

# When do GFlowNets Learn the Right Distribution?

Tiago da Silva, Rodrigo Alves, Eliezer Souza, Amauri Souza, Samuel Kaski, Vikas Garg, Diego Mesquita



**KEYWORDS** — GFlowNets, GNNs

TL;DR

- we demonstrate that the impact of an imbalanced transition on the GFlowNet’s correctness is unevenly distributed across the network,
- we delineate the expressiveness of a GFlowNet parameterized by a 1-WL GNN in terms of the distributions it can learn,
- we introduce FCS as a tractable and accurate goodness-of-fit measure for GFlowNets
- and we show that conventionally used diagnostic procedures often misrepresent the distributional correctness of a GFlowNet.

## I. BACKGROUND: WHAT IS A GFLOWNET?

**GFlowNets** are amortized algorithms that sample from distributions over discrete and compositional objects (such as graphs) by learning a Markov Decision Process (MDP) over an extension of the target space.

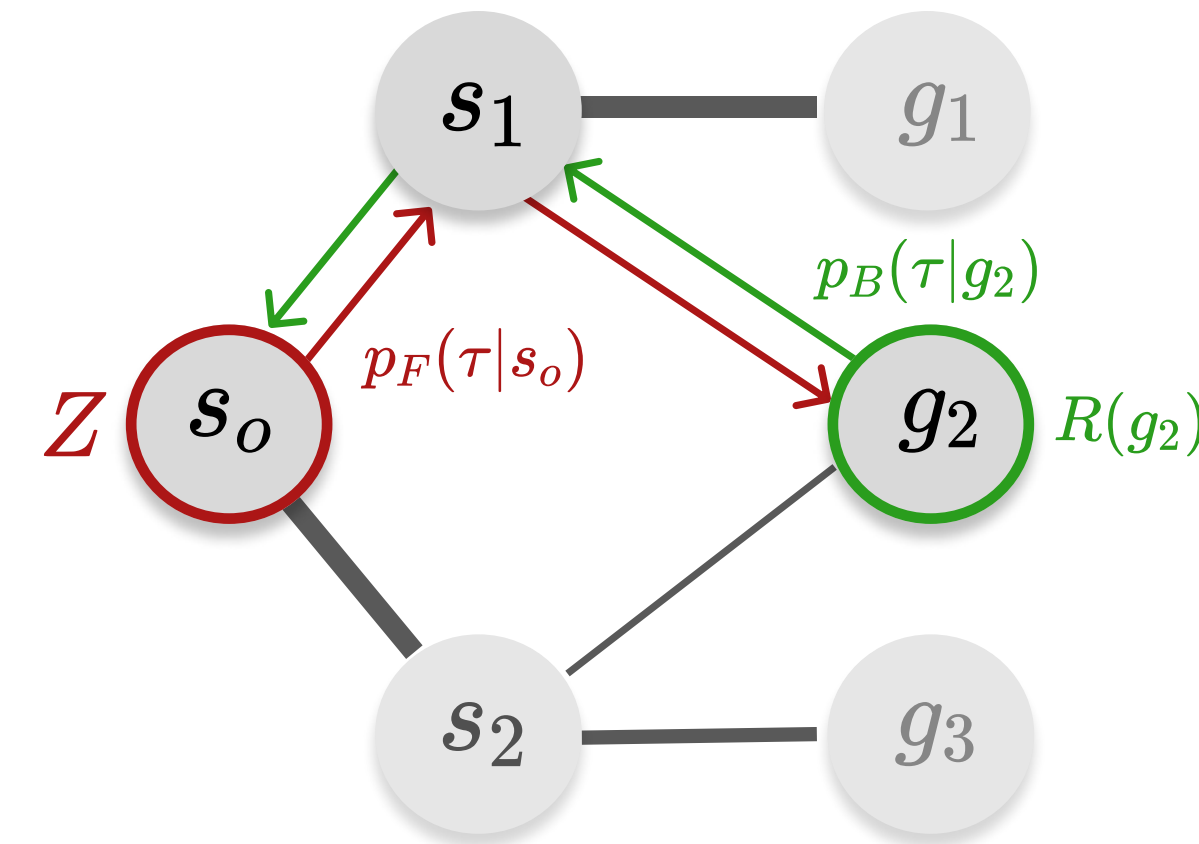


Figure 1: A GFlowNet learns a **forward policy** on a state graph.

