GENERATING π-FUNCTIONAL MOLECULES USING STGG+ WITH ACTIVE LEARNING

Alexia Jolicoeur-Martineau
Samsung SAIL Montréal
alexia.j@samsung.com

Yan Zhang
Samsung SAIL Montréal
y2.zhang@samsung.com

Boris Knyazev
Samsung SAIL Montréal
b.knyazev@samsung.com

Aristide Baratin
Samsung SAIL Montréal
a.baratin@samsung.com

Cheng-Hao Liu
California Institute of Technology
chl@caltech.edu

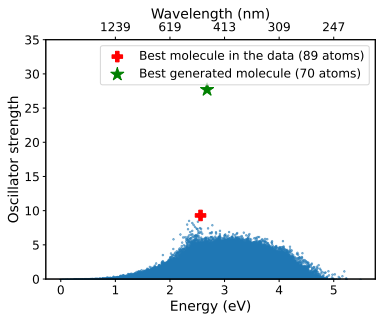
ABSTRACT

Generating novel molecules with out-of-distribution properties is a major challenge in molecular discovery. While supervised learning methods generate high-quality molecules similar to those in a dataset, they struggle to generalize to out-of-distribution properties. Reinforcement learning can explore new chemical spaces but often conducts 'reward-hacking' and generates non-synthesizable molecules.

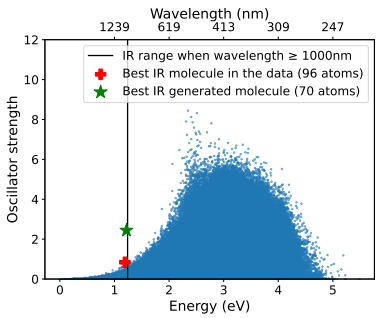
In this work, we address this problem by integrating a state-of-the-art supervised learning method, STGG+ (Jolicoeur-Martineau et al., 2024), in an active learning loop. Our approach iteratively generates, evaluates, and fine-tunes STGG+ to continuously expand its knowledge. We denote this approach STGG+AL.

We apply STGG+AL to the design of organic π-functional materials, specifically two challenging tasks: 1) generating highly absorptive molecules characterized by high oscillator strength and 2) designing absorptive molecules with reasonable oscillator strength in the near-infrared (NIR) range. The generated molecules are validated and rationalized *in-silico* with time-dependent density functional theory. Our results demonstrate that our method is highly effective in generating novel molecules with high oscillator strength, contrary to existing methods such as reinforcement learning (RL) methods.

We open-source our active-learning code along with our Conjugated-xTB dataset containing 2.9 million π-conjugated molecules and the function for approximating the oscillator strength and absorption wavelength (based on sTDA-xTB).
Code: https://github.com/SamsungSAILMontreal/STGG-AL.



Task 1: Maximizing f_{osc}. STGG+ with active learning generates strong out-of-distribution (OOD) molecules.



Task 2: Maximizing f_{osc} in the short-wave infrared range. STGG+ with active learning generates strong OOD molecules.

1 INTRODUCTION

Generating novel organic molecules with desirable, previously unseen optoelectronic properties holds transformative potential across many applications, from display technology to wearable electronics to biomedical imaging (Bilodeau et al., 2022; Fromer and Coley, 2023). Central to this pursuit are π-conjugated functional molecules, where their π-delocalized electrons enable functionalities such as in organic light-emitting diodes (OLED) and short-wave infrared (SWIR) absorbers. Traditional approaches to molecular design, however, face the persistent challenge of systematically exploring uncharted regions in the chemical space to identify out-of-distribution properties while remaining chemically reasonable.

Supervised learning methods typically address this problem by modeling the distribution of a given dataset, but extrapolating beyond the training set (i.e., out-of-distribution generalization) is difficult. Effective molecular generation requires generative models to capture meaningful patterns (e.g., chemical rules) that enable generalization.

Unsupervised methods such as Reinforcement learning (RL) (Olivecrona et al., 2017; Loeffler et al., 2024; Popova et al., 2018) do not rely on datasets and instead generate molecules and evaluate them progressively. Although powerful, with imperfect reward models in chemistry, RL can exploit the reward function and generate chemically non-viable molecules unless carefully regularized.

Active learning (Settles, 2009) holds promise in combing both worlds (supervised and unsupervised) by training a model using supervised learning and then iteratively generating new molecules, labeling them, and continuing training the model with them (Merchant et al., 2023; Korablyov et al., 2024; Kyro et al., 2024; Antoniuk et al., 2025). This approach allows the joint sampling from a strong base model and the reward model. It is more aligned with how humans learn: chemists learn about molecules from existing literature, then they generate new molecules, test them, and then rebuild their own priors about which aspects of the molecule lead to better properties.

STGG+ is an autoregressive generative model that uses spanning tree-based graph generation and is trained in a supervised manner with strong in-distribution and out-of-distribution capabilities (Jolicoeur-Martineau et al., 2024; Ahn et al., 2021b). In this work, we propose to combine STGG+ with active learning to design π-conjugated molecules with out-of-distribution optoelectronic properties, a challenging problem which current RL methods struggle with.

We explore two proof-of-concept yet application-oriented challenges. Specifically, we design:

- Molecules with exceptionally high oscillator strength (f_{osc}), which correlates with efficient photo-absorption/emission, relevant for designing OLED materials (Abroshan et al., 2022).
- Molecules with high f_{osc} and strong absorption in targeted spectral ranges, particularly in NIR for potential biomedical imaging applications (Privitera et al., 2023).

We constructed a computational dataset of 2.9 million π-conjugated molecules and pre-trained STGG+ on it, followed by active learning. Our results show that STGG+ combined with active learning can progressively move toward higher f_{osc} molecules much more effectively than baseline methods such as genetic algorithms and reinforcement learning. Active learning required only 30,000 additional data points to discover candidates with f_{osc} of 27.7, compared to a maximum of 9.3 found through traditional virtual screening. This is not only a great improvement, but also a significant speedup compared to virtual screening considering that reward evaluation requires expensive simulation. Furthermore, molecules generated by RL tend to have issues with chemical validity or synthesizability (e.g. exotic ring structures), while STGG+ generates chemically sound molecules by design. We validated the top-1 generated samples using time-dependent (TD) density functional theory (DFT), which explain the new scaffolds.

