9	19.2, 44.0	59	190	297	2.1e−4	143.9	4.57	0.966	0.970	0.911
Dc	Λ_8	161	5, 44	0.5, 17.11	25	69	130	2.5e−3	45.3	0.35	0.602	0.574	0.622
EDPA	Λ_1	52	11, 16	0.80, 3.76	9	33	64	2.6e−3	19.0	0.06	0.999	0.999	0.992
Fp	Λ_2	368	4, 67	-82.99, 300.0	20	131	183	6.0e−4	86.6	1.31	0.719	0.589	0.746
Fp	Λ_8	424	4, 67	-82.99, 300.0	25	161	229	2.1e−4	109.4	1.92	0.684	0.571	0.745
Gap	Λ_{10}	977	4, 9	0.15, 0.41	59	190	297	1.6e−4	145.2	4.77	0.755	0.784	0.795
Ha	Λ_3	115	4, 11	1100.6, 3009.6	8	83	115	1.1e−3	56.5	0.29	0.998	0.997	0.926
Hc	Λ_2	255	4, 63	49.6, 35099.6	17	106	154	1.0e−3	69.9	0.74	0.986	0.946	0.848
Hc	Λ_8	282	4, 63	49.6, 35099.6	21	118	177	1.6e−3	69.8	0.84	0.986	0.951	0.903
Hf	Λ_1	82	4, 16	30.2, 94.8	5	50	74	8.4e−4	5.8	0.05	0.982	0.987	0.928
Homo	Λ_{10}	977	4, 9	-0.11, 0.10	59	190	297	1.0e−4	158.7	4.95	0.689	0.841	0.689
Hv	Λ_2	95	4, 16	19.12,5193.1	12	63	105	1.6e−3	40.7	0.19	0.626	-13.7	-8.44
IhcL	Λ_2	770	4, 78	106.3, 1956.1	23	200	256	3.6e−5	126.3	3.24	0.987	0.986	0.974
IhcL	Λ_8	865	4, 78	106.3, 1956.1	29	246	316	6.0e−4	64.2	1.98	0.989	0.985	0.971
IhcS	Λ_6	581	5, 70	67.4, 1220.9	33	124	192	1.1e−5	86.5	1.72	0.971	0.985	0.971
IhcS	Λ_{11}	668	5, 70	67.4, 1220.9	40	140	228	1.0e−6	96.1	2.21	0.974	0.982	0.968
KvI	Λ_1	52	11, 16	1422.0, 1919.0	9	33	64	5.9e−3	18.8	0.05	0.838	0.677	0.727
Kow	Λ_2	684	4, 58	-7.5, 15.6	25	166	223	3.1e−4	119.2	3.13	0.964	0.953	0.952
Kow	Λ_9	899	4, 69	-7.5, 15.6	37	219	303	5.5e−4	141.4	4.95	0.952	0.927	0.937
Lp	Λ_2	615	6, 60	-3.62, 6.84	32	116	186	3.1e−4	85.6	1.81	0.844	0.856	0.867
Lp	Λ_9	936	6, 74	-3.62, 6.84	44	136	231	1.0e−4	108.4	3.37	0.807	0.840	0.859
Lumo	Λ_{10}	977	4, 9	-0.11, 0.10	59	190	297	6.0e−4	81.1	2.75	0.833	0.841	0.860
Mp	Λ_2	467	4, 122	-185.33, 300.0	23	142	197	8.0e−4	95.7	1.78	0.831	0.810	0.799
Mp	Λ_9	577	4, 122	-185.33, 300.0	32	176	255	1.1e−4	136.6	2.99	0.807	0.810	0.820
OptR	Λ_2	147	5, 44	-117.0, 165.0	21	55	107	1.0e−3	30.2	0.24	0.876	0.825	0.919
OptR	Λ_5	157	5, 69	-117.0, 165.0	25	62	123	1.1e−3	32.3	0.27	0.870	0.825	0.878
Rf	Λ_2	166	4, 26	1.3326, 1.613	14	98	142	3.5e−3	25.4	0.24	0.685	0.619	0.521
Sl	Λ_2	673	4, 55	-9.332, 1.11	27	154	217	1.0e−3	45.2	1.21	0.784	0.808	0.848
Sl	Λ_9	915	4, 55	-11.6, 1.11	42	207	300	6.0e−4	73.4	2.33	0.828	0.808	0.861
SfT	Λ_4	247	5, 33	12.3, 45.1	11	91	128	1.0e−3	63.1	0.67	0.847	0.927	0.859
U0	Λ_{10}	977	4, 9	-570.6, -272.84	59	190	297	1.1e−3	69.7	2.40	0.995	0.999	0.890
Vis	Λ_4	282	5, 36	-0.64, 1.63	12	88	126	2.1e−3	23.2	0.37	0.911	0.893	0.929
Vd	Λ_2	474	4, 30	0.7, 20.6	21	160	214	1.0e−4	119.1	2.24	0.985	0.927	0.912
Vd	Λ_{12}	551	4, 30	0.7, 20.6	24	191	256	6.0e−4	101.1	2.28	0.980	0.942	0.889

Tables 1 and 2 show the results on Stages 2 and 3 for the properties on monomers and polymers, respectively, where we denote the following:

- λ_π: the penalty value in the Lasso function selected for a property π, where a e b means $a \times 10^b$;
- K': the average of the number of descriptors selected in the linear regression over all 50 trials in ten cross-validations;
- L-time: the average time (sec.) to construct a prediction function with ALR by solving an LP with $O(|D_\pi| + K)$ variables and constraints over all 50 trials in ten cross-validations;
- ALR: the median of test R^2 over all 50 trials in ten cross-validations for prediction functions constructed with ALR;
- LLR: the median of test R^2 over all 50 trials in ten cross-validations for prediction functions constructed with Lasso linear regression; and
- ANN: the median of test R^2 over all 50 trials in ten cross-validations for prediction functions constructed with ANN (see [26] for the details of the implementation).

Table 2: Results in Phase 1 for polymers.

π	Λ	$\lvert D_\pi\rvert$	$\underline{n}$, $\overline{n}$	$\underline{a}$, $\overline{a}$	$\lvert\Gamma\rvert$	$\lvert\mathcal{F}\rvert$	K	λ_π	K'	L-time	ALR	LLR	ANN
AMD	Λ_2	86	4, 45	0.838, 1.34	28	25	83	1.5E−3	22.2	0.09	0.933	0.914	0.885
AMD	Λ_{15}	93	4, 45	0.838, 1.45	31	30	94	2.5E−3	24.4	0.10	0.917	0.918	0.823
CHAR	Λ_1	24	4, 18	5.5, 13.2	15	15	56	5.0E−5	15.0	0.02	0.904	0.650	0.616
CHAR	Λ_2	27	4, 18	5.5, 13.2	22	17	67	5.9E−3	17.8	0.03	0.835	0.431	0.641
DEC	Λ_{15}	37	4, 22	2.13, 3.4	22	19	72	3.1E−3	18.2	0.04	0.918	0.761	0.641
HCL	Λ_2	52	4, 25	105.7, 677.8	22	17	67	2.1E−3	19.3	0.06	0.996	0.990	0.969
HCL	Λ_8	55	4, 32	105.7, 678.1	27	20	81	2.6E−3	17.3	0.05	0.992	0.987	0.970
HCS	Λ_2	54	4, 45	84.5, 720.5	26	20	75	1.0E−3	23.5	0.07	0.963	0.968	0.893
HCS	Λ_8	59	4, 45	84.5, 720.5	32	24	92	4.1E−4	30.7	0.09	0.983	0.961	0.880
MLV	Λ_2	86	4, 45	60.7, 466.6	28	25	83	1.0E−3	22.2	0.10	0.998	0.996	0.931
MLV	Λ_{15}	93	4, 45	60.7, 466.6	31	30	94	5.6E−4	27.5	0.09	0.997	0.994	0.894
PRM	Λ_2	112	4, 45	2.23, 4.91	25	15	69	2.6E−4	14.5	0.09	0.505	0.801	0.801
PRM	Λ_{14}	131	4, 45	2.23, 4.91	25	17	73	5.9E−3	13.9	0.09	0.489	0.784	0.735
RFID	Λ_{13}	91	4, 29	1.4507, 1.683	26	35	96	3.1E−4	32.9	0.15	0.953	0.852	0.871
RFID	Λ_{16}	124	4, 29	1.339, 1.683	32	50	124	2.1E−3	37.5	0.21	0.956	0.832	0.891
TG	Λ_2	204	4, 58	171, 673	32	36	101	2.5E−3	26.9	0.23	0.923	0.902	0.883
TG	Λ_8	232	4, 58	171, 673	36	43	118	1.1E−3	38.6	0.54	0.927	0.894	0.881

From Tables 1 and 2, we see that ALR performs well for most of the properties in our experiments, The performance by ALR is inferior to that by LLR or ANN for some properities such as GAP, HOMO, LUMO, OPTR, SL, SFT and PRM, whereas ALR outperforms LLR and ANN for properties BP, CT, HV, KVI, VD, CHAR, RFID and TG. It should be noted that ALR drastically improves the result for properties CT and HV.

Results on Phase 2. To execute Stages 4 and 5 in Phase 2, we used a set of seven instances $I_{\rm a}$, $I_{\rm b}^i, i \in [1, 4]$, $I_{\rm c}$ and $I_{\rm d}$ based on the seed graphs prepared by Zhu et al. [20]. We here present their seed graphs $G_{\rm C}$ (see Appendix B for the details of $I_{\rm a}$ and Appendix C for the details of $I_{\rm b}^i, i \in [1, 4]$, $I_{\rm c}$ and $I_{\rm d}$). The seed graph $G_{\rm C}$ of $I_{\rm a}$ is given by the graph in Figure 5(a). The seed graph $G_{\rm C}^1$ of

$I_{\rm b}^1$ (resp., $G_{\rm C}^i, i = 2, 3, 4$ of $I_{\rm b}^i, i = 2, 3, 4$) is illustrated in Figure 6.

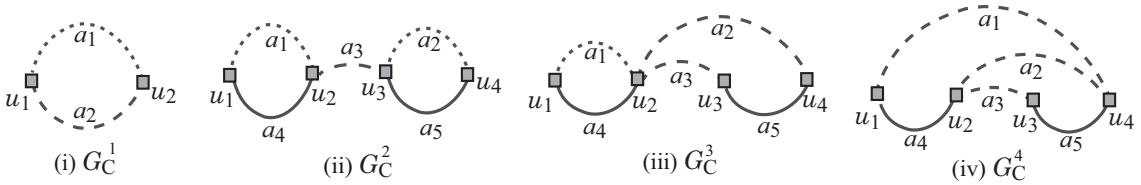


Figure 6: Seed graphs: (i) $G_{\rm C}^1$ for $I_{\rm b}^1$ and $I_{\rm d}$; (ii) $G_{\rm C}^2$ for $I_{\rm b}^2$; (iii) $G_{\rm C}^3$ for $I_{\rm b}^3$; (iv) $G_{\rm C}^4$ for $I_{\rm b}^4$.

Instance $I_{\rm c}$ has been introduced in order to infer a chemical graph $\mathbb{C}^\dagger$ such that the core of $\mathbb{C}^\dagger$ is equal to the core of chemical graph $\mathbb{C}_A$: CID 24822711 in Figure 7(a) and the frequency of each edge-configuration in the non-core of $\mathbb{C}^\dagger$ is equal to that of chemical graph $\mathbb{C}_B$: CID 59170444 in Figure 7(b). This means that the seed graph $G_{\rm C}$ of $I_{\rm c}$ is the core of $\mathbb{C}_A$ which is indicated by a shaded area in Figure 7(a).

Instance $I_{\rm d}$ has been introduced in order to infer a chemical monocyclic graph $\mathbb{C}^\dagger$ such that the frequency vector of edge-configurations in $\mathbb{C}^\dagger$ is a vector obtained by merging those of chemical graphs $\mathbb{C}_A$: CID 10076784 and $\mathbb{C}_B$: CID 44340250 in Figure 7(c) and (d), respectively. The seed graph $G_{\rm C}$ of $I_{\rm d}$ is given by $G_{\rm C}^1$ in Figure 6(i).

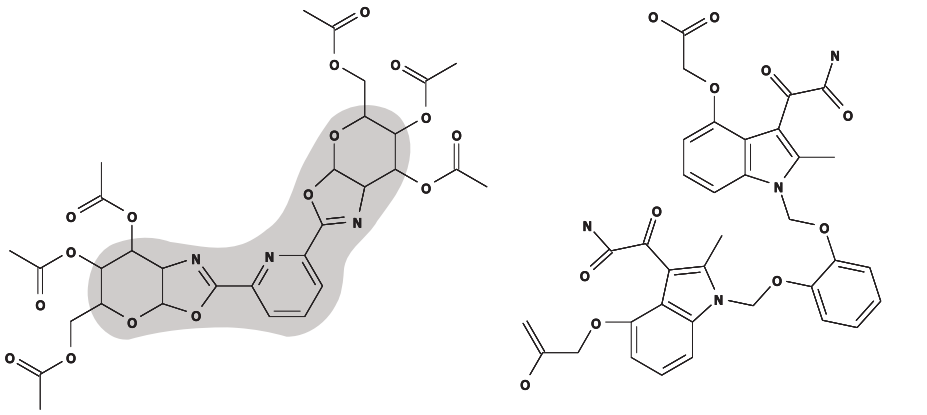
Stage 4. We executed Stage 4 for two properties $\pi \in \{\text{Ct}, \text{Hv}\}$.

For the MILP formulation $\mathcal{M}(x, y; \mathcal{C}_1)$ in the framework, we use the prediction function $\eta_{w,b}$ constructed in Stage 3. Tables 3 and 4 show the computational results of the experiment in Stage 4 for the two properties, where we denote the following:

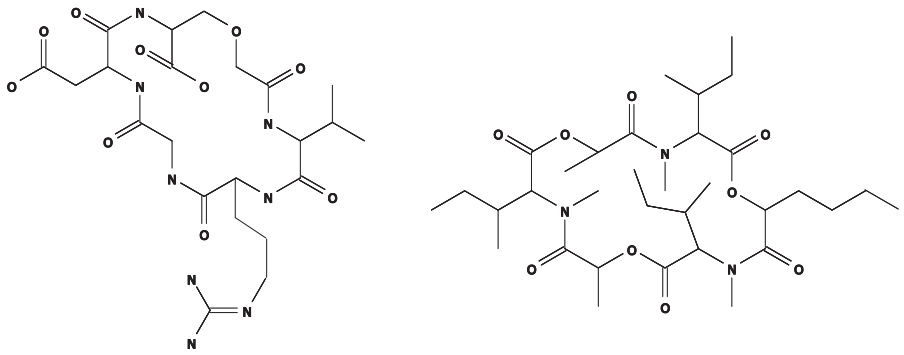
- $\underline{y}^*$, $\overline{y}^*$: lower and upper bounds $\underline{y}^*, \overline{y}^* \in \mathbb{R}$ on the value $a(\mathbb{C})$ of a chemical graph $\mathbb{C}$ to be inferred;
- #v (resp., #c): the number of variables (resp., constraints) in the MILP in Stage 4;
- I-time: the time (sec.) to solve the MILP in Stage 4;
- n: the number $n(\mathbb{C}^\dagger)$ of non-hydrogen atoms in the chemical graph $\mathbb{C}^\dagger$ inferred in Stage 4;
- $\mathrm{n}^{\rm int}$: the number $\mathrm{n}^{\rm int}(\mathbb{C}^\dagger)$ of interior-vertices in the chemical graph $\mathbb{C}^\dagger$ inferred in Stage 4; and
- η: the predicted property value $\eta(f(\mathbb{C}^\dagger))$ of the chemical graph $\mathbb{C}^\dagger$ inferred in Stage 4.

From Tables 3 and 4, we observe that an instance with around 50 non-hydrogen atoms is solved in at most around 1000 seconds. Instances with the size in Stage 4 were solved at most around 100 seconds in the experiments due to Azam et al. [26]. The computation time for Stage 4 increased possibly because we included a new set of constraints for representing activation functions ϕ_j in the MILP $\mathcal{M}(x, y; \mathcal{C}_1)$ of the framework (see Appendix D.11 for the details). Note that for the properties $\pi \in \{\text{Ct}, \text{Hv}\}$, no prediction functions constructed for the framework [20, 26] performed well.

Stage 5. We executed Stage 5 to generate a more number of target chemical graphs $\mathbb{C}^*$, where we call a chemical graph $\mathbb{C}^*$ a *chemical isomer* of a target chemical graph $\mathbb{C}^\dagger$ of a topological specification σ if $f(\mathbb{C}^*) = f(\mathbb{C}^\dagger)$ and $\mathbb{C}^*$ also satisfies the same topological specification σ. We computed chemical isomers $\mathbb{C}^*$ of each target chemical graph $\mathbb{C}^\dagger$ inferred in Stage 4. We execute an algorithm for generating chemical isomers of $\mathbb{C}^\dagger$ up to 100 when the number of all chemical



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



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

Figure 7: Chemical compounds: (a) CID 24822711; (b) CID 59170444; (c) CID 10076784; (d) CID 44340250, where hydrogens are omitted.

Table 3: Results of Stages 4 and 5 for Cᴛ.

inst.	$\underline{y}^*,\ \overline{y}^*$	#v	#c	I-time	n	$\mathrm{n}^{\mathrm{int}}$	η	D-time	$\mathbb{C}$-LB	#$\mathbb{C}$
I_{a}	1180, 1220	16190	16895	8.4	42	25	1197.0	0.0619	1	1
I_{b}^1	1080, 1120	86470	85993	119.2	35	14	1118.0	0.266	84	84
I_{b}^2	1980, 2020	113055	114138	365.5	45	30	1986.9	2.62	4.4×10^6	100
I_{b}^3	1930, 1970	112745	114086	555.4	45	25	1950.9	4.57	1.1×10^6	100
I_{b}^4	1630, 1670	112455	114054	661.2	47	29	1663.8	0.171	5.5×10^5	100
I_{c}	1980, 2020	14794	15226	7.2	50	34	2011.0	0.0161	1	1
I_{d}	1430, 1470	12790	14104	11.9	45	23	1447.5	0.164	5184	100

isomers exceeds 100. Such an algorithm can be obtained from the dynamic programming proposed by Tanaka et al. [19] with a slight modification. The algorithm first decomposes $\mathbb{C}^\dagger$ into a set of

Table 4: Results of Stages 4 and 5 for Hv.

inst.	$\underline{y}^*,\ \overline{y}^*$	#v	#c	I-time	n	$n^{\rm int}$	η	D-time	$\mathbb{C}$-LB	#$\mathbb{C}$
$I_{\rm a}$	145, 150	16190	16879	24.9	37	23	147.3	0.0632	2	2
$I_{\rm b}^1$	190, 195	14179	13334	146.6	35	12	190.2	0.121	30	30
$I_{\rm b}^2$	290, 295	17986	18547	188.8	46	25	294.2	0.154	604	100
$I_{\rm b}^3$	165, 170	17683	18495	1167.2	45	25	166.8	36.8	7.5×10^6	100
$I_{\rm b}^4$	250, 255	17400	18463	313.7	50	25	251.8	0.166	2208	100
$I_{\rm c}$	285, 290	14838	15254	102.5	50	32	286.8	0.016	1	1
$I_{\rm d}$	245, 250	12828	14126	351.9	40	23	249.2	5.53	3.9×10^5	100

acyclic chemical graphs, next replace each acyclic chemical graph T with another acyclic chemical graph T' that admits the same feature vector as that of T and finally assemble the resulting acyclic chemical graphs into a chemical isomer $\mathbb{C}^*$ of $\mathbb{C}^\dagger$. The algorithm can compute a lower bound on the total number of all chemical isomers $\mathbb{C}^\dagger$ without generating all of them.

