LING682 Final Project - Controlled Molecular Generation using Sequential Monte Carlo Sampling with Small Language Models

Chih-Chao Hsu
McGill University, Department of Electrical and Computer Engineering

Abstract

This report studies Sequential Monte Carlo (SMC) as an inference-time control mechanism for molecular generation with small chemistry-focused language models. By incorporating continuous, property-aware potential functions and SMILES validity constraints into the sampling process, SMC significantly improves controllability without modifying or retraining the underlying model. Experiments on GPT2-ZINC show that SMC-guided generation achieves substantially higher constraint adherence than prompt-based control in an 8B-parameter SmileyLlama model, while maintaining high validity and diversity at a fraction of the computational cost.

Code: `https://github.com/leohsuofnthu/Chem-SMC/`

1 Introduction

1.1 De Novo Molecular Discovery

De novo molecular discovery plays a crucial role in modern drug discovery and materials design. The ultimate goal is to be able to reveal novel, valid compounds that satisfy specified physicochemical or pharmacological properties, which could then possibly be potential precursors of ground-breaking new drugs and materials. Traditionally, rule-based, fragment-based design, combinatorial chemistry, and QSAR methods have been the major approaches. However, they are limited because predefined rules and fragment libraries restrict the accessible chemical space. With the rapid development of AI, various approaches including large language model pretraining Frey et al. (2023), reinforcement learning Loeffler et al. (2024), and more recently diffusion-based models Gong et al. (2024), have all been explored for this purpose. A persistent challenge across these methods is to achieve reliable and robust control over properties such as molecular weight (MW), lipophilicity (logP), synthetic accessibility, and drug-likeness (QED) for proposing useful candidates.

1.2 SMILES Representation and LLM-based Molecular Design

One popular direction of de novo molecular generation is to leverage the auto-regressive nature of large language models for generating SMILES strings. A SMILES string is a linear string representation of molecules, making it suitable to be modeled as a text generation task that follows certain grammar. However, SMILES generation is fragile: tokenization errors or invalid syntactic structures can easily produce chemically impossible molecules. Most of the previous works focus on pre-training the models Frey et al. (2023)Ross et al. (2022) with large amounts of SMILES strings in the scale of millions, so as to teach the model how to generate realistic molecules. The validness of generating molecules thus has higher guarantee, since the distribution of generated SMILES shares the similarity of the training dataset, but the ability of sampling molecules within specific range of properties remains a challenge.

1.3 LLM-Constrained Generation

While large-scale pre-trained models can generate valid and diverse molecules, the generation process still needs the ability to specify precise and desired property range targets, which will accelerate the discovery process and make it more efficient for filtering out unwanted molecules early. Prior works to tackle this challenge have incorporated the instruction-following philosophy by supervised fine-tuning and direct preference optimization Cavanagh et al. (2025). Although it manages to offer the ability to specify preferable ranges for generation, the effort for building datasets for fine-tuning and preference encoding is still time-consuming. On the other hand, the concept of LLM constraint decoding during inference time has become popular; various directions such as grammar-constrained decoding Beurer-Kellner et al. (2024)Dong et al. (2025), and probability logit modification Ahmed et al. (2024) have been proposed. Besides requiring less effort for computationally intensive pretraining, some even show a promising way for enhancing smaller models to have comparable performance to their bigger counterparts Loula et al. (2025).

1.4 Motivation

The motivation for this project comes from one simple research question:

> *Can a small model outperform a much larger model in range-specific molecular generation via a constraint decoding mechanism?*

Although the large pre-trained model SmileyLlama Cavanagh et al. (2025) has demonstrated strong ability for property-constrained generation through pure text prompts, the model itself is still huge, with 8 billion parameters. Based on the experiment in Loula et al. (2025), it shows the potential of sequential Monte Carlo (SMC) constrained generation in molecular generation with smaller models. This work will extend this idea to explore how much performance can be squeezed out from a relatively smaller model compared to SmileyLlama. The hypothesis is that SMC could empower the model to have more fine-grained control over chemical property ranges, and possibly better performance under stricter constraints.