2 METHOD

In this work, we seek to maximize f_{osc} while maintaining chemical reasonableness and some additional constraints. More generally, assume that we aim to generate molecules with out-of-distribution properties not seen in the dataset. Some properties need to be maximized (Λ), while others need to be constrained within some range of values (Ω). Our approach is described below.

First, we pre-train STGG+ on some dataset(s) with molecules similar to those desired conditioned on their properties. We fix the range of properties for the constraints $\Omega \sim \mathcal{U}(\Omega_{min}, \Omega_{max})$ and initialize the properties to be maximized $\Lambda \sim \mathcal{U}(\Lambda_{min}, \Lambda_{max})$ to be around (slightly-lower and slightly-above) the maximum values found in the dataset.

Then, we iterate through N steps of active learning:

1. Generate k molecules from STGG+ conditioned on the sampled properties. $\Lambda \sim \mathcal{U}(\Lambda_{min}, \Lambda_{max})$ and $\Omega \sim \mathcal{U}(\Omega_{min}, \Omega_{max})$
2. Remove invalid/duplicated molecules.
3. Evaluate the generated molecules to determine their properties using the pipeline in Sec. 3.
4. Update the range of Λ_{min}, and Λ_{max} to be respectively the top-1 and top-100 properties to maximize (Λ) of the generated molecules (slowly expanding the Pareto frontier).
5. Fine-tune STGG+ on the generated molecules conditional on their properties.

3 CONJUGATED-XTB DATASET

We present a dataset containing 2.9 million π-conjugated organic molecules. The molecules are constructed by sampling a set of 181 hand-curated π-conjugated molecular fragments and connecting them at allowed atomic indices. All molecules have between 4-8 fragments and a maximum of 100 heavy atoms. We did not consider solubility, but alkyl chains can be readily appended to the building blocks. The 181 fragments represent common, synthesizable building blocks which we classify into electronic donors, acceptors, and 'neutral' connecting bridges. While the dataset is not optimized for synthesizability, the combinations of these building blocks are expected to represent most motifs of optoelectronically-active molecules. On average, each fragment has 2.77 connections; for 4-8 fragments, the total number of molecules that can be constructed using these fragments (without atom limitation) are respectively $\sim 6 \times 10^{10}, 3 \times 10^{13}, 2 \times 10^{16}, 8 \times 10^{18},$ and 4×10^{21}.

For each sampled molecule, we generate 32 conformations using ETKDG as implemented in RDKit (Riniker and Landrum, 2015), and these geometries are optimized by MMFF94 forcefield (Halgren, 1996). The lowest-energy conformer is selected for further geometry optimization using the semiempirical quantum chemistry method GFN2-xTB (Bannwarth et al., 2019). We then approximate the optical properties of this conformer using sTDA-xTB (Grimme and Bannwarth, 2016), which calculates the vertical absorption energy and the corresponding f_{osc} (Grimme and Bannwarth, 2016). The dataset contains 2.9 millions rows and 3 columns (SMILES, f_{osc}, absorption wavelength (in nm)). We open-sourced the full dataset.

4 EXPERIMENTS

We run experiments on two problems. First, we seek to maximize f_{osc}, which is correlated to the intensity of absorption/emission processes. Second, we aim to maximize f_{osc} in the short-wave infrared absorption range (absorption wavelength $\lambda_{abs} \geq 1000$ nm), which is crucial for biomedical imaging as tissues exhibit relatively low absorption and scattering in NIR, allowing for deeper penetration of light (Wilson et al., 2015; Privitera et al., 2023).

To increase the chances of synthesizability, we also impose a maximum ring size of 6 and a maximum number of heavy atoms of 70. STGG+ also imposes proper valency (Ahn et al., 2021b) by its design.

STGG+ is first pre-trained on the Conjugated-xTB dataset for 5 epochs. Then, active learning is applied. To maximize diversity, we uniformly sample a temperature between 0.7 and 1.1 and classifier-free guidance (Ho and Salimans, 2022) between 0.5 and 1.5. We generate 2000 molecules per active learning step, and they are trimmed down (removing duplicates and invalid molecules). Fine-tuning is done for 100 epochs on the last batch of 2000 generated molecules. The other hyperparameters follow the default ones by Jolicoeur-Martineau et al. (2024). More training and architectures details can be found in Appendix A.1-A.3.

We compare STGG+ to two strong baselines (as mentioned by Tripp and Hernández-Lobato (2023)): GraphGA (Jensen, 2019; Brown et al., 2019), and REINVENT4 (Loeffler et al., 2024), version 4 of

the popular REIVENT (Olivecrona et al., 2017). For REINVENT4, we use the default settings. For GraphGA, we used the good choice of hyperparameters suggested by Tripp and Hernández-Lobato (2023) consisting of 5 new candidates per generation and running as many generations as possible. The baseline methods are given the top 100 molecules from the Conjugated-xTB dataset as initial starting points. We run each algorithm long enough to reach around 30K evaluations. Since RL can be quite noisy compared to supervised learning, we do 3 runs of the RL baselines using 3 different seeds. See Appendix A.2 for more details on the RL baselines.

The results are described in the subsections below. We also describe the top-10 molecules made by STGG+ in Appendix A.4 and show the top-1 molecules made by all methods in Appendix A.5.

The geometries of top-1 molecules are selected to be further optimized in DFT with the B3LYP hybrid functional and def2-SVP basis set. Single-point TD-DFT calculations are then computed to cross-check with sTDA-xTB vertical absorption energies/oscillator strength.

4.1 TASK 1: MAXIMIZING THE OSCILLATOR STRENGTH

Figure 3 shows the molecule with the highest f_{osc} from the Conjugated-xTB dataset, and the molecule with the highest f_{osc} generated by STGG+ with active learning given the constraints ($\leq$ 70 atoms, max ring size of 6). Figure 4 shows the progress over time. STGG+ learns to sample rigid and planar molecules, which can have high orbital overlap and hence high f_{osc}. We see that mini-batch diversity initially drops down at around 5K Oracle calls, then moves back up at 10K Oracle calls and stabilizes.