Tables 3 and 4 show the computational results of the experiment in Stage 5 for properties Ct and Hv, where we denote the following:

- D-time: the running time (sec.) to execute the dynamic programming algorithm in Stage 5 to compute a lower bound on the number of all chemical isomers $\mathbb{C}^*$ of $\mathbb{C}^\dagger$ and generate all (or up to 100) chemical isomers $\mathbb{C}^*$;
- $\mathbb{C}$-LB: a lower bound on the number of all chemical isomers $\mathbb{C}^*$ of $\mathbb{C}^\dagger$; and
- #$\mathbb{C}$: the number of all (or up to 100) chemical isomers $\mathbb{C}^*$ of $\mathbb{C}^\dagger$ generated in Stage 5.

From Tables 3 and 4, we observe that the running time for generating up to 100 target chemical graphs in Stage 5 is less than around 5 second for many cases. For some chemical graph $\mathbb{C}^\dagger$, no chemical isomer was found by our algorithm. For such a case, we may use a method of generating other chemical graphs $\mathbb{C}^\dagger$ in Stage 4 by solving the MILP again with additional linear constraints in a systematical way (see [26] for the details).

6 Concluding Remarks

In this paper, we proposed a new machine learning method, adjustive linear regression (ALR), which has the following feature: (i) ALR is an extension of the Lasso linear regression except for the definition of error functions; (ii) ALR is a special case of an ANN except that a choice of activation functions is also optimized differently from the standard ANNs and the definition of error functions; and (iii) ALR can be executed exactly by solving the equivalent linear program with $O(m+K)$ variables and constraints for a set of m data with K descriptors. Even though ALR is a special case of an ANN with non-linear activation functions, we still can read the relationship between cause and effect from a prediction function due to the simple structure of ALR.

When the size of the original data set is large, we can choose a smaller subset for constructing a prediction function. For example, we chose 1,000 compounds from the data set of 130,000

compounds provided by Ramakrishnan et al. [30] for conducting experiments in Stage 3.

In this paper, we used a quadratic function for a set Ψ of activation functions ϕ. We can use many different functions such as sigmoid function and ramp functions, where the non-linearity of a function does not affect to derive a linear program for ALR.

References

[1] Lo, Y-C., Rensi, S.E., Torng, W., Altman, R.B.: Machine learning in chemoinformatics and drug discovery. Drug Discovery Today **23**, 1538–1546 (2018)

[2] Tetko, I.V., Engkvist, O.: From Big Data to Artificial Intelligence: chemoinformatics meets new challenges. J. Cheminformatics **12**, 74 (2020)

[3] Ghasemi, F., Mehridehnavi, A., Pérez-Garrido, A., Pérez-Sánchez, H.: Neural network and deep-learning algorithms used in QSAR studies: merits and drawbacks. Drug Discovery Today **23**, 1784–1790 (2018)

[4] Miyao, T., Kaneko, H., Funatsu, K.: Inverse QSPR/QSAR analysis for chemical structure generation (from y to x). J. Chem. Inf. Model. **56**, 286–299 (2016)

[5] Ikebata, H., Hongo, K., Isomura, T., Maezono, R., Yoshida, R.: Bayesian molecular design with a chemical language model. J. Comput. Aided Mol. Des. **31**, 379–391 (2017)

[6] Rupakheti, C., Virshup, A., Yang, W., Beratan, D.N.: Strategy to discover diverse optimal molecules in the small molecule universe. J. Chem. Inf. Model. **55**, 529–537 (2015)

[7] Bohacek, R.S., McMartin, C., Guida, W.C.: The art and practice of structure-based drug design: A molecular modeling perspective. Med. Res. Rev. **16**, 3–50 (1996)

[8] Kipf, T. N., Welling, M.: Semi-supervised classification with graph convolutional networks, arXiv:1609.02907 (2016)

[9] Gómez-Bombarelli, R., Wei, J.N., Duvenaud, D., Hernández-Lobato, J.M., Sánchez-Lengeling, B., Sheberla, D., Aguilera-Iparraguirre, J., Hirzel, T.D., Adams, R.P., Aspuru-Guzik, A.: Automatic chemical design using a data-driven continuous representation of molecules. ACS Cent. Sci. **4**, 268–276 (2018)

[10] Segler, M.H.S., Kogej, T., Tyrchan, C., Waller, M.P.: Generating focused molecule libraries for drug discovery with recurrent neural networks. ACS Cent. Sci. **4**, 120–131 (2017)

[11] Yang, X., Zhang, J., Yoshizoe, K., Terayama, K., Tsuda, K.: ChemTS: an efficient python library for de novo molecular generation. STAM **18**, 972–976 (2017)

[12] Kusner, M.J., Paige, B., Hernández-Lobato, J.M.: Grammar variational autoencoder. Proc. of the 34th International Conference on Machine Learning-Volume 70, 1945–1954 (2017)

[13] De Cao, N., Kipf, T.: MolGAN: An implicit generative model for small molecular graphs. arXiv:1805.11973 (2018)

[14] Madhawa, K., Ishiguro, K., Nakago, K., Abe, M.: GraphNVP: an invertible flow model for generating molecular graphs. arXiv:1905.11600 (2019)

[15] Shi, C., Xu, M., Zhu, Z., Zhang, W., Zhang, M., Tang, J.: GraphAF: a flow-based autoregressive model for molecular graph generation. arXiv:2001.09382 (2020)

[16] Akutsu, T., Nagamochi, H.: A mixed integer linear programming formulation to artificial neural networks. Proc. of the 2nd Int. Conf. on Information Science and Systems, 215–220 (2019)

[17] Azam, N. A., Chiewvanichakorn, R., Zhang, F., Shurbevski, A., Nagamochi, H., Akutsu, T.: A method for the inverse QSAR/QSPR based on artificial neural networks and mixed integer linear programming. Proc. of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies – Volume 3: BIOINFORMATICS, 101–108 (2020)

[18] Shi, Y., Zhu, J., Azam, N. A., Haraguchi, K., Zhao, L., Nagamochi, H., Akutsu, T.: An inverse QSAR method based on a two-layered model and integer programming. International Journal of Molecular Sciences. **22**, 2847 (2021)

[19] Tanaka, K., Zhu, J., Azam, N. A., Haraguchi, K., Zhao, L., Nagamochi, H., Akutsu, T.: An inverse QSAR method based on decision tree and integer programming, The 17th International Conference on Intelligent Computing, August 12-15, 2021, in Shenzhen, China, In: Huang D.S., Jo K.H., Li J., Gribova V., Hussain A. (eds) Intelligent Computing Theories and Application, ICIC 2021, Lecture Notes in Computer Science, vol. 12837. Springer, Cham.

[20] Zhu, J., Azam, N. A., Haraguchi, K., Zhao, L., Nagamochi, H., Akutsu, T.: A method for molecular design based on linear regression and integer programming. arXiv: 2107.02381 (2021) http://arxiv.org/abs/2107.02381

[21] Hoerl, A., Kennard, R.: Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics, **12**(1), 55–67 (1970)

[22] Hoerl, A., Kennard, R.: Ridge Regression: Applications to Nonorthogonal Problems. Technometrics, **12**(1), 69–82 (1970)

[23] Tibshirani, R.: Regression shrinkage and selection via the lasso. J. R. Statist. Soc. B **58**, 267–288 (1996)

[24] Zou, H., and Hastie, T.: Regularization and variable selection via the elastic net. J. Royal Statistical Society: Series B (Statistical Methodology), **67**(2), 301-320 (2005)

[25] Ido, R., Cao, S., Zhu, J., Azam, N. A., Haraguchi, K., Zhao, L., Nagamochi, H., Akutsu, T.: A method for inferring polymers based on linear regression and integer programming. Department of Applied Mathematics and Physics, Kyoto University, Technical Report, TR: 2021-001 `http://www.amp.i.kyoto-u.ac.jp/tecrep/` (2021)

[26] Azam, N. A., Zhu, J., Haraguchi, K., Zhao, L., Nagamochi, H., Akutsu, T.: Molecular design based on artificial neural networks, integer programming and grid neighbor search. arXiv: 2108.10266 (2021)

[27] Annotations from HSDB (on pubchem): https://pubchem.ncbi.nlm.nih.gov/. Accessed: 2021-8-26.

[28] Jalali-Heravi, M., Fatemi, M. : H.Artificial neural network modeling of Kovats retention indices for noncyclic and monocyclic terpenes (2001) https://doi.org/10.1016/S0021-9673(00)01274-7/

[29] Roy, K., Saha, A.: Comparative QSPR studies with molecular connectivity, molecular negentropy and TAU indices (2003) https://doi.org/10.1007/s00894-003-0135-z/

[30] Ramakrishnan, R., Dral, P., Rupp, M., Anatole von Lilienfeld, O.: Quantum chemistry structures and properties of 134 kilo molecules (2014) https://doi.org/10.6084/m9.figshare.c.978904.v5. Accessed: 2021-8-26.

[31] Goussard, V., Duprat, F., Gerbaud, V., Ploix, J.-J., Dreyfus, G., Nardello-Rataj, V., Aubry, J.-M.: Predicting the surface tension of liquids: comparison of four modeling approaches and application to cosmetic oils, J. Chem. Inf. Model., 57, 12, 29862995 (2017) https://pubs.acs.org/doi/full/10.1021/acs.jClm.7b00512

[32] Goussard, V., Franois Duprat F., Ploix, J.-L., Dreyfus, G., Nardello-Rataj, V., Aubry, J.-M.: A new machine-learning tool for fast estimation of liquid viscosity. application to cosmetic oils, J. Chem. Inf. Model., 60, 4, 20122023 (2020) https://pubs.acs.org/doi/10.1021/acs.jcim.0c00083

[33] Naef, R.: Calculation of the isobaric heat capacities of the liquid and solid phase of organic compounds at and around 298.15 K based on their “true” molecular volume. Molecules, 24 (8) (2019), https://www.mdpi.com/1420-3049/24/8/1626/

[34] Xiao, N.: Lipophilicity Dataset - logD7.4 of 1,130 Compounds (2017) https://doi.org/10.6084/m9.figshare.5596750. Accessed: 2021-8-26.

[35] Delaney, J.S.: ESOL: Estimating Aqueous Solubility Directly from Molecular Structure (2019) https://figshare.com/articles/dataset/ESOL_Estimating_Aqueous_Solubility_Directly_from_Molecular_Structure/7944677/1. Accessed: 2021-8-26.

[36] Bicerano, J.: Prediction of Polymer Properties. 3rd Edition, Revised and Expanded. CRC Press (2002)

Appendix

A A Full Description of Descriptors

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

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

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

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

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

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

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

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

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

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

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

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

We assume that a chemical graph $\mathbb{C}$ treated in this paper satisfies $\deg_{\langle\mathbb{C}\rangle}(v) \leq 4$ in the hydrogen-suppressed graph $\langle\mathbb{C}\rangle$.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

B Specifying Target Chemical Graphs

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

We first describe how to reduce a chemical graph $\mathbb{C} = (H, \alpha, \beta)$ into an abstract form based on which our specification rules will be defined. To illustrate the reduction process, we use the chemical graph $\mathbb{C} = (H, \alpha, \beta)$ such that $\langle \mathbb{C} \rangle$ is given in Figure 3.

R1 **Removal of all ρ-fringe-trees:** The interior $H^{\rm int} = (V^{\rm int}(\mathbb{C}), E^{\rm int}(\mathbb{C}))$ of $\mathbb{C}$ is obtained by removing the non-root vertices of each ρ-fringe-trees $\mathbb{C}[u] \in \mathcal{T}(\mathbb{C}), u \in V^{\rm int}(\mathbb{C})$. Figure 8 illustrates the interior $H^{\rm int}$ of chemical graph $\mathbb{C}$ with $\rho = 2$ in Figure 3.

R2 **Removal of some leaf paths:** We call a u, v-path Q in $H^{\rm int}$ a *leaf path* if vertex v is a leaf-vertex of $H^{\rm int}$ and the degree of each internal vertex of Q in $H^{\rm int}$ is 2, where we regard that Q is rooted at vertex u. A connected subgraph S of the interior $H^{\rm int}$ of $\mathbb{C}$ is called a *cyclical-base* if S is obtained from H by removing the vertices in $V(Q_u) \setminus \{u\}, u \in X$ for a subset X of interior-vertices and a set $\{Q_u \mid u \in X\}$ of leaf u, v-paths Q_u such that no two paths Q_u and $Q_{u'}$ share a vertex. Figure 9(a) illustrates a cyclical-base $S = H^{\rm int} - \bigcup_{u \in X}(V(Q_u) \setminus \{u\})$ of the interior $H^{\rm int}$ for a set $\{Q_{u_5} = (u_5, u_{24}), Q_{u_{18}} = (u_{18}, u_{25}, u_{26}, u_{27}), Q_{u_{22}} = (u_{22}, u_{28})\}$ of leaf paths in Figure 8.

R3 **Contraction of some pure paths:** A path in S is called *pure* if each internal vertex of the path is of degree 2. Choose a set $\mathcal{P}$ of several pure paths in S so that no two paths share vertices except for their end-vertices. A graph S' is called a *contraction* of a graph S (with respect to $\mathcal{P}$) if S' is obtained from S by replacing each pure u, v-path with a single edge $a = uv$, where S' may contain multiple edges between the same pair of adjacent vertices. Figure 9(b) illustrates a contraction S' obtained from the chemical graph S by contracting each uv-path $P_a \in \mathcal{P}$ into a new edge $a = uv$, where $a_1 = u_1u_2, a_2 = u_1u_3, a_3 = u_4u_7, a_4 = u_{10}u_{11}$ and $a_5 = u_{11}u_{12}$ and $\mathcal{P} = \{P_{a_1} = (u_1, u_{13}, u_2), P_{a_2} = (u_1, u_{14}, u_3), P_{a_3} = (u_4, u_{15}, u_{16}, u_7), P_{a_4} = (u_{10}, u_{17}, u_{18}, u_{19}, u_{11}), P_{a_5} = (u_{11}, u_{20}, u_{21}, u_{22}, u_{12})\}$ of pure paths in Figure 9(a).

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

Seed Graph

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

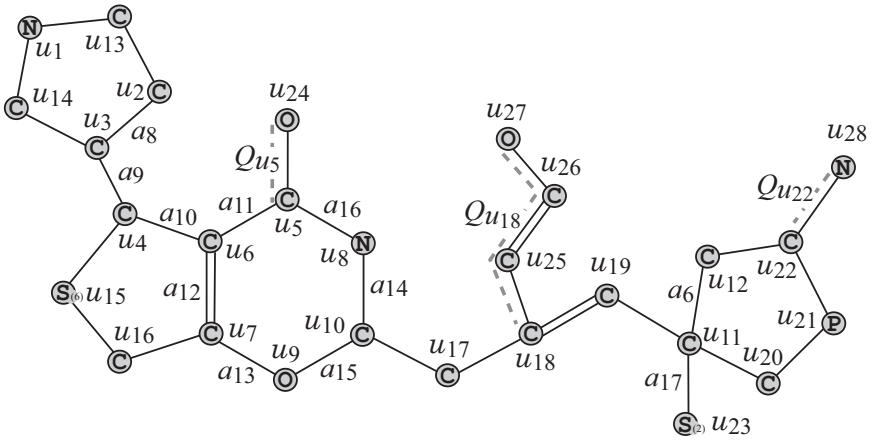


Figure 8: The interior H^{int} of chemical graph $\mathbb{C}$ with $\langle\mathbb{C}\rangle$ in Figure 3 for $\rho = 2$.

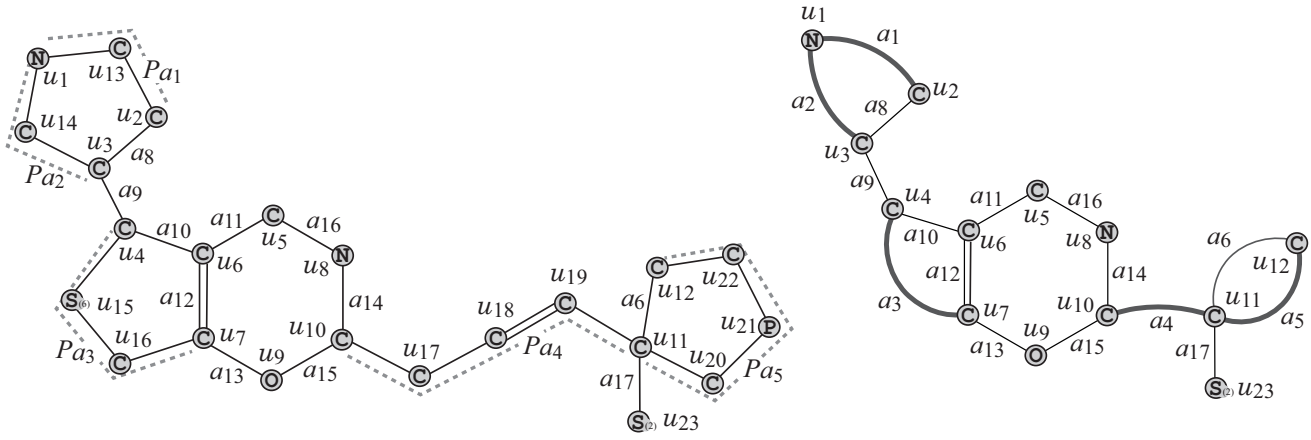


Figure 9: (a) A cyclical-base $S = H^{\mathrm{int}} - \bigcup_{u\in\{u_5,u_{18},u_{22}\}}(V(Q_u)\setminus\{u\})$ of the interior H^{int} in Figure 8; (b) A contraction S' of S for a pure path set $\mathcal{P} = \{P_{a_1}, P_{a_2}, \ldots, P_{a_5}\}$ in (a), where a new edge obtained by contracting a pure path is depicted with a thick line.

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

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

- Each edge $e \in E_{(=1)}$ is always used directly.

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

Interior-specification

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

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

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

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

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

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

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

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

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

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

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

Table 5: Example 1 of an interior-specification σ_{int}.

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

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

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

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

Table 5 shows an example of an interior-specification σ_{int} to the seed graph G_{C} in Figure 5.

Figure 10 illustrates an example of an σ_{int}-extension H^* of seed graph G_{C} in Figure 5 under the interior-specification σ_{int} in Table 5.

