Preference Optimization for Molecule Synthesis with Conditional Residual Energy-based Models

Songtao Liu [1] **Hanjun Dai** [2] **Yue Zhao** [3] **Peng Liu** [1]

Abstract

Molecule synthesis through machine learning is one of the fundamental problems in drug discovery. Current data-driven strategies employ one-step retrosynthesis models and search algorithms to predict synthetic routes in a top-bottom manner. Despite their effective performance, these strategies face limitations in the molecule synthetic route generation due to a greedy selection of the next molecule set without any lookahead. Furthermore, existing strategies cannot control the generation of synthetic routes based on possible criteria such as material costs, yields, and step count. In this work, we propose a general and principled framework via conditional residual energy-based models (EBMs), that focus on the quality of the entire synthetic route based on the specific criteria. By incorporating an additional energy-based function into our probabilistic model, our proposed algorithm can enhance the quality of the most probable synthetic routes (with higher probabilities) generated by various strategies in a plug-and-play fashion. Extensive experiments demonstrate that our framework can consistently boost performance across various strategies and outperforms previous state-of-the-art top-1 accuracy by a margin of 2.5%. Code is available at `https://github.com/SongtaoLiu0823/CREBM`.

1. Introduction

Machine learning for molecule synthesis is crucial for drug discovery, as it ensures that the new molecule can be integrated into industrial-scale chemical manufacturing processes. Existing data-driven strategies employ retrosynthesis prediction, a concept introduced by Corey & Wipke (1969) and further developed in Corey (1991), to derive precursors from a product, which formalizes multi-step retrosynthetic planning. During each planning step, the one-step retrosynthesis model (Dai et al., 2019; Chen et al., 2020; Chen & Jung, 2021) generates multiple reactant sets. The search algorithm (Segler et al., 2018; Chen et al., 2020) then chooses promising sets to further expand synthetic routes. This process is repeated until all molecules at the leaf nodes of the synthetic route[1] are commercially available starting materials.

Retrosynthetic planning is still a very challenging problem, which requires us to consider the quality of synthetic routes based on specific criteria. According to Chapter 8 in Hoffmann (2009), possible criteria for evaluating synthetic routes include: "***1.*** *the shortest route (time involved);* ***2.*** *the cheapest route (cost of materials);* ***3.*** *the novelty of the route (patentability);* ***4.*** *the greenest route (avoidance of problematic waste)*". However, current evaluation metrics (Chen et al., 2020; Maziarz et al., 2023) often neglect these criteria, primarily focusing on assessing a strategy's performance by the proportion of molecules for which the shortest possible synthetic route, with all leaf nodes as starting materials, can be found within a specified number of iterations. Additionally, one of the most serious issues within these metrics is their failure to verify whether the predicted starting materials are actually capable of undergoing the required reactions to synthesize the target molecule (Liu et al., 2023b; Tripp et al., 2024). To tackle this challenge and ensure that evaluation metrics more accurately evaluate the model's ability to predict feasible synthetic routes, Liu et al. (2023b) has introduced the set-wise exact match of predicted starting materials with those in reference synthetic routes extracted from the reaction database (USPTO-full), while retro-fallback (Tripp et al., 2024) evaluates the probability that at least one route can be executed in the wet lab. These alternative metrics provide a more realistic reflection

[1]The Pennsylvania State University [2]Google DeepMind [3]University of Southern California. Correspondence to: Songtao Liu <skl5761@psu.edu>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024.

[1]Note that the definition of a synthetic route in chemistry is different from its counterpart in computer science. In organic synthesis, the term "route" typically denotes an organic synthesis flowchart in Figure 1. However, in graph theory, a route is defined as a continuous sequence of edges in a graph that connects one vertex to another. In this work, we follow the definition of route in chemistry to avoid misunderstanding.

of the strategy's performance.

While we recognize the significance of these criteria like the cost of starting materials, the step count, and whether the predicted synthetic routes are feasible, these crucial aspects are currently not incorporated into the predictions made by existing strategies. The generation of synthetic routes is guided by a product of conditional probabilities, each corresponding to a retrosynthesis prediction, where local normalization is applied. These strategies, while general, fail to adequately consider the above criteria and also suffer from the limitations of local normalization. A notable issue in local normalization is the exposure bias, a discrepancy that arises because, at training time, the model is conditioned on the actual synthetic route, whereas at test time, it depends on its own predictions (Ranzato et al., 2016; Deng et al., 2020).

Although advanced search algorithms like beam search (Retro*-0), MCTS (Segler et al., 2018), and Retro* (Chen et al., 2020) provide some improvement through the use of retrosynthesis prediction probability as a prior and by rescoring at the level of the entire synthetic route, the generation process that one step at a time still tends to lack long-range consideration. This shortcoming is mainly due to the reliance on pure probability for route prediction, without considering the cost and feasibility of starting materials in a forward-looking way. A similar issue also arises in recent large language models (Brown et al., 2020; Touvron et al., 2023). Without alignment (Ouyang et al., 2022; Rafailov et al., 2023), relying solely on the highest probability for response generation can lead to unsafe outputs. Incorporating alignment helps steer the generated content to conform to certain standards, improving safety. Moreover, some works (Tripp et al., 2022; Genheden & Bjerrum, 2022; Liu et al., 2023b; Tripp et al., 2024) report very little performance difference among some different search algorithms. Based on our investigation, we find it non-trivial to extend or retrofit existing strategies to include those criteria.

The above discussion raises an essential question: *Is it possible to improve the quality of synthetic routes generated by various strategies?* We start with an existing probabilistic model $P_{Retro}(\mathcal{T})$, which integrates a one-step retrosynthesis model with a search algorithm as one strategy and wants to improve the quality of synthetic routes. Energy-based models (Hinton, 2002; LeCun et al., 2006; Ranzato et al., 2007; Xie et al., 2016; Grathwohl et al., 2019; Sun et al., 2021) provide *compositionality* (Du et al., 2020; 2021; 2023), which means we can add up multiple energy functions to form a new probabilistic model that has the property of each component. Specifically, we want something like

$$p_\theta(\mathcal{T}) \propto \exp\left(\log P_{Retro}(\mathcal{T}) - E_\theta(\mathcal{T})\right). \quad (1)$$

By introducing $E_\theta(\mathcal{T})$, we can incorporate additional properties encoded within $E_\theta(\mathcal{T})$, which is defined on the synthetic route, allowing for route evaluation based on various criteria. This allows us to reformulate our approach as a residual EBM (Deng et al., 2020) and focus on training $E_\theta(\mathcal{T})$ exclusively. Note that our introduced framework directly operates on top of any strategy P_{Retro} without touching its training and our framework differs from Sun et al. (2021), as our energy function is based on specific criteria, whereas theirs is based on reaction prediction probability. Therefore, our model $P_\theta(\mathcal{T})$ can be guided by the energy function $E_\theta(\mathcal{T})$, enabling controllable synthetic route generation based on various criteria.

It is challenging to directly train our EBM via Maximum Likelihood Estimation (MLE) or score matching (Song & Kingma, 2021) due to the difficulty in computing the normalization term. A commonly used framework based on the contrastive divergence (Carreira-Perpinan & Hinton, 2005) requires the samples from the current model, either using gradient-based Markov Chain Monte Carlo methods (Du & Mordatch, 2019) in continuous spaces or Gibbs sampling and its improved versions (Sun et al., 2023) in discrete spaces. However, for the large space of synthetic routes, these methods would be hard to apply due to their slow convergence and difficulty in sampling. Noise Contrastive Estimation (NCE) (Gutmann & Hyvärinen, 2010; Wang & Ou, 2018; Parshakova et al., 2019) provides an easier way to train the EBM when the ground-truth data is available. However in our setting, we typically don't have the ground-truth synthetic routes as the supervision, but we can possibly get the preference comparisons between different routes. That motivates us to derive a new training paradigm that fits the application in this setting. Inspired by the recent advancements (Ouyang et al., 2022) in large language models (LLMs) that leverage reward models to steer pre-trained language models, we have the potential to derive comparative preferences among different routes tailored to particular criteria. Therefore, we utilize a preference-based loss function for training our model. Given that open-source datasets lack information on yield rates and starting material costs, we use the feasibility of the synthetic routes as our criteria to implement our EBM. Extensive experimental results show our proposed framework can consistently improve the performance of predictions made by a wide range of strategies. Our contribution can be summarized as follows:

- We explain retrosynthetic planning through the lens of probability and find that local normalization is present at each planning step. This insight also bridges retrosynthetic planning with sequence generation in LLMs, uncovering certain challenges inherent in retrosynthetic planning (Section 3.1, 3.2).
- Based on this new view, we propose a **general** and **principled** framework that can improve the quality of synthetic routes generated by various strategies in a plug-and-play

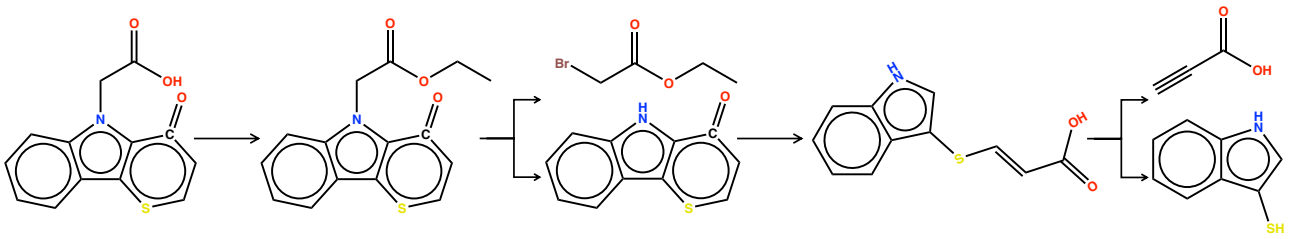

Figure 1. Illustration of a synthetic route. The target molecule we aim to synthesize is the one located on the extreme left, while the molecules positioned at the leaf nodes are the starting materials. The remaining molecules in the diagram are intermediates.

manner (Section 3.3).