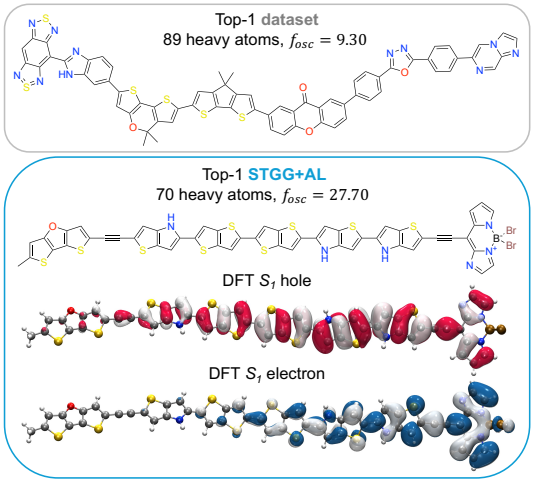


Figure 3: Case study of the top-1 molecule with the highest f_{osc}.

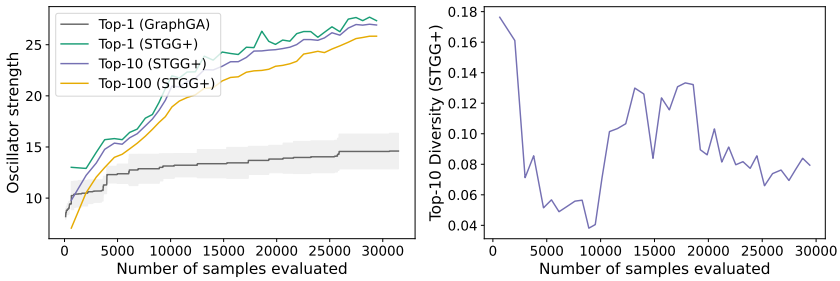


Figure 4: Maximizing f_{osc} using active learning with constraints: max 70 heavy atoms, max ring-size of 6. STGG+ (top-1, top-10, top-100; from a single run) vs GraphGA (top-1; average and 95% confidence interval over 3 runs).

TD-DFT calculations show a high $f_{\text{osc}} = 3.79$[1]. The natural transition orbital (NTO) of the first excited state confirms large hole/electron overlap over the rigid backbone.

4.2 TASK 2: MAXIMIZING THE OSCILLATOR STRENGTH IN THE SHORT-WAVE INFRARED RANGE OF ABSORPTION

Figure 3 shows the highest f_{osc} molecule in the short-wave infrared range from the Conjugated-xTB dataset, and the highest f_{osc} molecule in the NIR range generated by STGG+ with active learning given the constraints ($\leq$ 70 atoms, max ring size of 6). Here, STGG+ learns to generate different scaffolds of charge-transfer species, which can explain their lower absorption energy. We see that mini-batch diversity drops slowly over time, showing convergence toward some regions of the molecular space. Improvement in oscillator strength over time is somewhat linear, showing that STGG+AL could improve further if more Oracle calls were given.

The case study in Figure 3 demonstrates a semi-symmetric scaffold that does not have highly electron-rich nor -deficient fragments, hence it is not expected to bear low energy transitions.

TD-DFT calculations confirms the NIR absorption wavelength found in sTDA-xTB, where the $S_0 \leftrightarrow S_1$ transition is at 973 nm, with a small but not negligible $f_{\text{osc}} = 0.3$. NTO analysis showcases a charge-transfer behavior with a small orbital overlap.

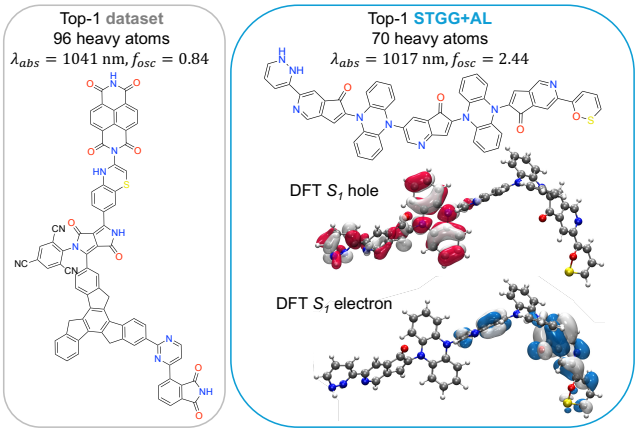


Figure 5: Case study of the top-1 molecule with NIR absorption but the highest f_{osc}.

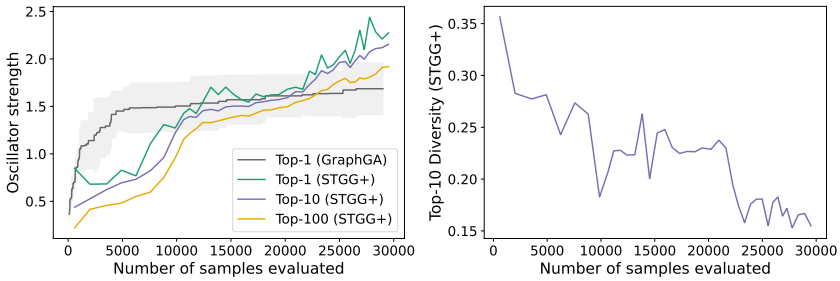


Figure 6: Maximizing f_{osc} using active learning with constraints: near-IR absorption ($\lambda_{abs} \geq 1000$ nm), max 70 heavy atoms, max ring-size of 6. STGG+ (top-1, top-10, top-100; from a single run) vs GraphGA (top-1; average and 95% confidence interval over 3 runs).

[1]sTDA-xTB shows a $f_{\text{osc}} = 27.70$. We note that f_{osc} from different quantum chemical methods may not be directly comparable.