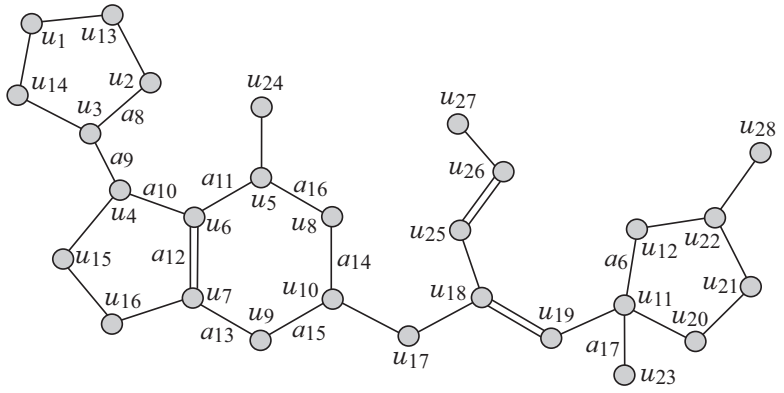


Figure 10: An illustration of a graph H^* that is obtained from the seed graph G_{C} in Figure 5 under the interior-specification σ_{int} in Table 5, where the vertices newly introduced by pure paths P_{a_i} and leaf paths Q_{v_i} are depicted with white squares and circles, respectively.

Chemical-specification

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Table 6 shows an example of a chemical-specification σ_{ce} to the seed graph G_{C} in Figure 5.

Table 6: Example 2 of a chemical-specification σ_{ce}.

$n_{\mathrm{LB}} = 30$, $n^* = 50$.
branch-parameter: $\rho = 2$

Each of sets $\mathcal{F}(v), v \in V_{\mathrm{C}}$ and $\mathcal{F}_E$ is set to be the set $\mathcal{F}$ of chemical rooted trees ψ with $\mathrm{ht}(\langle\psi\rangle) \leq \rho = 2$ in Figure 5(b).

$\Lambda = \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}, \mathtt{S}_{(2)}, \mathtt{S}_{(6)}, \mathtt{P} = \mathtt{P}_{(5)}\}$	$\Lambda_{\mathrm{dg}} = \{\mathtt{C2}, \mathtt{C3}, \mathtt{C4}, \mathtt{N2}, \mathtt{N3}, \mathtt{O2}, \mathtt{S}_{(2)}\mathtt{2}, \mathtt{S}_{(6)}\mathtt{3}, \mathtt{P4}\}$

$\Gamma^{\mathrm{int}}_{\mathrm{ac}}$	$\nu_1 = (\mathtt{C}, \mathtt{C}, 1), \nu_2 = (\mathtt{C}, \mathtt{C}, 2), \nu_3 = (\mathtt{C}, \mathtt{N}, 1), \nu_4 = (\mathtt{C}, \mathtt{O}, 1), \nu_5 = (\mathtt{C}, \mathtt{S}_{(2)}, 1), \nu_6 = (\mathtt{C}, \mathtt{S}_{(6)}, 1), \nu_7 = (\mathtt{C}, \mathtt{P}, 1)$
Γ^{int}	$\gamma_1 = (\mathtt{C2}, \mathtt{C2}, 1), \gamma_2 = (\mathtt{C2}, \mathtt{C3}, 1), \gamma_3 = (\mathtt{C2}, \mathtt{C3}, 2), \gamma_4 = (\mathtt{C2}, \mathtt{C4}, 1), \gamma_5 = (\mathtt{C3}, \mathtt{C3}, 1), \gamma_6 = (\mathtt{C3}, \mathtt{C3}, 2),$ $\gamma_7 = (\mathtt{C3}, \mathtt{C4}, 1), \gamma_8 = (\mathtt{C2}, \mathtt{N2}, 1), \gamma_9 = (\mathtt{C3}, \mathtt{N2}, 1), \gamma_{10} = (\mathtt{C3}, \mathtt{O2}, 1), \gamma_{11} = (\mathtt{C2}, \mathtt{C2}, 2), \gamma_{12} = (\mathtt{C2}, \mathtt{O2}, 1),$ $\gamma_{13} = (\mathtt{C3}, \mathtt{N3}, 1), \gamma_{14} = (\mathtt{C4}, \mathtt{S}_{(2)}\mathtt{2}, 2), \gamma_{15} = (\mathtt{C2}, \mathtt{S}_{(6)}\mathtt{3}, 1), \gamma_{16} = (\mathtt{C3}, \mathtt{S}_{(6)}\mathtt{3}, 1), \gamma_{17} = (\mathtt{C2}, \mathtt{P4}, 2),$ $\gamma_{18} = (\mathtt{C3}, \mathtt{P4}, 1)$

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

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

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

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

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

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

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

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

Figure 3 illustrates an example $\mathbb{C}$ of a $(\sigma_{\mathrm{int}}, \sigma_{\mathrm{ce}})$-extension of G_{C} obtained from the σ_{int}-extension H^* in Figure 10 under the chemical-specification σ_{ce} in Table 6. Note that $\mathrm{r}(\mathbb{C}) = \mathrm{r}(H^*) = \mathrm{r}(G_{\mathrm{C}}) - 1 = 4$ holds since the edge in $E_{(0/1)}$ is discarded in H^*.

C Test Instances for Stages 4 and 5

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

In Stages 4 and 5, we use five properties $\pi \in \{\text{Hc}, \text{Vd}, \text{OptR}, \text{IhcLiq}, \text{Vis}\}$ and define a set $\Lambda(\pi)$ of chemical elements as follows:

$\Lambda(\text{Hc}) = \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}, \mathtt{S}_{(2)}, \mathtt{S}_{(6)}, \mathtt{Cl}\}$, $\Lambda(\text{Vd}) = \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}, \mathtt{N}, \mathtt{Cl}, \mathtt{P}_{(3)}, \mathtt{P}_{(5)}\}$,
$\Lambda(\text{OptR}) = \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}, \mathtt{S}_{(2)}, \mathtt{F}\}$, $\Lambda(\text{IhcLiq}) = \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}, \mathtt{S}_{(2)}, \mathtt{S}_{(6)}, \mathtt{Cl}\}$ and
$\Lambda(\text{Vis}) = \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{Si}\}$.

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

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

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

The lower bounds ℓ_{LB}, bl_{LB}, ch_{LB}, $\text{bd}_{2,\text{LB}}$, $\text{bd}_{3,\text{LB}}$, na_{LB}, $\text{na}^{\text{int}}_{\text{LB}}$, $\text{ns}^{\text{int}}_{\text{LB}}$, $\text{ac}^{\text{int}}_{\text{LB}}$, $\text{ec}^{\text{int}}_{\text{LB}}$ and $\text{ac}^{\text{lf}}_{\text{LB}}$ are all set to be 0.

Set upper bounds $\text{na}_{\text{UB}}(\mathtt{a}) := n^*, \text{na} \in \{\mathtt{H}, \mathtt{C}\}$, $\text{na}_{\text{UB}}(\mathtt{a}) := 5, \text{na} \in \{\mathtt{O}, \mathtt{N}\}$, $\text{na}_{\text{UB}}(\mathtt{a}) := 2, \text{na} \in \Lambda \setminus \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{N}\}$. The other upper bounds ℓ_{UB}, bl_{UB}, ch_{UB}, $\text{bd}_{2,\text{UB}}$, $\text{bd}_{3,\text{UB}}$, $\text{na}^{\text{int}}_{\text{UB}}$, $\text{ns}^{\text{int}}_{\text{UB}}$, $\text{ac}^{\text{int}}_{\text{UB}}$, $\text{ec}^{\text{int}}_{\text{UB}}$ and $\text{ac}^{\text{lf}}_{\text{UB}}$ are all set to be an upper bound n^* on $n(G^*)$.

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

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

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

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

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

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

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

(c) $I_{\rm c} = (G_{\rm C}, \sigma_{\rm int}, \sigma_{\rm ce})$: An instance aimed to infer a chemical graph $G^\dagger$ such that the core of $G^\dagger$ is equal to the core of G_A and the frequency of each edge-configuration in the non-core of $G^\dagger$ is equal to that of G_B. We use chemical compounds CID 24822711 and CID 59170444 in Figure 7(a) and (b) for G_A and G_B, respectively.
Set a seed graph $G_{\rm C} = (V_{\rm C}, E_{\rm C} = E_{(=1)})$ to be the core of G_A.
Set $\Lambda := \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}\}$, and set $\Lambda_{\rm dg}^{\rm int}$ to be the set of all possible chemical symbols in $\Lambda \times [1, 4]$.
Set $\Gamma^{\rm int} := \Gamma_A^{\rm int} \cup \Gamma_B^{\rm int}$ and $\Lambda^*(v) := \{\alpha_A(v)\}$, $v \in V_{\rm C}$.
Set $\mathrm{n}_{\rm LB}^{\rm int} := \min\{\mathrm{n}^{\rm int}(G_A), \mathrm{n}^{\rm int}(G_B)\}$, $\mathrm{n}_{\rm UB}^{\rm int} := \max\{\mathrm{n}^{\rm int}(G_A), \mathrm{n}^{\rm int}(G_B)\}$,
$n_{\rm LB} := \min\{n(G_A), n(G_B)\} - 10$ and $n^* := \max\{n(G_A), n(G_B)\} + 5$.
Set lower bounds $\ell_{\rm LB}$, $\mathrm{bl}_{\rm LB}$, $\mathrm{ch}_{\rm LB}$, $\mathrm{bd}_{2,\rm LB}$, $\mathrm{bd}_{3,\rm LB}$, $\mathrm{na}_{\rm LB}$, $\mathrm{na}_{\rm LB}^{\rm int}$, $\mathrm{ns}_{\rm LB}^{\rm int}$, $\mathrm{ac}_{\rm LB}^{\rm int}$ and $\mathrm{ac}_{\rm LB}^{\rm lf}$ to be 0.
Set upper bounds $\mathrm{na}_{\rm UB}(\mathtt{a}) := n^*, \mathrm{na} \in \{\mathtt{H}, \mathtt{C}\}$, $\mathrm{na}_{\rm UB}(\mathtt{a}) := 5, \mathrm{na} \in \{\mathtt{O}, \mathtt{N}\}$, $\mathrm{na}_{\rm UB}(\mathtt{a}) := 2, \mathrm{na} \in \Lambda \setminus \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{N}\}$ and set the other upper bounds $\ell_{\rm UB}$, $\mathrm{bl}_{\rm UB}$, $\mathrm{ch}_{\rm UB}$, $\mathrm{bd}_{2,\rm UB}$, $\mathrm{bd}_{3,\rm UB}$, $\mathrm{na}_{\rm UB}^{\rm int}$, $\mathrm{ns}_{\rm UB}^{\rm int}$, $\mathrm{ac}_{\rm UB}^{\rm int}$ and $\mathrm{ac}_{\rm UB}^{\rm lf}$ to be n^*.
Set $\mathrm{ec}_{\rm LB}^{\rm int}(\gamma)$ to be the number of core-edges in G_A with $\gamma \in \Gamma^{\rm int}$ and $\mathrm{ec}_{\rm UB}^{\rm int}(\gamma)$ to be the number interior-edges in G_A and G_B with edge-configuration γ.
Let $\mathcal{F}_B^{(p)}, p \in [1, 2]$ denote the set of chemical rooted trees r-isomorphic p-fringe-trees in G_B;
Set $\mathcal{F}_E := \mathcal{F}(v) := \mathcal{F}_B^{(1)} \cup \mathcal{F}_B^{(2)}$, $v \in V_{\rm C}$ and $\mathrm{fc}_{\rm LB}(\psi) := 0, \mathrm{fc}_{\rm UB}(\psi) := 10, \psi \in \mathcal{F}_B^{(1)} \cup \mathcal{F}_B^{(2)}$.

(d) $I_{\rm d} = (G_{\rm C}^1, \sigma_{\rm int}, \sigma_{\rm ce})$: An instance aimed to infer a chemical monocyclic graph $G^\dagger$ such that the frequency vector of edge-configurations in $G^\dagger$ is a vector obtained by merging those of G_A and G_B. We use chemical monocyclic compounds CID 10076784 and CID 44340250 in Figure 7(c) and (d) for G_A and G_B, respectively. Set a seed graph to be the monocyclic seed graph $G_{\rm C}^1 = (V_{\rm C}, E_{\rm C} = E_{(\geq 2)} \cup E_{(\geq 1)})$ with $V_{\rm C} = \{u_1, u_2\}$, $E_{(\geq 2)} = \{a_1\}$ and $E_{(\geq 1)} = \{a_2\}$ in Figure 6(i).
Set $\Lambda := \{\mathtt{H}, \mathtt{C}, \mathtt{N}, \mathtt{O}\}$, $\Lambda_{\rm dg}^{\rm int} := \Lambda_{{\rm dg},A}^{\rm int} \cup \Lambda_{{\rm dg},B}^{\rm int}$ and $\Gamma^{\rm int} := \Gamma_A^{\rm int} \cup \Gamma_B^{\rm int}$.
Set $\mathrm{n}_{\rm LB}^{\rm int} := \min\{\mathrm{n}^{\rm int}(G_A), \mathrm{n}^{\rm int}(G_B)\}$, $\mathrm{n}_{\rm UB}^{\rm int} := \max\{\mathrm{n}^{\rm int}(G_A), \mathrm{n}^{\rm int}(G_B)\}$,
$n_{\rm LB} := \min\{n(G_A), n(G_B)\}$ and $n^* := \max\{n(G_A), n(G_B)\}$.

Set lower bounds ℓ_{LB}, $\mathrm{bl}_{\mathrm{LB}}$, $\mathrm{ch}_{\mathrm{LB}}$, $\mathrm{bd}_{2,\mathrm{LB}}$, $\mathrm{bd}_{3,\mathrm{LB}}$, $\mathrm{na}_{\mathrm{LB}}$, $\mathrm{na}^{\mathrm{int}}_{\mathrm{LB}}$, $\mathrm{ns}^{\mathrm{int}}_{\mathrm{LB}}$, $\mathrm{ac}^{\mathrm{int}}_{\mathrm{LB}}$ and $\mathrm{ac}^{\mathrm{lf}}_{\mathrm{LB}}$ to be 0.
Set upper bounds $\mathrm{na}_{\mathrm{UB}}(\mathtt{a}) := n^*, \mathrm{na} \in \{\mathtt{H}, \mathtt{C}\}$, $\mathrm{na}_{\mathrm{UB}}(\mathtt{a}) := 5, \mathrm{na} \in \{\mathtt{O}, \mathtt{N}\}$, $\mathrm{na}_{\mathrm{UB}}(\mathtt{a}) := 2, \mathrm{na} \in \Lambda \setminus \{\mathtt{H}, \mathtt{C}, \mathtt{O}, \mathtt{N}\}$ and set the other upper bounds ℓ_{UB}, $\mathrm{bl}_{\mathrm{UB}}$, $\mathrm{ch}_{\mathrm{UB}}$, $\mathrm{bd}_{2,\mathrm{UB}}$, $\mathrm{bd}_{3,\mathrm{UB}}$, $\mathrm{na}^{\mathrm{int}}_{\mathrm{UB}}$, $\mathrm{ns}^{\mathrm{int}}_{\mathrm{UB}}$, $\mathrm{ac}^{\mathrm{int}}_{\mathrm{UB}}$ and $\mathrm{ac}^{\mathrm{lf}}_{\mathrm{UB}}$ to be n^*.
For each edge-configuration $\gamma \in \Gamma^{\mathrm{int}}$, let $x^*_A(\gamma^{\mathrm{int}})$ (resp., $x^*_B(\gamma^{\mathrm{int}})$) denote the number of interior-edges with γ in G_A (resp., G_B), $\gamma \in \Gamma^{\mathrm{int}}$ and set
$x^*_{\min}(\gamma) := \min\{x^*_A(\gamma), x^*_B(\gamma)\}$, $x^*_{\max}(\gamma) := \max\{x^*_A(\gamma), x^*_B(\gamma)\}$,
$\mathrm{ec}^{\mathrm{int}}_{\mathrm{LB}}(\gamma) := \lfloor (3/4)x^*_{\min}(\gamma) + (1/4)x^*_{\max}(\gamma) \rfloor$ and
$\mathrm{ec}^{\mathrm{int}}_{\mathrm{UB}}(\gamma) := \lceil (1/4)x^*_{\min}(\gamma) + (3/4)x^*_{\max}(\gamma) \rceil$.
Set $\mathcal{F}_E := \mathcal{F}(v) := \mathcal{F}_A \cup \mathcal{F}_B$, $v \in V_{\mathrm{C}}$ and $\mathrm{fc}_{\mathrm{LB}}(\psi) := 0, \mathrm{fc}_{\mathrm{UB}}(\psi) := 10, \psi \in \mathcal{F}_A \cup \mathcal{F}_B$.
We include a linear constraint $\ell(a_1) \leq \ell(a_2)$ and $5 \leq \ell(a_1) + \ell(a_2) \leq 15$ as part of the side constraint.

