Analyzing GFlowNets: Stability, Expressiveness, and Assessment

Tiago da Silva [1] **Eliezer de Souza da Silva** [1] **Rodrigo Barreto Alves** [1] **Luiz Max Carvalho** [1] **Amauri H Souza** [2 3] **Samuel Kaski** [2 4] **Vikas Garg** [2] **Diego Mesquita** [1]

Abstract

Generative Flow Networks (GFlowNets) are powerful samplers for distributions over compositional objects (e.g., graphs). In this work, we analyze GFlowNets from three fundamental perspectives: stability, expressiveness, and assessment.

For stability, we analyze how fluctuations in balance conditions impact the accuracy of GFlowNets. Our theoretical results suggest that i) the effect of balance violations is heterogeneous across the state graph and ii) each node's influence on GFlowNet's accuracy is tied to the reward associated with its descendants. We leverage these insights to propose a weighted balance loss that leads to faster training convergence.

Regarding expressiveness, we consider GFlowNets for graph generation. We prove that, given a suitable state graph, GFlowNets can accurately learn any distribution supported over trees. Strikingly, however, we show simple combinations of state graphs and reward functions that cause GFlowNets to fail, i.e., for which balance is unattainable. We propose leveraging embeddings of children's states to circumvent this limitation and thus increase the expressiveness of GFlowNets, provably.

Lastly, we propose a theoretically sound and computationally tractable metric for assessing GFlowNets. We experimentally show it is a better proxy for distributional correctness than popular evaluation protocols.

[1]School of Applied Mathematics, Getulio Vargas Foundation, Brazil [2]Department of Computer Science, Aalto University [3]Federal Institute of Ceará, Brazil [4]Department of Computer Science, University of Manchester, England. Correspondence to: Tiago da Silva <tiago.henrique@fgv.br>, Diego Mesquita <diego.mesquita@fgv.br>.

Accepted by the Structured Probabilistic Inference & Generative Modeling workshop of ICML 2024, Vienna, Austria.

1. Introduction

Generative flow networks (GFlowNets, Bengio et al., 2021; 2023) are reward-driven generative models for compositional objects (e.g., sequences or graphs) that have been successfully employed in several scientific domains (Bengio et al., 2021; da Silva et al., 2023; Deleu et al., 2022; 2023; Jain et al., 2022; 2023; Zhang et al., 2023b). In essence, GFlowNets cast sampling from an unnormalized distribution as solving a network flow problem (Bazaraa et al., 2004). Starting from an initial state, GFlowNets create valid samples by drawing a series of actions according to a (forward) policy network.

While most works on GFlowNets are primarily empirical, developing a deeper theoretical understanding of GFlowNets is key to designing better models and assessment methodologies that are both theoretically sound and practically efficacious. In this regard, Bengio et al. (2021; 2023) laid out the technical foundations for GFlowNets, showing that a model satisfying the imposed balance conditions samples from the target discrete distribution. Lahlou (2023) extended this theory to the context of probability measures supported on arbitrary topological spaces. Also recently, the relationship of GFlowNets with variational inference (Malkin et al., 2023), reinforcement learning (Tiapkin et al., 2024), and diffusion models (Garipov et al., 2023) has been formally established. Despite these advances, important theoretical questions with practical implications remain an uncharted territory, including (i) *How do balance violations impact GFlowNets' accuracy?* (ii) *How does the parameterization of the policy networks affect the expressive power of GFlowNets?* (iii) *How to diagnose whether a GFlowNet correctly samples from a target distribution?*

This paper establishes a series of results to address these fundamental questions. Firstly, we provide bounds on the total variation of GFlowNets as functions of balance fluctuations/violations. By considering tree-structured state graphs with identical rewards, we show that flow imbalances at different depths have a non-uniform impact on the approximation capabilities of GFlowNets — more specifically, balance mismatches near the root state may have a higher impact than those near terminal states. We also extend our analysis to show that similar results hold for general directed acyclic state graphs (DAGs) and multimodal target distribu-

Table 1: Main contributions of this work. Highlighted items represent methodological advancements.

Section 3: Stability	
Sensitivity to local failures for tree-shaped SGs and uniform targets	Thm. 1
Sensitivity to local failures for general SGs and targets	Thm. 2
Formulation of weighted DB loss and comparison against DB	Eq. 7, Fig. 7
Section 4: Expressive power of GNN-based GFlowNets	
Universal approximation of distributions over trees	Thm. 3
Representational limits of 1-WL GFlowNets	Thm. 4
Formulation of Look-Ahead (LA) GFlowNets	Eq. 4
LA-GFlowNets $\succ$ Standard GFlowNets	Thm. 5, Fig. 4
Section 5: Assessment	
Definition of FCS as a tractable goodness-of-fit metric	Def. 1
Relationship between FCS and TV	Thm. 6, Cor. 1
Concentration inequality for FCS and TV	Cor. 2
FCS is a good approximation to TV	Sec. 5.1
Unsoundness of commonly used evaluation protocols	Thm. 7, Sec. 5.2

tions. To illustrate the pragmatic benefits of these insights, we devise a novel self-normalized importance sampling estimator for the detailed balance loss — GFlowNet's only transition-decomposable learning objective. Results in Section 3 suggest our approach often leads to faster convergence relative to the traditional one.

Secondly, we study the distributional limits of GFlowNets when sampling graph-structured objects. Notably, most applications of GFlowNets consist of sampling from distributions over graphs, which render graph neural networks (GNNs) (Gilmer et al., 2017; Gori et al., 2005; Xu et al., 2019a) particularly convenient to parameterize policy networks. In fact, GNNs are often used to parameterize the policies in practice (Bengio et al., 2021; Deleu et al., 2023; Zhang et al., 2023b; Zhu et al., 2023). With this in mind, we provide constructions exposing their shortcomings. While GNN-based GFlowNets can express any distribution over trees under mild conditions, we show that there are simple state graphs and target distributions that no GFlowNet can sample from. We leverage our analysis to introduce *look-ahead GFlowNets* (LA-GFlowNets), a simple yet effective scheme to provably boost the expressiveness of GFlowNets. In essence, LA-GFlowNet incorporates children-state embeddings as inputs to the forward policy. This allows LA-GFlowNets to distinguish actions that lead to distinguishable states but cannot be told apart by the Weisfeiler-Leman (WL) (Weisfeiler & Lehman, 1968) test.