Table 1: Comparing STGG+AL to current molecular design baselines

Method	Max f_{osc} mean (standard-deviation)	Oracle calls mean (standard-deviation)
Maximizing f_{osc}		
Dataset (no atoms limit)	9.30	0
STGG+	13.01	0
STGG+ with active learning	**27.69**	30.0K
REINVENT4	4.53 (0.17)	30.0K
GraphGA	14.56 (1.84)	29.6K (4.0K)
Maximizing f_{osc} in short-wave infrared range		
Dataset (no atoms limit)	0.84	0
STGG+	0.85	0
STGG+ with active learning	**2.44**	30.0K
REINVENT4	0.36 (0.03)	30.0K
GraphGA	1.70 (0.28)	30.9K (1.9K)

4.3 Summary

From Figures 4 and 6, we see that STGG+ already generates high f_{osc} molecules (similar to or above the best value found in the Conjugated-xTB dataset) before starting the active learning, which shows strong out-of-distribution generalization. Over the active learning duration, it learns to generate even high f_{osc}molecules by actively expanding its known region. STGG+ improves consistently over time. Meanwhile, GraphGA learns quickly, but then it saturates to a local optimum and cannot improve further. We note that the molecules produced by the baseline methods have undesirable chemical features such as exotic functional groups (e.g. carbonofluoridoimidic acid) or non-conjugated components (e.g. tetraalkylammonium salt), as shown in Appendix A.5.

The final results for both tasks are shown in Table 1. STGG+ with active learning obtains the molecules with the highest f_{osc}. REIVENT4 performs poorly, while GraphGA manages to obtain high f_{osc} molecules (albeit at a much lower value than STGG+).

5 Conclusion

STGG+ has strong out-of-distribution capabilities up to a certain limit. Active learning allows it to generate more out-of-distribution molecules with improved properties by expanding its realm of knowledge over time. We find that STGG+ with active learning is more sample-efficient compared to RL methods in search of π-conjugated molecules with high f_{osc} and NIR dyes as simulated by semiempirical quantum chemistry methods. The generated molecules are computationally validated and rationalized using the more accurate TD-DFT methods.

We plan to expand our approach to model more complex optoelectronic properties such as fluorescence, which currently remains computationally cost-prohibitive to screen. We note that our method is not without limitation; for example, reward models such as sTDA-xTB (or TD-DFT) often fail to accurately reflect experimental results when pushed to the boundaries, and the diversity of discovered scaffolds can be further improved.

REFERENCES

Hadi Abroshan, Paul Winget, H Shaun Kwak, Yuling An, Christopher T Brown, and Mathew D Halls. Machine learning for the design of novel oled materials. In *Machine Learning in Materials Informatics: Methods and Applications*, pages 33–49. ACS Publications, 2022. 2

Sungsoo Ahn, Binghong Chen, Tianzhe Wang, and Le Song. Spanning tree-based graph generation for molecules. In *International Conference on Learning Representations*, 2021a. 10

Sungsoo Ahn, Binghong Chen, Tianzhe Wang, and Le Song. Spanning tree-based graph generation for molecules. In *International Conference on Learning Representations*, 2021b. 2, 3

Evan R Antoniuk, Peggy Li, Nathan Keilbart, Stephen Weitzner, Bhavya Kailkhura, and Anna M Hiszpanski. Active learning enables extrapolation in molecular generative models. *arXiv preprint arXiv:2501.02059*, 2025. 2

Christoph Bannwarth, Sebastian Ehlert, and Stefan Grimme. Gfn2-xtb—an accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *Journal of chemical theory and computation*, 15 (3):1652–1671, 2019. 3

Camille Bilodeau, Wengong Jin, Tommi Jaakkola, Regina Barzilay, and Klavs F Jensen. Generative models for molecular discovery: Recent advances and challenges. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 12(5):e1608, 2022. 2

Nathan Brown, Marco Fiscato, Marwin HS Segler, and Alain C Vaucher. Guacamol: benchmarking models for de novo molecular design. *Journal of chemical information and modeling*, 59(3): 1096–1108, 2019. 3

Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, et al. Palm: Scaling language modeling with pathways. *Journal of Machine Learning Research*, 24(240):1–113, 2023. 9

Tri Dao. Flashattention-2: Faster attention with better parallelism and work partitioning. *arXiv preprint arXiv:2307.08691*, 2023. 9

Tri Dao, Dan Fu, Stefano Ermon, Atri Rudra, and Christopher Ré. Flashattention: Fast and memory-efficient exact attention with io-awareness. *Advances in Neural Information Processing Systems*, 35:16344–16359, 2022. 9

Jenna C Fromer and Connor W Coley. Computer-aided multi-objective optimization in small molecule discovery. *Patterns*, 4(2), 2023. 2

Stefan Grimme and Christoph Bannwarth. Ultra-fast computation of electronic spectra for large systems by tight-binding based simplified tamm-dancoff approximation (stda-xtb). *The Journal of chemical physics*, 145(5), 2016. 3

Thomas A Halgren. Merck molecular force field. i. basis, form, scope, parameterization, and performance of mmff94. *Journal of computational chemistry*, 17(5-6):490–519, 1996. 3

Dan Hendrycks and Kevin Gimpel. Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*, 2016. 9

Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022. 3, 9

Jan H Jensen. A graph-based genetic algorithm and generative model/monte carlo tree search for the exploration of chemical space. *Chemical science*, 10(12):3567–3572, 2019. 3

Alexia Jolicoeur-Martineau, Aristide Baratin, Kisoo Kwon, Boris Knyazev, and Yan Zhang. Any-property-conditional molecule generation with self-criticism using spanning trees. *arXiv preprint arXiv:2407.09357*, 2024. 1, 2, 3, 9, 10

Diederik P Kingma. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 9

Maksym Korablyov, Cheng-Hao Liu, Moksh Jain, Almer M van der Sloot, Eric Jolicoeur, Edward Ruediger, Andrei Cristian Nica, Emmanuel Bengio, Kostiantyn Lapchevskyi, Daniel St-Cyr, et al. Generative active learning for the search of small-molecule protein binders. *arXiv preprint arXiv:2405.01616*, 2024. 2

Gregory W Kyro, Anton Morgunov, Rafael I Brent, and Victor S Batista. Chemspaceal: An efficient active learning methodology applied to protein-specific molecular generation. *Biophysical Journal*, 123(3):283a, 2024. 2

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. 2, 3

I Loshchilov. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017. 9

Amil Merchant, Simon Batzner, Samuel S Schoenholz, Muratahan Aykol, Gowoon Cheon, and Ekin Dogus Cubuk. Scaling deep learning for materials discovery. *Nature*, 624(7990):80–85, 2023. 2