2 Methodology

We outline the Sequential Monte Carlo (SMC) framework underlying our approach and describe its implementation for controlled molecular generation using the GenLM framework.

2.1 Importance Sampling

Importance sampling estimates expectations under an intractable target distribution $\pi(x)$ by drawing samples from a proposal $q(x)$ and reweighting them. The self-normalized estimator is

$$\mathbb{E}_\pi[f(x)] \approx \sum_{i=1}^{N} \tilde{w}^{(i)} f(x^{(i)}), \qquad x^{(i)} \sim q(x),$$

with weights

$$w^{(i)} = \frac{\gamma(x^{(i)})}{q(x^{(i)})}, \qquad \tilde{w}^{(i)} = \frac{w^{(i)}}{\sum_j w^{(j)}},$$

where $\pi(x) \propto \gamma(x)$ is known up to a constant. In molecular generation, π represents a property-aligned molecule distribution, while q corresponds to the language model's proposal distribution.

2.2 Sequential Importance Sampling (SIS)

Sequential Importance Sampling extends importance sampling to sequence generation by updating particle weights as tokens are appended. We assume a factorized proposal distribution:

$$q(x_{1:t}) = q(x_{1:t-1})\, q_t(x_t \mid x_{1:t-1}).$$

The resulting weight update is given by

$$w_t^{(i)} = w_{t-1}^{(i)} \frac{\gamma_t(x_{1:t}^{(i)})}{\gamma_{t-1}(x_{1:t-1}^{(i)})\, q_t(x_t^{(i)} \mid x_{1:t-1}^{(i)})}.$$

Although SIS can in principle sample from an intractable target distribution using a proposal q, the recursive update causes the weight variance to grow rapidly with t, leading to severe weight degeneracy where most particles collapse to negligible weight. This limitation motivates the introduction of resampling.

2.3 Sequential Monte Carlo (SMC)

Sequential Monte Carlo uses the same sequential importance weighting mechanism as SIS, but augments it with **adaptive resampling** to prevent particle collapse. At each step, particles propose a new token $x_t^{(i)} \sim q_t(\cdot \mid x_{1:t-1}^{(i)})$. When the effective sample size (ESS) falls below a threshold, SMC resamples, replicating high-weight particles and discarding low-weight ones, to maintain diversity and stabilize long-horizon sequence generation.

2.4 Potential-Based Preference Encoding

In SMC, preferences can be encoded into the target density by adding a potential function Φ_t that scores how well a partial sequence matches the desired property range so far. We define

$$\gamma_t(x_{1:t}) \propto p_\theta(x_{1:t})\, \Phi_t(x_{1:t}),$$

so that the importance weights naturally become

$$w_t^{(i)} \propto w_{t-1}^{(i)} \frac{\Phi_t(x_{1:t}^{(i)})}{\Phi_{t-1}(x_{1:t-1}^{(i)})}.$$

Particles whose predicted properties move toward the target range receive larger weights, while those deviating from it are down-weighted and eventually discarded during resampling. This allows SMC to steer generation according to property constraints during inference time.

3 Experimental Setup

3.1 Models

Two models are chosen for experiments. One is GPT2-Zinc-87M, which was fine-tuned on around 480 million SMILES from ZINC dataset, and SmileyLlama-8B mentioned in the previous section, fine-tuned with around 2 million molecules from the ChEMBL Dataset v33. The reasons for choosing these two models are due to the big difference in parameter scale, similar trained molecules, and that both output SMILES strings, which is convenient to compare. Table 1 is the comparison between the two models.

Model	Parameters	Training Data	Prompt Format
GPT2-ZINC	87M	ZINC (SMILES)	Direct SMILES prefix
SmileyLlama-8B	8B	ChEMBL (SMILES)	Natural language property prompt

Table 1: Comparison of the two molecular language models and the formats of their prompts.