Finally, we provide a theoretically sound framework for the distributional assessment (i.e., goodness-of-fit) of GFlowNets in high-dimensional state spaces. To this extent, we propose the *flow consistency in sub-graphs* (FCS) metric. Put simply, FCS consists of a Monte Carlo estimate of the average L1 error wrt a distribution of "cuts" of the target's support. The FCS metric serves as a proxy for the absolute error between a GFlowNet's sampling distribution and its target. This is extremely valuable for real-world applications, in which the target's support is usually too large to be enumerated. We empirically show that FCS highly correlates with the (often intractable) L1 error. In contrast, popular evaluation metrics poorly capture distributional correctness, e.g., the number of high-reward states visited during training and the average reward of the top-k scoring states (Bengio et al., 2021; Jang et al., 2024; Kim et al., 2024; Pan et al., 2023a).

In Table 1, we summarize the main contributions of this work. Section 3 analyzes the distributional correctness of GFlowNets as a function of balance violations and leverages theoretical insights to propose the WDB loss. Section 4 discusses the representational limits of GFlowNets for graph domains and proposes LA-GFlowNets as a way to boost the expressive power of GNN-based GFlowNets. Finally, Section 5 proposes FCS as a theoretically grounded metric to assess the accuracy of GFlowNets well-suited to high-dimensional settings. Importantly, all sections of this work provide experiments to substantiate our theoretical analyses, illustrating the claims and demonstrating the practical relevance of the methodological contributions.

2. Background

Notations. Let $\mathcal{X}$ be a finite set of *terminal states*, R be an unnormalized distribution over $\mathcal{X}$. We define the set of states $\mathcal{S}$ as an extension of $\mathcal{X}$ comprising two distinctive

elements: an *initial state*, $s_o \in \mathcal{S}$, and a *final state*, $s_f \in \mathcal{S}$. We hence define a weakly connected DAG $\mathcal{G} = (\mathcal{S}, \mathcal{E})$, termed *state graph* (SG), such that (i) there are no incoming edges/transitions to s_o; (ii) for each $x \in \mathcal{X}$, there is a directed path from s_o to x; (iii) there is an edge from each x to s_f, which is not directly connected to any other state in $\mathcal{S}$; and (iv) there are no outgoing edges from s_f. A *forward policy* over $\mathcal{G}$ is a function $p_F \colon \mathcal{S} \times \mathcal{S} \to \mathbb{R}_+$ for which $p_F(s, \cdot)$ is a probability measure supported on s's children in $\mathcal{G}$, denoted by $\text{child}(s)$. A *backward policy* p_B is a forward policy over $\mathcal{G}$'s transpose. For a trajectory $\tau = (s_j)_{j=0}^{h}$ on $\mathcal{G}$, we write $p_F(\tau) = \prod_{i=1\ldots h} p_F(s_{i-1}, s_i)$ for the induced distribution over trajectories. Finally, a *flow* is a function $F \colon \mathcal{S} \to \mathbb{R}_+$ s.t. $F|_{\mathcal{X}} = R$. Throughout the work, we denote the cardinality operator as $\#$. For a trajectory τ, $\#\tau$ denotes its number of transitions.

GFlowNets. We denote a *GFlowNet* by a tuple $(\mathcal{G}, p_F, p_B, F)$ composed of a SG $\mathcal{G}$, parametric functions for p_F, F, and p_B. By definition (Bengio et al., 2021; 2023), the model's parameters are estimated to ensure that the marginal distribution over $\mathcal{X}$ induced by trajectories starting at s_o matches in proportion a given *reward function* $R \colon \mathcal{X} \to \mathbb{R}_+$ defined on the set $\mathcal{X}$, i.e.,

$$p_T(x) = \sum_{\tau \rightsquigarrow x} p_F(\tau) \propto R(x) \quad \forall x \in \mathcal{X}, \tag{1}$$

in which $\tau \rightsquigarrow x$ denotes that τ's last transition is (x, s_f). In practice, p_F is often parameterized by a GNN (Bengio et al., 2021; Zhang et al., 2023b) or a transformer (Deleu et al., 2022; Kim et al., 2024) and p_B is a fixed uniform policy (Deleu et al., 2022; Malkin et al., 2022; 2023; Shen et al., 2023; Zhou et al., 2024). Training GFlowNets aims to satisfying either the *detalied balance* (DB) (Bengio et al., 2023) condition for all $(s, s') \in \mathcal{E}$:

$$\begin{cases} F(s)p_F(s, s') = F(s')p_B(s', s) \text{ if } s' \neq s_f, \\ F(s)p_F(s, s') = R(s) \text{ otherwise}, \end{cases} \tag{2}$$

or the *trajectory balance* (TB) (Malkin et al., 2022) condition $F(s_o)p_F(\tau) = p_B(\tau|x)R(x)$ for all complete trajectories. Both of these balance conditions provably imply Equation (1). When one of such conditions is not satisfied by the GFlowNet, we say that the underlying flow network is *imbalanced*. To estimate the parameters of p_F, one commonly employs a variant of SGD to minimize the expectation of the log-squared difference between the left- and right-hand sides of balance conditions. When enforcing DB, the loss function — denoted $\mathcal{L}_{DB}(p_F, p_B, F)$ — becomes

$$\mathop{\mathbb{E}}_{\tau \sim p_\epsilon} \left[\frac{1}{\#\tau} \sum_{(s,s') \in \tau} \left(\log \frac{F(s)p_F(s, s')}{F(s')p_B(s', s)} \right)^2 \right], \tag{3}$$

with $F(s_f)p_B(s_f, x) := R(x)\ \forall x \in \mathcal{X}$. Here, p_ϵ is an exploratory policy that controls the trade-off between exploration and exploitation in GFlowNet training, and is conventionally set as $p_\epsilon = \epsilon p_F + (1 - \epsilon)p_U$ for a uniform policy p_U. Importantly, albeit $\mathcal{L}_{DB}$ frequently leads to slower convergence rates relatively to $\mathcal{L}_{TB}$ (Malkin et al., 2022), Zhang et al. (2023b) suggested the use of $\mathcal{L}_{DB}$ when modeling state spaces with very long trajectories that cannot be sampled in batches large enough for accurate estimation of $\mathcal{L}_{TB}$'s gradients. In this context, our analysis in Section 3 shows the DB objective can be significantly improved by appropriately choosing the distribution over transitions in Equation (3).