To achieve this, a **flow network** is defined over an extension  $\mathcal{S}$  of  $\mathcal{G}$ , which represents the sink nodes. To navigate through this network and sample from  $\mathcal{G}$  proportionally to a **reward function**  $R : \mathcal{G} \rightarrow \mathbb{R}_+$ , a forward (resp. backward) policy  $p_F(\tau)$  ( $p_B(\tau|x)$ ) is used.

$$p_F(\tau) = \prod_{(s,s') \in \tau} p_F(s'|s) \text{ and } \sum_{\tau \sim \mathcal{G}} p_F(\tau) = R(g). \quad (1)$$

To achieve this, we parameterize  $p_F(\tau)$  as a neural network trained by minimizing

$$\mathcal{L}_{TB}(p_F) = \mathbb{E} \left[ \left( \log \frac{p_F(\tau)Z}{p_B(\tau|x)R(x)} \right)^2 \right]. \quad (2)$$

for a given  $p_B(\tau|x)$ . GFlowNets can be trained in an **off-policy** fashion and the above expectation can be under any full-support distribution over trajectories. When  $\mathcal{L}_{TB}(p_F) = 0$ ,

$$p_\tau(x) = \sum_{\tau \text{ finishes at } x} p_F(\tau) \propto R(x). \quad (3)$$