D All Constraints in an MILP Formulation for Chemical Graphs

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

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

D.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 $\langle \mathbb{C} \rangle$;
- Each edge $a_i \in E_{(0/1)}$ is included in $\langle \mathbb{C} \rangle$ if necessary;
- For each edge $a_i \in E_{(\geq 2)}$, edge a_i is not included in $\langle \mathbb{C} \rangle$ and instead a path

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

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

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

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

Define

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

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

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

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

constants:

- $t_{\mathrm{C}} = |V_{\mathrm{C}}|$, $\widetilde{k_{\mathrm{C}}} = |E_{(\geq 2)}|$, $k_{\mathrm{C}} = |E_{(\geq 2)} \cup E_{(\geq 1)}|$, $t_{\mathrm{T}} = \mathrm{n}^{\mathrm{int}}_{\mathrm{UB}} - |V_{\mathrm{C}}|$, $m_{\mathrm{C}} = |E_{\mathrm{C}}|$. Note that $a_i \in E_{\mathrm{C}} \setminus (E_{(\geq 2)} \cup E_{(\geq 1)})$ holds $i \in [k_{\mathrm{C}} + 1, m_{\mathrm{C}}]$;
- $\ell_{\mathrm{LB}}(k), \ell_{\mathrm{UB}}(k) \in [1, t_{\mathrm{T}}]$, $k \in [1, k_{\mathrm{C}}]$: lower and upper bounds on the length of path P_k;
- $r_{G_{\mathrm{C}}} \in [1, m_{\mathrm{C}}]$: the rank $\mathrm{r}(G_{\mathrm{C}})$ of seed graph G_{C};

variables:

- $e^{\mathrm{C}}(i) \in [0, 1]$, $i \in [1, m_{\mathrm{C}}]$: $e^{\mathrm{C}}(i)$ represents edge $a_i \in E_{\mathrm{C}}$, $i \in [1, m_{\mathrm{C}}]$ ($e^{\mathrm{C}}(i) = 1$, $i \in I_{(=1)}$; $e^{\mathrm{C}}(i) = 0$, $i \in I_{(\geq 2)}$) ($e^{\mathrm{C}}(i) = 1 \Leftrightarrow$ edge a_i is used in $\langle \mathbb{C} \rangle$);
- $v^{\mathrm{T}}(i) \in [0, 1]$, $i \in [1, t_{\mathrm{T}}]$: $v^{\mathrm{T}}(i) = 1 \Leftrightarrow$ vertex ${v^{\mathrm{T}}}_i$ is used in $\langle \mathbb{C} \rangle$;
- $e^{\mathrm{T}}(i) \in [0, 1]$, $i \in [1, t_{\mathrm{T}} + 1]$: $e^{\mathrm{T}}(i)$ represents edge ${e^{\mathrm{T}}}_i = ({v^{\mathrm{T}}}_{i-1}, {v^{\mathrm{T}}}_i) \in E_{\mathrm{T}}$, where ${e^{\mathrm{T}}}_1$ and ${e^{\mathrm{T}}}_{t_{\mathrm{T}}+1}$ are fictitious edges ($e^{\mathrm{T}}(i) = 1 \Leftrightarrow$ edge ${e^{\mathrm{T}}}_i$ is used in $\langle \mathbb{C} \rangle$);
- $\chi^{\mathrm{T}}(i) \in [0, k_{\mathrm{C}}]$, $i \in [1, t_{\mathrm{T}}]$: $\chi^{\mathrm{T}}(i)$ represents the color assigned to vertex ${v^{\mathrm{T}}}_i$ ($\chi^{\mathrm{T}}(i) = k > 0 \Leftrightarrow$ vertex ${v^{\mathrm{T}}}_i$ is assigned color k; $\chi^{\mathrm{T}}(i) = 0$ means that vertex ${v^{\mathrm{T}}}_i$ is not used in $\langle \mathbb{C} \rangle$);
- $\mathrm{clr}^{\mathrm{T}}(k) \in [\ell_{\mathrm{LB}}(k) - 1, \ell_{\mathrm{UB}}(k) - 1]$, $k \in [1, k_{\mathrm{C}}]$, $\mathrm{clr}^{\mathrm{T}}(0) \in [0, t_{\mathrm{T}}]$: the number of vertices ${v^{\mathrm{T}}}_i \in V_{\mathrm{T}}$ with color c;

- $\delta^{\mathrm{T}}_{\chi}(k) \in [0,1]$, $k \in [0, k_{\mathrm{C}}]$: $\delta^{\mathrm{T}}_{\chi}(k) = 1 \Leftrightarrow \chi^{\mathrm{T}}(i) = k$ for some $i \in [1, t_{\mathrm{T}}]$;
- $\chi^{\mathrm{T}}(i,k) \in [0,1]$, $i \in [1, t_{\mathrm{T}}]$, $k \in [0, k_{\mathrm{C}}]$ ($\chi^{\mathrm{T}}(i,k) = 1 \Leftrightarrow \chi^{\mathrm{T}}(i) = k$);
- $\widetilde{\mathrm{deg}^{+}_{\mathrm{C}}}(i) \in [0,4]$, $i \in [1, t_{\mathrm{C}}]$: the out-degree of vertex ${v^{\mathrm{C}}}_i$ with the used edges e^{C} in E_{C};
- $\widetilde{\mathrm{deg}^{-}_{\mathrm{C}}}(i) \in [0,4]$, $i \in [1, t_{\mathrm{C}}]$: the in-degree of vertex ${v^{\mathrm{C}}}_i$ with the used edges e^{C} in E_{C};
- rank: the rank r($\mathbb{C}$) of a target chemical graph $\mathbb{C}$;

constraints:

$$\mathrm{rank} = r_{G_{\mathrm{C}}} - \sum_{i \in I_{(0/1)}} (1 - e^{\mathrm{C}}(i)), \tag{6}$$

$$e^{\mathrm{C}}(i) = 1, \quad i \in I_{(=1)}, \tag{7}$$

$$e^{\mathrm{C}}(i) = 0, \quad \mathrm{clr}^{\mathrm{T}}(i) \geq 1, \quad i \in I_{(\geq 2)}, \tag{8}$$

$$e^{\mathrm{C}}(i) + \mathrm{clr}^{\mathrm{T}}(i) \geq 1, \qquad \mathrm{clr}^{\mathrm{T}}(i) \leq t_{\mathrm{T}} \cdot (1 - e^{\mathrm{C}}(i)), \quad i \in I_{(\geq 1)}, \tag{9}$$

$$\sum_{c \in I^{-}_{(\geq 1)}(i) \cup I^{-}_{(0/1)}(i) \cup I^{-}_{(=1)}(i)} e^{\mathrm{C}}(c) = \widetilde{\mathrm{deg}^{-}_{\mathrm{C}}}(i), \qquad \sum_{c \in I^{+}_{(\geq 1)}(i) \cup I^{+}_{(0/1)}(i) \cup I^{+}_{(=1)}(i)} e^{\mathrm{C}}(c) = \widetilde{\mathrm{deg}^{+}_{\mathrm{C}}}(i), \quad i \in [1, t_{\mathrm{C}}], \tag{10}$$

$$\chi^{\mathrm{T}}(i,0) = 1 - v^{\mathrm{T}}(i), \qquad \sum_{k \in [0, k_{\mathrm{C}}]} \chi^{\mathrm{T}}(i,k) = 1, \qquad \sum_{k \in [0, k_{\mathrm{C}}]} k \cdot \chi^{\mathrm{T}}(i,k) = \chi^{\mathrm{T}}(i), \qquad i \in [1, t_{\mathrm{T}}], \tag{11}$$

$$\sum_{i \in [1, t_{\mathrm{T}}]} \chi^{\mathrm{T}}(i,k) = \mathrm{clr}^{\mathrm{T}}(k), \quad t_{\mathrm{T}} \cdot \delta^{\mathrm{T}}_{\chi}(k) \geq \sum_{i \in [1, t_{\mathrm{T}}]} \chi^{\mathrm{T}}(i,k) \geq \delta^{\mathrm{T}}_{\chi}(k), \qquad k \in [0, k_{\mathrm{C}}], \tag{12}$$

$$\begin{aligned} v^{\mathrm{T}}(i-1) &\geq v^{\mathrm{T}}(i), \\ k_{\mathrm{C}} \cdot (v^{\mathrm{T}}(i-1) - e^{\mathrm{T}}(i)) &\geq \chi^{\mathrm{T}}(i-1) - \chi^{\mathrm{T}}(i) \geq v^{\mathrm{T}}(i-1) - e^{\mathrm{T}}(i), \qquad i \in [2, t_{\mathrm{T}}]. \end{aligned} \tag{13}$$

D.2 Constraints for Including Leaf Paths

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

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

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

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

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

constants:

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

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

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

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

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

variables:

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

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

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

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

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

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

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

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

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

constraints:

$$\chi^{\rm F}(i,0) = 1 - v^{\rm F}(i), \quad \sum_{c \in [0, c_{\rm F}]} \chi^{\rm F}(i,c) = 1, \quad \sum_{c \in [0, c_{\rm F}]} c \cdot \chi^{\rm F}(i,c) = \chi^{\rm F}(i), \qquad i \in [1, t_{\rm F}], \tag{14}$$

$$\sum_{i \in [1, t_{\rm F}]} \chi^{\rm F}(i,c) = \mathrm{clr}^{\rm F}(c), \quad t_{\rm F} \cdot \delta^{\rm F}_\chi(c) \geq \sum_{i \in [1, t_{\rm F}]} \chi^{\rm F}(i,c) \geq \delta^{\rm F}_\chi(c), \qquad c \in [0, c_{\rm F}], \tag{15}$$

$$e^{\rm F}(1) = e^{\rm F}(t_{\rm F} + 1) = 0, \tag{16}$$

$$v^{\mathrm{F}}(i-1) \geq v^{\mathrm{F}}(i),$$
$$c_{\mathrm{F}} \cdot (v^{\mathrm{F}}(i-1) - e^{\mathrm{F}}(i)) \geq \chi^{\mathrm{F}}(i-1) - \chi^{\mathrm{F}}(i) \geq v^{\mathrm{F}}(i-1) - e^{\mathrm{F}}(i), \qquad i \in [2, t_{\mathrm{F}}], \tag{17}$$

$$\mathrm{bl}(k,i) \geq \delta^{\mathrm{F}}_{\chi}(\widetilde{t_{\mathrm{C}}} + i) + \chi^{\mathrm{T}}(i,k) - 1, \qquad k \in [1, k_{\mathrm{C}}], i \in [1, t_{\mathrm{T}}], \tag{18}$$

$$\sum_{k \in [1,k_{\mathrm{C}}], i \in [1,t_{\mathrm{T}}]} \mathrm{bl}(k,i) \leq \sum_{i \in [1,t_{\mathrm{T}}]} \delta^{\mathrm{F}}_{\chi}(\widetilde{t_{\mathrm{C}}} + i), \tag{19}$$

$$\mathrm{bl}_{\mathrm{LB}}(k) \leq \sum_{i \in [1,t_{\mathrm{T}}]} \mathrm{bl}(k,i) \leq \mathrm{bl}_{\mathrm{UB}}(k), \qquad k \in [1, k_{\mathrm{C}}], \tag{20}$$

$$t_{\mathrm{C}} + \sum_{i \in [1,t_{\mathrm{T}}]} v^{\mathrm{T}}(i) + \sum_{i \in [1,t_{\mathrm{F}}]} v^{\mathrm{F}}(i) = \mathrm{n}^{\mathrm{int}}_{G}. \tag{21}$$

D.3 Constraints for Including Fringe-trees

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

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

constants:

- n_{LB}: a lower bound on the number $n(\mathbb{C})$ of non-hydrogen atoms in $\mathbb{C}$, where $n_{\mathrm{LB}}, n^* \geq \mathrm{n}^{\mathrm{int}}_{\mathrm{LB}}$;
- $\mathrm{ch}_{\mathrm{LB}}(i), \mathrm{ch}_{\mathrm{UB}}(i) \in [0, n^*]$, $i \in [1, t_{\mathrm{T}}]$: lower and upper bounds on $\mathrm{ht}(\langle T_i \rangle)$ of the tree T_i rooted at a vertex ${v^{\mathrm{C}}}_i$;
- $\mathrm{ch}_{\mathrm{LB}}(k), \mathrm{ch}_{\mathrm{UB}}(k) \in [0, n^*]$, $k \in [1, k_{\mathrm{C}}] = I_{(\geq 2)} \cup I_{(\geq 1)}$: lower and upper bounds on the maximum height $\mathrm{ht}(\langle T \rangle)$ of the tree $T \in \mathcal{F}(P_k)$ rooted at an internal vertex of a path P_k for an edge $a_k \in E_{(\geq 1)} \cup E_{(\geq 2)}$;
- Prepare a coding of the set $\mathcal{F}(D_\pi)$ and let $[\psi]$ denote the coded integer of an element ψ in $\mathcal{F}(D_\pi)$;
- Sets $\mathcal{F}(v) \subseteq \mathcal{F}(D_\pi), v \in V_{\mathrm{C}}$ and $\mathcal{F}_E \subseteq \mathcal{F}(D_\pi)$ of chemical rooted trees T with $\mathrm{ht}(T) \in [1, \rho]$;
- Define $\mathcal{F}^* := \bigcup_{v \in V_{\mathrm{C}}} \mathcal{F}(v) \cup \mathcal{F}_E$, $\mathcal{F}^{\mathrm{C}}_i := \mathcal{F}({v^{\mathrm{C}}}_i)$, $i \in [1, t_{\mathrm{C}}]$, $\mathcal{F}^{\mathrm{T}}_i := \mathcal{F}_E$, $i \in [1, t_{\mathrm{T}}]$ and $\mathcal{F}^{\mathrm{F}}_i := \mathcal{F}_E$, $i \in [1, t_{\mathrm{F}}]$;

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

- $\mathcal{F}_i^{\mathrm{X}}[p], p \in [1, \rho], \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: the set of chemical rooted trees $T \in \mathcal{F}_i^{\mathrm{X}}$ with $\mathrm{ht}(\langle T \rangle) = p$;

- $n_{\overline{\mathtt{H}}}([\psi]) \in [0, 3^\rho], \psi \in \mathcal{F}^*$: the number $n(\langle \psi \rangle)$ of non-root hydrogen vertices in a chemical rooted tree ψ;

- $\mathrm{ht}_{\overline{\mathtt{H}}}([\psi]) \in [0, \rho], \psi \in \mathcal{F}^*$: the height $\mathrm{ht}(\langle \psi \rangle)$ of the hydrogen-suppressed chemical rooted tree $\langle \psi \rangle$;

- $\deg_{\mathrm{r}}^{\overline{\mathtt{H}}}([\psi]) \in [0, 3], \psi \in \mathcal{F}^*$: the number $\deg_{\mathrm{r}}(\langle \psi \rangle)$ of non-hydrogen children of the root r of a chemical rooted tree ψ;

- $\deg_{\mathrm{r}}^{\mathrm{hyd}}([\psi]) \in [0, 3], \psi \in \mathcal{F}^*$: the number $\deg_{\mathrm{r}}(\psi) - \deg_{\mathrm{r}}(\langle \psi \rangle)$ of hydrogen children of the root r of a chemical rooted tree ψ;

- $\mathrm{v}_{\mathrm{ion}}(\psi) \in [-3, +3], \psi \in \mathcal{F}^*$: the ion-valence of the root in ψ;

- $\mathrm{ac}_{\nu}^{\mathrm{lf}}(\psi), \nu \in \Gamma_{\mathrm{ac}}^{\mathrm{lf}}$: the frequency of leaf-edges with adjacency-configuration ν in ψ;

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

variables:

- $n_G \in [n_{\mathrm{LB}}, n^*]$: the number $n(\mathbb{C})$ of non-hydrogen atoms in $\mathbb{C}$;

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

- $\delta_{\mathrm{fr}}^{\mathrm{X}}(i, [\psi]) \in [0, 1], i \in [1, t_{\mathrm{X}}], \psi \in \mathcal{F}_i^{\mathrm{X}}, \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: $\delta_{\mathrm{fr}}^{\mathrm{X}}(i, [\psi]) = 1 \Leftrightarrow \psi$ is the ρ-fringe-tree rooted at vertex ${v^{\mathrm{X}}}_i$ in $\mathbb{C}$;

- $\mathrm{fc}([\psi]) \in [\mathrm{fc}_{\mathrm{LB}}(\psi), \mathrm{fc}_{\mathrm{UB}}(\psi)], \psi \in \mathcal{F}^*$: the number of interior-vertices v such that $\mathbb{C}[v]$ is r-isomorphic to ψ in $\mathbb{C}$;

- $\mathrm{ac}^{\mathrm{lf}}([\nu]) \in [\mathrm{ac}_{\mathrm{LB}}^{\mathrm{lf}}(\nu), \mathrm{ac}_{\mathrm{UB}}^{\mathrm{lf}}(\nu)], \nu \in \Gamma_{\mathrm{ac}}^{\mathrm{lf}}$: the number of leaf-edge with adjacency-configuration ν in $\mathbb{C}$;

- $\deg_{\mathrm{X}}^{\mathrm{ex}}(i) \in [0, 3], i \in [1, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: the number of non-hydrogen children of the root of the ρ-fringe-tree rooted at vertex ${v^{\mathrm{X}}}_i$ in $\mathbb{C}$;

- $\mathrm{hyddeg}^{\mathrm{X}}(i) \in [0, 4]$, $i \in [1, t_{\mathrm{X}}]$, $\mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: the number of hydrogen atoms adjacent to vertex ${v^{\mathrm{X}}}_i$ (i.e., $\mathrm{hyddeg}({v^{\mathrm{X}}}_i)$) in $\mathbb{C} = (H, \alpha, \beta)$;

- $\mathrm{eledeg}_{\mathrm{X}}(i) \in [-3, +3]$, $i \in [1, t_{\mathrm{X}}]$, $\mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: the ion-valence $\mathrm{v}_{\mathrm{ion}}(\psi)$ of vertex ${v^{\mathrm{X}}}_i$ (i.e., $\mathrm{eledeg}_{\mathrm{X}}(i) = \mathrm{v}_{\mathrm{ion}}(\psi)$ for the ρ-fringe-tree ψ rooted at ${v^{\mathrm{X}}}_i$) in $\mathbb{C} = (H, \alpha, \beta)$;

- $h^{\mathrm{X}}(i) \in [0, \rho]$, $i \in [1, t_{\mathrm{X}}]$, $\mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$: the height $\mathrm{ht}(\langle T \rangle)$ of the hydrogen-suppressed chemical rooted tree $\langle T \rangle$ of the ρ-fringe-tree T rooted at vertex ${v^{\mathrm{X}}}_i$ in $\mathbb{C}$;