Evaluation protocols for GFlowNets. Due to the intractability of enumerating the support of R to compare the learned and target distributions, one must rely on approximate and easy-to-evaluate approaches for assessing the accuracy of a GFlowNet. Under these conditions, many works (e.g., (Jang et al., 2024; Pan et al., 2023a;b; 2024; Zhang et al., 2023b)) define a threshold R_o and report the average of $R(x)$ for those x with $R(x) \geq R_o$ found during training as a convergence diagnostics, the intuition being that a properly fitted model will swiftly find high-probability regions of the target distribution. Often, R_o is based on an empirical quantile of the observed $R(x)$. Alternatively, the correlation between $\log p_T(x)$ and $\log R(x)$ is a popular metric for goodness-of-fit in the GFlowNet literature (Madan et al., 2022; Malkin et al., 2023; Nica et al., 2022). However, as pointed out by Shen et al. (2023), this correlation is perfect when $p_T(x) \propto R(x)^\alpha$ for any $\alpha > 0$ and does not necessarily reflect the adequacy of the trained model. In this setting, Shen et al. (2023) also suggested the computation of the relative averages of R under the learned and target distributions,

$$\text{Acc}(p_T|R) = \min \left\{ \frac{\mathbb{E}_{x \sim p_T}[R(x)]}{\mathbb{E}_{x \sim \pi \propto R}[R(x)]}, 1 \right\}, \tag{4}$$

as a measure of goodness-of-fit; A was referred to as the GFlowNet's *accuracy*, adopted in subsequent works (Jang et al., 2024; Kim et al., 2024). Remarkably, we show in Section 5.1 that these metrics based on empirical averages of R routinely fail to detect the distributional correctness of GFlowNets when the learned distribution is excessively concentrated in high-probability regions of the target.

3. Bounds on the TV of GFlowNets

In this section, we assess the impact of local violations of the balance conditions on the overall distributional approximation of GFlowNets. We start our analysis by considering uniform distributions and tree-structured SGs, which is a common setting in sequence generation tasks and strategic design in adversarial games (Jain et al., 2022; Jiralerspong et al., 2023). More specifically, Theorem 1 analyzes how a node's violation of the DB condition affects the total variation (TV) between the GFlowNet's sampling and target dis-

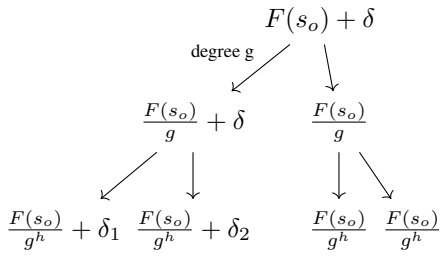

Figure 1: Tree-structured SG w/ excess flow δ from s_0 to left child. We omit node labels to show $F(\cdot)$.

tributions, providing tight lower- and upper-bounds. Henceforth, $\|p - q\|_{TV} = 1/2 \max_{S \subseteq \mathcal{X}} \left|\sum_{x \in S} p(x) - q(x)\right| = 1/2 \sum_{x \in \mathcal{X}} |p(x) - q(x)|$ denotes the TV distance between the probability measures p and q.

Theorem 1 (TV for tree-structured SGs)**.** *Let $(\mathcal{G}, p_F, p_B, F)$ be a GFlowNet balanced wrt to a reward R, where $\mathcal{G}$ is a directed regular tree with branching factor g and depth h, and R is unnormalized uniform. Also, consider the GFlowNet $(\mathcal{G}, \tilde{p}_F, p_B, \tilde{F})$ such that i) $\tilde{F}(s_o) = F(s_o) + \delta$ and $\tilde{F}(s^\star) = F(s^\star) + \delta$ for some $s^\star \in \text{child}(s_o)$ and $\delta \geq 0$; ii) $\tilde{F}(s) = F(s)$ for all s not reachable from $s^\star$; iii) $\tilde{F}(s) = \sum_{s' \in \text{child}(s)} \tilde{F}(s')$; and iv) $\tilde{p}_F(s, s') \propto \tilde{F}(s') \; \forall (s, s') \in \mathcal{E}(\mathcal{G})$ (see Figure 1). Let $\tilde{p}_T$ be the marginal distribution induced by $\tilde{p}_F$. The TV between $\tilde{p}_T$ and $\pi \propto R$ abides by:*

$$\epsilon\left(\delta, g, F(s_0)\right) \leq \|\tilde{p}_T - \pi\|_{TV} \leq \epsilon\left(\delta, g^h, F(s_0)\right), \quad (5)$$

with $\epsilon(\delta, x, t) \coloneqq (1 - 1/x) \frac{\delta}{t + \delta}$.

Naturally, both the upper and lower bounds are increasing functions of δ. Importantly, these bounds are tight, i.e., for any δ, there is an appropriate flow function for which the TV equals the stated bounds. Also, we note that the upper bound $\epsilon(\delta, g^h, F(s_o))$ increases monotonically with the number of leaves g^h, i.e., the further the imbalanced edge $(s_o, s^\star)$ is from the leaves, the higher the potential damage to accuracy. Notably, this suggests that the effect of balance violations on the distributional approximation is heterogeneously spread among the SG's edges.

We now extend the results of Theorem 1 to arbitrary SGs and target distributions. Theorem 2 suggests that the potential effect of an imbalance in $(s, s^\star)$ on the overall distributional approximation is proportional to the probability mass associated with $s^\star$'s terminal descendants. Intuitively, the more terminal descendants $s^\star$ has, the more likely it is for one of them to have low reward/target mass. In turn, the upper bound on the TV increases. For the tree-structured SGs analyzed in Theorem 1, having more terminal descendants is equivalent to having greater height. Overall, Theorem 2 reinforces the initial insight from Theorem 1 that the impact of local violations on the GFlowNet's accuracy is not homogeneous across the SG's edges. In the case of DB, they also serve as a theoretical justification for the importance of choosing appropriate measures to integrate the loss against, which has been shown to significantly influence the sample-efficiency of GFlowNets (Atanackovic & Bengio, 2024; Malkin et al., 2023; Rector-Brooks et al., 2023).