One of the concerns about comparing these two models could be from the fact that they are trained on different datasets. If the distribution of trained molecules is totally different, it would be hard to set a proper range of chemical properties for fair comparison. Therefore, a subset of molecules, around 170K, is randomly sampled from both datasets for visualization

to see if there are huge differences. As shown in Fig. 1, we can see that for molecular weight (MW), LogP, and number of rotatable bonds (RotB), both datasets have highly similar distributions except for QED. Therefore, we only use properties that have similar distributions as the constraint elements. This difference in QED is expected, since ZINC is a drug-like filtered dataset while ChEMBL contains more diverse bioactive scaffolds, leading to broader QED variation.

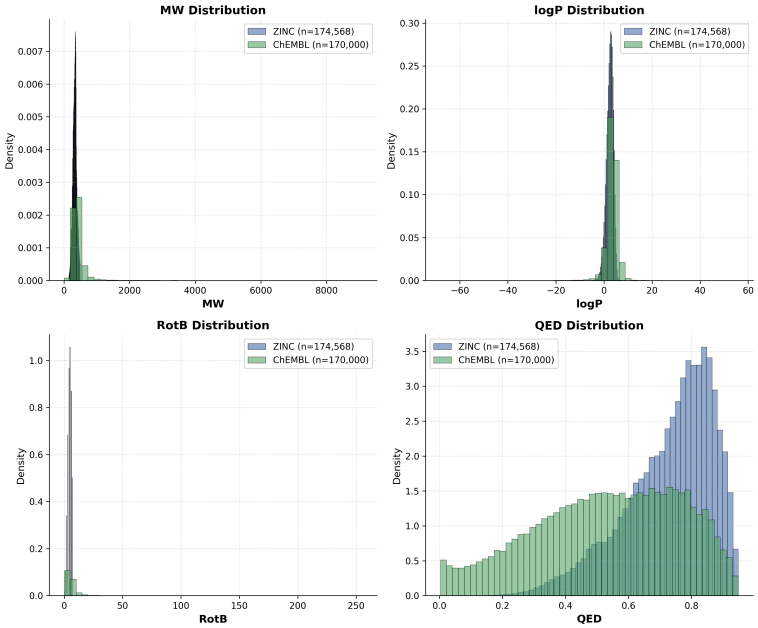

Figure 1: Property distribution comparison between **ZINC** and **ChEMBL**.

3.2 Constraints Design

We evaluate controlled molecular generation using two types of constraints. The first, **range-based constraints**, derives property bounds from the percentile distribution of the combined sampled ZINC and ChEMBL data analyzed in Fig. 1. We define three difficulty levels: loose (5th–95th percentile: MW 233–577, logP -0.07–5.73, RotB 2–10), tight (25th–75th percentile: MW 304–419, logP 1.88–4.00, RotB 3–6), and ultra-tight (40th–60th percentile: MW 336–372, logP 2.58–3.36, RotB 4–5). These data-driven ranges provide a fair basis for comparing SMC-guided controlled generation for GPT2-ZINC-87M with SmileyLlama under a unified set of acceptable ranges. The second paradigm, **gradual constraints**, uses upper-bound-only specifications designed to match SmileyLlama's instruction-following training format. We similarly define three levels; at each level, one additional constraint is introduced as shown in Table 2. The purpose of these two designs is to examine how well reward-based control and pure prompting can satisfy the specified constraints as the difficulty escalates.

Constraint Type	Level	MW	logP	RotB
Range-Based (Percentile)	Loose (5–95%)	233–577	-0.07–5.73	2–10
	Tight (25–75%)	304–419	1.88–4.00	3–6
	Ultra-tight (40–60%)	336–372	2.58–3.36	4–5
Gradual (Upper-Bound)	Loosen	$\leq$500	–	–
	Tight	$\leq$400	$\leq$4	–
	Ultra-tight	$\leq$350	$\leq$3.5	$\leq$8

Table 2: Specification of range-based and gradual constraints.