- $\sigma(k,i) \in [0,1]$, $k \in [1,k_{\rm C}] = I_{(\geq 2)} \cup I_{(\geq 1)}, i \in [1,t_{\rm T}]$: $\sigma(k,i) = 1 \Leftrightarrow$ the ρ-fringe-tree T_v rooted at vertex $v = v^{\rm T}{}_i$ with color k has the largest height $\mathrm{ht}(\langle \mathcal{T}_v \rangle)$ among such trees $T_v, v \in V_{\rm T}$;

constraints:

$$\begin{aligned} \sum_{\psi \in \mathcal{F}_i^{\rm C}} \delta_{\rm fr}^{\rm C}(i,[\psi]) &= 1, && i \in [1,t_{\rm C}], \\ \sum_{\psi \in \mathcal{F}_i^{\rm X}} \delta_{\rm fr}^{\rm X}(i,[\psi]) &= v^{\rm X}(i), && i \in [1,t_{\rm X}], {\rm X} \in \{{\rm T},{\rm F}\}, \end{aligned} \tag{22}$$

$$\begin{aligned} \sum_{\psi \in \mathcal{F}_i^{\rm X}} \deg_{\rm r}^{\overline{\rm H}}([\psi]) \cdot \delta_{\rm fr}^{\rm X}(i,[\psi]) &= \deg_{\rm X}^{\rm ex}(i), \\ \sum_{\psi \in \mathcal{F}_i^{\rm X}} \deg_{\rm r}^{\rm hyd}([\psi]) \cdot \delta_{\rm fr}^{\rm X}(i,[\psi]) &= \mathrm{hyddeg}^{\rm X}(i), \\ \sum_{\psi \in \mathcal{F}_i^{\rm X}} {\rm v}_{\rm ion}([\psi]) \cdot \delta_{\rm fr}^{\rm X}(i,[\psi]) &= \mathrm{eledeg}_{\rm X}(i), \qquad i \in [1,t_{\rm X}], {\rm X} \in \{{\rm C},{\rm T},{\rm F}\}, \end{aligned} \tag{23}$$

$$\sum_{\psi \in \mathcal{F}_i^{\rm F}[\rho]} \delta_{\rm fr}^{\rm F}(i,[\psi]) \geq v^{\rm F}(i) - e^{\rm F}(i+1), \qquad i \in [1,t_{\rm F}] \ (e^{\rm F}(t_{\rm F}+1) = 0), \tag{24}$$

$$\sum_{\psi \in \mathcal{F}_i^{\rm X}} \mathrm{ht}_{\overline{\rm H}}([\psi]) \cdot \delta_{\rm fr}^{\rm X}(i,[\psi]) = h^{\rm X}(i), \qquad i \in [1,t_{\rm X}], {\rm X} \in \{{\rm C},{\rm T},{\rm F}\}, \tag{25}$$

$$\sum_{\substack{\psi \in \mathcal{F}_i^{\rm X} \\ i \in [1,t_{\rm X}], {\rm X} \in \{{\rm C},{\rm T},{\rm F}\}}} n_{\overline{\rm H}}([\psi]) \cdot \delta_{\rm fr}^{\rm X}(i,[\psi]) + \sum_{i \in [1,t_{\rm X}], {\rm X} \in \{{\rm T},{\rm F}\}} v^{\rm X}(i) + t_{\rm C} = n_G, \tag{26}$$

$$\sum_{i \in [1,t_{\rm X}], {\rm X} \in \{{\rm C},{\rm T},{\rm F}\}} \delta_{\rm fr}^{\rm X}(i,[\psi]) = \mathrm{fc}([\psi]), \qquad \psi \in \mathcal{F}^*, \tag{27}$$

$$\sum_{\psi \in \mathcal{F}_i^{\rm X}, i \in [1,t_{\rm X}], {\rm X} \in \{{\rm C},{\rm T},{\rm F}\}} \mathrm{ac}_\nu^{\rm lf}(\psi) \cdot \delta_{\rm fr}^{\rm X}(i,[\psi]) = \mathrm{ac}^{\rm lf}([\nu]), \qquad \nu \in \Gamma_{\rm ac}^{\rm lf}, \tag{28}$$

$$\begin{aligned} h^{\rm C}(i) \geq \mathrm{ch}_{\rm LB}(i) - n^* \cdot \delta_\chi^{\rm F}(i), \quad \mathrm{clr}^{\rm F}(i) + \rho \geq \mathrm{ch}_{\rm LB}(i), \\ h^{\rm C}(i) \leq \mathrm{ch}_{\rm UB}(i), \quad \mathrm{clr}^{\rm F}(i) + \rho \leq \mathrm{ch}_{\rm UB}(i) + n^* \cdot (1 - \delta_\chi^{\rm F}(i)), \qquad i \in [1,\widetilde{t_{\rm C}}], \end{aligned} \tag{29}$$

$$\mathrm{ch}_{\rm LB}(i) \leq h^{\rm C}(i) \leq \mathrm{ch}_{\rm UB}(i), \qquad i \in [\widetilde{t_{\rm C}}+1, t_{\rm C}], \tag{30}$$

$$
\begin{aligned}
h^{\mathrm{T}}(i) &\leq \mathrm{ch}_{\mathrm{UB}}(k) + n^* \cdot (\delta^{\mathrm{F}}_{\chi}(\widetilde{t_{\mathrm{C}}} + i) + 1 - \chi^{\mathrm{T}}(i,k)),\\
\mathrm{clr}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}} + i) + \rho &\leq \mathrm{ch}_{\mathrm{UB}}(k) + n^* \cdot (2 - \delta^{\mathrm{F}}_{\chi}(\widetilde{t_{\mathrm{C}}} + i) - \chi^{\mathrm{T}}(i,k)), \qquad k \in [1, k_{\mathrm{C}}], i \in [1, t_{\mathrm{T}}],
\end{aligned} \tag{31}
$$

$$
\sum_{i \in [1, t_{\mathrm{T}}]} \sigma(k,i) = \delta^{\mathrm{T}}_{\chi}(k), \qquad k \in [1, k_{\mathrm{C}}], \tag{32}
$$

$$
\begin{aligned}
\chi^{\mathrm{T}}(i,k) &\geq \sigma(k,i),\\
h^{\mathrm{T}}(i) &\geq \mathrm{ch}_{\mathrm{LB}}(k) - n^* \cdot (\delta^{\mathrm{F}}_{\chi}(\widetilde{t_{\mathrm{C}}} + i) + 1 - \sigma(k,i)),\\
\mathrm{clr}^{\mathrm{F}}(\widetilde{t_{\mathrm{C}}} + i) + \rho &\geq \mathrm{ch}_{\mathrm{LB}}(k) - n^* \cdot (2 - \delta^{\mathrm{F}}_{\chi}(\widetilde{t_{\mathrm{C}}} + i) - \sigma(k,i)), \qquad k \in [1, k_{\mathrm{C}}], i \in [1, t_{\mathrm{T}}].
\end{aligned} \tag{33}
$$

D.4 Descriptor for the Number of Specified Degree

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

variables:

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

constraints:

$$
\sum_{k \in I^{+}_{(\geq 2)}(i) \cup I^{+}_{(\geq 1)}(i)} \delta^{\mathrm{T}}_{\chi}(k) = \deg_{\mathrm{CT}}(i), \quad \sum_{k \in I^{-}_{(\geq 2)}(i) \cup I^{-}_{(\geq 1)}(i)} \delta^{\mathrm{T}}_{\chi}(k) = \deg_{\mathrm{TC}}(i), \qquad i \in [1, t_{\mathrm{C}}], \tag{34}
$$

$$\widetilde{\deg}^{-}_{\mathrm{C}}(i) + \widetilde{\deg}^{+}_{\mathrm{C}}(i) + \deg_{\mathrm{CT}}(i) + \deg_{\mathrm{TC}}(i) + \delta^{\mathrm{F}}_{\chi}(i) = \deg^{\mathrm{int}}_{\mathrm{C}}(i), \qquad i \in [1, \widetilde{t_{\mathrm{C}}}], \tag{35}$$

$$\widetilde{\deg}^{-}_{\mathrm{C}}(i) + \widetilde{\deg}^{+}_{\mathrm{C}}(i) + \deg_{\mathrm{CT}}(i) + \deg_{\mathrm{TC}}(i) = \deg^{\mathrm{int}}_{\mathrm{C}}(i), \qquad i \in [\widetilde{t_{\mathrm{C}}} + 1, t_{\mathrm{C}}], \tag{36}$$

$$\deg^{\mathrm{int}}_{\mathrm{C}}(i) + \deg^{\mathrm{ex}}_{\mathrm{C}}(i) = \deg^{\mathrm{C}}(i), \qquad i \in [1, t_{\mathrm{C}}], \tag{37}$$

$$\sum_{\psi \in \mathcal{F}^{\mathrm{C}}_i[\rho]} \delta^{\mathrm{C}}_{\mathrm{fr}}(i, [\psi]) \geq 2 - \deg^{\mathrm{int}}_{\mathrm{C}}(i) \qquad i \in [1, t_{\mathrm{C}}], \tag{38}$$

$$\begin{aligned} 2v^{\mathrm{T}}(i) + \delta^{\mathrm{F}}_{\chi}(\widetilde{t_{\mathrm{C}}} + i) &= \deg^{\mathrm{int}}_{\mathrm{T}}(i), \\ \deg^{\mathrm{int}}_{\mathrm{T}}(i) + \deg^{\mathrm{ex}}_{\mathrm{T}}(i) &= \deg^{\mathrm{T}}(i), \qquad i \in [1, t_{\mathrm{T}}] \ (e^{\mathrm{T}}(1) = e^{\mathrm{T}}(t_{\mathrm{T}} + 1) = 0), \end{aligned} \tag{39}$$

$$\begin{aligned} v^{\mathrm{F}}(i) + e^{\mathrm{F}}(i+1) &= \deg^{\mathrm{int}}_{\mathrm{F}}(i), \\ \deg^{\mathrm{int}}_{\mathrm{F}}(i) + \deg^{\mathrm{ex}}_{\mathrm{F}}(i) &= \deg^{\mathrm{F}}(i), \qquad i \in [1, t_{\mathrm{F}}] \ (e^{\mathrm{F}}(1) = e^{\mathrm{F}}(t_{\mathrm{F}} + 1) = 0), \end{aligned} \tag{40}$$

$$\begin{aligned} \sum_{d \in [0,4]} \delta^{\mathrm{X}}_{\mathrm{dg}}(i,d) = 1, \quad \sum_{d \in [1,4]} d \cdot \delta^{\mathrm{X}}_{\mathrm{dg}}(i,d) &= \deg^{\mathrm{X}}(i) + \mathrm{hyddeg}^{\mathrm{X}}(i), \\ \sum_{d \in [0,4]} \delta^{\mathrm{int}}_{\mathrm{dg,X}}(i,d) = 1, \quad \sum_{d \in [1,4]} d \cdot \delta^{\mathrm{int}}_{\mathrm{dg,X}}(i,d) &= \deg^{\mathrm{int}}_{\mathrm{X}}(i), \quad i \in [1, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{T}, \mathrm{C}, \mathrm{F}\}, \end{aligned} \tag{41}$$

$$\begin{aligned} \sum_{i \in [1, t_{\mathrm{C}}]} \delta^{\mathrm{C}}_{\mathrm{dg}}(i,d) + \sum_{i \in [1, t_{\mathrm{T}}]} \delta^{\mathrm{T}}_{\mathrm{dg}}(i,d) + \sum_{i \in [1, t_{\mathrm{F}}]} \delta^{\mathrm{F}}_{\mathrm{dg}}(i,d) &= \mathrm{dg}(d), \\ \sum_{i \in [1, t_{\mathrm{C}}]} \delta^{\mathrm{int}}_{\mathrm{dg,C}}(i,d) + \sum_{i \in [1, t_{\mathrm{T}}]} \delta^{\mathrm{int}}_{\mathrm{dg,T}}(i,d) + \sum_{i \in [1, t_{\mathrm{F}}]} \delta^{\mathrm{int}}_{\mathrm{dg,F}}(i,d) &= \mathrm{dg}^{\mathrm{int}}(d), \qquad d \in [1,4]. \end{aligned} \tag{42}$$

D.5 Assigning Multiplicity

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

constants:

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

variables:

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

constraints:

$$e^{\mathrm{C}}(i) \leq \beta^{\mathrm{C}}(i) \leq 3e^{\mathrm{C}}(i), i \in [\widetilde{k_{\mathrm{C}}}+1, m_{\mathrm{C}}] = I_{(\geq 1)} \cup I_{(0/1)} \cup I_{(=1)}, \tag{43}$$

$$e^{\mathrm{X}}(i) \leq \beta^{\mathrm{X}}(i) \leq 3e^{\mathrm{X}}(i), \qquad i \in [2, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{T}, \mathrm{F}\}, \tag{44}$$

$$\delta^{\mathrm{T}}_{\chi}(k) \leq \beta^{\mathrm{CT}}(k) \leq 3\delta^{\mathrm{T}}_{\chi}(k), \quad \delta^{\mathrm{T}}_{\chi}(k) \leq \beta^{\mathrm{TC}}(k) \leq 3\delta^{\mathrm{T}}_{\chi}(k), \qquad k \in [1, k_{\mathrm{C}}], \tag{45}$$

$$\delta^{\mathrm{F}}_{\chi}(c) \leq \beta^{\mathrm{XF}}(c) \leq 3\delta^{\mathrm{F}}_{\chi}(c), \qquad c \in [1, c_{\mathrm{F}}], \tag{46}$$

$$\sum_{m \in [0,3]} \delta^{\mathrm{X}}_{\beta}(i,m) = 1, \quad \sum_{m \in [0,3]} m \cdot \delta^{\mathrm{X}}_{\beta}(i,m) = \beta^{\mathrm{X}}(i), \qquad i \in [2, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{T}, \mathrm{F}\}, \tag{47}$$

$$\sum_{m\in[0,3]} \delta^{\mathrm{C}}_{\beta}(i,m) = 1, \quad \sum_{m\in[0,3]} m\cdot\delta^{\mathrm{C}}_{\beta}(i,m) = \beta^{\mathrm{C}}(i), \qquad i\in[\widetilde{k_{\mathrm{C}}}+1, m_{\mathrm{C}}], \tag{48}$$

$$\begin{aligned}
\sum_{m\in[0,3]} \delta^{\mathrm{CT}}_{\beta}(k,m) = 1, \quad \sum_{m\in[0,3]} m\cdot\delta^{\mathrm{CT}}_{\beta}(k,m) = \beta^{\mathrm{CT}}(k), \qquad & k\in[1,k_{\mathrm{C}}],\\
\sum_{m\in[0,3]} \delta^{\mathrm{TC}}_{\beta}(k,m) = 1, \quad \sum_{m\in[0,3]} m\cdot\delta^{\mathrm{TC}}_{\beta}(k,m) = \beta^{\mathrm{TC}}(k), \qquad & k\in[1,k_{\mathrm{C}}],\\
\sum_{m\in[0,3]} \delta^{*\mathrm{F}}_{\beta}(c,m) = 1, \quad \sum_{m\in[0,3]} m\cdot\delta^{*\mathrm{F}}_{\beta}(c,m) = \beta^{*\mathrm{F}}(c), \qquad & c\in[1,c_{\mathrm{F}}],
\end{aligned} \tag{49}$$

$$\sum_{\psi\in\mathcal{F}^{\mathrm{X}}_i} \beta_{\mathrm{r}}([\psi])\cdot\delta^{\mathrm{X}}_{\mathrm{fr}}(i,[\psi]) = \beta^{\mathrm{X}}_{\mathrm{ex}}(i), \qquad i\in[1,t_{\mathrm{X}}], \mathrm{X}\in\{\mathrm{C},\mathrm{T},\mathrm{F}\}, \tag{50}$$

$$\begin{aligned}
\sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]} \delta^{\mathrm{C}}_{\beta}(i,m) = \mathrm{bd}_{\mathrm{C}}(m), \quad \sum_{i\in[2,t_{\mathrm{T}}]} \delta^{\mathrm{T}}_{\beta}(i,m) = \mathrm{bd}_{\mathrm{T}}(m),\\
\sum_{k\in[1,k_{\mathrm{C}}]} \delta^{\mathrm{CT}}_{\beta}(k,m) = \mathrm{bd}_{\mathrm{CT}}(m), \quad \sum_{k\in[1,k_{\mathrm{C}}]} \delta^{\mathrm{TC}}_{\beta}(k,m) = \mathrm{bd}_{\mathrm{TC}}(m),\\
\sum_{i\in[2,t_{\mathrm{F}}]} \delta^{\mathrm{F}}_{\beta}(i,m) = \mathrm{bd}_{\mathrm{F}}(m), \quad \sum_{c\in[1,\widetilde{t_{\mathrm{C}}}]} \delta^{*\mathrm{F}}_{\beta}(c,m) = \mathrm{bd}_{\mathrm{CF}}(m),\\
\sum_{c\in[\widetilde{t_{\mathrm{C}}}+1,c_{\mathrm{F}}]} \delta^{*\mathrm{F}}_{\beta}(c,m) = \mathrm{bd}_{\mathrm{TF}}(m),\\
\mathrm{bd}_{\mathrm{C}}(m)+\mathrm{bd}_{\mathrm{T}}(m)+\mathrm{bd}_{\mathrm{F}}(m)+\mathrm{bd}_{\mathrm{CT}}(m)+\mathrm{bd}_{\mathrm{TC}}(m)+\mathrm{bd}_{\mathrm{TF}}(m)+\mathrm{bd}_{\mathrm{CF}}(m) = \mathrm{bd}^{\mathrm{int}}(m),\\
m\in[1,3].
\end{aligned} \tag{51}$$

D.6 Assigning Chemical Elements and Valence Condition

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

constants:

- Subsets $\Lambda^{\mathrm{int}} \subseteq \Lambda\setminus\{\mathtt{H}\}, \Lambda^{\mathrm{ex}} \subseteq \Lambda$ of chemical elements, where we denote by $[\mathtt{e}]$ (resp., $[\mathtt{e}]^{\mathrm{int}}$ and $[\mathtt{e}]^{\mathrm{ex}}$) of a standard encoding of an element $\mathtt{e}$ in the set Λ (resp., $\Lambda^{\mathrm{int}}_{\epsilon}$ and $\Lambda^{\mathrm{ex}}_{\epsilon}$);
- A valence function: $\mathrm{val}:\Lambda\to[1,6]$;
- 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 $\alpha(v) = \mathtt{a}$;

- $\alpha_{\mathrm{r}}([\psi]) \in [\Lambda^{\mathrm{ex}}], \in \mathcal{F}^*$: 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, where possibly $\mathtt{a} = \mathtt{H}$;

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

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 $\mathbb{C}$);