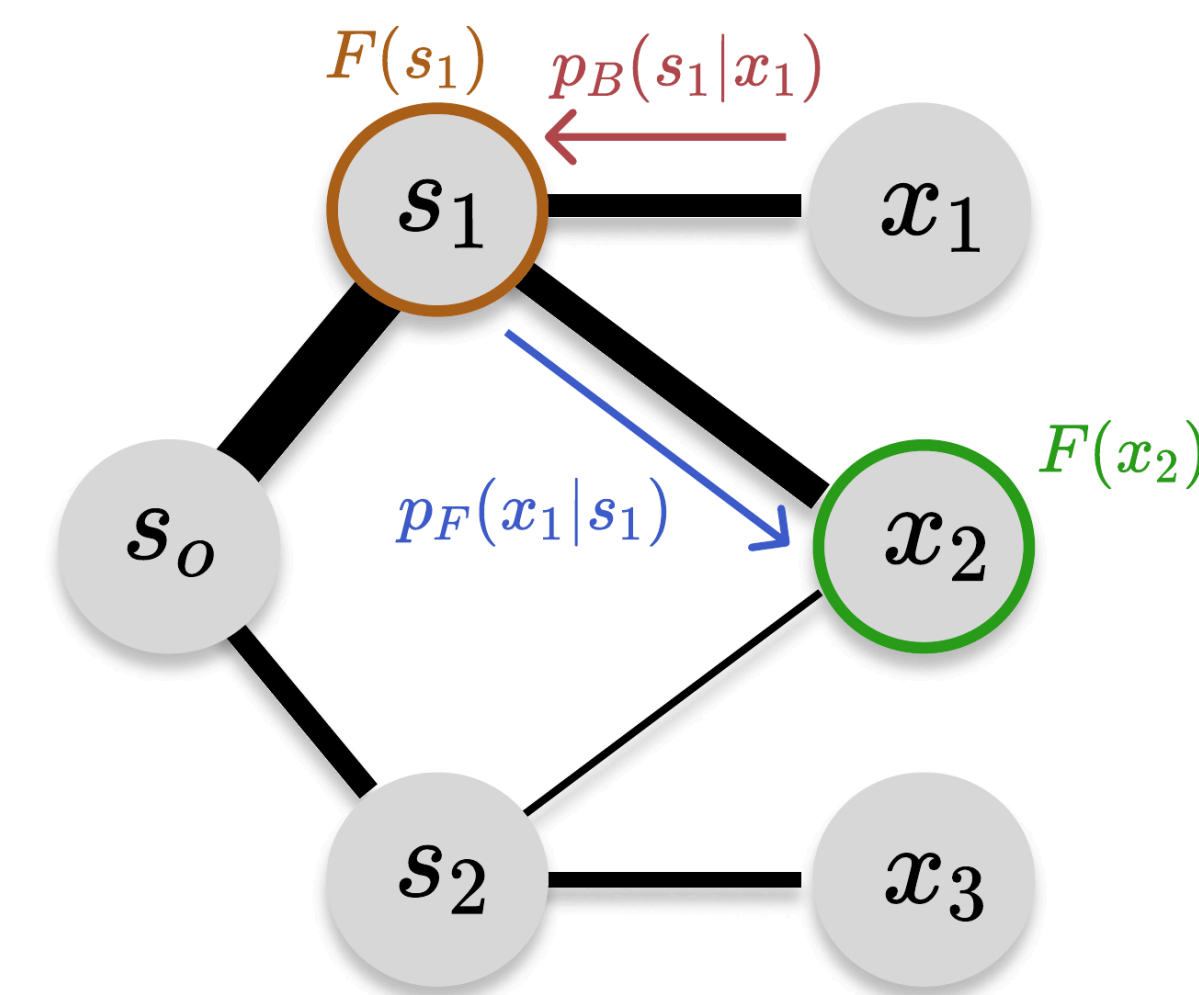


Figure 2: A GFlowNet may also learn a **forward policy** and a **flow function** on a state graph.

When trajectories are long and memory is constrained, we instead parameterize a flow function  $F : \mathcal{G} \rightarrow \mathbb{R}$  and minimize the **detailed balance** (DB) loss

$$\mathcal{L}_{DB}(s, s') = \left( \log \frac{F(s)p_F(s'|s)}{F(s')p_B(s|s')} \right). \quad (4)$$

Conventionally,  $F$  and  $p_F$  are learned by optimizing a uniformly weighted averaged of  $\mathcal{L}_{DB}(s, s')$  over a trajectory, i.e.,  $\frac{1}{|\tau|} \sum_{(s,s') \in \tau} \mathcal{L}_{DB}(s, s')$ , in which  $|\tau|$  is the trajectory’s length.

## II. QUESTION: WHAT ARE THE LIMITS OF GFLOWNETS?

Our work addresses three major challenges in the GFlowNet literature:

1. How do an insufficiently minimized  $\mathcal{L}_{DB}$  affect the distributional accuracy of a GFlowNet?
2. What functions  $R$  cannot be sampled from when  $p_F$  is parameterized by a GNN with 1-WL expressivity?
3. How can we tractably measure the proximity of a GFlowNet to its target distribution?

Jointly, we seek a deeper understanding on the consequences (1), causes (2), and assessment (3) of GFlowNets.

## III. WHAT ARE THE CONSEQUENCES OF AN IMBALANCED TRANSITION?

We estimate the extent to which a lack of balance within an edge of the flow network affect the marginal distribution over the terminal states of the corresponding MDP.