3.3 Reward (Potential) Design

To guide SMC toward molecules that satisfy the desired property ranges, we use a **constraint-based molecular potential** that scores each SMILES according to how well its RDKit-computed properties match the target constraints. For each property (MW, logP, RotB), we apply a soft distance-based penalty that rewards values near the center of the target interval and smoothly decays toward the boundaries. Additional mild bonuses (e.g., ring count) encourage reasonable, drug-like structures. Compared to the the molecular generation formulation used in Loula et al. (2025), as implemented in the GenLM molecular synthesis evaluation codebase[1] ,which relied primarily on a simple validity potential that checks whether a completion corresponds to a chemically valid molecule, our approach is inspired by this framework and extends it to incorporate continuous, property-aware scoring, yielding an overall molecular potential of the form

$$\Phi(x) \propto \exp\left(\sum_k \phi_k(x)\right)$$

where each $\phi_k(x)$ represents the alignment of a single molecular property to its target range. In practice, each $\phi_k(x)$ may be implemented using property-specific and possibly non-smooth scoring functions, while the exponential form provides a unified potential for SMC weighting. This richer potential enables SMC to favor molecules that are both valid and well aligned with the specified property constraints, while allowing properties to be added or removed easily without modifying or retraining the underlying language model.

3.4 Evaluation Metrics

Three key quantitative metrics are reported to assess the quality and controllability of molecule generation.

- **Validity (%)** — the percentage of generated SMILES strings that correspond to chemically valid molecules as determined by RDKit. This metric reflects the model's ability to produce syntactically and chemically feasible structures.
- **Adherence (%)** — the percentage of valid molecules whose physicochemical properties (e.g., molecular weight, logP, rotatable bonds) fall within the predefined constraint ranges for the current generation condition. This measures how effectively the model satisfies the desired property constraints.
- **Diversity** — the average pairwise Tanimoto distance between valid molecules, computed using RDKit fingerprints. Higher values indicate a more structurally diverse set of generated molecules, with less redundancy among samples.

3.5 Prompts

The prompts used for both models are different due to their nature. For GPT2-ZINC-87M, the prompt is a SMILES prefix. We run SMC separately for each prefix in Table 3, using each prefix to initialize diverse generation.

As for SmileyLlama-8B, we use the official prompt template from the model, which contains three sections:

Instruction: You love and excel at generating SMILES strings of drug-like molecules.
Input: Output a SMILES string for a molecule with: MW 300–400, logP 2–4, and $\leq$ 7 rotatable bonds.
Response:

[1] `https://github.com/genlm/genlm-eval/blob/main/genlm/eval/domains/molecular_synthesis.py`

Name	Prefix (SMILES)	Chemical Intent
aromatic_core	C1=CC=	Aromatic ring growth; encourages ring closure.
amino_aliphatic	CCN	Early N introduction for amines.
polar_ether	COC	Adds O; biases toward ethers/alcohols.
carbonyl_anchor	CC(=O)	Seeds esters / amides.
heterocycle_friendly	C1CN	N-containing ring fragment.
sulfur_motif	CCS	Thioether / sulfur-containing motifs.
extended_polar	CCOCCN	Polar fragment encouraging moderate lipophilicity.
fallback_generic	CCC	Neutral aliphatic baseline.
aliphatic_ring	C1CCCCC1	Cyclohexane seed; saturated rings.
aromatic_ring	c1ccccc1	Benzene core.
hetero_aromatic	n1ccccc1	Pyridine; hetero-aromatic bias.
saturated_n_heterocycle	N1CCCCC1	Piperidine scaffold.
morpholine	O1CCNCC1	Morpholine O/N heterocycle.
amide_starter	NC(=O)	Amide initiation.
sulfonamide_starter	NS(=O)(=O)	Sulfonamide seed.
urea	NC(=O)N	Urea / guanidine starter.
nitrile	C#N	Nitrile handle.
aryl_chloride	Clc1ccccc1	Aryl chloride substitution.
phosphonate	CP(=O)(O)O	Phosphonate moiety.

Table 3: SMILES prefix prompts for GPT2-ZINC-87M.

3.6 SMC Configuration

We use a standardized SMC setup across all experiments. Each run uses 20 particles, a sampling temperature of 1.0, and an ESS resampling threshold of $0.3N$. Generation is performed with nucleus sampling (top-$p = 0.9$ and top-$k = 30$), and sequences are truncated at a maximum of 128 tokens.