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

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

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

constraints:

$$\begin{aligned}\beta^{\mathrm{CT}}(k) - 3(e^{\mathrm{T}}(i) - \chi^{\mathrm{T}}(i,k) + 1) \leq \beta^{\mathrm{CT}}(i) \leq \beta^{\mathrm{CT}}(k) + 3(e^{\mathrm{T}}(i) - \chi^{\mathrm{T}}(i,k) + 1), i \in [1, t_{\mathrm{T}}],\\ \beta^{\mathrm{TC}}(k) - 3(e^{\mathrm{T}}(i+1) - \chi^{\mathrm{T}}(i,k) + 1) \leq \beta^{\mathrm{TC}}(i) \leq \beta^{\mathrm{TC}}(k) + 3(e^{\mathrm{T}}(i+1) - \chi^{\mathrm{T}}(i,k) + 1), i \in [1, t_{\mathrm{T}}],\\ k \in [1, k_{\mathrm{C}}],\end{aligned} \tag{52}$$

$$\begin{aligned}\beta^{*\mathrm{F}}(c) - 3(e^{\mathrm{F}}(i) - \chi^{\mathrm{F}}(i,c) + 1) \leq \beta^{\mathrm{CF}}(i) \leq \beta^{*\mathrm{F}}(c) + 3(e^{\mathrm{F}}(i) - \chi^{\mathrm{F}}(i,c) + 1), i \in [1, t_{\mathrm{F}}], \quad c \in [1, \widetilde{t_{\mathrm{C}}}],\\ \beta^{*\mathrm{F}}(c) - 3(e^{\mathrm{F}}(i) - \chi^{\mathrm{F}}(i,c) + 1) \leq \beta^{\mathrm{TF}}(i) \leq \beta^{*\mathrm{F}}(c) + 3(e^{\mathrm{F}}(i) - \chi^{\mathrm{F}}(i,c) + 1), i \in [1, t_{\mathrm{F}}], \quad c \in [\widetilde{t_{\mathrm{C}}} + 1, c_{\mathrm{F}}],\end{aligned} \tag{53}$$

$$\sum_{\mathsf{a}\in\Lambda^{\rm int}}\delta^{\rm C}_{\alpha}(i,[\mathsf{a}]^{\rm int})=1,\quad \sum_{\mathsf{a}\in\Lambda^{\rm int}}[\mathsf{a}]^{\rm int}\cdot\delta^{\rm X}_{\alpha}(i,[\mathsf{a}]^{\rm int})=\alpha^{\rm C}(i),\qquad i\in[1,t_{\rm C}],$$
$$\sum_{\mathsf{a}\in\Lambda^{\rm int}}\delta^{\rm X}_{\alpha}(i,[\mathsf{a}]^{\rm int})=v^{\rm X}(i),\quad \sum_{\mathsf{a}\in\Lambda^{\rm int}}[\mathsf{a}]^{\rm int}\cdot\delta^{\rm X}_{\alpha}(i,[\mathsf{a}]^{\rm int})=\alpha^{\rm X}(i),\qquad i\in[1,t_{\rm X}],{\rm X}\in\{{\rm T},{\rm F}\},\tag{54}$$

$$\sum_{\psi\in\mathcal{F}^{\rm X}_i}\alpha_{\rm r}([\psi])\cdot\delta^{\rm X}_{\rm fr}(i,[\psi])=\alpha^{\rm X}(i),\qquad i\in[1,t_{\rm X}],{\rm X}\in\{{\rm C},{\rm T},{\rm F}\},\tag{55}$$

$$\sum_{j\in I_{\rm C}(i)}\beta^{\rm C}(j)+\sum_{k\in I^{+}_{(\geq 2)}(i)\cup I^{+}_{(\geq 1)}(i)}\beta^{\rm CT}(k)+\sum_{k\in I^{-}_{(\geq 2)}(i)\cup I^{-}_{(\geq 1)}(i)}\beta^{\rm TC}(k)$$
$$+\beta^{*{\rm F}}(i)+\beta^{\rm C}_{\rm ex}(i)-{\rm eledeg}_{\rm C}(i)=\sum_{\mathsf{a}\in\Lambda^{\rm int}}{\rm val}(\mathsf{a})\delta^{\rm C}_{\alpha}(i,[\mathsf{a}]^{\rm int}),\qquad i\in[1,\widetilde{t_{\rm C}}],\tag{56}$$

$$\sum_{j\in I_{\rm C}(i)}\beta^{\rm C}(j)+\sum_{k\in I^{+}_{(\geq 2)}(i)\cup I^{+}_{(\geq 1)}(i)}\beta^{\rm CT}(k)+\sum_{k\in I^{-}_{(\geq 2)}(i)\cup I^{-}_{(\geq 1)}(i)}\beta^{\rm TC}(k)$$
$$+\beta^{\rm C}_{\rm ex}(i)-{\rm eledeg}_{\rm C}(i)=\sum_{\mathsf{a}\in\Lambda^{\rm int}}{\rm val}(\mathsf{a})\delta^{\rm C}_{\alpha}(i,[\mathsf{a}]^{\rm int}),\qquad i\in[\widetilde{t_{\rm C}}+1,t_{\rm C}],\tag{57}$$

$$\beta^{\rm T}(i)+\beta^{\rm T}(i{+}1)+\beta^{\rm T}_{\rm ex}(i)+\beta^{\rm CT}(i)+\beta^{\rm TC}(i)$$
$$+\beta^{*{\rm F}}(\widetilde{t_{\rm C}}+i)-{\rm eledeg}_{\rm T}(i)=\sum_{\mathsf{a}\in\Lambda^{\rm int}}{\rm val}(\mathsf{a})\delta^{\rm T}_{\alpha}(i,[\mathsf{a}]^{\rm int}),$$
$$i\in[1,t_{\rm T}]\ (\beta^{\rm T}(1)=\beta^{\rm T}(t_{\rm T}+1)=0),\tag{58}$$

$$\beta^{\rm F}(i)+\beta^{\rm F}(i{+}1)+\beta^{\rm CF}(i)+\beta^{\rm TF}(i)$$
$$+\beta^{\rm F}_{\rm ex}(i)-{\rm eledeg}_{\rm F}(i)=\sum_{\mathsf{a}\in\Lambda^{\rm int}}{\rm val}(\mathsf{a})\delta^{\rm F}_{\alpha}(i,[\mathsf{a}]^{\rm int}),$$
$$i\in[1,t_{\rm F}]\ (\beta^{\rm F}(1)=\beta^{\rm F}(t_{\rm F}+1)=0),\tag{59}$$

$$\sum_{i\in[1,t_{\rm X}]}\delta^{\rm X}_{\alpha}(i,[\mathsf{a}]^{\rm int})={\rm na}_{\rm X}([\mathsf{a}]^{\rm int}),\qquad \mathsf{a}\in\Lambda^{\rm int},{\rm X}\in\{{\rm C},{\rm T},{\rm F}\},\tag{60}$$

$$\sum_{\psi\in\mathcal{F}^{\rm X}_i,i\in[1,t_{\rm X}]}{\rm na}^{\rm ex}_{\mathsf{a}}([\psi])\cdot\delta^{\rm X}_{\rm fr}(i,[\psi])={\rm na}^{\rm ex}_{\rm X}([\mathsf{a}]^{\rm ex}),\qquad \mathsf{a}\in\Lambda^{\rm ex},{\rm X}\in\{{\rm C},{\rm T},{\rm F}\},\tag{61}$$

$$
\begin{aligned}
\mathrm{na_C}([\mathsf{a}]^{\mathrm{int}}) + \mathrm{na_T}([\mathsf{a}]^{\mathrm{int}}) + \mathrm{na_F}([\mathsf{a}]^{\mathrm{int}}) &= \mathrm{na}^{\mathrm{int}}([\mathsf{a}]^{\mathrm{int}}), && \mathsf{a} \in \Lambda^{\mathrm{int}},\\
\sum_{\mathrm{X}\in\{\mathrm{C,T,F}\}} \mathrm{na}^{\mathrm{ex}}_{\mathrm{X}}([\mathsf{a}]^{\mathrm{ex}}) &= \mathrm{na}^{\mathrm{ex}}([\mathsf{a}]^{\mathrm{ex}}), && \mathsf{a} \in \Lambda^{\mathrm{ex}},\\
\mathrm{na}^{\mathrm{int}}([\mathsf{a}]^{\mathrm{int}}) + \mathrm{na}^{\mathrm{ex}}([\mathsf{a}]^{\mathrm{ex}}) &= \mathrm{na}([\mathsf{a}]), && \mathsf{a} \in \Lambda^{\mathrm{int}} \cap \Lambda^{\mathrm{ex}},\\
\mathrm{na}^{\mathrm{int}}([\mathsf{a}]^{\mathrm{int}}) &= \mathrm{na}([\mathsf{a}]), && \mathsf{a} \in \Lambda^{\mathrm{int}} \setminus \Lambda^{\mathrm{ex}},\\
\mathrm{na}^{\mathrm{ex}}([\mathsf{a}]^{\mathrm{ex}}) &= \mathrm{na}([\mathsf{a}]), && \mathsf{a} \in \Lambda^{\mathrm{ex}} \setminus \Lambda^{\mathrm{int}},
\end{aligned} \tag{62}
$$

$$
\sum_{\mathsf{a}\in\Lambda^*(i)} \delta^{\mathrm{C}}_{\alpha}(i, [\mathsf{a}]^{\mathrm{int}}) = 1, \qquad i \in [1, t_{\mathrm{C}}]. \tag{63}
$$

D.7 Constraints for Bounds on the Number of Bonds

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

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 :

- $\mathrm{bd_T}(k, i, m) \in [0, 1]$, $k \in [1, k_{\mathrm{C}}]$, $i \in [2, t_{\mathrm{T}}]$, $m \in [2, 3]$: $\mathrm{bd_T}(k, i, m) = 1 \Leftrightarrow$ 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:

$$
\mathrm{bd}_{m,\mathrm{LB}}(i) \leq \delta^{\mathrm{C}}_{\beta}(i, m) \leq \mathrm{bd}_{m,\mathrm{UB}}(i), i \in I_{(=1)} \cup I_{(0/1)}, m \in [2, 3], \tag{64}
$$

$$
\mathrm{bd_T}(k, i, m) \geq \delta^{\mathrm{T}}_{\beta}(i, m) + {\chi^{\mathrm{T}}}(i, k) - 1, \quad k \in [1, k_{\mathrm{C}}], i \in [2, t_{\mathrm{T}}], m \in [2, 3], \tag{65}
$$

$$
\sum_{j\in[2,t_{\mathrm{T}}]} \delta^{\mathrm{T}}_{\beta}(j, m) \geq \sum_{k\in[1,k_{\mathrm{C}}], i\in[2,t_{\mathrm{T}}]} \mathrm{bd_T}(k, i, m), \quad m \in [2, 3], \tag{66}
$$

$$
\mathrm{bd}_{m,\mathrm{LB}}(k) \leq \sum_{i\in[2,t_{\mathrm{T}}]} \mathrm{bd_T}(k, i, m) + \delta^{\mathrm{CT}}_{\beta}(k, m) + \delta^{\mathrm{TC}}_{\beta}(k, m) \leq \mathrm{bd}_{m,\mathrm{UB}}(k),
$$
$$
k \in [1, k_{\mathrm{C}}], m \in [2, 3]. \tag{67}
$$

D.8 Descriptor for the Number of Adjacency-configurations

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

constants:

- A set Γ^{int} of edge-configurations $\gamma = (\mu, \mu', m)$ with $\mu \leq \mu'$;
- Let $\overline{\gamma}$ of an edge-configuration $\gamma = (\mu, \mu', m)$ denote the edge-configuration (μ', μ, m);
- Let $\Gamma^{\mathrm{int}}_{<} = \{(\mu, \mu', m) \in \Gamma^{\mathrm{int}} \mid \mu < \mu'\}$, $\Gamma^{\mathrm{int}}_{=} = \{(\mu, \mu', m) \in \Gamma^{\mathrm{int}} \mid \mu = \mu'\}$ 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;
- $\delta^{\mathrm{C}}_{\mathrm{ac}}(i, [\nu]^{\mathrm{int}}) \in [0, 1], i \in [\widetilde{k_{\mathrm{C}}} + 1, m_{\mathrm{C}}] = I_{(\geq 1)} \cup I_{(0/1)} \cup I_{(=1)}, \nu \in \widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ac}}$, $\delta^{\mathrm{T}}_{\mathrm{ac}}(i, [\nu]^{\mathrm{int}}) \in [0, 1], i \in [2, t_{\mathrm{T}}], \nu \in \widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ac}}$, $\delta^{\mathrm{F}}_{\mathrm{ac}}(i, [\nu]^{\mathrm{int}}) \in [0, 1], i \in [2, t_{\mathrm{F}}], \nu \in \widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}$: $\delta^{\mathrm{X}}_{\mathrm{ac}}(i, [\nu]^{\mathrm{int}}) = 1 \Leftrightarrow$ edge ${e^{\mathrm{X}}}_i$ has adjacency-configuration ν;

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

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

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

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

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

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

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

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

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

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

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

constraints:

$$
\begin{aligned}
\mathrm{ac}_{\mathrm{C}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma^{\mathrm{int}}_{\mathrm{ac}} \setminus \widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ac}},\\
\mathrm{ac}_{\mathrm{T}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma^{\mathrm{int}}_{\mathrm{ac}} \setminus \widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ac}},\\
\mathrm{ac}_{\mathrm{F}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma^{\mathrm{int}}_{\mathrm{ac}} \setminus \widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}},\\
\mathrm{ac}_{\mathrm{CT}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma^{\mathrm{int}}_{\mathrm{ac}} \setminus \widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}},\\
\mathrm{ac}_{\mathrm{TC}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma^{\mathrm{int}}_{\mathrm{ac}} \setminus \widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}},\\
\mathrm{ac}_{\mathrm{CF}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma^{\mathrm{int}}_{\mathrm{ac}} \setminus \widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ac}},\\
\mathrm{ac}_{\mathrm{TF}}([\nu]^{\mathrm{int}}) &= 0, & \nu &\in \Gamma^{\mathrm{int}}_{\mathrm{ac}} \setminus \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}},
\end{aligned}
\tag{68}
$$

$$
\begin{aligned}
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{C}}([\nu]^{\mathrm{int}}) &= \sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]} \delta_{\beta}^{\mathrm{C}}(i,m), && m\in[1,3],\\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{T}}([\nu]^{\mathrm{int}}) &= \sum_{i\in[2,t_{\mathrm{T}}]} \delta_{\beta}^{\mathrm{T}}(i,m), && m\in[1,3],\\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{F}}([\nu]^{\mathrm{int}}) &= \sum_{i\in[2,t_{\mathrm{F}}]} \delta_{\beta}^{\mathrm{F}}(i,m), && m\in[1,3],\\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{CT}}([\nu]^{\mathrm{int}}) &= \sum_{k\in[1,k_{\mathrm{C}}]} \delta_{\beta}^{\mathrm{CT}}(k,m), && m\in[1,3],\\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{TC}}([\nu]^{\mathrm{int}}) &= \sum_{k\in[1,k_{\mathrm{C}}]} \delta_{\beta}^{\mathrm{TC}}(k,m), && m\in[1,3],\\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{CF}}([\nu]^{\mathrm{int}}) &= \sum_{c\in[1,\widetilde{t_{\mathrm{C}}}]} \delta_{\beta}^{*\mathrm{F}}(c,m), && m\in[1,3],\\
\sum_{(\mathsf{a},\mathsf{b},m)=\nu\in\Gamma_{\mathrm{ac}}^{\mathrm{int}}} \mathrm{ac}_{\mathrm{TF}}([\nu]^{\mathrm{int}}) &= \sum_{c\in[\widetilde{t_{\mathrm{C}}}+1,c_{\mathrm{F}}]} \delta_{\beta}^{*\mathrm{F}}(c,m), && m\in[1,3],
\end{aligned}
\tag{69}
$$

$$
\begin{aligned}
\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{C}}} m\cdot\delta_{\mathrm{ac}}^{\mathrm{C}}(i,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{C}}(i),\\
\Delta_{\mathrm{ac}}^{\mathrm{C}+}(i)+\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{C}}} [\mathsf{a}]^{\mathrm{int}}\delta_{\mathrm{ac}}^{\mathrm{C}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{tail}(i)),\\
\Delta_{\mathrm{ac}}^{\mathrm{C}-}(i)+\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{C}}} [\mathsf{b}]^{\mathrm{int}}\delta_{\mathrm{ac}}^{\mathrm{C}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{head}(i)),\\
\Delta_{\mathrm{ac}}^{\mathrm{C}+}(i)+\Delta_{\mathrm{ac}}^{\mathrm{C}-}(i) &\le 2|\Lambda^{\mathrm{int}}|(1-e^{\mathrm{C}}(i)), && i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}],\\
\sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]} \delta_{\mathrm{ac}}^{\mathrm{C}}(i,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{C}}([\nu]^{\mathrm{int}}), && \nu\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{C}},
\end{aligned}
\tag{70}
$$

$$
\begin{aligned}
\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{T}}} m\cdot\delta_{\mathrm{ac}}^{\mathrm{T}}(i,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{T}}(i),\\
\Delta_{\mathrm{ac}}^{\mathrm{T}+}(i)+\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{T}}} [\mathsf{a}]^{\mathrm{int}}\delta_{\mathrm{ac}}^{\mathrm{T}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{T}}(i-1),\\
\Delta_{\mathrm{ac}}^{\mathrm{T}-}(i)+\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{T}}} [\mathsf{b}]^{\mathrm{int}}\delta_{\mathrm{ac}}^{\mathrm{T}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{T}}(i),\\
\Delta_{\mathrm{ac}}^{\mathrm{T}+}(i)+\Delta_{\mathrm{ac}}^{\mathrm{T}-}(i) &\le 2|\Lambda^{\mathrm{int}}|(1-e^{\mathrm{T}}(i)), && i\in[2,t_{\mathrm{T}}],\\
\sum_{i\in[2,t_{\mathrm{T}}]} \delta_{\mathrm{ac}}^{\mathrm{T}}(i,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{T}}([\nu]^{\mathrm{int}}), && \nu\in\widetilde{\Gamma}_{\mathrm{ac}}^{\mathrm{T}},
\end{aligned}
\tag{71}
$$

$$
\begin{aligned}
\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}} m\cdot\delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{F}}(i),\\
\Delta^{\mathrm{F}+}_{\mathrm{ac}}(i)+\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}} [\mathsf{a}]^{\mathrm{int}}\delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{F}}(i-1),\\
\Delta^{\mathrm{F}-}_{\mathrm{ac}}(i)+\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}} [\mathsf{b}]^{\mathrm{int}}\delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{F}}(i),\\
\Delta^{\mathrm{F}+}_{\mathrm{ac}}(i)+\Delta^{\mathrm{F}-}_{\mathrm{ac}}(i) &\le 2|\Lambda^{\mathrm{ex}}|(1-e^{\mathrm{F}}(i)), && i\in[2,t_{\mathrm{F}}],\\
\sum_{i\in[2,t_{\mathrm{F}}]}\delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{F}}([\nu]^{\mathrm{int}}), && \nu\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}},
\end{aligned}
\tag{72}
$$

$$
\begin{aligned}
\alpha^{\mathrm{T}}(i)+|\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i)) &\ge \alpha^{\mathrm{CT}}(k),\\
\alpha^{\mathrm{CT}}(k) &\ge \alpha^{\mathrm{T}}(i)-|\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i)), && i\in[1,t_{\mathrm{T}}],\\
\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}} m\cdot\delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{CT}}(k),\\
\Delta^{\mathrm{CT}+}_{\mathrm{ac}}(k)+\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}} [\mathsf{a}]^{\mathrm{int}}\delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{tail}(k)),\\
\Delta^{\mathrm{CT}-}_{\mathrm{ac}}(k)+\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}} [\mathsf{b}]^{\mathrm{int}}\delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{CT}}(k),\\
\Delta^{\mathrm{CT}+}_{\mathrm{ac}}(k)+\Delta^{\mathrm{CT}-}_{\mathrm{ac}}(k) &\le 2|\Lambda^{\mathrm{int}}|(1-\delta^{\mathrm{T}}_{\chi}(k)), && k\in[1,k_{\mathrm{C}}],\\
\sum_{k\in[1,k_{\mathrm{C}}]}\delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{CT}}([\nu]^{\mathrm{int}}), && \nu\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}},
\end{aligned}
\tag{73}
$$

$$
\begin{aligned}
\alpha^{\mathrm{T}}(i)+|\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i+1)) &\ge \alpha^{\mathrm{TC}}(k),\\
\alpha^{\mathrm{TC}}(k) &\ge \alpha^{\mathrm{T}}(i)-|\Lambda^{\mathrm{int}}|(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i+1)), && i\in[1,t_{\mathrm{T}}],\\
\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}} m\cdot\delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \beta^{\mathrm{TC}}(k),\\
\Delta^{\mathrm{TC}+}_{\mathrm{ac}}(k)+\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}} [\mathsf{a}]^{\mathrm{int}}\delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{TC}}(k),\\
\Delta^{\mathrm{TC}-}_{\mathrm{ac}}(k)+\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}} [\mathsf{b}]^{\mathrm{int}}\delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{C}}(\mathrm{head}(k)),\\
\Delta^{\mathrm{TC}+}_{\mathrm{ac}}(k)+\Delta^{\mathrm{TC}-}_{\mathrm{ac}}(k) &\le 2|\Lambda^{\mathrm{int}}|(1-\delta^{\mathrm{T}}_{\chi}(k)), && k\in[1,k_{\mathrm{C}}],\\
\sum_{k\in[1,k_{\mathrm{C}}]}\delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}) &= \mathrm{ac}_{\mathrm{TC}}([\nu]^{\mathrm{int}}), && \nu\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}},
\end{aligned}
\tag{74}
$$