- Our work can inspire future research in developing more advanced datasets that consider a broader range of criteria and leverage controllable generation techniques for molecule synthesis (Section 3.3).
- Extensive experimental results demonstrate the effectiveness of our proposed framework, which can increase previous state-of-the-art top-1 accuracy by 2.5% (Section 4.2).

2. Preliminary

In this section, we describe some foundational concepts such as synthetic route and introduce the problem.

2.1. Synthetic Route

The chemical molecule space is denoted by $\mathcal{M}$. Starting materials, defined as commercially available molecules, form a subset of $\mathcal{M}$ and are represented by $\mathcal{S} \subset \mathcal{M}$. Note that the term "building block" in some literature refers to what we consider starting material. As illustrated in Figure 1, a synthetic route comprises a tuple with four elements:

$$\mathcal{T} = (m_{tar}, \tau, \mathcal{I}, \mathcal{B}) . \tag{2}$$

The target molecule is represented by $m_{tar} \in \mathcal{M} \backslash \mathcal{S}$. The starting materials, denoted by $\mathcal{B} \subseteq \mathcal{S}$, undergo a series of reactions τ to synthesize m_{tar}. The intermediates are marked as $\mathcal{I} \subseteq \mathcal{M} \backslash \mathcal{S}$.

2.2. Reaction & One-step Retrosynthesis

We use one injection to denote a one-step reaction as

$$\Phi : \mathcal{R} \to m_p. \tag{3}$$

For a product molecule m_p, and a set of reactants $\mathcal{R} = \left\{ m_r^{(j)} \right\}_{j=1}^{n} \subseteq \mathcal{M}$ capable of synthesizing m_p, with $\Phi(\mathcal{R}) = m_p$, we focus on the main product despite potential multiple byproducts in a reaction. The objective of one-step retrosynthesis is to identify a reactant set $\mathcal{R}$ for synthesizing m_p. Given that multiple reactant sets may achieve this, indicating a one-to-many relationship, we define Ψ as the mapping from m_p to its various reactant sets $\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_n$, each capable of producing m_p.

$$\Psi : m_p \to \{\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_n\} . \tag{4}$$

Therefore, $\mathcal{R} \in \{\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_n\}$ in Eq. (3). More details about the difference between reaction and retrosynthesis prediction can be found in Appendix A.

2.3. Retrosynthetic Planning

The goal of retrosynthetic planning is to find a series of chemical reactions to transform the starting material set $\mathcal{B} \in \{\mathcal{B}_1, \mathcal{B}_2, \cdots, \mathcal{B}_n\}$ to a target molecule $m_{tar} \in \mathcal{M} \backslash \mathcal{S}$. Current strategies employ backward chaining, beginning with the target molecule m_{tar} and performing a series of one-step retrosynthesis predictions $\left\{ \Psi_i^{(t)} \right\}_{t=1}^{T_i} (\forall i \in \{1, 2, \cdots, n\})$ until all molecules at leaf nodes are from $\mathcal{S}$, which can be formulated as

$$\Gamma : m_{tar} \to \{\mathcal{T}_1, \mathcal{T}_2, \cdots, \mathcal{T}_n\} , \tag{5}$$

where Γ denotes the one-to-many mapping.

2.4. Energy-based Models

Energy-based models define the distribution via an energy function. For $x \in \mathbb{R}^D$, its probability density can be expressed as

$$P_\theta(x) = \frac{\exp\left(-E_\theta(x)\right)}{Z(\theta)}, \tag{6}$$

where $E_\theta(x) : \mathbb{R}^D \to \mathbb{R}$ is the energy function, mapping the data point x to a scalar, and $Z(\theta) = \int_x \exp\left(-E_\theta(x)\right)$ is the normalization constant. An EBM can be trained using any function that receives data points as input and outputs a scalar value.

3. Conditional Residual Energy-based Models for Molecule Synthetic Route Generation

In this section, we discuss the details of our proposed framework. We first provide a deep understanding of retrosyn-

thetic planning in Section 3.1 and 3.2, and then discuss the details of our framework in Section 3.3.

3.1. Retrosynthetic Planning is a Conditional Generation Task

Consider a continuous sequence of edges that connects the target molecule to a starting material in a synthetic route. This sequence consists of a series of retrosynthesis predictions, where a reactant in one reaction becomes the product in the subsequent one. This process can thus be conceptualized as a Markov chain, which can be formulated as follows:

$$\begin{aligned} & P_{Retro}\left(m_1, \cdots, m_T \mid m_{tar}\right) \\ &= P\left(\Psi\left(m_{tar}\right)\right) \prod_{t=1}^{T-1} P\left(\Psi\left(m_t\right)\right) \\ &= P\left(\mathcal{R}_{tar} \mid m_{tar}\right) \prod_{t=1}^{T-1} P\left(\mathcal{R}_t \mid m_t\right) \\ &= P\left(m_1 \mid m_{tar}\right) \prod_{t=1}^{T-1} P\left(m_{t+1} \mid m_t\right), \end{aligned} \tag{7}$$

where m_{tar} is the target molecule we aim to synthesize, m_t (for $1 \leq t \leq T-1$) is the intermediate molecule for which we need to predict its reactant set $\mathcal{R}_t$, and m_T is the starting material. In one-step retrosynthesis prediction, all reactants are generated simultaneously during both training and inference, instead of one at a time, so the probability $P\left(\mathcal{R}_t \mid m_t\right)$ is equivalent to $P\left(m_{t+1} \mid m_t\right)$, with $m_{t+1} \in \mathcal{R}_t$. Although $\mathcal{R}_{tar}$ and $\mathcal{R}_t$ $(1 \leq t \leq T-2)$ may include starting materials, we exclude these starting materials from the Markov chain for the simplicity of notation since the search ends at these points. In contrast to text generation in an autoregressive manner, where each next token is generated based on all the previous tokens, our one-step retrosynthesis models are trained on individual reactions and only rely on the preceding precursor for the next molecule set generation at each planning step as shown in Eq. (7). Note that a recent work (Liu et al., 2023b) has conditioned predictions on all predecessor molecules and makes training and inference in an autoregressive manner.

$$P\left(m_{t+1} \mid m_t\right) = P\left(m_{t+1} \mid m_{tar}, m_1, \cdots, m_t\right). \tag{8}$$

3.2. Your Retrosynthesis Model is Secretly a Locally Normalized Model in Retrosynthetic Planning

The term "local normalization" in sequence generation refers to the way probabilities are assigned. To understand this better, consider a language model generating a sequence of tokens. At each step, when the model predicts the next token in the sequence, it considers a set of possible next tokens and assigns probabilities to each of these tokens such that the sum of these probabilities equals one. This process is considered "local" because the normalization, which ensures that the probabilities sum to one, is performed independently at each step without considering the entire sequence.

Indeed, in retrosynthetic planning, our retrosynthesis model is locally normalized with a focus on predicting the next reactant set given the current product in the synthetic route. At each planning step, the model evaluates a set of possible reactant sets, normalizing the probabilities of these options so that they sum to one. This normalization happens independently for each retrosynthesis prediction, without considering the entire synthetic route. We explain the details by examining the categorization of retrosynthesis models.

Template-based Model. In the generation of reactants given a molecule, the widely used template-based model Neuralsym (Segler & Waller, 2017) operates by first predicting the template class, which is then applied to the product molecule to derive the reactants. These templates are extracted from the training dataset. Therefore, the retrosynthesis prediction in template-based models can be formulated as follows:

$$P\left(\mathcal{R} \mid m_p\right) = P\left(Tem | m_p\right), \tag{9}$$

where Tem denotes the template. At each step of retrosynthetic planning, Neuralsym assigns probabilities to these pre-defined template options, ensuring they sum to one. This probability reflects how suitable each template is for the current retrosynthesis step, but this assignment is done without taking the entire synthetic route into account. Besides, two other template-based models, GLN (Dai et al., 2019) and RetroComposer (Yan et al., 2022), also apply a softmax function to the scores of all candidate reactants associated with the product. Therefore, some of the widely used template-based models can be considered locally normalized.

Semi-template-based Model. Semi-template-based (two-stage) models (Shi et al., 2020; Yan et al., 2020; Somnath et al., 2021) first predict reaction centers and then disconnect the product molecule into synthons which are finally converted into chemically valid molecules. We can consider breaking reaction centers and attaching atoms or leaving groups as actions. Accordingly, we define two action spaces: $\mathcal{E}$ and $\mathcal{A}$. Here, $\mathcal{E}$ denotes the action space for breaking reaction centers and $\mathcal{A}$ represents the action space for attaching atoms or leaving groups. The two-stage model can be formulated as follows:

$$\begin{aligned} P\left(\mathcal{R} \mid m_p\right) &= P\left(\mathcal{G}_{\mathcal{R}} \mid \mathcal{G}_p\right) \\ &= \sum_{\mathcal{E}, \mathcal{A}} P(\mathcal{E} \mid \mathcal{G}_p) P\left(\mathcal{A} \mid \mathcal{G}_p, \mathcal{G}_s\right), \end{aligned} \tag{10}$$

where $\mathcal{G}_p = (V_p, E_p)$ denotes the product molecule graph with atom set V_p and chemical bond set E_p, $\mathcal{G}_{\mathcal{R}}$ represents

the reactants graph, and $\mathcal{G}_s$ denotes the synthons graph. $P(\mathcal{R} \mid m_p)$ incorporates the combined probabilities of actions in both $\mathcal{E}$ and $\mathcal{A}$ spaces. In the first stage, we predict the reaction centers, based on which we disconnect the product molecule graph. We use a binary label $y_b \in \{0, 1\}$ for each bond $b \in E_p$ in the product molecule graph $\mathcal{G}_p$, which indicates whether the bond is a reaction center. So the prediction of the reaction centers can be formulated as

$$P(\mathcal{E} \mid \mathcal{G}_p) = \prod_{b \in E_p} P(y_b | \mathcal{G}_p). \tag{11}$$