4 Results and Discussion

Model	Constraint Type	Level	Valid (%)	Distinct (%)	Adherence (%)	Diversity	Runtime (s)
GPT2-ZINC + SMC	Range-based	Loose	100	100	**50.8**	0.9	5m 57s
		Tight	100	100	**19.7**	0.899	5m 6s
		Ultra-tight	100	100	**2.7**	0.899	6m 42s
	Gradual	Loosen	100	100	**100**	0.91	6m 18s
		Tight	100	100	**91.4**	0.91	5m 9s
		Ultra-tight	100	100	**73.65**	0.91	5m 34s
SmileyLlama-8B	Range-based	Loose	100	100	27.9	0.88	9m 52s
		Tight	100	100	2.3	0.889	9m 50s
		Ultra-tight	100	100	0.5	0.893	7m 46s
	Gradual	Loosen	100	100	95.1	0.876	5m 44s
		Tight	100	100	80.4	0.875	4m 13s
		Ultra-tight	100	100	53.3	0.871	3m 59s

Table 4: Comparison of GPT2-ZINC and SmileyLlama-8B under both range-based and gradual constraint paradigms. Metrics include validity, diversity, constraint adherence, and runtime.

The results in Table 4 and Figure 2 reveal a clear and consistent trend: **SMC improves controllability for small molecular language models**. First, GPT2-ZINC+SMC achieves **100%** validity across all settings. Second, under *gradual* (upper-bound) constraints, which align with SmileyLlama's instruction-tuning format, GPT2-ZINC+SMC consistently outperforms the 8B SmileyLlama model. For example, in the ultra-tight setting, GPT2-ZINC+SMC attains **73.65%** adherence compared with SmileyLlama's **53.3%**, and maintains strong margins at loosen and tight levels as well. These findings highlight a key insight: **a small, chemical-specialized model equipped with constrained decoding can rival or surpass a much larger model relying solely on prompt conditioning**. On the other hand, two-sided *range* constraints (percentile-based) remain extremely difficult for both models. Adherence

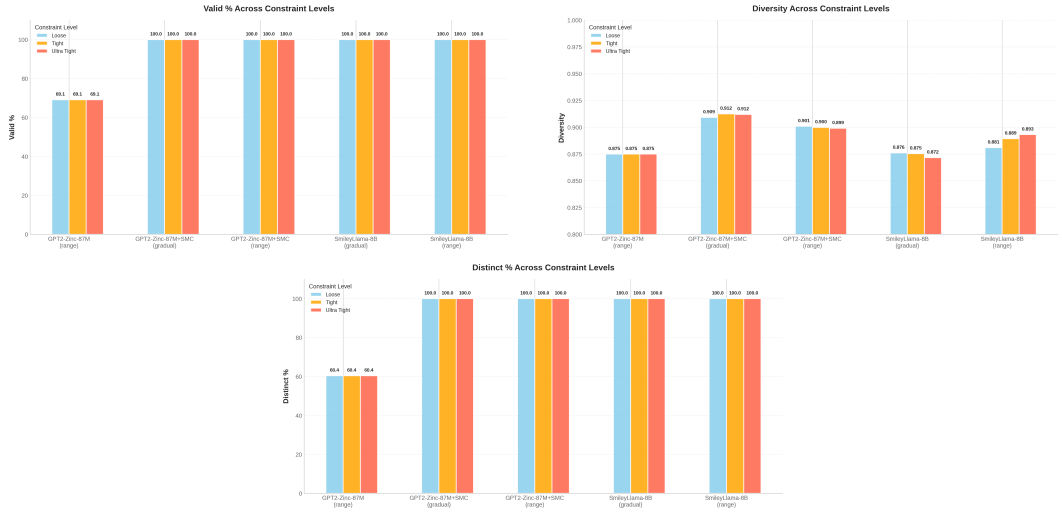

Figure 2: Constraint validity, diversity, and adherence across constraint levels.

collapses below 3% with ultra-tight constraint, illustrating that narrow bidirectional property intervals are inherently challenging for autoregressive molecular decoders, even with SMC guidance. Furthermore, diversity remains high in all SMC settings (∼0.89–0.91), which possibly arises from the use of distinct SMILES prefixes offering structural variety. However, SmileyLlama's performance may be slightly understated because all experiments were run using its 4-bit quantized version due to limited computational resource availability, which is known to exhibit mild degradation compared to full-precision inference.