Theorem 2 (TV bounds for arbitrary distributions)**.** *Let $(\mathcal{G}, p_F, p_B, F)$ be a GFlowNet with arbitrary state graph $\mathcal{G}$ satisfying the DB condition wrt an arbitrary reward R. Similarly to Theorem 1, define $(\mathcal{G}, \tilde{p}_F, p_B, \tilde{F})$ by increasing the flow $F(s)$ in some node s by δ and redirecting the extra flow to a direct child $s^\star$ by properly adjusting $p_F(s, \cdot)$. Likewise, $\tilde{F}$ is defined by propagating the extra flows to all states reachable from $s^\star$. Also, let $\mathcal{D}_{s^\star} \subseteq \mathcal{X}$ be the set of terminal states reachable from $s^\star$. Then, the TV between the distribution $\tilde{p}_T$ over $\mathcal{X}$ induced by $\tilde{p}_F$ and the normalized target $\pi \propto R$ satisfies*

$$\frac{\delta}{F(s_0) + \delta}\left(1 - \sum_{x \in \mathcal{D}_{s^\star}} \pi(x)\right) \leq \|\tilde{p}_T - \pi\|_{TV} \leq \frac{\delta}{F(s_0) + \delta}\left(1 - \min_{x \in \mathcal{D}_{s^\star}} \pi(x)\right). \quad (6)$$

Application to the training of GFlowNets. To underline the consequences of our theoretical analysis, we consider training a GFlowNet by optimizing a transition-weighted DB loss, denoted $\mathcal{L}_{WDB}(\tau)$ and defined as

$$\frac{1}{\sum_{(s,s') \in \tau} \gamma(s, s')} \sum_{(s,s') \in \tau} \gamma(s, s') \left(\log \frac{F(s) p_F(s, s')}{F(s') p_B(s', s)}\right)^2, \quad (7)$$

with $F(s_f) p_B(s_f, x) \coloneqq R(x)$ for all $x \in \mathcal{X}$. To define an appropriate weighting function γ, we note from Theorem 1 and Theorem 2 that the maximum disruption on TV enacted by a lack of balance at an edge (s, s') generally increases as a function of the number $d(s')$ of the terminal descendants of s'. Thus, when optimizing the standard DB loss, one may hope that a variant of the SGD algorithm would be predominantly guided by the dominating effect of early-state imbalances, which have the most pronounced impact on the flow network's balance in terms of magnitude, and be lesser affected by late-state transitions, where the supervised signal provided by the reward function is bestowed. Indeed, this intuition is confirmed by our experiments in Figure 8 (details in the supplement). To mitigate this issue, we consider setting $\gamma(s, s')$ as $1/d(s')$, thereby assigning a relatively smaller weight to early transitions. Remarkably, the concept of homogenizing transition-decomposable loss functions in iterative generative models was previously explored in the implementation of adaptive noise schedules for diffusion

probabilistic models (Kingma & Gao, 2023; Kingma et al., 2021), which are closely connected to GFlowNets (Lahlou, 2023; Zhang et al., 2023a), and on the design of the SubTB objective (Madan et al., 2022), which assigns exponentially smaller weights to longer trajectories — which are generally connected to a larger number of terminal states.

Empirical illustration. To illustrate the difference between optimization of Equation 7 and of the standard DB objective, and confirm our theoretically-backed intuitions, we consider three distinct generative tasks: (i) set generation (Bengio et al., 2023; Pan et al., 2023a), (ii) phylogenetic inference (Zhou et al., 2024), and (iii) design of sequences (Jain et al., 2022). For (i), $\mathcal{X}$ is a set of fixed-size subsets of $\{1, \dots, S\}$ for a given $S \in \mathbb{N}$. For (ii), $\mathcal{X}$ contains all phylogenetic trees describing the evolutionary history of a collection of biological species. For (iii), $\mathcal{X}$ consists of discrete sequences of size up to a given constant S. We provide further details in the supplement. Notably, Figure 7 in the supplement shows that optimizing $\mathcal{L}_{WDB}$ often leads to faster training convergence than optimizing $\mathcal{L}_{DB}$, empirically supporting the heuristic derived from Theorem 1 and Theorem 2. As a concluding remark, we note that Equation 7 may be seen as a self-normalized importance estimator of $\mathbb{E}_{s,s' \sim p(\cdot|\gamma)}[\mathcal{L}_{DB}(s, s')]$ for a distribution $p(\cdot|\gamma)$ depending on γ. Interestingly, then, the problem of finding an optimal weighting scheme for $\mathcal{L}_{WDB}$ is equivalent to the design of optimal sampling policies for off-policy training of GFlowNets.

4. Distributional limits of GNN-based policy networks

We now analyze GFlowNets for graphs. In principle, one could use MLP policy networks. However, if the target is supported over graphs without unique node IDs, this choice incurs a factorial increase in SG size. Thus, GFlowNets often use GNNs to parameterize policies (Bengio et al., 2021; Zhang et al., 2023b; Zhu et al., 2023).

Our analysis starts with a positive result (Theorem 3) — for any reward supported over trees, there is a GNN-based GFlowNet using a suitable SG capable of sampling proportional to that reward.

Theorem 3 (Universality of GNN-based GFlowNets for trees.)**.** *If $\mathcal{S}$ is a collection of trees such that $(s, s') \in \mathcal{E}$ implies that $s \subset s'$ (s is a proper subtree of s') with $\#E(s') = \#E(s) + 1$, then there is a GFlowNet equipped with 1-WL GNNs can approximate any distribution π over $\mathcal{X} \subseteq \mathcal{S}$.*

Despite Theorem 3, the expressiveness of a variety of popular GNNs is bounded by the 1-WL isomorphism test. A natural question ensues: *how is this limitation reflected in GNN-based GFlowNets?* To that extent, Theorem 4 shows that there is a broad family of cases (i.e., combinations of SGs and reward functions) for which GNN-based GFlowNets are bound to fail. Intuitively, this result rests on the fact that a state must distribute its flow evenly to children if the actions leading to them are not 1-WL distinguishable. To illustrate this, Figure 3 provides a construction in which GNN-based GFlowNets fail. As 1-WL cannot distinguish between the initial state's children, GFlowNets inevitably learn the same forward conditional distribution at the highlighted states. Thus, GFlowNets cannot learn distributions over the leaves if they are endowed with different rewards.