It follows that $\sum \prod_{b \in E_p} P(y_b \mid \mathcal{G}_p) = 1$. Note that in G2Gs, the center identification is treated as a multi-class classification problem, where $P(\mathcal{E} \mid \mathcal{G}_p) = 1$. In the second stage, we transform synthons into valid molecules. GraphRetro (Somnath et al., 2021) attaches leaving groups with

$$P(\mathcal{A} \mid \mathcal{G}_p, \mathcal{G}_s) = \prod_{s \in S} P(q_{l_s} \mid \mathcal{G}_p, \mathcal{G}_s), \tag{12}$$

where S is the number of connected components (synthons) and the leaving group q_{l_s} is selected from a pre-computed vocabulary from the training dataset. Therefore, we have $\sum \prod_{s \in S} P(q_{l_s} \mid \mathcal{G}_p, \mathcal{G}_s) = 1$. G2Gs modifies the second stage by replacing the attachment of a leaving group with a sequential attachment of multiple atoms, which is not difficult to demonstrate that $P(\mathcal{A} \mid \mathcal{G}_p, \mathcal{G}_s) = 1$. With $\sum P(\mathcal{A} \mid \mathcal{G}_p, \mathcal{G}_s) = 1$ and $\sum P(\mathcal{E} \mid \mathcal{G}_p) = 1$, it is evident that each stage of the two-stage model operates with local normalization. Based on the conclusion that any model that is locally normalized is also globally normalized (Variani et al., 2022). We can conclude

$$\sum P(\mathcal{R} \mid m_p) = 1, \tag{13}$$

and these two semi-template-based models are also locally normalized in retrosynthetic planning.

Template-free Model. Template-free models often approach retrosynthesis as a sequence-to-sequence task by representing molecules with SMILES string.

$$P(\mathcal{R} \mid m_p) = \prod_{t=1}^{T} P(x_t \mid X_p, x_1, x_2, \cdots, x_{t-1}), \tag{14}$$

where x_t denotes the t-th token in SMILES string of reactants and X_p is the SMILES string of product. Obviously, we have $\sum P(x_t \mid X_p, x_1, x_2, \cdots, x_{t-1}) = 1$ and template-free models are also locally normalized.

For more retrosynthesis models, readers are encouraged to check their implementation and can find that the softmax function is applied to the probability scores of the predicted candidate reactant sets associated with the product. Therefore, the discrete generation nature in retrosynthesis prediction ensures the retrosynthesis model is locally normalized during retrosynthetic planning.

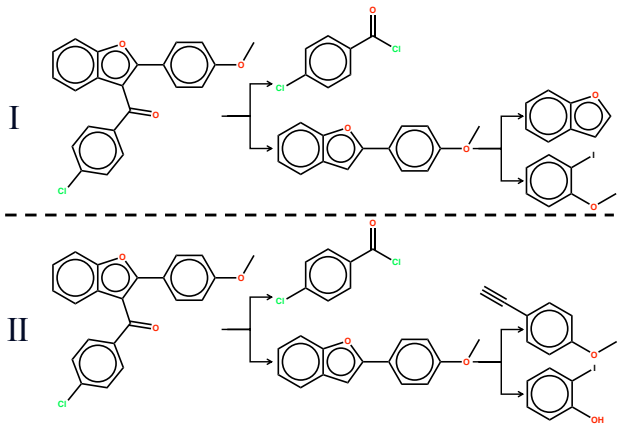

Figure 2. For a given target molecule, we find two synthetic routes that can synthesize it in the dataset.

As discussed before, locally normalized models operate on the molecule instead of the whole synthetic route. This step-by-step generation process often fails to account for long-range factors, primarily because it relies on pure probability for predicting routes without forward-thinking various criteria such as starting material costs and the feasibility of synthetic routes. In this work, we propose a general and principled framework via conditional residual energy-based models to improve the quality of synthetic routes.

3.3. Conditional Residual Energy-based Models

EBMs offer *compositionality* (Du et al., 2020; 2021; 2023), enabling the integration of extra energy functions to evaluate synthetic routes based on multiple criteria and thus develop a new probabilistic model. This formulates our method as a conditional residual EBM (Deng et al., 2020).

$$\begin{aligned} & P_\theta(\mathcal{T} \mid m_{tar}, c) \\ &= P_{Retro}(\mathcal{T} \mid m_{tar}) \frac{\exp(-E_\theta(\mathcal{T} \mid m_{tar}, c))}{Z_\theta(m_{tar}, c)} \\ &\propto P_{Retro}(\mathcal{T} \mid m_{tar}) \exp(-E_\theta(\mathcal{T} \mid m_{tar}, c)), \end{aligned} \tag{15}$$

where $P_{Retro}(\mathcal{T} \mid m_{tar})$ is a strategy via the combination of the retrosynthesis model and search algorithm, c denotes specific criteria (condition), $Z_\theta(m_{tar}, c)$ is a normalizing factor, P_θ is the joint model, and E_θ is the conditional residual energy function for evaluating the quality given c. During the training of our energy function E_θ, $P_{Retro}(\mathcal{T} \mid m_{tar})$ is fixed. Therefore, our approach is a post-training method, freezing the base model when training additional components, and can be applied on top of any existing strategy, while its black-box nature makes it convenient to adopt. With such CREBM, we aim to improve the quality of the most probable synthetic routes, those with higher likelihoods, generated by existing strategies to meet specific criteria and achieve controllable generation.

3.3.1. TRAINING

As discussed in Section 1, maximizing Eq. (15) through MLE is challenging due to the normalization factor. Deng et al. (2020) address this by sampling sequences from the training dataset and generated sequences from an autoregressive LM as positive and negative pairs, respectively. They then train the energy function as a binary classifier to distinguish between ground truth and generated texts.

However, in retrosynthetic planning, we usually do not have fixed ground truth synthetic routes. Given a product, we can find multiple routes within the dataset to synthesize it as shown in Figure 2. Preference for different synthetic routes varies according to different criteria. For example, consider two synthetic routes: one with a high yield but environmentally unfriendly, and the other with a low yield but eco-friendly. In this case, our preference for either route varies based on our priorities regarding energy conservation and emission reduction. We use a real-world example in Appendix E to illustrate the trade-off between yield and environmental impact in molecule synthesis. Therefore, we can't train our model via NCE. Inspired by the recent successes of alignment algorithms (Ouyang et al., 2022; Rafailov et al., 2023) in LLMs, we can potentially obtain preference comparisons among various routes based on specific criteria. Then, we train our model using a preference-based loss function. Given a target molecule m_{tar}, we have n synthetic routes $\{\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_n\}$ that can synthesize it. We define a reward function based on a specific criteria c.

$$\varphi\left(\cdot \mid m_{tar}, c\right) : \mathcal{X}_{tar} \rightarrow \mathcal{V}, \tag{16}$$

where $\mathcal{X}_{tar}$ is the space of synthetic routes for molecule m_{tar} and $\mathcal{V} \in \mathbb{R}$ is a scalar value. The higher the value $\mathcal{V}$ of a synthetic route $\mathcal{T}$, the more it meets our criteria c. So given two synthetic routes $\mathcal{T}_1$ and $\mathcal{T}_2$, we can have preference. Bradley-Terry (BT) (Bradley & Terry, 1952) model is a common choice for modeling preferences. The likelihood of preferring route $\mathcal{T}_1$ over $\mathcal{T}_2$ is as follows:

$$\begin{aligned} & P^*\left(\mathcal{T}_1 \succ \mathcal{T}_2 \mid m_{tar}, c\right) \\ = & \frac{\exp\left(\psi^*\left(\mathcal{T}_1 \mid m_{tar}, c\right)\right)}{\exp\left(\psi^*\left(\mathcal{T}_1 \mid m_{tar}, c\right)\right) + \exp\left(\psi^*\left(\mathcal{T}_2 \mid m_{tar}, c\right)\right)} \\ & \text{s.t. } \varphi\left(\mathcal{T}_1 \mid m_{tar}, c\right) > \varphi\left(\mathcal{T}_2 \mid m_{tar}, c\right), \end{aligned} \tag{17}$$

where ψ^* is the ground truth reward function. $\mathcal{T}_1 \succ \mathcal{T}_2$ means $\varphi\left(\mathcal{T}_1 \mid m_{tar}, c\right) > \varphi\left(\mathcal{T}_2 \mid m_{tar}, c\right)$, and $\mathcal{T}_1$ is preferred than $\mathcal{T}_2$. Thus, the loss function for our model is

$$\begin{aligned} & \mathcal{L} = -\mathbb{E}\left[\log \sigma\left(-E_\theta\left(\mathcal{T}_w \mid m_{tar}, c\right) + E_\theta\left(\mathcal{T}_l \mid m_{tar}, c\right)\right)\right] \\ & \text{s.t. } \left(m_{tar}, \mathcal{T}_w, \mathcal{T}_l\right) \sim \mathcal{D}, \end{aligned} \tag{18}$$

where σ is the sigmoid function and $\varphi\left(\mathcal{T}_w \mid m_{tar}, c\right) > \varphi\left(\mathcal{T}_l \mid m_{tar}, c\right)$ given c. We use a reference synthetic route and multiple routes sampled from P_{Retro} for each molecule in the training dataset to construct our synthetic preference dataset $\mathcal{D}$. In our implementation, we only use one strategy (Neuralsym+Retro*-0) as our P_{Retro} for sampling multiple synthetic routes to update θ during training.