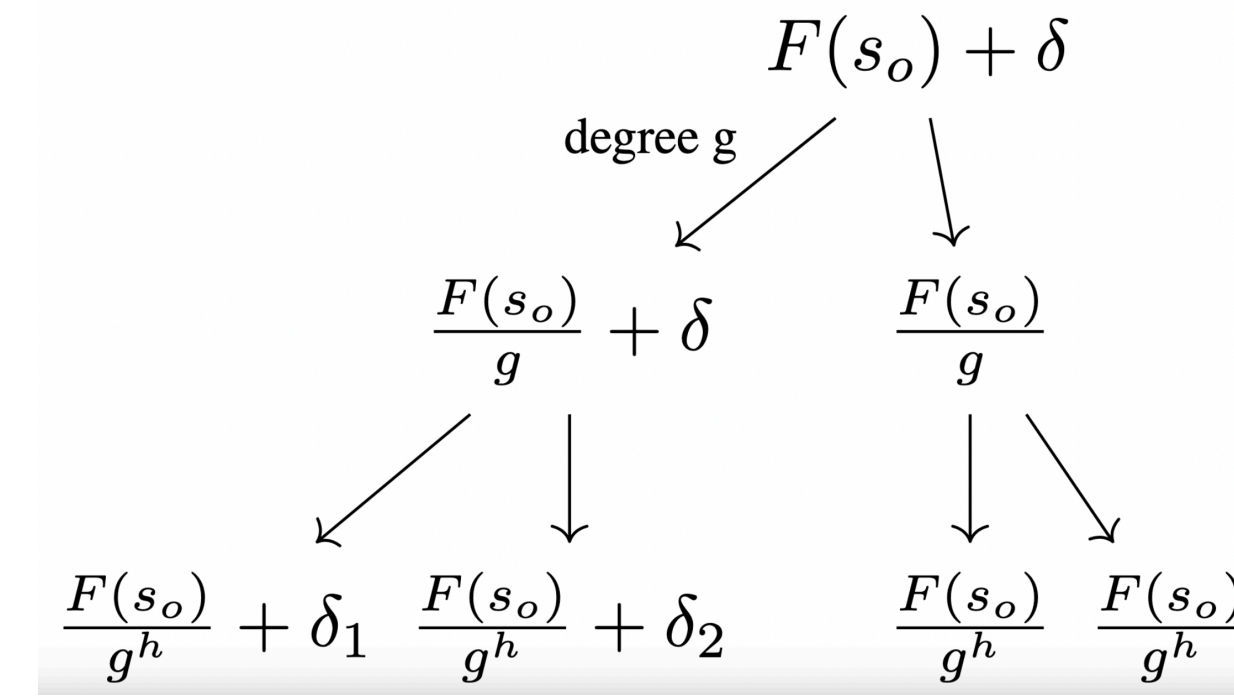


Figure 3: An imbalanced flow network.

**Error propagation in GFlowNets.** Figure 3 presents a scenario in which an extra volume of  $\delta$  flows through one of the flow network’s edges. Under these conditions, the perturbed marginal distribution  $\tilde{p}_\tau$  over  $\mathcal{G}$  satisfies

$$\varepsilon(\delta, g, F(s_o)) \leq \text{TV}(\tilde{p}_\tau, \pi) \leq \varepsilon(\delta, g^h, F(s_o)), \text{ with } \varepsilon(\delta, x, t) = \left(1 - \frac{1}{x}\right) \frac{\delta}{t + \delta}, \quad (5)$$

$\pi(x) \propto R(x)$ , and  $\text{TV}(\tilde{p}_\tau, \pi) = \frac{1}{2} \sum_{x \in \mathcal{G}} |\pi(x) - \tilde{p}_\tau(x)|$  is the total variation distance.

In doing so, we observe that the impact of each transition over the downstream distribution is heterogeneously distributed throughout the network.

Drawing on this result, we repurpose the DB loss to allow for non-uniform weighting of the transitions within a trajectory. More precisely, we let  $\gamma : \mathcal{S} \rightarrow \mathbb{R}$  be state-weighting and define

$$\mathcal{L}_{WDB}(\tau) = \frac{1}{|\tau|} \sum_{s,s'} \gamma(s) \mathcal{L}_{DB}(s, s') \quad (6)$$

as the weighted detailed balance. During training, we let  $\gamma(s) \propto (\text{depth of } s)$  linearly anneal  $\gamma(s)$  to 1. Empirical results show that our  $\mathcal{L}_{WDB}$  often leads to faster or comparable convergence than the traditional DB loss.

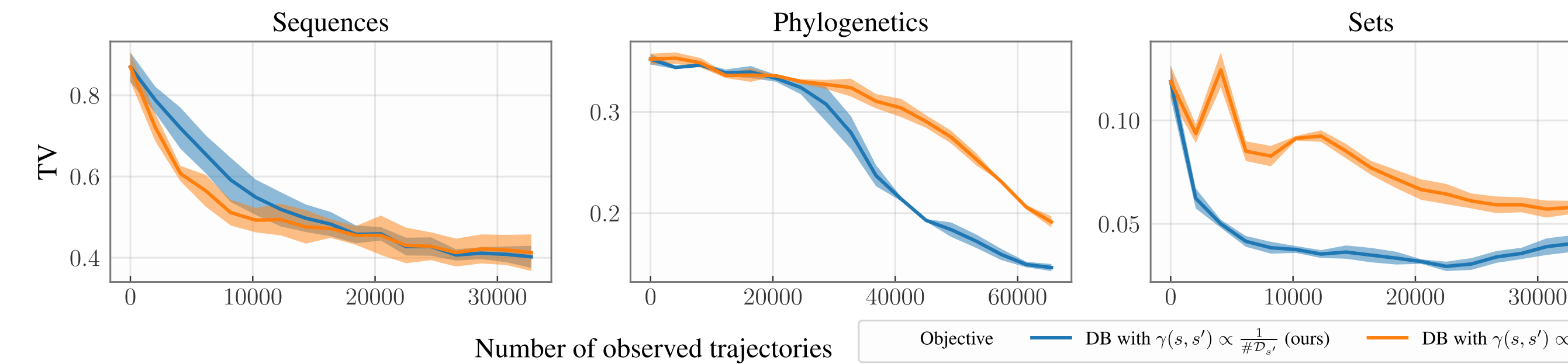


Figure 4:  $\mathcal{L}_{WDB}$  often leads to faster learning convergence than  $\mathcal{L}_{DB}$  in benchmark tasks.