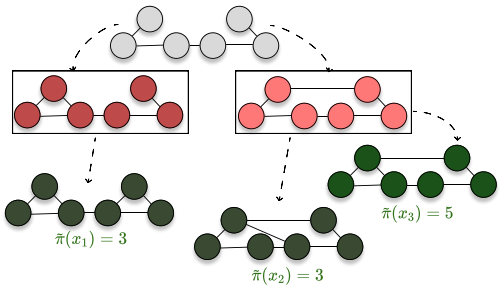

Figure 3: Simple state graph + reward that causes GNN-based GFlowNet to fail.

Theorem 4 (Limitations of GNN-based GFlowNets)**.** *Let $\mathcal{G} = (\mathcal{S}, A)$ be a state graph and $R : \mathcal{X} \subseteq \mathcal{S} \to \mathbb{R}^+$ be a reward function. Suppose $\mathcal{G}$ is a directed tree. Let $T(s) \subseteq \mathcal{X}$ for $s \in \mathcal{S}$ denote the set of terminal states reachable by a directed path starting at s. If there is a state $s = (V, E) \in \mathcal{S}$ and two pairs of nodes $(a, b) \neq (b, c) \in V^2 \setminus E$ that are not 1-WL distinguishable and $\sum_{x \in T(s')} R(x) \neq \sum_{x \in T(s'')} R(x)$ with $s' = (V, E \cup \{(a, b)\})$ and $s'' = (V, E \cup \{(c, d)\})$, then there is no 1-WL GFlowNet capable of approximating $\pi \propto R$ with TV zero.*

We now leverage these insights to propose a more powerful GNN-based GFlowNets: *Look-ahead GFlowNets* (LA-GFlowNets). The rationale of LA-GFlowNets is to incorporate children's graph embeddings as inputs to the forward policy. This allows LA-GFlowNets to disambiguate between children states obtained from 1-WL equivalent actions, enabling assignment of uneven probabilities to non-distinguishable actions as long as the embeddings of corresponding children states differ.

More formally, let s' and s be two neighboring nodes in the SG, differing only by an edge (u, v) not in s — note that s and s' are graphs themselves. Let also $\phi_{v|G}$ be the 1-WL embedding of a node v within a graph G. Then, LA-GFlowNets' forward policy $p_F(s, s')$ can be described as proportional to

$$\exp\left\{\text{MLP}\left(\psi_1\left(\{\phi_{u|s}, \phi_{v|s}\}\right) \,\|\, \psi_2\left(\{\phi_{w|s'}\}_{w \in V(s')}\right)\right)\right\},$$

where ϕ_1 and ϕ_2 are order-invariant functions implemented as deep sets (Zaheer et al., 2017). Since child embed-

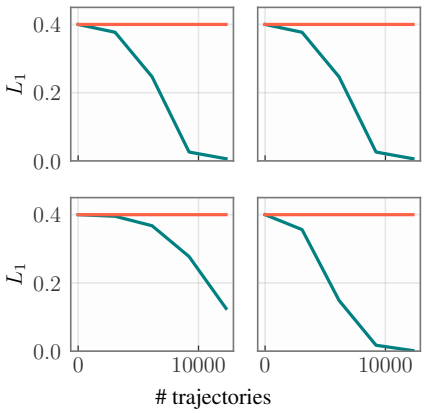

Figure 4: Illustration of cases in which LA-GFlowNets succeed but standard GNN-based GFlowNet fail.

dings are added (concatenated) to the original action embedding, there is naturally no loss of expressiveness wrt 1-WL GFlowNets. On the other hand, LA-GFlowNets can perfectly approximate cases like the one depicted in Figure 3. Theorem 5 states the superior expressiveness of LA-GFlowNets.

Theorem 5 (LA-GFlowNet is more powerful than 1-WL GFlowNet)**.** *If there is a 1-WL forward policy inducing a sampling distribution proportional to a reward R, there is an LA-GFlowNet over the same SG with a sampling distribution proportional to R. The converse does not hold.*

Empirical illustration. To demonstrate the limitations of GNN-parameterized GFlowNets, we define next a group $\mathcal{G}$ of SGs for which there are actions that, despite leading to non-isomorphic states, cannot be distinguished by a GNN-based policy. In this scenario, let $\mathcal{R}_{n,k}$ be the set of regular graphs with n nodes of degree k. Then, let $\mathcal{G}$ be the set of SGs $C_1 \leftarrow P \rightarrow C_2$ such that $P \in \mathcal{R}_{n,k}$ and $C_1 \not\equiv C_2$ differ from P by a single additional edge; see the Appendix for an illustration. Note that, due to the regularity of P, $p_F(P, C_1) = p_F(P, C_2)$ for any GNN p_F. Thus, the corresponding GFlowNet is inherently unable to learn a non-uniform distribution on $\{C_1, C_2\}$. LA-GFlowNets, in contrast, are not constrained by such limited expressivity. As an example, we create four triples (C_1, P, C_2) with $n = 8$, $k = 3$, $R(C_1) = 0.1$ and $R(C_2) = 0.9$. Under these conditions, Figure 4 shows LA-GFlowNet can accurately sample from the target distribution. However, a standard GNN-based GFlowNet can only sample from a uniform, attaining a (constant) L_1 error of 0.4 throughout training.

5. Convergence diagnostics for GFNs

In this section, we propose a provably correct and computationally tractable metric for verifying the distributional incorrectness of GFlowNets, along with probably approximately correct (PAC) statistical guarantees for the accompanying estimators (Section 5.1). Then, we apply this metric to two recently published methods for training GFlowNets, namely, LED- and FL-GFlowNets, and note that they are generally incapable of learning to sample from the target distribution (Section 5.2).