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

$$
\begin{aligned}
\alpha^{\mathrm{F}}(j) + |\Lambda^{\mathrm{int}}|(1 - \chi^{\mathrm{F}}(j, i + \widetilde{t_{\mathrm{C}}}) + e^{\mathrm{F}}(j)) &\geq \alpha^{\mathrm{TF}}(i), \\
\alpha^{\mathrm{TF}}(i) \geq \alpha^{\mathrm{F}}(j) - |\Lambda^{\mathrm{int}}|(1 - \chi^{\mathrm{F}}(j, i + \widetilde{t_{\mathrm{C}}}) + e^{\mathrm{F}}(j)), &\qquad j \in [1, t_{\mathrm{F}}], \\
\sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}} m \cdot \delta^{\mathrm{TF}}_{\mathrm{ac}}(i, [\nu]^{\mathrm{int}}) &= \beta^{*\mathrm{F}}(i + \widetilde{t_{\mathrm{C}}}), \\
\Delta^{\mathrm{TF}+}_{\mathrm{ac}}(i) + \sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}} [\mathsf{a}]^{\mathrm{int}} \delta^{\mathrm{TF}}_{\mathrm{ac}}(i, [\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{T}}(i), \\
\Delta^{\mathrm{TF}-}_{\mathrm{ac}}(i) + \sum_{\nu=(\mathsf{a},\mathsf{b},m)\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}} [\mathsf{b}]^{\mathrm{int}} \delta^{\mathrm{TF}}_{\mathrm{ac}}(i, [\nu]^{\mathrm{int}}) &= \alpha^{\mathrm{TF}}(i), \\
\Delta^{\mathrm{TF}+}_{\mathrm{ac}}(i) + \Delta^{\mathrm{TF}-}_{\mathrm{ac}}(i) \leq 2\max\{|\Lambda^{\mathrm{int}}|, |\Lambda^{\mathrm{int}}|\}(1 - \delta^{\mathrm{F}}_{\chi}(i + \widetilde{t_{\mathrm{C}}})), &\qquad i \in [1, t_{\mathrm{T}}], \\
\sum_{i\in[1,t_{\mathrm{T}}]} \delta^{\mathrm{TF}}_{\mathrm{ac}}(i, [\nu]^{\mathrm{int}}) = \mathrm{ac}_{\mathrm{TF}}([\nu]^{\mathrm{int}}), &\qquad \nu \in \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}, \qquad (76)
\end{aligned}
$$

$$
\begin{aligned}
\sum_{\mathrm{X}\in\{\mathrm{C,T,F,CT,TC,CF,TF}\}} (\mathrm{ac}_{\mathrm{X}}([\nu]^{\mathrm{int}}) + \mathrm{ac}_{\mathrm{X}}([\overline{\nu}]^{\mathrm{int}})) &= \mathrm{ac}^{\mathrm{int}}([\nu]^{\mathrm{int}}), &\qquad \nu \in \Gamma^{\mathrm{int}}_{\mathrm{ac},<}, \\
\sum_{\mathrm{X}\in\{\mathrm{C,T,F,CT,TC,CF,TF}\}} \mathrm{ac}_{\mathrm{X}}([\nu]^{\mathrm{int}}) &= \mathrm{ac}^{\mathrm{int}}([\nu]^{\mathrm{int}}), &\qquad \nu \in \Gamma^{\mathrm{int}}_{\mathrm{ac},=}. \qquad (77)
\end{aligned}
$$

D.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 an interior-vertex v in a chemical graph $\mathbb{C}$ to be inferred; i.e., $\mathrm{cs}(v) = \mu = \mathsf{a}d \in \Lambda_{\mathrm{dg}}$ such that $\alpha(v) = \mathsf{a}$ and $\deg_{\langle\mathbb{C}\rangle}(v) = \deg_H(v) - \deg^{\mathrm{hyd}}_{\mathbb{C}}(v) = d$ in $\mathbb{C} = (H, \alpha, \beta)$.

constants:

- A set $\Lambda^{\mathrm{int}}_{\mathrm{dg}}$ 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 $\widetilde{\Lambda}_{\mathrm{dg}}^{\mathrm{C}}, \widetilde{\Lambda}_{\mathrm{dg}}^{\mathrm{T}}, \widetilde{\Lambda}_{\mathrm{dg}}^{\mathrm{F}} \subseteq \Lambda_{\mathrm{dg}}^{\mathrm{int}}$: To compute the frequency of chemical symbols exactly, set $\widetilde{\Lambda}_{\mathrm{dg}}^{\mathrm{C}} := \widetilde{\Lambda}_{\mathrm{dg}}^{\mathrm{T}} := \widetilde{\Lambda}_{\mathrm{dg}}^{\mathrm{F}} := \Lambda_{\mathrm{dg}}^{\mathrm{int}}$;

variables:

- $\mathrm{ns}^{\mathrm{int}}([\mu]^{\mathrm{int}}) \in [0, \mathrm{n}_{\mathrm{UB}}^{\mathrm{int}}]$, $\mu \in \Lambda_{\mathrm{dg}}^{\mathrm{int}}$: the number of interior-vertices v with $\mathrm{cs}(v) = \mu$;

- $\delta_{\mathrm{ns}}^{\mathrm{X}}(i, [\mu]^{\mathrm{int}}) \in [0, 1]$, $i \in [1, t_{\mathrm{X}}], \mu \in \Lambda_{\mathrm{dg}}^{\mathrm{int}}$, $\mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}$;

constraints:

$$\sum_{\mu \in \widetilde{\Lambda}_{\mathrm{dg}}^{\mathrm{X}} \cup \{\epsilon\}} \delta_{\mathrm{ns}}^{\mathrm{X}}(i, [\mu]^{\mathrm{int}}) = 1, \quad \sum_{\mu = \mathtt{a}d \in \widetilde{\Lambda}_{\mathrm{dg}}^{\mathrm{X}}} [\mathtt{a}]^{\mathrm{int}} \cdot \delta_{\mathrm{ns}}^{\mathrm{X}}(i, [\mu]^{\mathrm{int}}) = \alpha^{\mathrm{X}}(i),$$
$$\sum_{\mu = \mathtt{a}d \in \widetilde{\Lambda}_{\mathrm{dg}}^{\mathrm{X}}} d \cdot \delta_{\mathrm{ns}}^{\mathrm{X}}(i, [\mu]^{\mathrm{int}}) = \deg^{\mathrm{X}}(i),$$
$$i \in [1, t_{\mathrm{X}}], \mathrm{X} \in \{\mathrm{C}, \mathrm{T}, \mathrm{F}\}, \tag{78}$$

$$\sum_{i \in [1, t_{\mathrm{C}}]} \delta_{\mathrm{ns}}^{\mathrm{C}}(i, [\mu]^{\mathrm{int}}) + \sum_{i \in [1, t_{\mathrm{T}}]} \delta_{\mathrm{ns}}^{\mathrm{T}}(i, [\mu]^{\mathrm{int}}) + \sum_{i \in [1, t_{\mathrm{F}}]} \delta_{\mathrm{ns}}^{\mathrm{F}}(i, [\mu]^{\mathrm{int}}) = \mathrm{ns}^{\mathrm{int}}([\mu]^{\mathrm{int}}), \qquad \mu \in \Lambda_{\mathrm{dg}}^{\mathrm{int}}. \tag{79}$$

D.10 Descriptor for the Number of Edge-configurations

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

constants:

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

- Let $\Gamma_{<}^{\mathrm{int}} = \{(\mu, \mu', m) \in \Gamma^{\mathrm{int}} \mid \mu < \mu'\}$, $\Gamma_{=}^{\mathrm{int}} = \{(\mu, \mu', m) \in \Gamma^{\mathrm{int}} \mid \mu = \mu'\}$ and $\Gamma_{>}^{\mathrm{int}} = \{(\mu', \mu, m) \mid (\mu, \mu', m) \in \Gamma_{<}^{\mathrm{int}}\}$;

- Prepare a coding of the set $\Gamma^{\mathrm{int}} \cup \Gamma_{>}^{\mathrm{int}}$ and let $[\gamma]^{\mathrm{int}}$ denote the coded integer of an element γ in $\Gamma^{\mathrm{int}} \cup \Gamma_{>}^{\mathrm{int}}$;

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

- $\mathrm{ec}_{\mathrm{LB}}^{\mathrm{int}}(\gamma), \mathrm{ec}_{\mathrm{UB}}^{\mathrm{int}}(\gamma) \in [0, 2\mathrm{n}_{\mathrm{UB}}^{\mathrm{int}}], \gamma = (\mu, \mu', m) \in \Gamma^{\mathrm{int}}$: lower and upper bounds on the number of interior-edges $e = uv$ with $\mathrm{cs}(u) = \mu$, $\mathrm{cs}(v) = \mu'$ and $\beta(e) = m$;

variables:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

constraints:

$$\begin{aligned}
\mathrm{ec}_{\mathrm{C}}([\gamma]^{\mathrm{int}}) &= 0, && \gamma \in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ec}},\\
\mathrm{ec}_{\mathrm{T}}([\gamma]^{\mathrm{int}}) &= 0, && \gamma \in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ec}},\\
\mathrm{ec}_{\mathrm{F}}([\gamma]^{\mathrm{int}}) &= 0, && \gamma \in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ec}},\\
\mathrm{ec}_{\mathrm{CT}}([\gamma]^{\mathrm{int}}) &= 0, && \gamma \in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ec}},\\
\mathrm{ec}_{\mathrm{TC}}([\gamma]^{\mathrm{int}}) &= 0, && \gamma \in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}},\\
\mathrm{ec}_{\mathrm{CF}}([\gamma]^{\mathrm{int}}) &= 0, && \gamma \in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}},\\
\mathrm{ec}_{\mathrm{TF}}([\gamma]^{\mathrm{int}}) &= 0, && \gamma \in \Gamma^{\mathrm{int}} \setminus \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}},
\end{aligned} \tag{80}$$

$$\begin{aligned}
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{C}}([\gamma]^{\mathrm{int}}) &= \sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]} \delta^{\mathrm{C}}_{\beta}(i,m), && m\in[1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{T}}([\gamma]^{\mathrm{int}}) &= \sum_{i\in[2,t_{\mathrm{T}}]} \delta^{\mathrm{T}}_{\beta}(i,m), && m\in[1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{F}}([\gamma]^{\mathrm{int}}) &= \sum_{i\in[2,t_{\mathrm{F}}]} \delta^{\mathrm{F}}_{\beta}(i,m), && m\in[1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{CT}}([\gamma]^{\mathrm{int}}) &= \sum_{k\in[1,k_{\mathrm{C}}]} \delta^{\mathrm{CT}}_{\beta}(k,m), && m\in[1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{TC}}([\gamma]^{\mathrm{int}}) &= \sum_{k\in[1,k_{\mathrm{C}}]} \delta^{\mathrm{TC}}_{\beta}(k,m), && m\in[1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{CF}}([\gamma]^{\mathrm{int}}) &= \sum_{c\in[1,\widetilde{t_{\mathrm{C}}}]} \delta^{*\mathrm{F}}_{\beta}(c,m), && m\in[1,3],\\
\sum_{(\mu,\mu',m)=\gamma\in\Gamma^{\mathrm{int}}} \mathrm{ec}_{\mathrm{TF}}([\gamma]^{\mathrm{int}}) &= \sum_{c\in[\widetilde{t_{\mathrm{C}}}+1,c_{\mathrm{F}}]} \delta^{*\mathrm{F}}_{\beta}(c,m), && m\in[1,3],
\end{aligned} \tag{81}$$

$$\begin{aligned}
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ec}}} [(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}}\cdot\delta^{\mathrm{C}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}}) &= \sum_{\nu\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ac}}} [\nu]^{\mathrm{int}}\cdot\delta^{\mathrm{C}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}),\\
\Delta^{\mathrm{C}+}_{\mathrm{ec}}(i) + \sum_{\gamma=(\mathtt{a}d,\mu',m)\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ec}}} d\cdot\delta^{\mathrm{C}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{C}}(\mathrm{tail}(i)),\\
\Delta^{\mathrm{C}-}_{\mathrm{ec}}(i) + \sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ec}}} d\cdot\delta^{\mathrm{C}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{C}}(\mathrm{head}(i)),\\
\Delta^{\mathrm{C}+}_{\mathrm{ec}}(i) + \Delta^{\mathrm{C}-}_{\mathrm{ec}}(i) &\le 8(1-e^{\mathrm{C}}(i)), && i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}],\\
\sum_{i\in[\widetilde{k_{\mathrm{C}}}+1,m_{\mathrm{C}}]} \delta^{\mathrm{C}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{C}}([\gamma]^{\mathrm{int}}), && \gamma\in\widetilde{\Gamma}^{\mathrm{C}}_{\mathrm{ec}},
\end{aligned} \tag{82}$$

$$\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ec}}}[(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}}\cdot\delta^{\mathrm{T}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}})=\sum_{\nu\in\widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ac}}}[\nu]^{\mathrm{int}}\cdot\delta^{\mathrm{T}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}),$$
$$\Delta^{\mathrm{T}+}_{\mathrm{ec}}(i)+\sum_{\gamma=(\mathtt{a}d,\mu',m)\in\widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ec}}}d\cdot\delta^{\mathrm{T}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}})=\deg^{\mathrm{T}}(i-1),$$
$$\Delta^{\mathrm{T}-}_{\mathrm{ec}}(i)+\sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ec}}}d\cdot\delta^{\mathrm{T}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}})=\deg^{\mathrm{T}}(i),$$
$$\Delta^{\mathrm{T}+}_{\mathrm{ec}}(i)+\Delta^{\mathrm{T}-}_{\mathrm{ec}}(i)\le 8(1-e^{\mathrm{T}}(i)),\qquad i\in[2,t_{\mathrm{T}}],$$
$$\sum_{i\in[2,t_{\mathrm{T}}]}\delta^{\mathrm{T}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}})=\mathrm{ec}_{\mathrm{T}}([\gamma]^{\mathrm{int}}),\qquad \gamma\in\widetilde{\Gamma}^{\mathrm{T}}_{\mathrm{ec}},\tag{83}$$

$$\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ec}}}[(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}}\cdot\delta^{\mathrm{F}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}})=\sum_{\nu\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ac}}}[\nu]^{\mathrm{int}}\cdot\delta^{\mathrm{F}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}),$$
$$\Delta^{\mathrm{F}+}_{\mathrm{ec}}(i)+\sum_{\gamma=(\mathtt{a}d,\mu',m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ec}}}d\cdot\delta^{\mathrm{F}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}})=\deg^{\mathrm{F}}(i-1),$$
$$\Delta^{\mathrm{F}-}_{\mathrm{ec}}(i)+\sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ec}}}d\cdot\delta^{\mathrm{F}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}})=\deg^{\mathrm{F}}(i,0),$$
$$\Delta^{\mathrm{F}+}_{\mathrm{ec}}(i)+\Delta^{\mathrm{F}-}_{\mathrm{ec}}(i)\le 8(1-e^{\mathrm{F}}(i)),\qquad i\in[2,t_{\mathrm{F}}],$$
$$\sum_{i\in[2,t_{\mathrm{F}}]}\delta^{\mathrm{F}}_{\mathrm{ec}}(i,[\gamma]^{\mathrm{int}})=\mathrm{ec}_{\mathrm{F}}([\gamma]^{\mathrm{int}}),\qquad \gamma\in\widetilde{\Gamma}^{\mathrm{F}}_{\mathrm{ec}},\tag{84}$$

$$\deg^{\mathrm{T}}(i)+4(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i))\ge\deg^{\mathrm{CT}}_{\mathrm{T}}(k),$$
$$\deg^{\mathrm{CT}}_{\mathrm{T}}(k)\ge\deg^{\mathrm{T}}(i)-4(1-\chi^{\mathrm{T}}(i,k)+e^{\mathrm{T}}(i)),\qquad i\in[1,t_{\mathrm{T}}],$$
$$\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ec}}}[(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}}\cdot\delta^{\mathrm{CT}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}})=\sum_{\nu\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ac}}}[\nu]^{\mathrm{int}}\cdot\delta^{\mathrm{CT}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}),$$
$$\Delta^{\mathrm{CT}+}_{\mathrm{ec}}(k)+\sum_{\gamma=(\mathtt{a}d,\mu',m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ec}}}d\cdot\delta^{\mathrm{CT}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}})=\deg^{\mathrm{C}}(\mathrm{tail}(k)),$$
$$\Delta^{\mathrm{CT}-}_{\mathrm{ec}}(k)+\sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ec}}}d\cdot\delta^{\mathrm{CT}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}})=\deg^{\mathrm{CT}}_{\mathrm{T}}(k),$$
$$\Delta^{\mathrm{CT}+}_{\mathrm{ec}}(k)+\Delta^{\mathrm{CT}-}_{\mathrm{ec}}(k)\le 8(1-\delta^{\mathrm{T}}_{\chi}(k)),\qquad k\in[1,k_{\mathrm{C}}],$$
$$\sum_{k\in[1,k_{\mathrm{C}}]}\delta^{\mathrm{CT}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}})=\mathrm{ec}_{\mathrm{CT}}([\gamma]^{\mathrm{int}}),\qquad \gamma\in\widetilde{\Gamma}^{\mathrm{CT}}_{\mathrm{ec}},\tag{85}$$

$$
\begin{aligned}
\deg^{\mathrm{T}}(i) + 4(1 - \chi^{\mathrm{T}}(i,k) + e^{\mathrm{T}}(i+1)) &\geq \deg^{\mathrm{TC}}_{\mathrm{T}}(k),\\
\deg^{\mathrm{TC}}_{\mathrm{T}}(k) &\geq \deg^{\mathrm{T}}(i) - 4(1 - \chi^{\mathrm{T}}(i,k) + e^{\mathrm{T}}(i+1)), \qquad i \in [1, t_{\mathrm{T}}],\\
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}}} [(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}} \cdot \delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}}) &= \sum_{\nu\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ac}}} [\nu]^{\mathrm{int}} \cdot \delta^{\mathrm{TC}}_{\mathrm{ac}}(k,[\nu]^{\mathrm{int}}),\\
\Delta^{\mathrm{TC}+}_{\mathrm{ec}}(k) + \sum_{\gamma=(\mathtt{a}d,\mu',m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{TC}}_{\mathrm{T}}(k),\\
\Delta^{\mathrm{TC}-}_{\mathrm{ec}}(k) + \sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{C}}(\mathrm{head}(k)),\\
\Delta^{\mathrm{TC}+}_{\mathrm{ec}}(k) + \Delta^{\mathrm{TC}-}_{\mathrm{ec}}(k) &\leq 8(1 - \delta^{\mathrm{T}}_{\chi}(k)), \qquad k \in [1, k_{\mathrm{C}}],\\
\sum_{k\in[1,k_{\mathrm{C}}]} \delta^{\mathrm{TC}}_{\mathrm{ec,C}}(k,[\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{TC}}([\gamma]^{\mathrm{int}}), \qquad \gamma \in \widetilde{\Gamma}^{\mathrm{TC}}_{\mathrm{ec}},
\end{aligned} \tag{86}
$$