**Open question.** While our findings demonstrate the suboptimality of an unweighted detailed balance loss, the search for an optimal weighting scheme for  $\mathcal{L}_{DB}$  remains an open challenge in the GFlowNet literature.

## IV. WHAT ARE THE CAUSES FOR AN IMBALANCED TRANSITION?

The flow network will only be balanced if the corresponding GFlowNet perfectly minimizes its learning objective. When a GNN is used to parameterize the  $p_F$ , as it commonly is, this minimization may not be achievable.

**Limitations of GFlowNets parameterized by 1-WL GNNs.** We consider distributions over graph-structured objects. Assume that a GFlowNet is parameterized as

$$p_F(s, s') \propto \exp\{\text{MLP}(\varphi_{(u|s)} \parallel \varphi_{(v|s)})\}, \quad (7)$$

in which  $\varphi_{u|s}$  (resp.  $\varphi_{v|s}$ ) is a representation of the node  $u$  (resp.  $v$ ) computed by a GNN with 1-WL expressivity;  $u$  and  $v$  are the endpoints of the added edge in the transition  $s \rightarrow s'$ .

Then, there are specific combinations of state graphs and reward functions for which a GFlowNet cannot solve the underlying flow assignment problem. Figure 7 and Figure 5 shows an example.

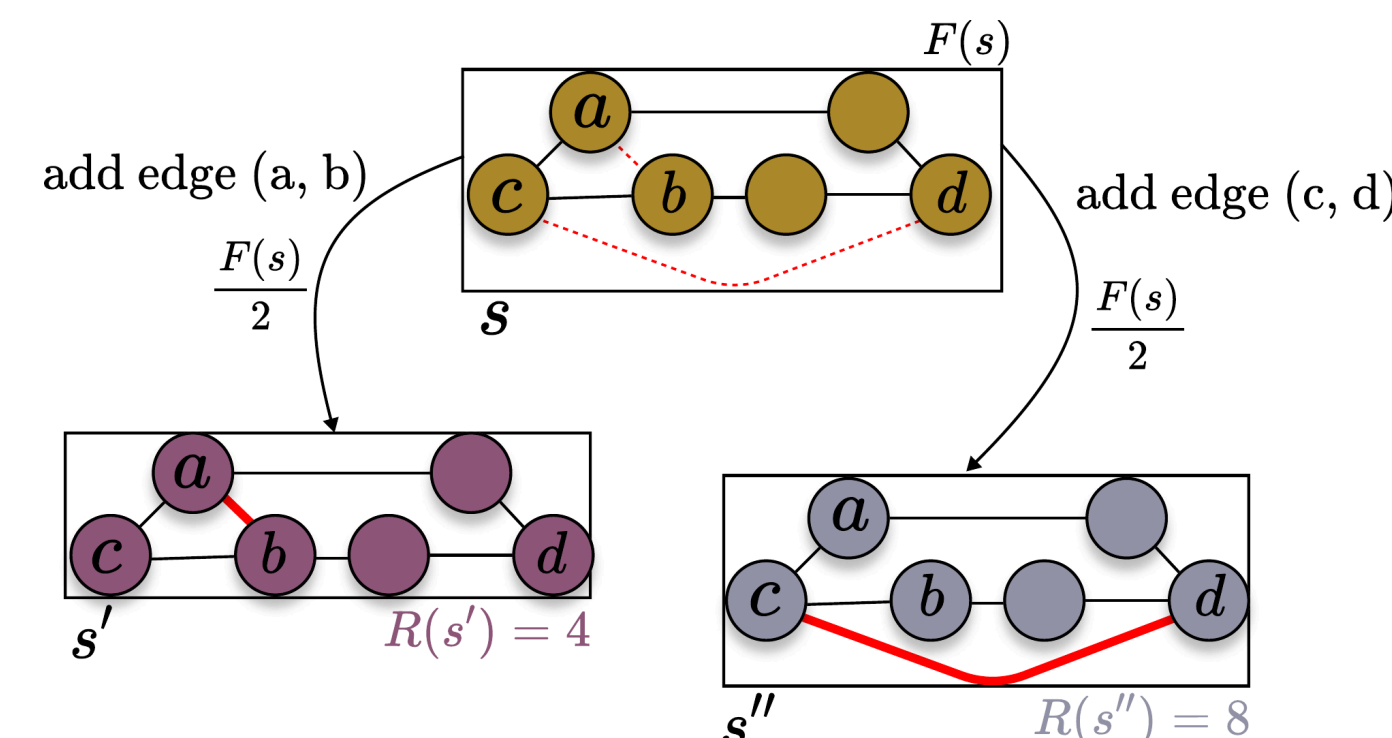
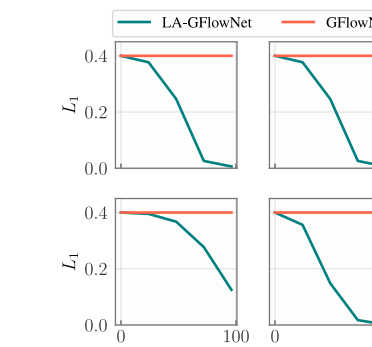