5.1. Assessing GFlowNets

Flow Consistency in Sub-graphs (FCS). Let $(\mathcal{G}, p_F, p_B, F)$ be a GFlowNet trained to sample from a distribution $\pi \propto R$ over $\mathcal{X}$. For each $x \in \mathcal{X}$, we can obtain an unbiased estimate of p_T at x induced by p_F using an importance sampling estimator for $\sum_{\tau \rightsquigarrow x} p_F(\tau)$:

$$p_T(x) = \sum_{\tau \rightsquigarrow x} p_F(\tau) = \mathbb{E}_{\tau \sim p_B(\cdot|x)} \left[\frac{p_F(\tau)}{p_B(\tau)} \right]. \quad (8)$$

For widespread benchmark tasks — such as grid world (Bengio et al., 2021; Malkin et al., 2023; Pan et al., 2023a), sequence design (Jain et al., 2022; Jiralerspong et al., 2023; Zhang et al., 2022), and set generation (Bengio et al., 2023; Pan et al., 2023a; Shen et al., 2023) —, the expectation in Equation 8 may be directly computed by enumerating the relatively few trajectories leading to x. Indeed, for some applications, there is a single τ terminating in each x (Jain et al., 2022; Jiralerspong et al., 2023; Zhou et al., 2024). However, this is intractable beyond toy environments. To circumvent this limitation, we propose the *flow consistency in sub-graphs* (FCS) metric, formally introduced in Definition 1. In summary, the FCS is the average TV between many "cuts" of the marginal p_T and the target $\pi \propto R$, defined by restraining their supports to fixed-size subsets.

Importantly, Theorem 6 shows FCS is closely related to the TV distance between the learned and target distributions. However, FCS is computationally tractable since it does not require instantiating the entire state graph — which is the major caveat of computing the TV.

Definition 1 (Flow Consistency in Sub-graphs)**.** Given a GFlowNet's marginal distribution over terminal states p_T, a reward function R, and a fully-supported probability distribution P_S over all subsets of $\mathcal{X}$ with fixed size $\beta > 0$, we define the FCS metric as:

$$\mathbb{E}_{S \sim P_S}\left[e(S, \theta)\right] := \mathbb{E}_{S \sim P_S}\left[\frac{1}{2} \sum_{x \in S} |p_T^{(S)}(x; \theta) - \pi^{(S)}(x)|\right],$$

with $p_T^{(S)}(\cdot; \theta)$ defined by conditioning p_T on x being in S, and $\pi^{(S)}$ defined similarly by conditioning $\pi(x) \propto R(x)$.

Theorem 6 (Equivalence between TV & FCS)**.** *Let P_S be any full-support distribution over $\{S \subseteq \mathcal{X} : \#S = \beta\}$ for some $\beta \geq 2$. Also, let $d_{TV} = e(\mathcal{X}, \theta)$ be the TV distance between p_T and π for a GFlowNet parameterized by θ. Then, $d_{TV} = 0$ if and only if $\mathbb{E}_{S \sim P_S}[e(S, \theta)] = 0$.*

To highlight the role of β in Theorem 6, Corollary 1 bounds the d_{TV} as a function of both the FCS and β. Note, in particular, that the statement also opens the possibility of using P_S support over differently-sized subsets without giving away the correctness of FCS.

Corollary 1 (Connection between TV & FCS). Let $\#\mathcal{X} = n$ and $P_S(S;\beta) = \binom{n-1}{\beta-1}^{-1} \sum_{x\in S} p_T(x)$ be a distribution on $\{S \subseteq \mathcal{X} \colon \#S = \beta\}$ for $\beta \geq 2$. Then,

$$\left|d_{TV} - \mathbb{E}_{S\sim P_S(\cdot;\beta)}\left[e(S,\theta)\right]\right| \leq \frac{1}{2}\cdot\frac{n}{\beta}\cdot\Delta(\beta) \qquad (9)$$

in which $\Delta(\beta) = \max_{S\subseteq\mathcal{X},\,\#S=\beta} |p_T(S) - \pi(S)|$, $\pi(S) = \sum_{x\in S}\pi(x)$, and $p_T(S) = \sum_{x\in S} p_T(x)$. Moreover, let Q be a probability mass function on $\{2,\dots,n\}$ and consider the distribution $P(S) = P_S(S;\#S)Q(\#S)$ over arbitrarily sized subsets of $\mathcal{X}$. Also, denote $\Delta(\beta) = \max_{S\subset\mathcal{X},\#S=\beta} |p_T(S) - \pi(S)|$. Then,

$$|d_{TV} - \mathbb{E}_{S\sim P}[e(S;\theta)]| \leq \frac{n}{2}\mathbb{E}_{\beta\sim Q}\left[\frac{\Delta(\beta)}{\beta}\right]. \qquad (10)$$

PAC statistical guarantees for $e(S;\theta)$. In practice, we use a Monte Carlo approximation of the intractable quantity $\mathbb{E}_{S\sim P_S}[e(S,\theta)]$ to assess the accuracy of the learned GFlowNet due to the large size of $\mathcal{X}$. In this sense, the corollary below of Corollary 1 underlines that our estimator of the FCS is a good approximation to the TV.

Corollary 2 (PAC bound for FCS). Let P_S as in Theorem 6 and $\#\mathcal{X} = n$. Then, for any $\delta \in (0,1)$, with probability at least $1-\delta$ over choosing m i.i.d. samples $S_1,\dots,S_m \sim P_S$, the following holds:

$$d_{TV} \leq \frac{1}{m}\sum_{1\leq i\leq m} e(S_i,\theta) + \frac{n}{2\beta}\cdot\Delta(\beta) + \sqrt{\frac{2\log\frac{1}{\delta}}{m}}.$$

Empirical illustration. To illustrate the adequacy of FCS as a relatively easy-to-compute proxy for the TV distance, we consider the task of set generation described in Section 3 for varying set and source sizes. For each experiment, we compute both the FCS and TV. For the latter, we use Equation 8 to evaluate the learned marginal distributions for every element in $\mathcal{X}$, which is very costly. Importantly, the result in Figure 5 shows that our FCS is strongly correlated to the TV distance. This observation empirically corroborates Corollary 2 and suggests that FCS is an appropriate tractable surrogate for the TV distance.

