## A Metric Details

In this section we provide details for the metrics reported in Table 2.

A few of the similarity measures (SNN and IntDiv) are based on the *Tanimoto coefficient*. In order to compute the Tanimoto coefficient, the molecules are mapped to a vector of fingerprints where each bit in the vector represents the presence (or absence) of a specific fragment.[6] For molecules $A, B$, denote their fingerprints by $m_A$ and $m_B$ respectively, the Tanimoto coefficient is then calculated as the Jaccard index of the two vectors,

$$J(m_A, m_B) = \frac{|m_A \cap m_B|}{|m_A \cup m_B|} = \frac{|m_A \cap m_B|}{|m_A| + |m_B| - |m_A \cap m_B|}. \tag{A.1}$$

We denote the Tanimoto coefficient of molecules $A, B$ by $T(A, B)$.

**Unique@K** report the fraction of uniquely generated valid SMILES strings amongst the $K$ molecules generated (validity is determined by the RDKit library). We generate $30,000$ molecules for each model and report for $K = 1,000$ and $K = 10,000$. High uniqueness values ensure the models do not collapse into repeatedly producing the same set of molecules.

**Fréchet ChemNet Distance (FCD)** is a metric for evaluating generative models in the chemical context, the method is based on the well established *Fréchet Inception Distance* (FID) metric used to evaluate the performance of generative models in computer vision (Heusel et al., 2017).

Fréchet distance measure the Wasserstein-2 distance (Vaserstein, 1969) from the distributions induced by taking the activations of the last layer of a relevant deep neural net. In the case of FCD, molecule activations are probed from ChemNet (Mayr et al., 2018). Given a set of generated molecules, denote by $G$ the set of vectors as obtained by the activations of ChemNet, one can calculate the mean and covariance $\mu_G$ and $\Sigma_G$. Similarly, denote $\mu_R$ and $\Sigma_R$ the mean and covariance of the set of molecules in the reference set, the FCD is calculated as follows,

$$FCD(G, R) = \|\mu_G - \mu_R\|^2 + Tr\left(\Sigma_G + \Sigma_R - 2(\Sigma_G \Sigma_R)^{1/2}\right). \tag{A.2}$$

where $Tr(M)$ denotes the trace of the matrix $M$. Low FCD values indicate that the generated molecules distribute similarly to the reference set.

**Similarity to Nearest Neighbor (SNN)** is the average of the Tanimoto coefficient of the generated molecule set denoted by $G$ and their respective nearest neighbor in a reference set of molecules denote by $R$. High SNN indicates the generated molecules have similar structures to those in the reference set. This metric is in the range of $[0, 1]$.

**Fragment similarity (Frag)** is a fragment similarity measure based on the BRICS fragments (Degen et al., 2008). Denote the set of BRICS fingerprints vectors of the generated molecules by $G$ and similarly $R$ for the reference molecules. The fragment similarity is defined as the cosine similarity of the sum vectors,

$$Frag(G, R) = cosine\left(\sum_{g \in G} g, \sum_{r \in R} r\right) \tag{A.3}$$

The Frag measure is in the range of $[0, 1]$, values closer to 1 indicate that the generated and reference molecule set have a similar distribution of BRICS fragment.

[6] The molecular fingerprints are obtained from RDKit (Landrum, 2006) and are based on the extended-connectivity fingerprints (Rogers & Hahn, 2010).

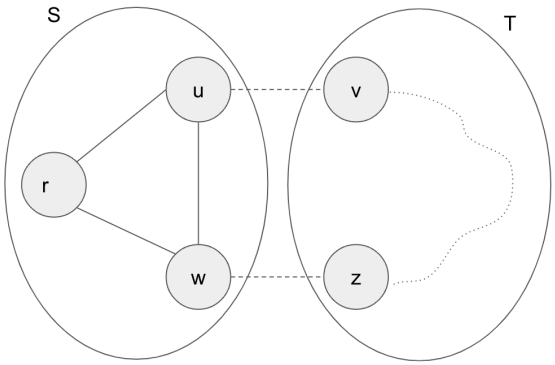


Figure 3: Proof illustration - $S$ has a cycle and two different trajectories starting from $u$ and ending with $w$ ($urw$ and $uw(r)$. Concatenating with the trajectory from $z$ to $v$ we obtain two different DFS trajectories with a shared suffix.

**Scaffold similarity (Scaff)** is similar to the fragment similarity, instead of the BRICS fragment, Scaff is based on mapping molecules to their Bemis–Murcko scaffolds (Bemis & Murcko, 1996).[7] The measure also has a range of $[0, 1]$, values closer to 1 indicate that the generated molecule set has a similar distribution of scaffold to the reference set.

**Internal diversity (IntDiv)** is a mesure of the chemical diversity within a generated set of molecules $G$. This metric indicates

$$IntDiv_p = 1 - \left( \frac{1}{|G|^2} \sum_{A,B \in G} T(A, B)^p \right)^{1/p} \tag{A.4}$$

We report the internal diversity for $p = 1, 2$. This measure is in the range $[0, 1]$. Low values indicate a lack of diversity in the generated molecules, i.e. that the model outputs molecules with similar fingerprints.

**Filters** is the fraction of generated molecules that pass a certain filtering that has been applied to the training data. The metric is in the range of $[0, 1]$, high values indicate that the model has learnt to generate molecules which avoid the structures omitted by the filtering process.

**Novelty** is the fraction of generated molecules that does not appear in the training set. This measure is in the range of $[0, 1]$ and is an indication of the whether the model overfits the training data.

## B Missing Proofs

In this section we show how to construct distinct DFS trajectories with common end vertex for a $2-$edge connected graph conditioned that the graph is not a cycle.

*Proof.* From our assumption that the graph is not a cycle, there exists at least two nodes with degree $\geqslant 3$. Denote by $C = (S, T)$ a minimal cut of size 2 (such a cut exists from our assumption that the graph is 2-connected). Denote the edges of the minimal cut by $e_1 = (u, v)$ and $e_2 = (w, z)$ such that $u, w \in S$ and $v, z \in T$. Next, we claim that at least one of the partitions contains a cycle, otherwise there is a path connecting $S$ and $T$ since there are nodes in the graph which have a degree of 3 in the original graph with a path between them. Assume with out loss of generality that $S$ is the partition with a cycle, therefore there are at least 2 different traversals of $S$ that start with $u$ and end with $w$. There is also a trajectory between

[7] Bemis–Murcko scaffold is the ring structure of a molecule along with the bonds connecting the rings, i.e. the molecule without the side chains.

$z$ and $v$. Putting together, there are at least 2 trajectories of the entire graph with a common suffix which is the traversal of $T$. Figure 3 illustrates the proof concept. □