Figure 5: A GNN GFlowNet cannot distinguish the left and right children of  $s$  - in spite of their different rewards. To circumvent this issue, we reparameterize  $p_F$  to incorporate the **children state** into the current transition  $s \rightarrow s'$ .

$$p_F(s, s') \propto \exp\{\text{MLP}\left(\psi_1\left(\{\varphi_{(u|s)}, \varphi_{(v|s)}\}\right) \parallel \psi_2\left(\{\varphi_{(w|s')}\}_{w \in V(s')}\right)\right)\}, \quad (8)$$



in which  $\psi_1$  and  $\psi_2$  are permutation-invariant functions and  $V(s')$  is the set of nodes of  $s'$ . The resulting method, termed Look Ahead (LA) GFlowNets, is strictly more expressive than a standard GFlowNet.

**Open question.** Although we have established the first results regarding the limited expressivity of a GFlowNet, a complete characterization of the model’s expressivity is open.

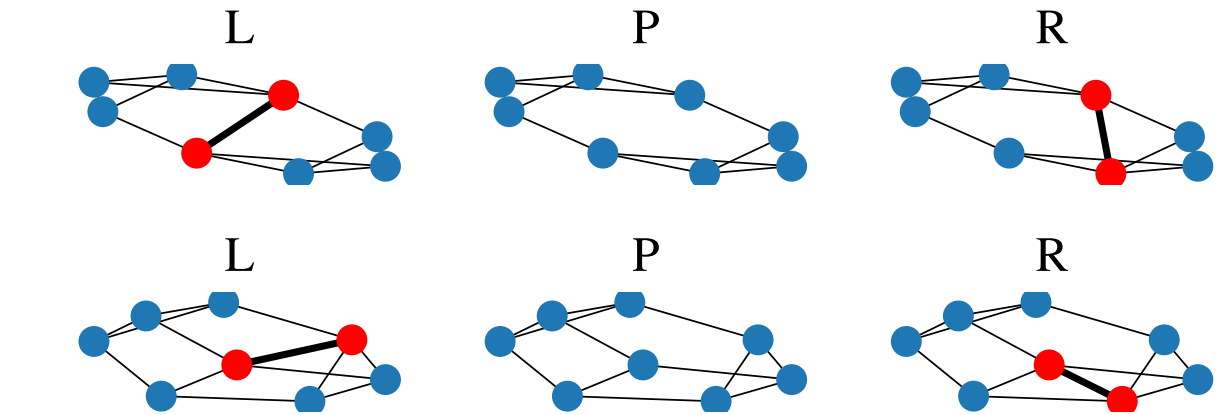


Figure 7: A GFlowNet parameterized as in Equation 7 cannot distinguish the actions  $P \rightarrow R$  and  $P \rightarrow L$ .