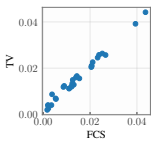

Figure 5: FCS vs. TV

5.2. Revisiting LED- and FL-GFlowNets

LED- and FL-GFlowNets. The accuracy of a GFlowNet heavily depends on the trajectories seen during training (Atanackovic & Bengio, 2024; Rector-Brooks et al., 2023), and the challenge of scoring each transition (s,s') according to its relevance to learning the target is known as the *credit assignment problem*. Notably, Learning Energy Decompositions (LED) (Jang et al., 2024) and Forward-Looking (FL) (Pan et al., 2023a) GFlowNets tackle this issue by reparametrizing the log-flow $\log F(s,s')$ as the residual of a (possibly learnable) potential function $\phi_\theta(s,s')$. More specifically, both models minimize

$$\mathcal{L}_{LED}(s,s') = \left(\log\frac{p_F(s,s')F(s)}{p_B(s',s)F(s')} + \phi_\theta(s,s')\right)^2 \qquad (11)$$

for every transition (s,s'). For FL-GFlowNet, $\phi_\theta(s,s') = \xi(s') - \xi(s)$ is fixed as the difference between hand-crafted state-level functions satisfying $\xi(x) = -\log R(x)$ for $x \in \mathcal{X}$ (Pan et al., 2023a). For LED-GFlowNet, $\phi_\theta(s,s')$ is parameterized as a neural network taking a concatenation of s and s''s as input and trained through the minimization of the least-squares loss, denoted $\mathcal{L}_{LS}(\tau)$,

$$\mathop{\mathbb{E}}_{(m_{s,s'})_{(s,s')\in\tau}}\left[\left(\frac{1}{\#\tau}\xi(x) - \frac{1}{C}\sum_{(s,s')\in\tau} m_{s,s'}\phi_\theta(s,s')\right)^2\right]$$

over all τ, where $m_{s,s'}$ is a dropout mask, $m_{s,s'}$ are independently sampled from a Bernoulli distribution with parameter γ, and $C = \sum_{(s,s')\in\tau} m_{s,s'}$. Remarkably, Theorem 7 shows that the learning problem defined by both FL- and LED-GFlowNets do not generally lead to a generative model that samples from R, despite their apparent empirical success in finding high-probability regions of the target.

Theorem 7 (Unpredictability of FL- and LED-GFlowNets)**.** *Consider a FL- or LED-GFlowNet achieving $\mathcal{L}_{LED}(s,s') = 0$ and $\mathcal{L}_{LS}(\tau) = 0$ for all transitions (s,s') and trajectories τ. Then, the learned marginal distribution over $\mathcal{X}$ satisfies $p_T(x) \propto R(x)F(x)$.*

Theorem 7 means that, unless the flow function F is constant as a function on $\mathcal{X}$, minimizing both $\mathcal{L}_{LED}$ and $\mathcal{L}_{LS}$ is not enough to ensure the GFlowNet samples proportionally to R. In the following, we provide empirical evidence highlighting the effectiveness of FCS in detecting the distributional incorrectness of both FL- and LED-GFlowNets.

Experimental setup. We consider the tasks of set generation, described in Section 3, and generation of *bags* (Shen et al., 2023), both of which were considered by Jang et al. (2024). In this setting, recall that a *bag* corresponds to a multiset of elements sampled from a fixed source $\mathcal{D}$ and that the log-reward of a bag $\mathcal{B}$ corresponds to the sum of utilities

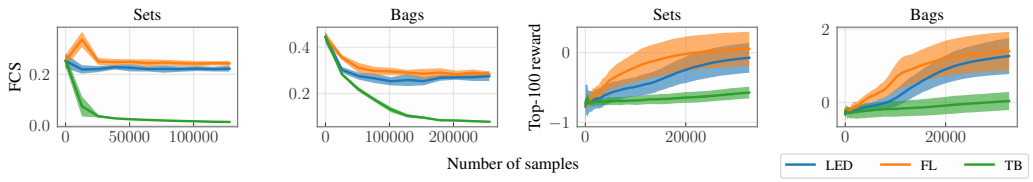

Figure 6: **FCS correctly detects** that neither FL- nor LED-GFlowNets learn to sample from the target (left and mid-left plots), contrasting with the performance of a standard GFlowNet. On the contrary, the average reward of the sampled objects during training incorrectly suggests that FL- and LED-GFlowNets achieve a faster convergence than a standardly trained model (mid-right and right).

Table 2: **Shen et al. (2023)'s accuracy metric** is perfect for FL and LED GFlowNet's formulations, failing to detect their unsoundness.

	LED	FL	TB
Sets	**100.00**±0.00	**100.00**±0.00	93.74±0.98
Bags	**100.00**±0.00	**100.00**±0.00	81.38±6.86

of elements in $\mathcal{B}$. We provide further details regarding these experiments. As a baseline, we train a GFlowNet by minimizing the TB loss (Malkin et al., 2022).

Results. Figure 6 (left) shows that the training of both FL- and LED-GFlowNets converge to a distribution that is significantly more distant from the target than the one learned by a standard GFlowNet is. This result, consistent with Theorem 7, highlights the unpredictability of FL- and LED-GFlowNets, and the potential of FCS to identify such pathological behaviors. Notably, alternative diagnostic methods, such as computing the accuracy metric of (Kim et al., 2024; Shen et al., 2023) and tracking the number of high-reward states found during training (Bengio et al., 2021; Pan et al., 2023a; Zhang et al., 2023b), are unable to detect the unsoundness of FL- and LED-GFlowNets, as we respectively show in Table 2 and Figure 6 (right). The reason for this is that both FL- and LED-GFlowNets learn, for these tasks, a highly right-skewed distribution, thereby exhibiting a very large score according to reward-based performance metrics. In light of this, we advocate for FCS as a standard metric for diagnosing the GFlowNets' learned distributions.