3.3.2. IMPLEMENTATION OF φ

The key is the implementation of φ, which ranks synthetic routes that align with diverse criteria. Some criteria, such as the novelty and reliability of a synthetic route, are not easily quantifiable through a reward function. While some criteria like material costs can be quantified, a pharmaceutical company's production equipment may not necessarily be compatible with the cheapest route predicted by the machine learning model. Therefore, it is crucial to consider whether the available equipment can accommodate the reaction conditions in the predicted synthetic routes. However, incorporating production equipment specifications into the reward function design poses a significant challenge. Ideally, chemists or manufacturers would evaluate and rank the routes according to their expertise, but it is expensive for our research. Therefore, we design a heuristic reward function for ranking purposes, rather than estimating rewards. This approach aligns with recent developments in LLMs, which also rank responses instead of assigning accurate scores when constructing preference datasets.

In this work, we focus on assessing the feasibility of starting materials for synthesizing the target molecule. This emphasis comes from the necessity for machine learning models to first ensure that the predicted synthetic route can be executed in the wet lab before considering other criteria. We use a forward model to simulate the synthesis process to replace the wet lab experiments and compare the similarity of the produced product with the target molecule. Besides, we have multiple routes for synthesizing the same target molecule. Sometimes a new synthetic route can be obtained by simply substituting a chlorine atom in a material with a different halogen atom. Therefore, we also consider routes whose materials are similar to those of the reference route as a better choice. So our φ can be rewritten as

$$\varphi\left(\mathcal{T} \mid m_{tar}, c\right) = \operatorname{sim}\left(f(\mathcal{B}), m_{tar}\right) + \operatorname{sim}\left(\mathcal{B}, \mathcal{B}_{ref}\right), \tag{19}$$

where $\mathcal{B}_{ref}$ is the starting material set of the reference synthetic route, f denotes the forward model, and $\operatorname{sim}(\cdot, \cdot)$ is the Tanimoto similarity function. Note that the Tanimoto similarity function provides a measure of similarity between two sets of fingerprint bits. Therefore, it can be used for computing the similarity between two sets of molecules.

3.3.3. INFERENCE

After completing the training, our energy function can be applied on top of any strategy in a plug-and-play man-

Algorithm 1 CREBM Framework

[Train Phase]: Learning:
Define reward function $\varphi(\cdot \mid m_{tar}, c)$ as Eq. (16)
$\varphi(\cdot \mid m_{tar}, c) : \mathcal{X}_{tar} \to \mathcal{V}$
Rank synthetic routes $\mathcal{X}_{tar} \sim P_{Retro}$ based on φ
Train Conditional Residual Energy-based Models:
$\theta^* = \arg\min_\theta \mathcal{L} = \arg\max_\theta \mathbb{E}_{(m_{tar}, \mathcal{T}_w, \mathcal{T}_l) \sim \mathcal{D}}$
[Test Phase]: Inference:
Input: θ^*, m_{tar}, Proposal $\mathcal{X}_{tar} \sim P_{Retro}(\cdot \mid m_{tar})$.
$L \leftarrow -\log P_{Retro}(\mathcal{T} \mid m_{tar}) + E_\theta(\mathcal{T} \mid m_{tar}, c)$
$\mathcal{T}^* = \arg\min_{\mathcal{T} \in \mathcal{X}_{tar}} L$
Return $\mathcal{T}^*$

ner. With the well-learned energy function E_{θ^*} where $\theta^* = \arg\min_\theta \mathcal{L}$, the inference process aims to find the best synthetic route $\mathcal{T}^*$ that minimizes the negative log-likelihood for a specific target molecule m_{tar} and a given criteria c, i.e.

$$\mathcal{T}^* = \arg\min_{\mathcal{T} \in \mathcal{X}_{tar}} \left(E_\theta(\mathcal{T} \mid m_{tar}, c) - \log P_{Retro}(\mathcal{T} \mid m_{tar})\right). \tag{20}$$

In other words, we find $\mathcal{T}^*$ by maximizing $P_{\theta^*}(\mathcal{T} \mid m_{tar}, c)$ among all routes sampled from $P_{Retro}(\mathcal{T} \mid m_{tar})$. We illustrate the overall process in Algorithm 1.

3.3.4. Discussion

In this section, we answer some questions that are related to our work.

Q1: With the reward function, why use Bradley-Terry model instead of reinforcement learning for CREBM training?
A1: As discussed in Section 3.3.2, it's very difficult to build an accurate point-wise reward function (Bhattacharyya et al., 2020; Haroutunian et al., 2023) for reinforcement learning (RL). We attempted to use the reward function in Eq. (16) to train an RL model. We generated synthetic routes from the RL model and used the product of its negative log-likelihood and the reward function as the training loss. However, experimental results indicated a significant drop in accuracy compared to the base model. Training and fine-tuning an RL model proved to be very challenging. In contrast, we did not adjust any hyperparameters when training our CREBM. Therefore, we believe training with a preference-based loss function is better, or at least much easier to make things work.

Q2: Why not use the parameter-efficient fine-tuning for preference alignment?
A2: Directly modifying the base model can be challenging because: **1.** Route generation involves a single-step retrosynthesis model and a search algorithm, both of which are currently trained independently. Integrating these components for joint training is necessary for parameter-efficient fine-tuning, but this integration requires significant modifications to both parts, which we believe is inefficient. For instance, it is difficult to fine-tune a template-based classifier for preference alignment. **2.** While specific criteria necessitate fine-tuning all strategies, with CREBM, we only need to train the model for these criteria once. Afterward, it can be applied to any strategy in a plug-and-play fashion. As discussed in Section 3.3, we can conclude our CREBM approach is more efficient.

Q3: Is a fair comparison since the baseline methods don't use an additional reward function?
A3: Our method is complementary to most existing strategies, as we are building an adapter that can improve most existing methods by treating them as a base model in the CREBM framework. Experimental results in Section 4.2 demonstrate that our method consistently enhances the performance across all base methods and is a general model-agnostic approach that can bring additional benefits with this lightweight post-training.

Q4: Comparison with an explicit search policy function like PDVN (Liu et al., 2023a)?
A4: PDVN requires training different explicit search policy functions based on the specific retrosynthesis models employed. Our approach begins by sampling routes from Neuralsym and Retro*-0. These sampled routes are then utilized to train our CREBM. Subsequently, we employ our CREBM as an adapter to collaborate with any base model and search algorithm during inference, without the need for any additional tuning.

4. Experiments

In this section, we evaluate the performance of our proposed framework in retrosynthetic planning.

4.1. Experimental Setup

Dataset. We use the public dataset RetroBench (Liu et al., 2023b) for evaluation. The target molecules associated with synthetic routes are split into training, validation, and test datasets in an 80%/10%/10% ratio. We have 46,458 data points for training, 5,803 for validation, and 5,838 for testing. Synthetic routes for each target molecule are extracted from the reaction network which is constructed with all reactions in USPTO-full. More details about the dataset can be found in Appendix C.

Evaluation Protocol. We use the set-wise exact match between the predicted starting material set and that of the reference route in the test dataset as our evaluation metric. Note that for a given target, there might be several synthetic routes available in the test set. We consider the prediction to be accurate when the set of starting materials predicted matches with any one of the various reference options.

Table 1. Summary of retrosynthetic planning results in terms of exact match accuracy (%). Our framework (CREBM) can consistently improve the performance of existing strategies.

One-step Model \ Search Algorithm	Retro*					Retro*-0					Greedy DFS
	Top-1	Top-2	Top-3	Top-4	Top-5	Top-1	Top-2	Top-3	Top-4	Top-5	Top-1
Template-based											
Retrosim (Coley et al., 2017)	35.1	40.5	42.9	44.0	44.6	35.0	40.5	43.0	44.1	44.6	31.5
Neuralsym (Segler & Waller, 2017)	41.7	49.2	52.1	53.6	54.4	42.0	49.3	52.0	53.6	54.3	**39.2**
Neuralsym+CREBM	**44.2**	50.8	53.6	54.6	55.4	**44.5**	**51.0**	53.5	54.5	55.2	-
GLN (Dai et al., 2019)	39.6	48.9	52.7	54.6	55.7	39.5	48.7	52.6	54.5	55.6	38.0
GLN+CREBM	43.3	**51.1**	**53.9**	**55.5**	**56.4**	43.2	**51.0**	**53.8**	**55.5**	**56.3**	-
Semi-template-based											
G2Gs (Shi et al., 2020)	5.4	8.3	9.9	10.9	11.7	4.2	6.5	7.6	8.3	8.9	3.8
GraphRetro (Somnath et al., 2021)	15.3	19.5	21.0	21.9	22.4	15.3	19.5	21.0	21.9	22.2	14.4
GraphRetro+CREBM	**16.3**	**20.1**	**21.6**	**22.3**	**22.7**	**16.3**	**20.2**	**21.6**	**22.3**	**22.7**	-
Template-free											
Transformer (Karpov et al., 2019)	31.3	40.4	44.7	47.2	48.9	31.2	40.5	45.1	47.3	48.7	26.7
Transformer+CREBM	35.0	43.4	46.7	48.5	49.7	34.9	43.5	46.6	48.4	49.6	-
Megan (Sacha et al., 2021)	18.8	29.7	37.2	42.6	45.9	19.5	28.0	33.2	36.4	38.5	32.9
FusionRetro (Liu et al., 2023b)	37.5	45.0	48.2	50.0	50.9	37.5	45.0	48.3	50.2	51.2	**33.8**
FusionRetro+CREBM	**39.4**	**46.6**	**49.3**	**50.7**	**51.5**	**39.6**	**46.7**	**49.5**	**51.0**	**51.7**	-

Baselines. A retrosynthetic planning strategy is a combination of a retrosynthesis model and a search algorithm. We consider template-based models: Retrosim (Coley et al., 2017), Neuralsym (Segler & Waller, 2017) and GLN (Dai et al., 2019); template-free models: Transformer (Karpov et al., 2019) and Megan (Sacha et al., 2021); semi-template-based models: G2Gs (Shi et al., 2020) and GraphRetro (Somnath et al., 2021); as our retrosynthesis models. For the algorithm, we use Retro* (Chen et al., 2020) and Retro*-0. Retro* utilizes a neural architecture to evaluate the future score in retrosynthetic planning, while Retro*-0 is indeed a beam search.