5 Conclusion and Future Work

This project demonstrates SMC's ability to transform a pretrained chemical language model into a well-guided de novo molecular generator. When equipped with SMC, GPT2-ZINC-87M achieves **perfect validity**, **high diversity**, and **strong compliance across different constraints**, outperforming the larger 8-billion-parameter SmileyLlama, which relies solely on prompt-based guidance. The main limitation arises under dual-bound (range-style) constraints, where performance drops sharply for narrow intervals, indicating that current potential designs lack sufficient flexibility for two-sided control. To address this limitation, we identify three promising directions for future work:

- **Non-Linear Restrictions** — Refining SMC potentials to support more complex chemical constraints beyond simple single range specifications.
- **Multi-Property Balancing** — Developing adaptive weighting mechanisms to balance multiple, possibly conflicting molecular objectives.
- **Core-Driven Synthesis** — Exploring scaffold-based generation in which molecules are constructed around a fixed chemical core instead of being generated from unconstrained prefixes.

References

Kareem Ahmed, Kai-Wei Chang, and Guy Van den Broeck. Controllable generation via locally constrained resampling, 2024. URL https://arxiv.org/abs/2410.13111.

Luca Beurer-Kellner, Marc Fischer, and Martin Vechev. Guiding llms the right way: Fast, non-invasive constrained generation, 2024. URL https://arxiv.org/abs/2403.06988.

Joseph M. Cavanagh, Kunyang Sun, Andrew Gritsevskiy, Dorian Bagni, Yingze Wang, Thomas D. Bannister, and Teresa Head-Gordon. Smileyllama: Modifying large language

models for directed chemical space exploration, 2025. URL https://arxiv.org/abs/2409.02231.

Yixin Dong, Charlie F. Ruan, Yaxing Cai, Ruihang Lai, Ziyi Xu, Yilong Zhao, and Tianqi Chen. Xgrammar: Flexible and efficient structured generation engine for large language models, 2025. URL https://arxiv.org/abs/2411.15100.

Nathan C. Frey, Ryan Soklaski, Simon Axelrod, Siddharth Samsi, Rafael Gómez-Bombarelli, Connor W. Coley, and Vijay Gadepally. Neural scaling of deep chemical models. *Nature Machine Intelligence*, 5(11):1297–1305, 2023. ISSN 2522-5839. doi: 10.1038/s42256-023-00740-3. URL https://doi.org/10.1038/s42256-023-00740-3.

Haisong Gong, Qiang Liu, Shu Wu, and Liang Wang. Text-guided molecule generation with diffusion language model, 2024. URL https://arxiv.org/abs/2402.13040.

Hannes H. Loeffler, Jiazhen He, Alessandro Tibo, Jon Paul Janet, Alexey Voronov, Lewis H. Mervin, and Ola Engkvist. Reinvent 4: Modern ai–driven generative molecule design. *Journal of Cheminformatics*, 16(1):20, 2024. ISSN 1758-2946. doi: 10.1186/s13321-024-00812-5. URL https://doi.org/10.1186/s13321-024-00812-5.

João Loula, Benjamin LeBrun, Li Du, Ben Lipkin, Clemente Pasti, Gabriel Grand, Tianyu Liu, Yahya Emara, Marjorie Freedman, Jason Eisner, Ryan Cotterell, Vikash Mansinghka, Alexander K. Lew, Tim Vieira, and Timothy J. O'Donnell. Syntactic and semantic control of large language models via sequential monte carlo, 2025. URL https://arxiv.org/abs/2504.13139.

Jerret Ross, Brian Belgodere, Vijil Chenthamarakshan, Inkit Padhi, Youssef Mroueh, and Payel Das. Large-scale chemical language representations capture molecular structure and properties, 2022. URL https://arxiv.org/abs/2106.09553.