## V. HOW TO MEASURE IMBALANCE?

Due to the combinatorial nature of the state space, we cannot directly compare the distribution learned by a GFlowNet against the target distribution cannot be tractably computed, i.e.,

$$\text{TV}(p_\tau, \pi) = \frac{1}{2} \sum_{x \in \mathcal{G}} |p_\tau(x) - \pi(x)| \text{ is intractable.} \quad (9)$$

Inspired by the Reinforcement Learning literature, prior works have computed quantities such as the top-quantile of the reward (TopQuantile) and the number of modes (NumModes) to decide upon which GFlowNet converges faster, which are respectively defined as

$$\text{TopQuantile}(R, \alpha, \mathcal{X}_n) = \text{Avg}(\{R(x) : x \in \mathcal{X}_n \wedge R(x) \geq \text{Quantile}(R(\mathcal{X}_n), \alpha)\}) \quad (10)$$

$$\text{NumModes}(R, \mathcal{X}_n, \xi) = |\{x \in \mathcal{X}_n : R(x) \geq \xi\}| \quad (11)$$

in which  $\mathcal{X}_n = \{x_1, \dots, x_n\} \subset \mathcal{G}$  is a set of samples that a GFlowNet found during training and  $\xi > 0$  is an arbitrary threshold. We show that neither of these quantities accurately reflect the distributional accuracy of a GFlowNet.

**Flow Consistency in Subgraphs.** Instead, we propose measuring the Flow Consistency in Subgraphs (FCS) for assessing GFlowNets. In a nutshell, FCS is the expected TV between  $p_\tau$  and  $\pi$  in random subsets of  $\mathcal{G}$ ,

$$\text{FCS}(\pi, p_\tau) = \mathbb{E}_{\mathcal{B}} [\text{TV}(\pi^{\mathcal{B}}, p_\tau^{\mathcal{B}})]. \quad (12)$$

Here,  $p_\tau^{\mathcal{B}}$  and  $\pi^{\mathcal{B}}$  are the restrictions of the corresponding distributions to  $\mathcal{B}$ , and the expectation is under a full-support distribution over subsets of  $\mathcal{G}$ .

Our experiments show that FCS is both sound and easy to evaluate (Figure 8), while prior approaches fail at correctly measuring the distributional accuracy of a GFlowNet (Figure 9).

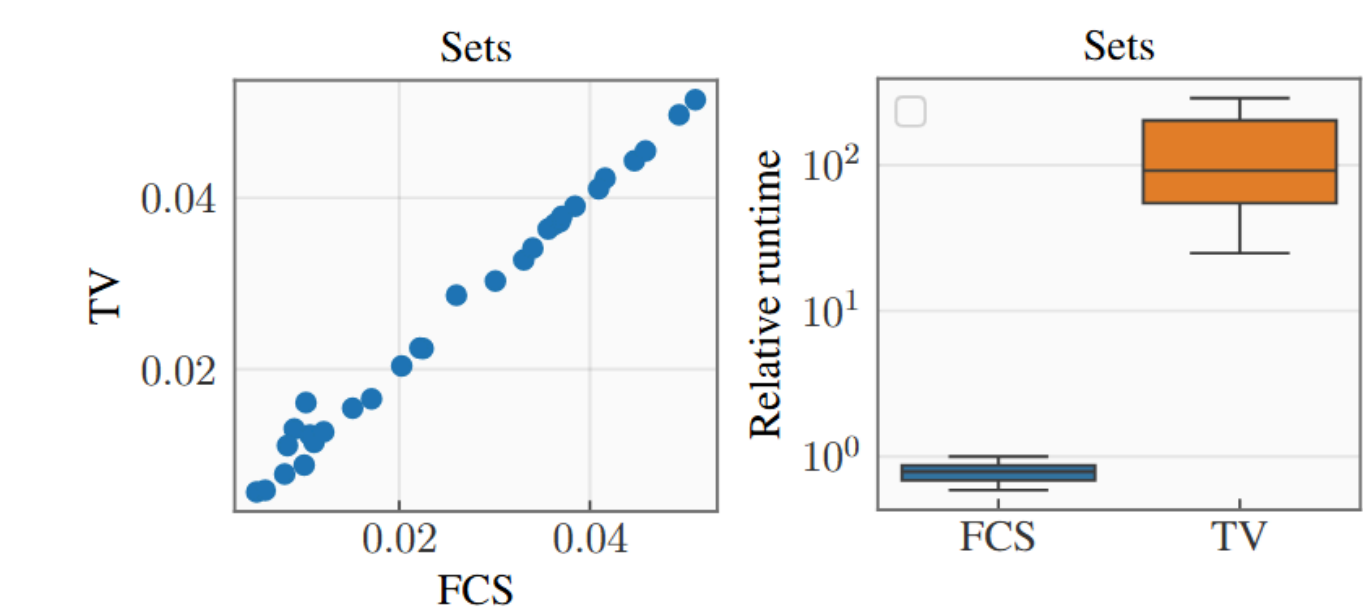


Figure 8: FCS is strongly correlated and drastically faster to evaluate than the TV distance.