Implementation Details. In this work, our focus is on the feasibility of the starting materials. Therefore, we disregard intermediate molecules for our E_θ. In a more general case, the preferences between routes can depend on intermediate molecules as well. We leave this exploration to future work, which is flexible in our framework. We employ a standard Transformer (Vaswani et al., 2017) architecture to implement $E_\theta(\mathcal{T} \mid m_{tar}, c)$, with the target molecule serving as the input for the encoder and the starting material (right shift) as the input for the decoder. The output is the logits of the starting material (left shift) for computing E_θ. One thing we'd like to point out is that E_θ is pretrained first on the target-to-starting material task, so we naturally deploy this for modeling, instead of training an encoder-only one from scratch. We also employ the standard Transformer architecture to implement the forward model, framing the task of predicting a product from starting materials as a sequence-to-sequence task. For constructing our preference dataset $\mathcal{D}$, we sample 10 synthetic routes for each molecule in the training dataset. All the models in the work are trained on the NVIDIA Tesla A100 GPU. More details about hyperparameters can be found in Appendix D.

Table 2. Summary of results with our CREBM in terms of top-1 accuracy (%) on routes of different depths.

Model	Retro*-0				
	2	3	4	5	6
Neuralsym	46.1	42.0	33.7	37.3	40.8
+CREBM	+2.2	+2.1	+4.9	+2.9	+5.0
GLN	46.4	39.0	32.6	25.3	21.8
+CREBM	+2.7	+3.1	+4.7	+7.1	+14.5
Transformer	39.3	30.1	20.9	15.2	16.2
+CREBM	+2.4	+5.3	+3.0	+5.6	+10.6

4.2. Results

We report top-k accuracy in Table 1. We reuse the metrics of the baselines already reported in Liu et al. (2023b). The results demonstrate that the strategies equipped with our

Table 3. Ablation study on our energy function.

Metric for Ranking	Top-1	Δ_1	Top-5	Δ_2
$-\log P_{Retro}(\mathcal{T} \mid m_{tar})$	42.0	0	54.3	0
$-\log P_{Retro}(\mathcal{T} \mid m_{tar}) + E_\theta(\mathcal{T} \mid m_{tar}, c)$	44.5	+2.5	55.2	+0.9
$E_\theta(\mathcal{T} \mid m_{tar}, c)$	34.1	-7.9	53.1	-1.2
$-\log P_{Retro}(\mathcal{T} \mid m_{tar}) - E_\theta(\mathcal{T} \mid m_{tar}, c)$	19.2	-22.8	44.6	-9.7

framework, CREBM, achieve better performance. Specifically, our framework can achieve state-of-the-art top-1 accuracy with Neuralsym+Retro*-0. We can conclude our framework can consistently improve the performance of all strategies, and predict more feasible synthetic routes. As shown in Table 2, our framework can also improve performance no matter how long the synthetic route is. When the length of routes is larger, our amplification becomes more pronounced, which demonstrates our framework can improve the performance of long route prediction.

4.3. Ablation Study

We conduct an ablation study with Retro*-0+Neuralsym to verify the effectiveness of our proposed energy function. As shown in Table 3, only using the energy function for route ranking results in a significant drop in top-1 accuracy and a marginal decrease in top-5 accuracy. This indicates that our energy function still performs well within the top-5 predictions. Adjusting the probability by adding or subtracting the energy function leads to corresponding increases or decreases in accuracy, thereby confirming the effectiveness of our energy function from both positive and negative perspectives.

4.4. Visualization of Synthetic Routes used to Train the Energy Function

We provide a visual example (Figure 3) of the synthetic routes that are used to train the energy function and ranked based on Eq. (19). We can find that higher-ranked routes have a greater likelihood of successfully synthesizing the target molecule, and their starting materials are more closely aligned with the reference materials. The starting materials of the second-ranked route only require replacing the iodine atom in one of the starting materials of the reference route with a bromine atom.

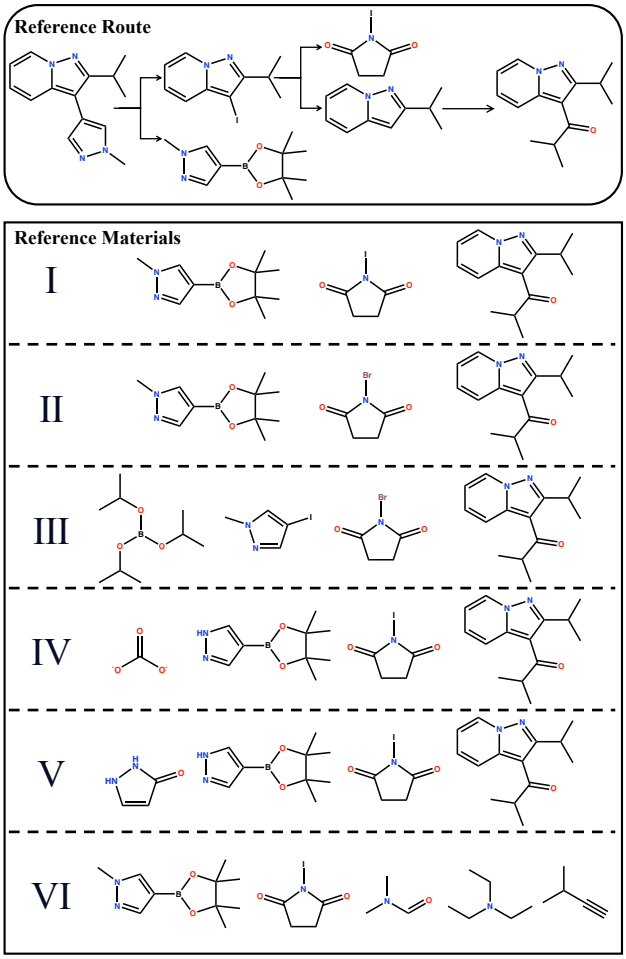

Figure 3. A visual example of a target molecule's synthetic routes used to train the energy function, with these routes ranked according to Eq. (19).

5. Related Work

Retrosynthesis Model. Current one-step retrosynthesis models fall into three distinct categories: template-based (Coley et al., 2017; Segler & Waller, 2017; Dai et al., 2019; Chen et al., 2020; Chen & Jung, 2021; Seidl et al., 2021), semi-template-based (Shi et al., 2020; Yan et al., 2020; Somnath et al., 2021), and template-free (Liu et al., 2017; Zheng et al., 2019; Chen et al., 2019; Karpov et al., 2019; Ishiguro et al., 2020).

Search Algorithm. Search algorithms (Segler et al., 2018; Han et al., 2022) select the most promising candidate at each step of the retrosynthetic planning. More details about related work can be found in Appendix B.

6. Conclusion and Future Work

In this work, we propose a framework based on conditional residual energy-based models to improve the quality of synthetic routes generated by existing retrosynthetic planning strategies. We implement one CREBM to enhance the feasibility of routes for synthesizing the target molecule. Extensive experimental results show our proposed framework can improve the accuracy of existing strategies. Our work can inspire future research to develop compositional CREBMs based on multiple criteria for molecule synthesis.

Acknowledgements

We thank all the anonymous reviewers for their helpful comments and suggestions. Songtao Liu thanks Zhengkai Tu for his helpful discussions and comments. Songtao Liu also acknowledges the research assistantship provided by Prof. Dongwon Lee during the summer of 2023. Although the project on geometric deep learning, partially conducted during that summer, resulted in failure, Songtao Liu believes Prof. Dongwon Lee's support helps him to continue his Ph.D. studies. Songtao Liu is deeply grateful to Prof. Dongwon Lee, the director of Penn State's IST doctoral programs, for his kindness which makes international students at IST feel at home. Therefore, this work is partially supported by the National Science Foundation under Grant No. 1820609.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

References

Baker, F. N., Chen, Z., and Ning, X. Rlsync: Offline-online reinforcement learning for synthon completion. *arXiv preprint arXiv:2309.02671*, 2023.

Bhattacharyya, S., Rooshenas, A., Naskar, S., Sun, S., Iyyer, M., and McCallum, A. Energy-based reranking: Improving neural machine translation using energy-based models. *arXiv preprint arXiv:2009.13267*, 2020.

Bradley, R. A. and Terry, M. E. Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. Language models are few-shot learners. In *Advances in Neural Information Processing Systems*, 2020.

Carreira-Perpinan, M. A. and Hinton, G. On contrastive divergence learning. In *International Workshop on Artificial Intelligence and Statistics*, 2005.

Chen, B., Shen, T., Jaakkola, T. S., and Barzilay, R. Learning to make generalizable and diverse predictions for retrosynthesis. *arXiv preprint arXiv:1910.09688*, 2019.

Chen, B., Li, C., Dai, H., and Song, L. Retro*: learning retrosynthetic planning with neural guided a* search. In *International Conference on Machine Learning*, 2020.

Chen, S. and Jung, Y. Deep retrosynthetic reaction prediction using local reactivity and global attention. *JACS Au*, 1(10):1612–1620, 2021.