$$
\begin{aligned}
\deg^{\mathrm{F}}(i) + 4(1 - \chi^{\mathrm{F}}(i,c) + e^{\mathrm{F}}(i)) &\geq \deg^{\mathrm{CF}}_{\mathrm{F}}(c),\\
\deg^{\mathrm{CF}}_{\mathrm{F}}(c) &\geq \deg^{\mathrm{F}}(i) - 4(1 - \chi^{\mathrm{F}}(i,c) + e^{\mathrm{F}}(i)), \qquad i \in [1, t_{\mathrm{F}}],\\
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}}} [(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}} \cdot \delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c,[\gamma]^{\mathrm{int}}) &= \sum_{\nu\in\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ac}}} [\nu]^{\mathrm{int}} \cdot \delta^{\mathrm{CF}}_{\mathrm{ac}}(c,[\nu]^{\mathrm{int}}),\\
\Delta^{\mathrm{CF}+}_{\mathrm{ec}}(c) + \sum_{\gamma=(\mathtt{a}d,\mu',m)\in\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{C}}(c),\\
\Delta^{\mathrm{CF}-}_{\mathrm{ec}}(c) + \sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{CF}}_{\mathrm{F}}(c),\\
\Delta^{\mathrm{CF}+}_{\mathrm{ec}}(c) + \Delta^{\mathrm{CF}-}_{\mathrm{ec}}(c) &\leq 8(1 - \delta^{\mathrm{F}}_{\chi}(c)), \qquad c \in [1, \widetilde{t_{\mathrm{C}}}],\\
\sum_{c\in[1,\widetilde{t_{\mathrm{C}}}]} \delta^{\mathrm{CF}}_{\mathrm{ec,C}}(c,[\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{CF}}([\gamma]^{\mathrm{int}}), \qquad \gamma \in \widetilde{\Gamma}^{\mathrm{CF}}_{\mathrm{ec}},
\end{aligned} \tag{87}
$$

$$
\begin{aligned}
\deg^{\mathrm{F}}(j) + 4(1 - \chi^{\mathrm{F}}(j, i+\widetilde{t_{\mathrm{C}}}) + e^{\mathrm{F}}(j)) &\geq \deg^{\mathrm{TF}}_{\mathrm{F}}(i),\\
\deg^{\mathrm{TF}}_{\mathrm{F}}(i) &\geq \deg^{\mathrm{F}}(j) - 4(1 - \chi^{\mathrm{F}}(j, i+\widetilde{t_{\mathrm{C}}}) + e^{\mathrm{F}}(j)), \qquad j \in [1, t_{\mathrm{F}}],\\
\sum_{\gamma=(\mathtt{a}d,\mathtt{b}d',m)\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}}} [(\mathtt{a},\mathtt{b},m)]^{\mathrm{int}} \cdot \delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i,[\gamma]^{\mathrm{int}}) &= \sum_{\nu\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ac}}} [\nu]^{\mathrm{int}} \cdot \delta^{\mathrm{TF}}_{\mathrm{ac}}(i,[\nu]^{\mathrm{int}}),\\
\Delta^{\mathrm{TF}+}_{\mathrm{ec}}(i) + \sum_{\gamma=(\mathtt{a}d,\mu',m)\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{T}}(i),\\
\Delta^{\mathrm{TF}-}_{\mathrm{ec}}(i) + \sum_{\gamma=(\mu,\mathtt{b}d,m)\in\widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}}} d \cdot \delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i,[\gamma]^{\mathrm{int}}) &= \deg^{\mathrm{TF}}_{\mathrm{F}}(i),\\
\Delta^{\mathrm{TF}+}_{\mathrm{ec}}(i) + \Delta^{\mathrm{TF}-}_{\mathrm{ec}}(i) &\leq 8(1 - \delta^{\mathrm{F}}_{\chi}(i+\widetilde{t_{\mathrm{C}}})), \qquad i \in [1, t_{\mathrm{T}}],\\
\sum_{i\in[1,t_{\mathrm{T}}]} \delta^{\mathrm{TF}}_{\mathrm{ec,T}}(i,[\gamma]^{\mathrm{int}}) &= \mathrm{ec}_{\mathrm{TF}}([\gamma]^{\mathrm{int}}), \qquad \gamma \in \widetilde{\Gamma}^{\mathrm{TF}}_{\mathrm{ec}},
\end{aligned} \tag{88}
$$

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

$$\sum_{\mathrm{X}\in\{\mathrm{C,T,F,CT,TC,CF,TF}\}} \mathrm{ec}_{\mathrm{X}}([\gamma]^{\mathrm{int}}) = \mathrm{ec}^{\mathrm{int}}([\gamma]^{\mathrm{int}}), \qquad \gamma \in \Gamma^{\mathrm{int}}_{=}. \tag{89}$$

D.11 Constraints for Prediction Functions

This section introduces an MILP that simulates the computation process of a prediction function constructed with elastic linear regression.

Let $x = (x(1), x(2), \ldots, x(K)) \in \mathbb{R}^K$ denote the feature function $f(\mathbb{C})$ of a chemical graph $\mathbb{C}$. Let $c_{\min}(j)$ (resp., $c_{\max}(j)$) denote the minimum (resp., maximum) values of the j-th descriptor in a data set D_π for a chemical property π. Let $\mathrm{atm}_{\mathrm{LB}} \in \mathbb{Z}_+$ (resp., $\mathrm{atm}_{\mathrm{UB}} \in \mathbb{Z}_+$) be a lower bound (resp., an upper bound) on the number of atoms in a chemical graph $\mathbb{C}$ to be inferred. Let $\mathrm{mass}(\mathtt{a})$ denote the observed mass of a chemical element $\mathtt{a} \in \Lambda$, and define $\mathrm{mass}^*(\mathtt{a}) \triangleq \lfloor 10 \cdot \mathrm{mass}(\mathtt{a}) \rfloor$. Let $\mathrm{Ms}_{\mathrm{LB}} \in \mathbb{Z}_+$ (resp., $\mathrm{Ms}_{\mathrm{UB}} \in \mathbb{Z}_+$) be a lower bound (resp., an upper bound) on the sum $\frac{1}{|V(H)|}\sum_{v\in V(H)} \mathrm{mass}^*(\alpha(v))$ in a chemical graph $\mathbb{C}$ to be inferred. Let j_{ms} denote the index $j \in [1, K]$ such that the j-th descriptor $\mathrm{dcp}_j(\mathbb{C})$ is the average mass $\overline{\mathrm{ms}}(\mathbb{C}) = \frac{1}{|V(H)|}\sum_{v\in V(H)} \mathrm{mass}^*(\alpha(v))$. Assume that all other descriptors $\mathrm{dcp}_j(\mathbb{C}), j \in [1, K] \setminus \{j_{\mathrm{ms}}\}$ are integers.

Let $\eta_{\Psi,w,b}$ be a prediction function obtained by elastic linear regression, where $\Psi = \{\phi_j \mid j \in [0, K]\}$ and (w, b) is a hyperplane.

We first normalize each of the sets of descriptors $x(j), j \in [1, K]$ and the set of observed values $a(\mathbb{C})$ before we apply the prediction function to compute a predicted value $\eta_{\Psi,w,b}(f(\mathbb{C}))$ of a chemical graph $\mathbb{C}$, where the set $\{a_i \mid i \in [1, m]\}$ of observed values in the data set D_π is converted into a set $\{\phi_0^{-1}(\frac{a_i - \underline{a}}{\overline{a} - \underline{a}}) \mid i \in [1, m]\}$, where $\underline{a}$ (resp., $\overline{a}$) denotes the minimum (resp., maximum) value of $a(\mathbb{C})$ over the chemical graphs $\mathbb{C} \in D_\pi$.

Let $\underline{y}^*$ and $\overline{y}^*$ be lower and upper bounds on the predicted value
$eta_{\Psi,w,b}(f(\mathbb{C}))$ of a target chemical graph $\mathbb{C}$, respectively.

We first converted them into $\phi_0^{-1}(\frac{\underline{y}^* - \underline{a}}{\overline{a} - \underline{a}})$ and $\phi_0^{-1}(\frac{\overline{y}^* - \underline{a}}{\overline{a} - \underline{a}})$. We denote by $\psi_j(t)$ the function $\phi_j((t - c_{\min}(j))/(c_{\max}(j) - c_{\min}(j)))$. We pre-compute the values $\psi_j(s)$ for all integers $s \in [c_{\min}(j), c_{\max}(j)]$, $j \in [1, K] \setminus \{j_{\mathrm{ms}}\}$ (resp., $\phi_{j_{\mathrm{ms}}}(\frac{s/i - c_{\min}(j_{\mathrm{ms}})}{c_{\max}(j_{\mathrm{ms}}) - c_{\min}(j_{\mathrm{ms}})})$ for all integers $s \in [\mathrm{Ms}_{\mathrm{LB}}, \mathrm{Ms}_{\mathrm{UB}}]$ and $i \in [\mathrm{atm}_{\mathrm{LB}}, \mathrm{atm}_{\mathrm{UB}}]$) as constants.

An MILP that simulates the computation process of a prediction function $\eta_{\Psi,w,b}$ is described as follows.

$\mathcal{M}(x, y; \mathcal{C}_1)$**:**
constants:

- A hyperplane (w, b) with $w \in \mathbb{R}^K$ and $b \in \mathbb{R}$;
- Activation functions $\phi_j : \mathbb{R} \to \mathbb{R}, j \in [0, K]$;
- Real values $\underline{y}^*, \overline{y}^* \in \mathbb{R}$ such that $\underline{y}^* < \overline{y}^*$; Set $\underline{y}^{**} := \phi_0^{-1}(\frac{\underline{y}^* - \underline{a}}{\overline{a} - \underline{a}})$ and $\overline{y}^{**} := \phi_0^{-1}(\frac{\overline{y}^* - \underline{a}}{\overline{a} - \underline{a}})$.
- $c_{\min}(j), c_{\max}(j) \in \mathbb{R}, j \in [1, K]$: the minimum and maximum values of the j-th descriptor in the data set D_π, respectively;

- Reals $\Delta(j,s) \in \mathbb{R}, j \in [1,K] \setminus \{j_{\rm ms}\}, s \in [c_{\min}(j), c_{\max}(j)]$: $\Delta(j, c_{\min}(j)) := \psi_j(c_{\min}(j))$ and $\Delta(j,s) := \psi_j(s) - \psi_j(s-1), s \in [c_{\min}(j)+1, c_{\max}(j)]$;
- $\mathrm{atm}_{\rm LB}, \mathrm{atm}_{\rm UB} \in \mathbb{Z}_+$: lower and upper bounds on the number of atoms in a chemical graph $\mathbb{C}$ to be inferred; Set $\mathrm{atm}_{\rm LB} := n_{\rm LB} + \mathrm{na}_{\rm LB}(\mathtt{H})$ and $\mathrm{atm}_{\rm UB} := n^* + \mathrm{na}_{\rm UB}(\mathtt{H})$;
- $\mathrm{Ms}_{\rm LB}, \mathrm{Ms}_{\rm UB} \in \mathbb{Z}_+$: lower and upper bounds on the sum $\sum_{v \in V(H)} \mathrm{mass}^*(\alpha(v))$; For example, set $\mathrm{Ms}_{\rm LB} := \lfloor \min\{\mathrm{mass}^*([\mathtt{a}]) \mid \mathtt{a} \in \Lambda, \mathrm{val}(a) = 1\} \cdot (3n_{\rm LB}/4) + \min\{\mathrm{mass}^*([\mathtt{a}]) \mid \mathtt{a} \in \Lambda, \mathrm{val}(a) \geq 2\} \cdot (n_{\rm LB}/4) + \mathrm{mass}^*([\mathtt{H}])\mathrm{na}_{\rm LB}(\mathtt{H}) \rfloor$ and $\mathrm{Ms}_{\rm UB} := n^* \max\{\mathrm{mass}^*([\mathtt{a}]) \mid \mathtt{a} \in \Lambda\} + \mathrm{mass}^*([\mathtt{H}])\mathrm{na}_{\rm UB}(\mathtt{H})$;
- $M \in \mathbb{R}_+$: an upper bound on $\zeta(x(j_{\rm ms}))$; For example, set $M := 2\phi_{j_{\rm ms}}(1)$;
- Define $\zeta(s,i) \triangleq \phi_{j_{\rm ms}}(\frac{s/i - c_{\min}(j_{\rm ms})}{c_{\max}(j_{\rm ms}) - c_{\min}(j_{\rm ms})}), s \in [\mathrm{Ms}_{\rm LB}, \mathrm{Ms}_{\rm UB}], i \in [\mathrm{atm}_{\rm LB}, \mathrm{atm}_{\rm UB}]$, where $\psi_{j_{\rm ms}}(s/i) = \zeta(s,i)$;
- Reals $\Delta_{\rm Ms}(s,i) \in \mathbb{R}, s \in [\mathrm{Ms}_{\rm LB}, \mathrm{Ms}_{\rm UB}], i \in [\mathrm{atm}_{\rm LB}, \mathrm{atm}_{\rm UB}]$: $\Delta_{\rm Ms}(\mathrm{Ms}_{\rm LB}, i) := \zeta(\mathrm{Ms}_{\rm LB}, i)$ and $\Delta_{\rm Ms}(s,i) := \zeta(s,i) - \zeta(s-1,i), s \in [\mathrm{Ms}_{\rm LB}+1, \mathrm{Ms}_{\rm UB}], i \in [\mathrm{atm}_{\rm LB}, \mathrm{atm}_{\rm UB}]$;
- A real $\varepsilon(j) > 0, j \in [1,K]$: a tolerance. For example, set $\varepsilon(j) := \frac{1}{10^5}\min\{\Delta(j,s) \mid s \in [c_{\min}(j), c_{\max}(j)]\}, j \in [1,K] \setminus \{j_{\rm ms}\}$ and $\varepsilon(j_{\rm ms}) := \frac{1}{10^5}\min\{\Delta_{\rm Ms}(s,i) \mid s \in [\mathrm{Ms}_{\rm LB}, \mathrm{Ms}_{\rm UB}], i \in [\mathrm{atm}_{\rm LB}, \mathrm{atm}_{\rm UB}]\}$:

variables:

- Real variables $\widehat{x}(j) \in \mathbb{R}, j \in [1,K]$: $\widehat{x}(j)$ represents $\psi_j(x(j))$;
- Integer variables $x(j) \in [c_{\min}(j), c_{\max}(j)], j \in [1,K] \setminus \{j_{\rm ms}\}$: $x(j)$ represents the j-th descriptor in an MILP $\mathcal{M}(x, g; \mathcal{C}_2)$;
- A real variable $x(j_{\rm ms}) \in \mathbb{R}_+$ with $c_{\min}(j_{\rm ms}) \leq x(j_{\rm ms}) \leq c_{\max}(j_{\rm ms})$: $x(j_{\rm ms})$ represents the average mass $\overline{\rm ms}(\mathbb{C})$ in an MILP $\mathcal{M}(x, g; \mathcal{C}_2)$;
- Binary variables $\delta(j,s) \in [0,1], j \in [1,K] \setminus \{j_{\rm ms}\}, s \in [c_{\min}(j), c_{\max}(j)]$: $\delta(j,s) = 1 \Leftrightarrow x(j) \geq s$;
- Binary variables $\delta_{\rm atm}(i) \in [0,1], i \in [\mathrm{atm}_{\rm LB}, \mathrm{atm}_{\rm UB}]$: $\delta_{\rm atm}(i) = 1 \Leftrightarrow |V(H)| = i$;
- Binary variables $\delta_{\rm Ms}(s) \in [0,1], s \in [\mathrm{Ms}_{\rm LB}, \mathrm{Ms}_{\rm UB}]$ $\delta_{\rm Ms}(i) = 1 \Leftrightarrow \sum_{v \in V(H)} \mathrm{mass}^*(\alpha(v)) \geq s$;

constraints:

$$\underline{y}^{**} \leq \sum_{j \in [1,K]} w(j)\widehat{x}(j) + b \leq \overline{y}^{**}, \tag{90}$$

$$\begin{gathered}
\sum_{s \in [c_{\min}(j), c_{\max}(j)]} \delta(j,s) + c_{\min}(j) - 1 = x(j), \\
\delta(j,s) \geq \delta(j,s+1), \quad s \in [c_{\min}(j), c_{\max}(j) - 1], \\
\sum_{s \in [c_{\min}(j), c_{\max}(j)]} \Delta(j,s)\delta(j,s) - \epsilon(j) \leq \widehat{x}(j) \leq \sum_{s \in [c_{\min}(j), c_{\max}(j)]} \Delta(j,s)\delta(j,s) + \epsilon(j), \quad j \in [1,K] \setminus \{j_{\rm ms}\},
\end{gathered} \tag{91}$$

$$\sum_{i\in[\mathrm{atm_{LB}},\mathrm{atm_{UB}}]} \delta_{\mathrm{atm}}(i) = 1,$$

$$\sum_{i\in[\mathrm{atm_{LB}},\mathrm{atm_{UB}}]} i\cdot\delta_{\mathrm{atm}}(i) = n_G + \mathrm{na}^{\mathrm{ex}}([\mathtt{H}]^{\mathrm{ex}}),$$

$$\sum_{\mathtt{a}\in\Lambda} \mathrm{mass}^*(\mathtt{a})\cdot\mathrm{na}([\mathtt{a}]) = \sum_{s\in[\mathrm{Ms_{LB}},\mathrm{Ms_{UB}}]} \delta_{\mathrm{Ms}}(s) + \mathrm{Ms_{LB}} - 1, \tag{92}$$

$$\delta_{\mathrm{Ms}}(s) \geq \delta_{\mathrm{Ms}}(s+1), \quad s\in[\mathrm{Ms_{LB}}, \mathrm{Ms_{UB}}-1], \tag{93}$$

$$\sum_{s\in[\mathrm{Ms_{LB}},\mathrm{Ms_{UB}}]} \Delta_{\mathrm{Ms}}(s,i)\delta_{\mathrm{Ms}}(s) - M\cdot(1–\delta_{\mathrm{atm}}(i)) - \epsilon(j_{\mathrm{ms}}) \leq \widehat{x}(j_{\mathrm{ms}}) \leq$$

$$\sum_{s\in[\mathrm{Ms_{LB}},\mathrm{Ms_{UB}}]} \Delta_{\mathrm{Ms}}(s,i)\delta_{\mathrm{Ms}}(s) + M\cdot(1–\delta_{\mathrm{atm}}(i)) + \epsilon(j_{\mathrm{ms}}), \qquad i\in[\mathrm{atm_{LB}}, \mathrm{atm_{UB}}]. \tag{94}$$