Marcus Olivecrona, Thomas Blaschke, Ola Engkvist, and Hongming Chen. Molecular de-novo design through deep reinforcement learning. *Journal of cheminformatics*, 9:1–14, 2017. 2, 4

Mariya Popova, Olexandr Isayev, and Alexander Tropsha. Deep reinforcement learning for de novo drug design. *Science advances*, 4(7):eaap7885, 2018. 2

Laura Privitera, Dale J Waterhouse, Alessandra Preziosi, Irene Paraboschi, Olumide Ogunlade, Chiara Da Pieve, Marta Barisa, Olumide Ogunbiyi, Gregory Weitsman, J Ciaran Hutchinson, et al. Shortwave infrared imaging enables high-contrast fluorescence-guided surgery in neuroblastoma. *Cancer Research*, 83(12):2077–2089, 2023. 2, 3

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9, 2019. 9

Sereina Riniker and Gregory A Landrum. Better informed distance geometry: using what we know to improve conformation generation. *Journal of chemical information and modeling*, 55(12): 2562–2574, 2015. 3

Burr Settles. Active learning literature survey. 2009. 2

Noam Shazeer. Glu variants improve transformer. *arXiv preprint arXiv:2002.05202*, 2020. 9

Teague Sterling and John J Irwin. Zinc 15–ligand discovery for everyone. *Journal of chemical information and modeling*, 55(11):2324–2337, 2015. 9

Jianlin Su, Murtadha Ahmed, Yu Lu, Shengfeng Pan, Wen Bo, and Yunfeng Liu. Roformer: Enhanced transformer with rotary position embedding. *Neurocomputing*, 568:127063, 2024. 9

Austin Tripp and José Miguel Hernández-Lobato. Genetic algorithms are strong baselines for molecule generation. *arXiv preprint arXiv:2310.09267*, 2023. 3, 4, 9

Robert H Wilson, Kyle P Nadeau, Frank B Jaworski, Bruce J Tromberg, and Anthony J Durkin. Review of short-wave infrared spectroscopy and imaging methods for biological tissue characterization. *Journal of biomedical optics*, 20(3):030901–030901, 2015. 3

Biao Zhang and Rico Sennrich. Root mean square layer normalization. *Advances in Neural Information Processing Systems*, 32, 2019. 9

A APPENDIX

A.1 ARCHITECTURE AND TRAINING DETAILS

Our STGG+ model uses mainly the same settings as Jolicoeur-Martineau et al. (2024) with some exceptions.

The model is a 3-layer transformer with 16 attention heads, SwiGLU (Hendrycks and Gimpel, 2016; Shazeer, 2020) with expansion scale of 2, no bias term (Chowdhery et al., 2023), Flash Attention (Dao et al., 2022; Dao, 2023), RMSNorm (Zhang and Sennrich, 2019), Rotary embeddings (Su et al., 2024), residual-path weight initialization (Radford et al., 2019).

We use min-max normalization for pre-processing the properties. We train the model using AdamW (Kingma, 2014; Loshchilov, 2017) with $\beta_1 = 0.9$, $\beta_2 = 0.95$, and weight decay 0.1. Since the molecules are large, we use a batch size of 128, learning rate of $2.5e-4$, and max sequence length of 700.

Jolicoeur-Martineau et al. (2024) trained for 50 epochs on Zinc (Sterling and Irwin, 2015) which has 250K molecules; this amounts to 12.5M total training samples seen. Since xTB has around 2.9M molecules, we pre-train for 5 epochs in order to process a similar amount of training samples (14.5M).

Fine-tuning is done for 100 epochs at whichever number of molecules is available ($\leq$ 2000 since we generate 2000 molecules before applying the constraints). This is effectively equivalent to training on $\leq$ 200K samples, which is around 1.4% if the pre-training time. Given our 40 active learning steps, around 55% of the training time that is spent on fine-tuning.

For generation of molecules, we sample uniformly from a range of hyperparameters in order to get diversity. While Jolicoeur-Martineau et al. (2024) only does this for the guidance term in the classifier-free guidance (Ho and Salimans, 2022), we also do it for temperature. We sample a temperature uniformly between 0.6 and 1.1 and guidance between 0.5 and 1.5. This range of values as not been tuned so its possible that there are better choices. We always sample 2000 molecules and remove duplicates and those not respecting the given constraints (max ring size of 6, maximum of 70 atoms).

For the RL baselines, we scaled the oscillator strength to become a reward between 0 and 1 in the following way: $R = min(max(f_{\text{osc}}/27.70, 0), 1)$ for task 1 and $R = min(max(f_{\text{osc}}/2.44, 0), 1)$ for the task 2. 27.7 and 2.44 are respectively the maximum f_{osc} obtained in task 1 and 2 by STGG+. None of the RL baselines reached 1.0 or above (otherwise, we would have rescaled them differently). We also enforced the ring size maximum of 6, number of heavy atoms $\leq$ 70, and NIR range by setting the reward to 0 when any of these constraints are violated.

A.2 RL BASELINE DETAILS

For REINVENT4, we use a standard config as provided by the authors with minimal modifications: using the reinvent prior and agent, batch-size=100, remove duplicates, randomizing smiles, a maximum number of steps of 300 (to reach around 30K Oracle calls), using the default DAP with sigma=128 and learning rate 0.0001, the default diversity filter (Identical Murcko Scaffold), and the default unwanted SMARTS penalty (to penalized wonky molecules).

For GraphGA, to maximize performance, we followed the good choice of hyperparameters by (Tripp and Hernández-Lobato, 2023) which consist in only generating 5 offspring by generation, but generating as many generation as desired (in our case 7500 generations to reach around 30K Oracle calls).

To give an head-start to the RL baselines, we fed the top-100 molecules for each task from the Conjugated-xTB dataset as prior molecules. For REINVENT4, at each iteration, 10 of these 100 molecules where randomly picked and added to the mini-batch to aid learning. For GraphGA these top-100 molecules formed the initial population.

A.3 ARCHITECTURE DIAGRAM

The architecture of STGG+ is shown in Figure 7.