Chen, Z., Ayinde, O. R., Fuchs, J. R., Sun, H., and Ning, X. G2retro as a two-step graph generative models for retrosynthesis prediction. *Communications Chemistry*, 6 (1):102, 2023.

Coley, C. W., Rogers, L., Green, W. H., and Jensen, K. F. Computer-assisted retrosynthesis based on molecular similarity. *ACS Central Science*, 3(12):1237–1245, 2017.

Corey, E. J. The logic of chemical synthesis: multistep synthesis of complex carbogenic molecules (nobel lecture). *Angewandte Chemie International Edition in English*, 30 (5):455–465, 1991.

Corey, E. J. and Wipke, W. T. Computer-assisted design of complex organic syntheses: Pathways for molecular synthesis can be devised with a computer and equipment for graphical communication. *Science*, 166(3902):178–192, 1969.

Dai, H., Li, C., Coley, C., Dai, B., and Song, L. Retrosynthesis prediction with conditional graph logic network. In *Advances in Neural Information Processing Systems*, 2019.

Deng, Y., Bakhtin, A., Ott, M., Szlam, A., and Ranzato, M. Residual energy-based models for text generation. In *International Conference on Learning Representations*, 2020.

Du, Y. and Mordatch, I. Implicit generation and generalization in energy-based models. *arXiv preprint arXiv:1903.08689*, 2019.

Du, Y., Li, S., and Mordatch, I. Compositional visual generation with energy based models. In *Advances in Neural Information Processing Systems*, 2020.

Du, Y., Li, S., Sharma, Y., Tenenbaum, J., and Mordatch, I. Unsupervised learning of compositional energy concepts. In *Advances in Neural Information Processing Systems*, 2021.

Du, Y., Durkan, C., Strudel, R., Tenenbaum, J. B., Dieleman, S., Fergus, R., Sohl-Dickstein, J., Doucet, A., and Grathwohl, W. S. Reduce, reuse, recycle: Compositional generation with energy-based diffusion models and mcmc. In *International Conference on Machine Learning*, 2023.

Fang, L., Li, J., Zhao, M., Tan, L., and Lou, J.-G. Leveraging reaction-aware substructures for retrosynthesis analysis. *arXiv preprint arXiv:2204.05919*, 2022.

Gao, Z., Tan, C., Wu, L., and Li, S. Z. Semiretro: Semi-template framework boosts deep retrosynthesis prediction. *arXiv preprint arXiv:2202.08205*, 2022.

Gao, Z., Chen, X., Tan, C., and Li, S. Z. Motifretro: Exploring the combinability-consistency trade-offs in retrosynthesis via dynamic motif editing. *arXiv preprint arXiv:2305.15153*, 2023.

Genheden, S. and Bjerrum, E. Paroutes: towards a framework for benchmarking retrosynthesis route predictions. *Digital Discovery*, 1(4):527–539, 2022.

Grathwohl, W., Wang, K.-C., Jacobsen, J.-H., Duvenaud, D., Norouzi, M., and Swersky, K. Your classifier is secretly an energy based model and you should treat it like one. *arXiv preprint arXiv:1912.03263*, 2019.

Guo, Z., Wu, S., Ohno, M., and Yoshida, R. Bayesian algorithm for retrosynthesis. *Journal of Chemical Information and Modeling*, 60(10):4474–4486, 2020.

Gutmann, M. and Hyvärinen, A. Noise-contrastive estimation: A new estimation principle for unnormalized statistical models. In *International Conference on Artificial Intelligence and Statistics*, 2010.

Han, P., Zhao, P., Lu, C., Huang, J., Wu, J., Shang, S., Yao, B., and Zhang, X. Gnn-retro: Retrosynthetic planning with graph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022.

Haroutunian, L., Li, Z., Galescu, L., Cohen, P., Tumuluri, R., and Haffari, G. Reranking for natural language generation from logical forms: A study based on large language models. *arXiv preprint arXiv:2309.12294*, 2023.

Hassen, A. K., Torren-Peraire, P., Genheden, S., Verhoeven, J., Preuss, M., and Tetko, I. Mind the retrosynthesis gap: Bridging the divide between single-step and multi-step retrosynthesis prediction. *arXiv preprint arXiv:2212.11809*, 2022.

He, H.-R., Wang, J., Liu, Y., and Wu, F. Modeling diverse chemical reactions for single-step retrosynthesis via discrete latent variables. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, 2022.

Heifets, A. and Jurisica, I. Construction of new medicines via game proof search. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2012.

Hinton, G. E. Training products of experts by minimizing contrastive divergence. *Neural computation*, 14(8):1771–1800, 2002.

Hoffmann, R. W. *Elements of synthesis planning*, volume 307. Springer, 2009.

Hong, S., Zhuo, H. H., Jin, K., and Zhou, Z. Retrosynthetic planning with experience-guided monte carlo tree search. *arXiv preprint arXiv:2112.06028*, 2021.

Igashov, I., Schneuing, A., Segler, M., Bronstein, M., and Correia, B. Retrobridge: Modeling retrosynthesis with markov bridges. In *International Conference on Learning Representations*, 2024.

Ishiguro, K., Ujihara, K., Sawada, R., Akita, H., and Kotera, M. Data transfer approaches to improve seq-to-seq retrosynthesis. *arXiv preprint arXiv:2010.00792*, 2020.

Jiang, Y., Ying, W., Wu, F., Huang, Z., Kuang, K., and Wang, Z. Learning chemical rules of retrosynthesis with pre-training. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023.

Karpov, P., Godin, G., and Tetko, I. V. A transformer model for retrosynthesis. In *International Conference on Artificial Neural Networks*, 2019.

Kim, J., Ahn, S., Lee, H., and Shin, J. Self-improved retrosynthetic planning. In *International Conference on Machine Learning*, 2021.

Kishimoto, A., Buesser, B., Chen, B., and Botea, A. Depth-first proof-number search with heuristic edge cost and application to chemical synthesis planning. In *Advances in Neural Information Processing Systems*, 2019.

Krause, B., Gotmare, A. D., McCann, B., Keskar, N. S., Joty, S., Socher, R., and Rajani, N. F. Gedi: Generative discriminator guided sequence generation. *arXiv preprint arXiv:2009.06367*, 2020.

Lan, Z., Zeng, Z., Hong, B., Liu, Z., and Ma, F. Rcsearcher: Reaction center identification in retrosynthesis via deep q-learning. *arXiv preprint arXiv:2301.12071*, 2023.

Lan, Z., Hong, B., Zhu, J., Zeng, Z., Liu, Z., Yu, L., and Ma, F. Retrosynthesis prediction via search in (hyper) graph. *arXiv preprint arXiv:2402.06772*, 2024.

LeCun, Y., Chopra, S., Hadsell, R., Ranzato, M., and Huang, F. A tutorial on energy-based learning. *Predicting structured data*, 1(0), 2006.

Lee, H., Ahn, S., Seo, S.-W., Song, Y. Y., Yang, E., Hwang, S.-J., and Shin, J. Retcl: A selection-based approach for retrosynthesis via contrastive learning. *arXiv preprint arXiv:2105.00795*, 2021.

Lee, S., Kim, T., Choi, M.-S., Kwak, Y., Park, J., Hwang, S. J., and Kim, S.-G. Readretro: Natural product biosynthesis planning with retrieval-augmented dual-view retrosynthesis. *bioRxiv*, pp. 2023–03, 2023.

Li, J., Fang, L., and Lou, J.-G. Retro-bleu: Quantifying chemical plausibility of retrosynthesis routes through reaction template sequence analysis. *arXiv preprint arXiv:2311.06304*, 2023a.

Li, J., Fang, L., and Lou, J.-G. Retroranker: leveraging reaction changes to improve retrosynthesis prediction through re-ranking. *Journal of Cheminformatics*, 15(1): 58, 2023b.

Li, X., Thickstun, J., Gulrajani, I., Liang, P. S., and Hashimoto, T. B. Diffusion-lm improves controllable text generation. In *Advances in Neural Information Processing Systems*, 2022.

Lin, M. H., Tu, Z., and Coley, C. W. Improving the performance of models for one-step retrosynthesis through re-ranking. *Journal of cheminformatics*, 14(1):1–13, 2022.

Lin, Z., Yin, S., Shi, L., Zhou, W., and Zhang, Y. J. G2gt: Retrosynthesis prediction with graph-to-graph attention neural network and self-training. *Journal of Chemical Information and Modeling*, 63(7):1894–1905, 2023.

Liu, A., Sap, M., Lu, X., Swayamdipta, S., Bhagavatula, C., Smith, N. A., and Choi, Y. Dexperts: Decoding-time controlled text generation with experts and anti-experts. *arXiv preprint arXiv:2105.03023*, 2021.

Liu, B., Ramsundar, B., Kawthekar, P., Shi, J., Gomes, J., Luu Nguyen, Q., Ho, S., Sloane, J., Wender, P., and Pande, V. Retrosynthetic reaction prediction using neural sequence-to-sequence models. *ACS Central Science*, 3 (10):1103–1113, 2017.

Liu, G., Xue, D., Xie, S., Xia, Y., Tripp, A., Maziarz, K., Segler, M., Qin, T., Zhang, Z., and Liu, T.-Y. Retrosynthetic planning with dual value networks. In *International Conference on Machine Learning*, 2023a.

Liu, J., Yan, C., Yu, Y., Lu, C., Huang, J., Ou-Yang, L., and Zhao, P. Mars: A motif-based autoregressive model for retrosynthesis prediction. *arXiv preprint arXiv:2209.13178*, 2022.

Liu, S., Tu, Z., Xu, M., Zhang, Z., Lin, L., Ying, R., Tang, J., Zhao, P., and Wu, D. Fusionretro: Molecule representation fusion via in-context learning for retrosynthetic planning. In *International Conference on Machine Learning*, 2023b.

