

1 We would like to thank the reviewers for their thoughtful comments. Detailed responses are below:

2 **Reviewer 1** [*..on clarity..*] We appreciate your compliments on the writing and background section. We will
3 follow your suggestion and add a figure showing the full Molecule Chef model as well as adding more details of the
4 architecture and hyperparameters to the appendix (so that they can be found more conveniently than from looking
5 through our code). Thank you for this suggestion.

6 [*..on another Fig. 6..*] Thanks, we will add to the appendix a third example figure, with a different initial molecule.

7 [*..on difference between correlations between reachable and unreachable products..*] We believe this difference is
8 because our latent space is only trained on reachable product reactions and so is better able to model these reactions.
9 Furthermore, some of the unreachable products may also require reactants that are not available in our pool of easily
10 available reactants, at least when considering one-step reactions. Including larger reactant vocabularies and predicting
11 of multi-stage reactions are two things we wish to look into in future work. However, given that unreachable products
12 have at least one reactant which is not in Molecule Chef's vocabulary, we think it is very encouraging that there still is
13 some, albeit smaller, correlation with the true QED. This is because it shows that our model can suggest molecules with
14 similar properties made from reactants that are available. Thanks for bringing this up, we will clarify this in the paper.

15 [*..on rearranging some of the dataset details from the appendix..*] Thanks for the suggestion, we will do this.

16 **Reviewer 2** Thank you for your positive comments and for saying that the 'paper should be accepted as it is'. We
17 are happy to hear that you found the provided code useful. We also appreciate you highlighting the importance of the
18 problem our model addresses and for recognizing our contributions, in particular the built-in generation of synthesis
19 routes for suggested products and the generation of chemically sensible samples, 'in contrast with previous works
20 showing undesirable molecular features'.

21 [*..on your 'minor comment' regarding our section 4.4..*] We obtained these annotations from an expert organic chemist.
22 For example, in the first row of Figure 6, the 4th compound is unstable because it contains an enol, which will
23 tautomerize (transform itself) to an aldehyde quickly. The 6th molecule in this row contains an aminal, which will
24 rapidly decompose into the free acid, formaldehyde and ammonia. We will have them make their annotation criteria
25 explicit in the paper. Thanks for this suggestion.

26 **Reviewer 3** [*..on comparing to the Kajino ICML 2019 paper..*] We think this is missing the point of our work. We
27 are happy to cite this paper, but ultimately our goal is not to compare with every last molecular generative model (there
28 are currently more than 15). Instead we wanted a comprehensive comparison with strong baseline methods (LSTM,
29 CVAE, GVAE) as well as state-of-the-art methods (AAE, CGVAE). In our paper we argued that our method has two
30 key advantages over these previous methods, which also holds over the one you cite.

31 First, it ensures that the generated molecules *can be synthesized* (and provides these synthetic routes), while remaining
32 competitive with current methods that do not have this restriction. To our knowledge, our model is the first to ensure
33 synthesizability whilst jointly optimizing. Suggesting molecules that can actually be made is key in practical drug and
34 materials design. Second, many of the previous models often generate chemically unstable molecules, which is not
35 captured in validity measures. The CGVAE that we compare against, as well as the MHG-VAE from the paper you cite,
36 are guaranteed to generate 100% valid molecules. However, this just ensures that the molecules can be parsed by the
37 cheminformatics library RDKit, e.g. by ensuring correct valencies. These models do not guarantee that the suggested
38 molecules are *chemically sensible* (ie are stable enough to exist on earth) or are non-toxic. Whereas we show that the
39 Molecule Chef is able to generate realistic, sensible and safe molecules (see Figure 6). Additional evidence of this is
40 the 'Quality' GuacaMol score [6] in Table 1 which gives low scores to molecules that are "potentially unstable, reactive,
41 laborious to synthesize, or simply unpleasant to the eye of medicinal chemists", and also in Appendix A.2.

42 [*..on providing further details of our method in the supplement..*] We will add further details of our architecture to the
43 supplement.

44 [*..on the Segler et al.'s method..*] We believe there is some confusion here: we did not use the above method. The
45 Segler model is only able to provide a plan towards a given molecule. It cannot perform any optimization of molecules.

46 [*..on adding a detailed readme and datasets to the provided code..*] Our plan was always to open-source the code
47 on publication with a detailed README, links to the datasets/weights, and a Docker Image (defining our package
48 environment) to make it easy to immediately use our code. The data was too large to upload at submission time, but we
49 uploaded the code to get feedback on its organization, and for additional clarification on architecture/hyperparameters.