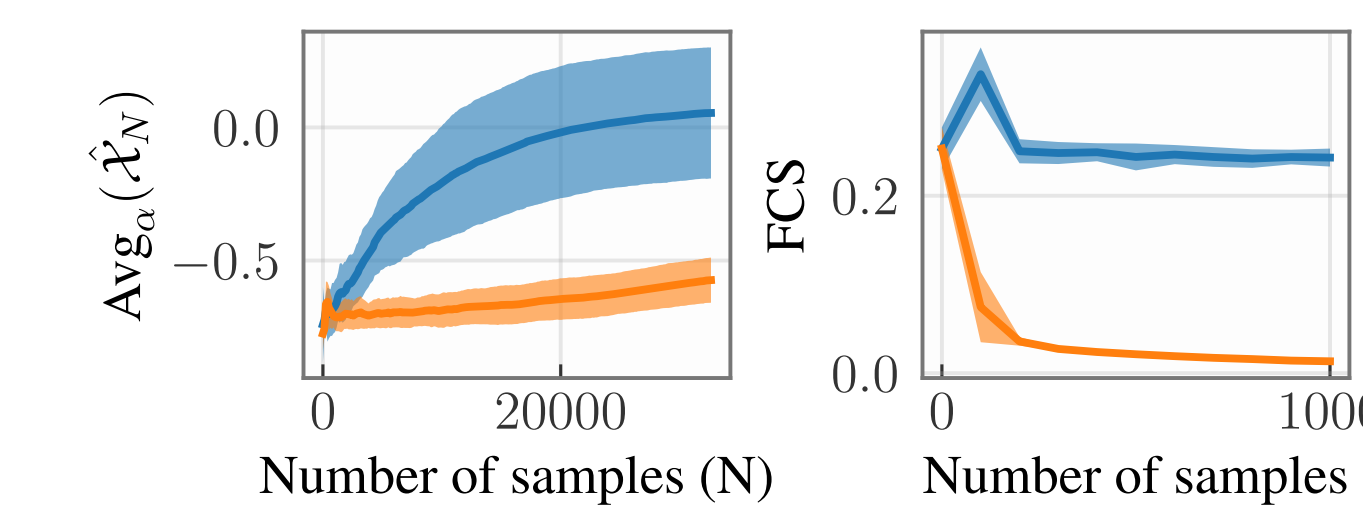


Figure 9: TopQuantile (left) assigns a high-score to an incorrectly learned GFlowNet (according to FCS; right). **Case study.** To illustrate the effectiveness of FCS, we evaluate its performance on variants of LED- and FL- GFlowNets to demonstrate that, despite their impressive mode-finding capability, they do not correctly sample from the target distribution (Figure 10).

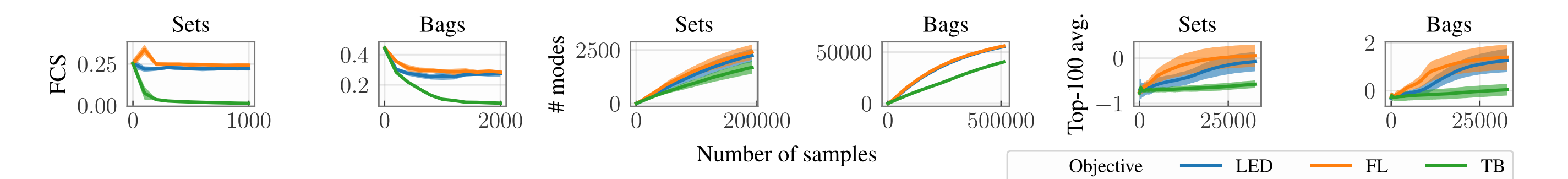


Figure 10: Despite finding modes very quickly (3-6), LED- and FL- GFlowNets fail at sampling from  $\pi(1-2)$ .

**Open question.** An understanding of the impressive mode-finding rate of LED- and FL- GFlowNets is lacking.

## VI. TAKE-HOME MESSAGE

Our work provides a principled approach for analyzing GFlowNets through the perspective of flow imbalance (I), characterizing its potential failure modes under the light of the limited expressivity of GNNs (II), and diagnosing its learned distribution with a sound and computationally tractable metric that may help standardize evaluation and accelerate progress in the field (III).