Liu, Y., Xu, H., Fang, T., Xi, H., Liu, Z., Zhang, S., Poon, H., and Wang, S. T-rex: Text-assisted retrosynthesis prediction. *arXiv preprint arXiv:2401.14637*, 2024.

Maziarz, K., Tripp, A., Liu, G., Stanley, M., Xie, S., Gaiński, P., Seidl, P., and Segler, M. Re-evaluating retrosynthesis algorithms with syntheseus. *arXiv preprint arXiv:2310.19796*, 2023.

Meng, Z., Zhao, P., Yu, Y., and King, I. A unified view of deep learning for reaction and retrosynthesis prediction: Current status and future challenges. In *Proceedings of the International Joint Conference on Artificial Intelligence*, 2023.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., et al. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems*, 2022.

Parshakova, T., Andreoli, J.-M., and Dymetman, M. Global autoregressive models for data-efficient sequence learning. *arXiv preprint arXiv:1909.07063*, 2019.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*, 2019.

Qian, Y., Li, Z., Tu, Z., Coley, C. W., and Barzilay, R. Predictive chemistry augmented with text retrieval. *arXiv preprint arXiv:2312.04881*, 2023.

Rafailov, R., Sharma, A., Mitchell, E., Ermon, S., Manning, C. D., and Finn, C. Direct preference optimization: Your language model is secretly a reward model. In *Advances in Neural Information Processing Systems*, 2023.

Ranzato, M., Boureau, Y.-L., Chopra, S., and LeCun, Y. A unified energy-based framework for unsupervised learning. In *International Conference on Artificial Intelligence and Statistics*, 2007.

Ranzato, M., Chopra, S., Auli, M., and Zaremba, W. Sequence level training with recurrent neural networks. In *International Conference on Learning Representations*, 2016.

Sacha, M., Błaz, M., Byrski, P., Dabrowski-Tumanski, P., Chrominski, M., Loska, R., Włodarczyk-Pruszynski, P., and Jastrzebski, S. Molecule edit graph attention network: modeling chemical reactions as sequences of graph edits. *Journal of Chemical Information and Modeling*, 61(7): 3273–3284, 2021.

Sacha, M., Sadowski, M., Kozakowski, P., van Workum, R., and Jastrzebski, S. Molecule-edit templates for efficient and accurate retrosynthesis prediction. *arXiv preprint arXiv:2310.07313*, 2023.

Segler, M. H. and Waller, M. P. Neural-symbolic machine learning for retrosynthesis and reaction prediction. *Chemistry–A European Journal*, 23(25):5966–5971, 2017.

Segler, M. H., Preuss, M., and Waller, M. P. Planning chemical syntheses with deep neural networks and symbolic ai. *Nature*, 555(7698):604, 2018.

Seidl, P., Renz, P., Dyubankova, N., Neves, P., Verhoeven, J., Wegner, J. K., Hochreiter, S., and Klambauer, G. Modern hopfield networks for few-and zero-shot reaction prediction. *arXiv preprint arXiv:2104.03279*, 2021.

Seo, S.-W., Song, Y. Y., Yang, J. Y., Bae, S., Lee, H., Shin, J., Hwang, S. J., and Yang, E. Gta: Graph truncated attention for retrosynthesis. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021.

Shi, C., Xu, M., Guo, H., Zhang, M., and Tang, J. A graph to graphs framework for retrosynthesis prediction. In *International Conference on Machine Learning*, 2020.

Somnath, V. R., Bunne, C., Coley, C., Krause, A., and Barzilay, R. Learning graph models for retrosynthesis prediction. In *Advances in Neural Information Processing Systems*, 2021.

Song, Y. and Kingma, D. P. How to train your energy-based models. *arXiv preprint arXiv:2101.03288*, 2021.

Sun, H., Dai, H., Dai, B., Zhou, H., and Schuurmans, D. Discrete langevin samplers via wasserstein gradient flow. In *International Conference on Artificial Intelligence and Statistics*, 2023.

Sun, R., Dai, H., Li, L., Kearnes, S., and Dai, B. Towards understanding retrosynthesis by energy-based models. In *Advances in Neural Information Processing Systems*, 2021.

Tetko, I. V., Karpov, P., Van Deursen, R., and Godin, G. State-of-the-art augmented nlp transformer models for direct and single-step retrosynthesis. *Nature communications*, 11(1):1–11, 2020.

Torren-Peraire, P., Hassen, A. K., Genheden, S., Verhoeven, J., Clevert, D.-A., Preuss, M., and Tetko, I. Models matter: The impact of single-step retrosynthesis on synthesis planning. *arXiv preprint arXiv:2308.05522*, 2023.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., et al. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 2023.

Tripp, A., Maziarz, K., Lewis, S., Liu, G., and Segler, M. Re-evaluating chemical synthesis planning algorithms. In *NeurIPS 2022 AI for Science: Progress and Promises*, 2022.

Tripp, A., Maziarz, K., Lewis, S., Segler, M., and Hernández-Lobato, J. M. Retro-fallback: retrosynthetic planning in an uncertain world. In *International Conference on Learning Representations*, 2024.

Tu, H., Shorewala, S., Ma, T., and Thost, V. Retrosynthesis prediction revisited. In *NeurIPS 2022 AI for Science: Progress and Promises*, 2022.

Tu, Z. and Coley, C. W. Permutation invariant graph-to-sequence model for template-free retrosynthesis and reaction prediction. *Journal of chemical information and modeling*, 62(15):3503–3513, 2022.

Variani, E., Wu, K., Riley, M. D., Rybach, D., Shannon, M., and Allauzen, C. Global normalization for streaming speech recognition in a modular framework. In *Advances in Neural Information Processing Systems*, 2022.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. Attention is all you need. In *Advances in Neural Information Processing Systems*, 2017.

Wan, Y., Hsieh, C.-Y., Liao, B., and Zhang, S. Retroformer: Pushing the limits of end-to-end retrosynthesis transformer. In *International Conference on Machine Learning*, 2022.

Wang, B. and Ou, Z. Learning neural trans-dimensional random field language models with noise-contrastive estimation. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2018.

Wang, Y., Song, Y., Xu, M., Wang, R., Zhou, H., and Ma, W. Retrodiff: Retrosynthesis as multi-stage distribution interpolation. *arXiv preprint arXiv:2311.14077*, 2023.

Xie, J., Lu, Y., Zhu, S.-C., and Wu, Y. A theory of generative convnet. In *International Conference on Machine Learning*, 2016.

Xie, S., Yan, R., Han, P., Xia, Y., Wu, L., Guo, C., Yang, B., and Qin, T. Retrograph: Retrosynthetic planning with graph search. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2022.

Xie, S., Yan, R., Guo, J., Xia, Y., Wu, L., and Qin, T. Retrosynthesis prediction with local template retrieval. *arXiv preprint arXiv:2306.04123*, 2023.

Xiong, J., Zhang, W., Fu, Z., Huang, J., Kong, X., Wang, Y., Xiong, Z., and Zheng, M. Improve retrosynthesis planning with a molecular editing language. *chemrxiv*, 2023.

Yan, C., Ding, Q., Zhao, P., Zheng, S., Yang, J., Yu, Y., and Huang, J. Retroxpert: Decompose retrosynthesis prediction like a chemist. In *Advances in Neural Information Processing Systems*, 2020.

Yan, C., Zhao, P., Lu, C., Yu, Y., and Huang, J. Retrocomposer: Composing templates for template-based retrosynthesis prediction. *Biomolecules*, 12(9):1325, 2022.

Yang, K. and Klein, D. Fudge: Controlled text generation with future discriminators. *arXiv preprint arXiv:2104.05218*, 2021.

Yao, L., Wang, Z., Guo, W., Xiang, S., Liu, W., and Ke, G. Node-aligned graph-to-graph generation for retrosynthesis prediction. *arXiv preprint arXiv:2309.15798*, 2023.

Yu, Y., Wei, Y., Kuang, K., Huang, Z., Yao, H., and Wu, F. Grasp: Navigating retrosynthetic planning with goal-driven policy. In *Advances in Neural Information Processing Systems*, 2022.

Yu, Y., Yuan, L., Wei, Y., Gao, H., Ye, X., Wang, Z., and Wu, F. Retroood: Understanding out-of-distribution generalization in retrosynthesis prediction. *arXiv preprint arXiv:2312.10900*, 2023.

Yuan, L., Yu, Y., Wei, Y., Wang, Y., Wang, Z., and Wu, F. Active retrosynthetic planning aware of route quality. In *International Conference on Learning Representations*, 2024.

Zhang, L., Rao, A., and Agrawala, M. Adding conditional control to text-to-image diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023a.

Zhang, Q., Liu, J., Zhang, W., Yang, F., Yang, Z., and Zhang, X. A multi-stream network for retrosynthesis prediction. *Frontiers of Computer Science*, 18(2):182906, 2024a.

Zhang, X., Mo, Y., Wang, W., and Yang, Y. Retrosynthesis prediction enhanced by in-silico reaction data augmentation. *arXiv preprint arXiv:2402.00086*, 2024b.

Zhang, Y., Hao, H., He, X., Gao, S., and Zhou, A. Evolutionary retrosynthetic route planning. *arXiv preprint arXiv:2310.05186*, 2023b.

Zheng, S., Rao, J., Zhang, Z., Xu, J., and Yang, Y. Predicting retrosynthetic reactions using self-corrected transformer neural networks. *Journal of Chemical Information and Modeling*, 60(1):47–55, 2019.

Zhong, W., Yang, Z., and Chen, C. Y.-C. Retrosynthesis prediction using an end-to-end graph generative architecture for molecular graph editing. *Nature Communications*, 14 (1):3009, 2023a.