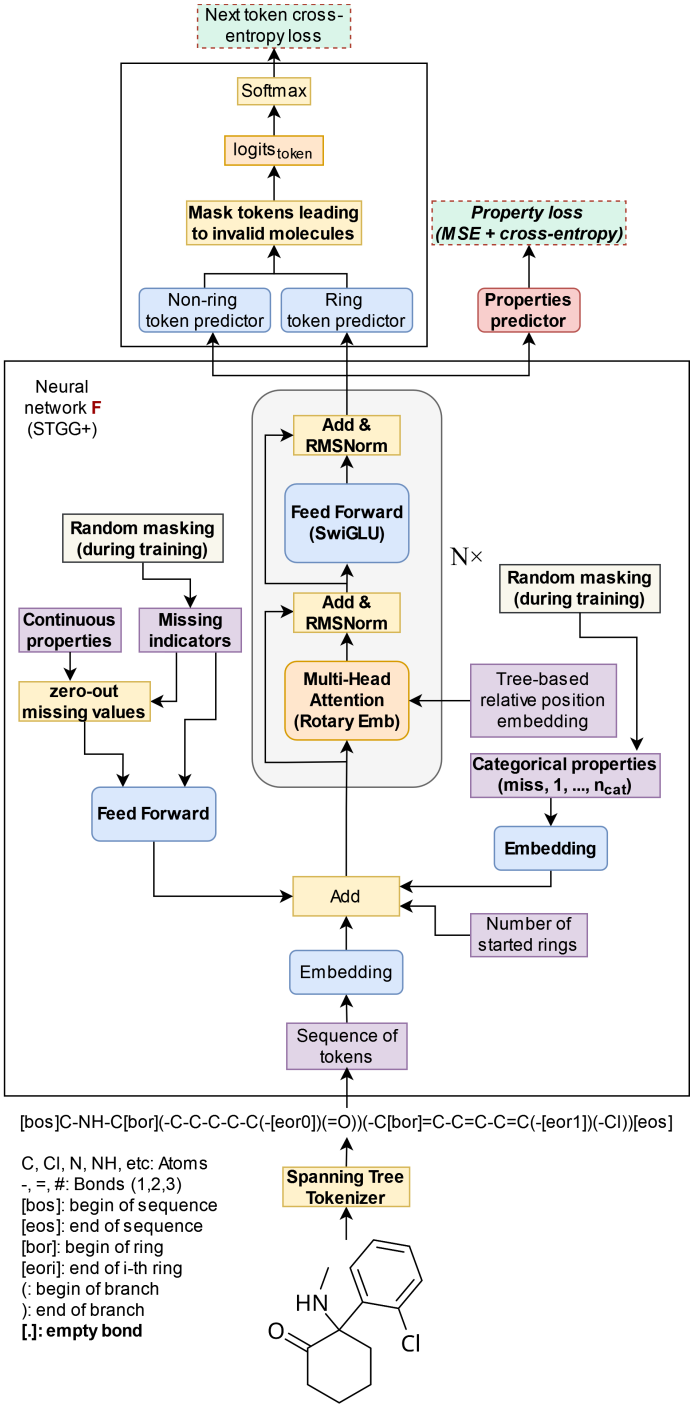


Figure 7: STGG+ architecture. The molecule is tokenized and embedded. The number of started rings and embeddings of continuous and categorical properties are added, and the output is passed to the Transformer. The Transformer output is then split to produce 1) the predicted property and 2) the token predictions (masked to prevent invalid tokens). Novel components compared to STGG (Ahn et al., 2021a) are in **bold**. The figure was taken from Jolicoeur-Martineau et al. (2024).