Zhong, Z., Song, J., Feng, Z., Liu, T., Jia, L., Yao, S., Wu, M., Hou, T., and Song, M. Root-aligned smiles: a tight representation for chemical reaction prediction. *Chemical Science*, 13(31):9023–9034, 2022.

Zhong, Z., Song, J., Feng, Z., Liu, T., Jia, L., Yao, S., Hou, T., and Song, M. Recent advances in artificial intelligence for retrosynthesis. *arXiv preprint arXiv:2301.05864*, 2023b.

Zhong, Z., Song, J., Feng, Z., Liu, T., Jia, L., Yao, S., Hou, T., and Song, M. Recent advances in deep learning for retrosynthesis. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, pp. e1694, 2024.

Zhu, J., Hong, B., Lan, Z., and Ma, F. Single-step retrosynthesis via reaction center and leaving groups prediction. In *2023 16th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics*, 2023a.

Zhu, J., Xia, Y., Wu, L., Xie, S., Zhou, W., Qin, T., Li, H., and Liu, T.-Y. Dual-view molecular pre-training. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2023b.

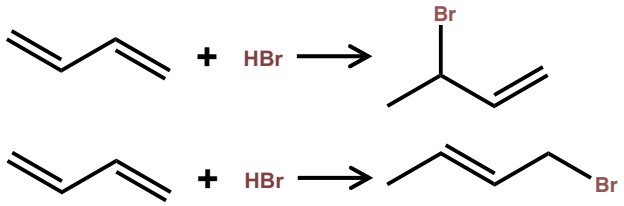

Figure 4. Hydrogen bromide (HBr) can undergo 1,2-addition and 1,4-addition with 1,3-butadiene under different reaction conditions.

A. Discriminative Reaction Prediction vs. Generative Retrosynthesis Prediction

Reaction prediction and retrosynthetic prediction are two different types of tasks. With given reactants and reaction conditions, the output product is definite. Although the reactants can undergo 1-2 and 1-4 addition reactions to synthesize different products, as shown in 4, the required reaction conditions are different. However, given a product molecule, the specific reaction conditions for its synthesis are not always known, and there could be multiple sets of reactants capable of synthesizing it.

B. Related Work

One-step Retrosynthesis Model. Current one-step retrosynthesis models (Meng et al., 2023; Zhong et al., 2023b; 2024) fall into three distinct categories: template-based (Coley et al., 2017; Segler & Waller, 2017; Dai et al., 2019; Chen et al., 2020; Chen & Jung, 2021; Seidl et al., 2021; Yan et al., 2022; Xie et al., 2023; Sacha et al., 2023; Zhang et al., 2024a), semi-template-based (Shi et al., 2020; Yan et al., 2020; Somnath et al., 2021; Gao et al., 2022; Zhong et al., 2023a; Baker et al., 2023; Wang et al., 2023; Gao et al., 2023; Zhu et al., 2023a; Lan et al., 2023; Chen et al., 2023; Lan et al., 2024), and template-free (Liu et al., 2017; Zheng et al., 2019; Chen et al., 2019; Karpov et al., 2019; Ishiguro et al., 2020; Guo et al., 2020; Tetko et al., 2020; Seo et al., 2021; Lee et al., 2021; Sun et al., 2021; Fang et al., 2022; Wan et al., 2022; He et al., 2022; Liu et al., 2022; Tu & Coley, 2022; Lin et al., 2022; Zhong et al., 2022; Yu et al., 2023; Li et al., 2023b; Zhu et al., 2023b; Jiang et al., 2023; Qian et al., 2023; Xiong et al., 2023; Yao et al., 2023; Lin et al., 2023; Liu et al., 2024; Zhang et al., 2024b). Template-based methods model retrosynthesis as either a classification or a template retrieval problem. These approaches involve extracting templates based on chemical rules, followed by training a model to identify the specific template that allows the product to be decomposed into reactants. During the inference stage, the process begins by predicting the class of the template rule, which is then applied to the product to deduce the corresponding reactants. Semi-template-based methods, often referred to as two-stage models, initially use templates to identify the reaction centers in a product. Subsequently, they break these centers down into several disconnected subgraphs, known as synthons. These synthons are then transformed into chemically valid molecules through generative modeling or by attaching leaving groups. Template-free approaches model retrosynthesis prediction as either a sequence-to-sequence task (Karpov et al., 2019), using SMILES strings to represent molecules, or as a graph-edit problem (Sacha et al., 2021; Igashov et al., 2024), where molecules are depicted as graphs.

Search Algorithm in Retrosynthetic Planning. Search algorithms (Segler et al., 2018; Hong et al., 2021; Kishimoto et al., 2019; Chen et al., 2020; Heifets & Jurisica, 2012; Kim et al., 2021; Yu et al., 2022; Han et al., 2022; Hassen et al., 2022; Li et al., 2023a; Zhang et al., 2023b; Lee et al., 2023; Yuan et al., 2024) select the most promising candidate from multiple predictions offered by the one-step retrosynthesis model at each step of the retrosynthetic planning. This process is repeated until all reactants are commercially available. MCTS (Segler et al., 2018; Hong et al., 2021) enhance search with policy networks for multi-step planning, while DFPN-E (Kishimoto et al., 2019) combines Depth-First Proof-Number with heuristic techniques for chemical synthesis planning. Neural-based A*-like algorithms in Retro* (Chen et al., 2020) estimate solution costs to identify promising routes. GRASP (Yu et al., 2022) applies RL in guiding the search, and both GNN-Retro (Han et al., 2022) and RetroGraph (Xie et al., 2022) use graph neural networks for better cost estimation in A* search. However, we observe that Retro* does not outperform trivial beam search (Retro*-0) in generating feasible synthetic routes as reported in Liu et al. (2023b) and we explore a new direction to improve the feasibility of synthetic routes in this work.

Evaluation of Retrosynthetic Planning. Some of the existing evaluation (Tu et al., 2022; Torren-Peraire et al., 2023) metrics fall short of verifying whether the starting materials can actually synthesize the given product. Liu et al. (2023b) addresses this issue by extracting synthetic routes from a reaction network and introducing a set-wise exact match evaluation metric. This metric involves comparing the predicted starting materials with the reference starting materials extracted from the reaction database for assessment.

Controllable Generation. In controllable text generation (Krause et al., 2020; Yang & Klein, 2021; Liu et al., 2021; Li et al., 2022), the language model (LM) remains fixed, while a potential function is used to steer its generation process. These methods utilize various plug-and-play techniques to manage the text generation, ensuring that the produced text meets diverse requirements. Controllable image generation leverages a control network (Zhang et al., 2023a), to introduce a spatial control signal into large, pre-trained text-to-image diffusion models. This approach enables the generation of images that meet various specific requirements.

C. Dataset Details

Table 4. Dataset Statistics.

#Molecules / **Depth** / **Dataset**	2	3	4	5	6	7	8	9	10	11	12	13
Training	22,903	12,004	5,849	3,268	1,432	594	276	107	25	0	0	0
Validation	2,862	1,500	731	408	179	74	34	13	2	0	0	0
Test	2,862	1,500	731	408	179	74	34	13	2	32	2	1

D. Reproducibility

D.1. Implementation Details

We use Pytorch (Paszke et al., 2019) to implement our models. The codes of all baselines are implemented referring to the implementation of FusionRetro (Liu et al., 2023b). All the experiments of baselines are conducted on a single NVIDIA Tesla A100 with 80GB memory size. The softwares that we use for experiments are Python 3.6.8, CUDA 10.2.89, CUDNN 7.6.5, einops 0.4.1, pytorch 1.9.0, pytorch-scatter 2.0.9, pytorch-sparse 0.6.12, numpy 1.19.2, torchvision 0.10.0, and torchdrug 0.1.3.

D.2. Hyperparameter Details

Table 5. The hyper-parameters for the forward model.

max length	300
embedding size	64
encoder layers	6
decoder layers	6
attention heads	8
FFN hidden	1024
dropout	0.1
epochs	1200
batch size	256
warmup	16000
lr factor	20
scheduling	$lr = \frac{\text{lr factor} \times \min\left(1.0, \frac{\text{global\_step}}{\text{warmup}}\right)}{\max(\text{global\_step}, \text{warmup})}$

Table 6. The hyper-parameters for the energy-based model.

max length	310
embedding size	64
encoder layers	6
decoder layers	6
attention heads	8
FFN hidden	1024
dropout	0.1
epochs	40
batch size	256
warmup	10000
lr	5e-5
scheduling	linear

E. A Real-world Example about the Preference of Two Synthetic Routes

Here's a real-world example (generated by ChatGPT) illustrating the trade-off between yield and environmental impact in chemical synthesis:

Example: Synthesis of Acetic Acid.

High-Yield but Less Eco-Friendly Route: Methanol Carbonylation

- The industrial production of acetic acid often involves the carbonylation of methanol, catalyzed by iodine and rhodium or cobalt.
- Yield: This process, known as the Monsanto or Cativa process, is highly efficient and yields a large quantity of acetic acid.
- Environmental Impact: However, it uses significant amounts of metals as catalysts and can produce harmful by-products like iodomethane. The use of these materials and the disposal of by-products can be environmentally damaging.

Lower-Yield but More Eco-Friendly Route: Fermentation

- An alternative method for acetic acid production is through bacterial fermentation, using species like Acetobacter. This process converts ethanol to acetic acid under aerobic conditions.
- Yield: The yield of this method is generally lower compared to the carbonylation of methanol.
- Environmental Impact: However, it's considered more environmentally friendly. It uses renewable resources (like ethanol), involves less hazardous chemicals, and generally has a smaller carbon footprint.

Decision Making: Although a company prioritizing production efficiency and cost-effectiveness might choose the carbonylation method for its high yield, despite the greater environmental concerns, we focus on environmental sustainability and prefer the fermentation route despite its lower yield, due to its lesser environmental impact.