A.4 THE BEST SMILES GENERATED BY STGG+

Table 2: Top-10 molecules found using STGG+ with active-learning

f_{osc}	Similarity	SMILES
		maximizing f_{osc}
27.69	0.88	Cc1cc2oc3cc(C#Cc4cc5[nH]c(-c6cc7sc(-c8cc9sc(-c%10cc%11sc(-c%12cc%13sc(C#CC%14=C%15N=CC=[N+]%15[B-](Br)(Br)n%15cccc%15%14)cc%13[nH]%12)cc%11[nH]%10)cc9s8)cc7s6)cc5s4)sc3c2s1
27.66	0.88	Brc1cc2c(s1)-c1sc(C#Cc3cc4[nH]c(-c5cc6sc(-c7cc8sc(-c9cc%10sc(-c%11cc%12sc(C#CC%13=C%14N=CC=[N+]%14[B-](Br)(Br)n%14cccc%14%13)cc%12[nH]%11)cc%10[nH]9)cc8s7)cc6s5)cc4s3)cc1C2
27.55	0.87	Clc1cc2[nH]c3cc(C#Cc4cc5sc(-c6cc7sc(-c8cc9sc(-c%10cc%11sc(-c%12cc%13sc(C#CC%14=C%15N=CC=[N+]%15[B-](Br)(Br)n%15cccc%15%14)cc%13[nH]%12)cc%11[nH]%10)cc9s8)cc7s6)cc5s4)sc3c2s1
27.50	0.85	Cc1cc2[nH]c3cc(C#Cc4cc5[nH]c(-c6cc7sc(-c8cc9sc(-c%10cc%11sc(-c%12cc%13sc(C#CC%14=C%15N=CC=[N+]%15[B-](Br)(Br)n%15cccc%15%14)cc%13[nH]%12)cc%11[nH]%10)cc9s8)cc7s6)cc5s4)sc3c2s1
27.38	0.89	Brc1cc2c(s1)-c1sc(C#Cc3cc4[nH]c(-c5cc6sc(-c7cc8sc(-c9cc%10sc(-c%11cc%12[nH]c(C#CC%13=C%14N=CC=[N+]%14[B-](Br)(Br)n%14cccc%14%13)cc%12s%11)cc%10[nH]9)cc8s7)cc6s5)cc4s3)cc1C2
27.35	0.86	Clc1cc2[nH]c3cc(C#Cc4cc5[nH]c(-c6cc7sc(-c8cc9sc(-c%10cc%11sc(-c%12cc%13sc(C#CC%14=C%15N=CC=[N+]%15[B-](Br)(Br)n%15cccc%15%14)cc%13[nH]%12)cc%11[nH]%10)cc9s8)cc7s6)cc5s4)sc3c2s1
27.35	0.88	Brc1cc2sc3cc(C#Cc4cc5[nH]c(-c6cc7sc(-c8cc9sc(-c%10cc%11sc(-c%12cc%13[nH]c(C#CC%14=C%15N=CC=[N+]%15[B-](Br)(Br)n%15cccc%15%14)cc%13s%12)cc%11[nH]%10)cc9s8)cc7s6)cc5s4)[nH]c3c2s1
27.27	0.87	Clc1cc2[nH]c3cc(C#Cc4cc5[nH]c(-c6cc7sc(-c8cc9sc(-c%10cc%11[nH]c(-c%12cc%13sc(C#CC%14=C%15N=CC=[N+]%15[B-](Br)(Br)n%15cccc%15%14)cc%13[nH]%12)cc%11s%10)cc9s8)cc7s6)cc5s4)sc3c2s1
27.22	0.92	Cc1cc2c(o1)-c1[nH]c(C#Cc3cc4[nH]c(-c5cc6sc(-c7cc8sc(-c9cc%10sc(-c%11cc%12[nH]c(C#CC%13=C%14N=CC=[N+]%14[B-](Br)(Br)n%14cccc%14%13)cc%12[nH]%11)cc%10[nH]9)cc8s7)cc6s5)cc4s3)cc1C2
27.09	0.87	Br[B-]1(Br)n2cccc2C(C#Cc2cc3[nH]c(-c4cc5[nH]c(-c6cc7sc(-c8cc9sc(-c%10cc%11sc(C#Cc%12cc%13sc%14cc[nH]c%14c%13[nH]%12)cc%11[nH]%10)cc9s8)cc7s6)cc5s4)cc3s2)=C2N=CC=[N+]21
		maximizing f_{osc} in short-wave infrared range
2.44	0.79	O=C1C(N2c3ccccc3N(C3=Cc4ncc(N5c6ccccc6N(C6=Cc7cnc(C8=CC=CNN8)cc7C6=O)c6ccccc65)cc4C3=O)c3ccccc32)=Cc2cnc(C3=CC=CSO3)cc21
2.30	0.74	O=C1C(N2c3ccccc3N(C3=Cc4cnc(N5c6ccccc6N(C6=Cc7cnc(C8=CC=CSN8)cc7C6=O)c6ccccc65)cc4C3=O)c3ccccc32)=Cc2cnc(C3=CC=CNN3)cc21
2.28	0.75	O=C1C(N2c3ccccc3N(C3=Cc4cnc(C5=CC=CSN5)cc4C3=O)c3ccccc32)=Cc2ccc(N3c4ccccc4N(C4=Cc5cc(C6=CC=CNN6)ncc5C4=O)c4ccccc43)cc21
2.27	0.81	O=C1C(N2c3ccccc3N(C3=Cc4cc(N5c6ccccc6N(C6=Cc7cnc(C8=CC=CSN8)nc7C6=O)c6ccccc65)cnc4C3=O)c3ccccc32)=Cc2cnc(C3=CC=CNO3)cc21
2.22	0.83	O=C1C(N2c3ccccc3N(c3cnc4c(c3)C(=O)C(N3c5ccccc5N(C5=Cc6ccc(C7=CC=CSO7)nc6C5=O)c5ccccc53)=C4)c3ccccc32)=Cc2cnc(C3=CC=CSN3)cc21
2.21	0.80	O=C1C(N2c3ccccc3N(C3=Cc4ccc(C5=CC=CSN5)nc4C3=O)c3ccccc32)=Cc2ncc(N3c4ccccc4N(C4=Cc5cnc(C6=CC=CSN6)nc5C4=O)c4ccccc43)cc21
2.20	0.79	O=C1C(N2c3ccccc3N(C3=Cc4cc(N5c6ccccc6N(C6=Cc7cnc(C8=CC=CSO8)cc7C6=O)c6ccccc65)cnc4C3=O)c3ccccc32)=Cc2cnc(C3=CC=CNS3)cc21
2.21	0.81	O=C1C(N2c3ccccc3N(C3=Cc4cc(N5c6ccccc6N(C6=Cc7cnc(C8=CC=CNN8)nc7C6=O)c6ccccc65)cnc4C3=O)c3ccccc32)=Cc2cnc(C3=CC=CSN3)cc21
2.19	0.83	O=C1C(N2c3ccccc3N(c3ccc4c(n3)C=C(N3c5ccccc5N(C5=Cc6cnc(C7=CC=CNN7)nc6C5=O)c5ccccc53)C4=O)c3ccccc32)=Cc2cnc(C3=CC=CSN3)cc21
2.18	0.78	O=C1C(N2c3ccccc3N(C3=Cc4cnc(C5=CC=CNN5)nc4C3=O)c3ccccc32)=Cc2ccc(N3c4ccccc4N(C4=Cc5ncc(C6=CC=CNN6)nc5C4=O)c4ccccc43)cc21

A.5 THE BEST MOLECULES GENERATED BY BASELINE METHODS

A.5.1 STGG+

The top-1 molecules generated by STGG+ are shown in Figures 8-9. The molecules are generally sensible and plausible, respecting proper valency. We note that certain structures can still be exotic (e.g. 1,2-oxathiine, dipyrromethene borondibromide), but they are nevertheless previously known compounds and do not affect the core scaffolds.

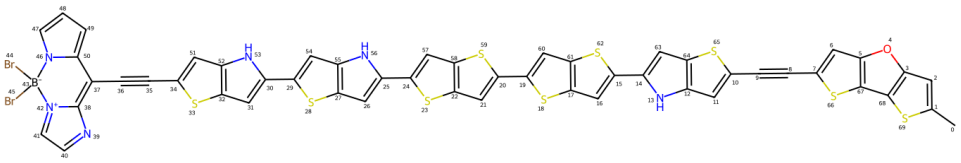


Figure 8: STGG+ Top-1 molecule with the highest f_{osc} out of a single run ($f_{\text{osc}} = 27.69$).

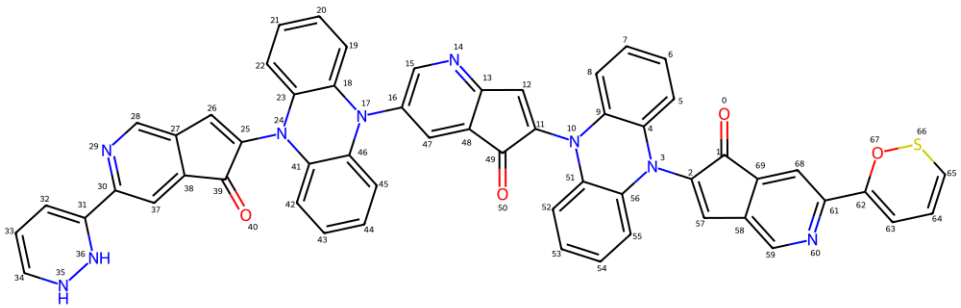


Figure 9: STGG+ Top-1 molecule with the highest f_{osc} and near-IR absorption out of a single run ($f_{\text{osc}} = 2.44$).

A.5.2 REINVENT4

The top-1 molecules generated by REINVENT4 are shown in Figures 10-11.

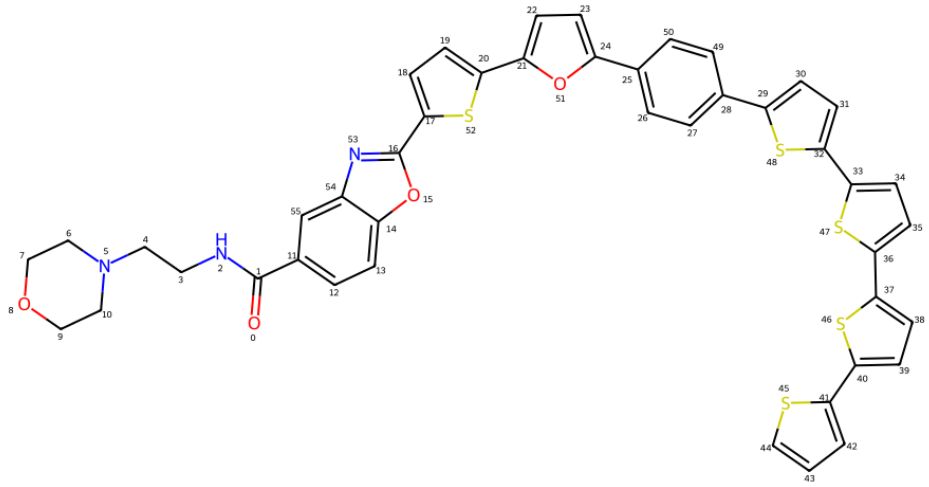


Figure 10: REINVENT4 Top-1 molecule with the highest f_{osc} out of 3 runs ($f_{\text{osc}} = 4.65$). This polythiophene derivative has a long non-conjugated group that does not contribute to f_{osc}.

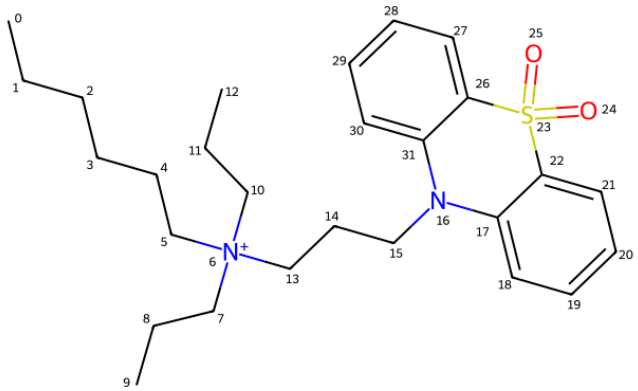


Figure 11: REINVENT4 Top-1 molecule with the highest f_{osc} and near-IR absorption out of 3 runs ($f_{\text{osc}} = 0.40$). This phenothiazine dioxide lacks conjugation, is not unexpected to be absorptive in NIR, and has an non-conjugated tetralkylammonium salt pendant group.

A.5.3 GRAPHGA

The top-1 molecules generated by GraphGA are shown in Figures 12-13. The best molecule for maximizing f_{osc} without constraint is extremely implausible and unlikely to be synthesizable.

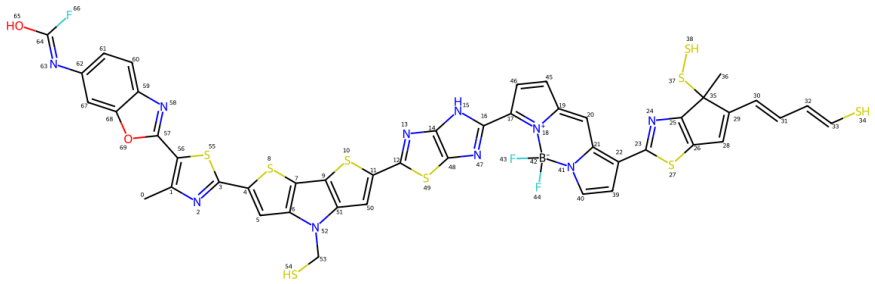


Figure 12: GraphGA Top-1 molecule with the highest f_{osc} out of 3 runs ($f_{\text{osc}} = 15.81$). This molecule has several undesirable functional groups including carbonofluoridoimidic acid and disulfaneylmethylcyclopenta[d]thiazole.

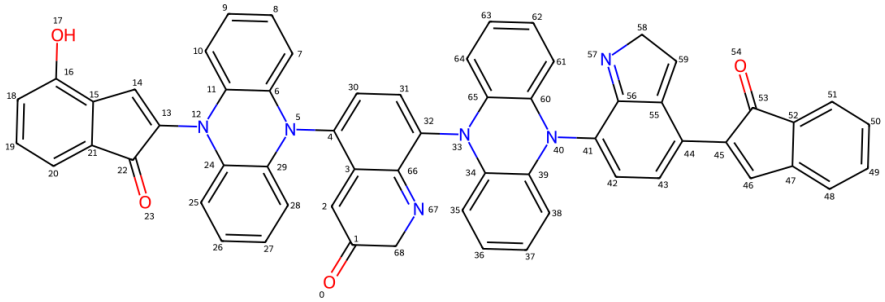


Figure 13: GraphGA Top-1 molecule with the highest f_{osc} and near-IR absorption out of 3 runs ($f_{\text{osc}} = 2.03$). This molecule is largely chemically sound, with the exception of the quinolin-3(2H)-one.