6. Conclusions and limitations

Our contributions are three-fold. Firstly, our sensitivity analyses in Section 3 quantified the effect of transition-level violations to the balance on the TV of the GFlowNet's learned distribution, highlighting the heterogeneous contribution of different node/edges of the SG to the overall distributional approximation. Inspired by these theoretical considerations, we derived a self-normalized importance-sampling estimator for the DB loss that often achieves faster convergence than the traditional approaches. Secondly, we laid out in Section 4 the capabilities and limits of GFlowNets parameterized by 1-WL GNNs in terms of which distributions they can (or cannot) learn. To improve the representational power of GFlowNets, we proposed *LA-GFlowNets* as a parameterization that incorporates the latent representations of a state s's children when computing the forward policy at s. Finally, we outlined in Section 5.1 the *flow-consistency in sub-networks* (FCS) as a provably correct and easy-to-compute metric for assessing the fitness of a GFlowNet. Using the FCS, we demonstrated that popular variants of GFlowNets based on intermediate credit assignment techniques (Jang et al., 2024; Pan et al., 2023a) may not necessarily learn to sample from the correct target distribution, a pathology that previously proposed assessment procedures such as Shen et al. (2023)'s accuracy failed to detect. All in all, this is the first theoretical work to provide a stability analysis, establish the distributional limitations, and design assessment tools for GFlowNets.

Note that our bounds on the TV of GFlowNets do not explicitly provide an optimal weighing scheme, which may be data-dependent. To this effect, we believe an extensive evaluation of different weighing functions γ could help develop even better algorithms to train GFlowNets, and merits a work of its own. Moreover, our analysis of GNN-based GFlowNets focused on 1-WL GNNs due to their popularity and widespread use. However, extending it more expressive (e.g., higher-order) GNNs is a promising direction. Additionally, LA-GFlowNets may fail when both a state's actions and children states are 1-WL indistinguishable. We could also leverage discriminative features (e.g., from substructures (Zeng et al., 2023) or pairwise node distances (Li et al., 2020)) to further boost LA-GFlowNets.

We believe our work paves the road for many advancements in the GFlowNet literature. Firstly, the development of more effective weighting schemes for the detailed balance objective in Section 3 may enhance GFlowNets' scalability. Similarly, the limited expressivity of GNN-based GFlowNets laid out in Section 4 is a cautionary tale for the use of equivariant neural networks when parameterizing the model's policies and may be a useful tool for explaining the difficulty in approximating certain distributions, as well as developing more expressive models. Lastly, we expect FCS to have a major impact on the assessment of GFlowNets and the validation of novel learning objectives.

Acknowledgments

This work was supported by the Research Council of Finland (Flagship programme: Finnish Center for Artificial Intelligence FCAI), EU Horizon 2020 (European Network of AI Excellence Centres ELISE, grant agreement 951847), UKRI Turing AI World-Leading Researcher Fellowship (EP/W002973/1), Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro FAPERJ (SEI-260003/000709/2023), São Paulo Research Foundation (FAPESP, grant 2023/00815-6), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, grant 404336/2023-0). We also acknowledge the computational resources provided by the Aalto Science-IT Project from Computer Science IT.

References

Alquier, P. User-friendly introduction to pac-bayes bounds. *arXiv preprint arXiv:2110.11216*, 2021.

Atanackovic, L. and Bengio, E. Investigating generalization behaviours of generative flow networks. *arXiv preprint arXiv:2402.05309*, 2024.

Bazaraa, M. S., Jarvis, J. J., and Sherali, H. D. *Linear Programming and Network Flows*. Wiley-Interscience, 2004.

Bengio, E., Jain, M., Korablyov, M., Precup, D., and Bengio, Y. Flow network based generative models for non-iterative diverse candidate generation. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

Bengio, Y., Lahlou, S., Deleu, T., Hu, E. J., Tiwari, M., and Bengio, E. Gflownet foundations. *Journal of Machine Learning Research (JMLR)*, 2023.

da Silva, T., Silva, E., Ribeiro, A., Góis, A., Heider, D., Kaski, S., and Mesquita, D. Human-in-the-loop causal discovery under latent confounding using ancestral gflownets. *arXiv preprint:2309.12032*, 2023.

Deleu, T., Góis, A., Emezue, C. C., Rankawat, M., Lacoste-Julien, S., Bauer, S., and Bengio, Y. Bayesian structure learning with generative flow networks. In *Uncertainty in Artificial Intelligence (UAI)*, 2022.

Deleu, T., Nishikawa-Toomey, M., Subramanian, J., Malkin, N., Charlin, L., and Bengio, Y. Joint Bayesian inference of graphical structure and parameters with a single generative flow network. In *Advances in Neural Processing Systems (NeurIPS)*, 2023.

Felsenstein, J. Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, 1981.

Garipov, T., Peuter, S. D., Yang, G., Garg, V., Kaski, S., and Jaakkola, T. S. Compositional sculpting of iterative generative processes. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., and Dahl, G. E. Neural message passing for quantum chemistry. In *International conference on machine learning*, pp. 1263–1272. PMLR, 2017.

Gori, M., Monfardini, G., and Scarselli, F. A new model for learning in graph domains. In *IEEE International Joint Conference on Neural Networks (IJCNN)*, 2005.

Jain, M., Bengio, E., Hernandez-Garcia, A., Rector-Brooks, J., Dossou, B. F. P., Ekbote, C. A., Fu, J., Zhang, T., Kilgour, M., Zhang, D., Simine, L., Das, P., and Bengio, Y. Biological sequence design with GFlowNets. In *International Conference on Machine Learning (ICML)*, 2022.

Jain, M., Raparthy, S. C., Hernandez-Garcia, A., Rector-Brooks, J., Bengio, Y., Miret, S., and Bengio, E. Multi-objective GFlowNets. In *International Conference on Machine Learning (ICML)*, 2023.

Jang, H., Kim, M., and Ahn, S. Learning energy decompositions for partial inference in GFlownets. In *The Twelfth International Conference on Learning Representations*, 2024.

Jiralerspong, M., Sun, B., Vucetic, D., Zhang, T., Bengio, Y., Gidel, G., and Malkin, N. Expected flow networks in stochastic environments and two-player zero-sum games, 2023.

Jukes, T. H. and Cantor, C. R. Evolution of protein molecules. In *Mammalian Protein Metabolism*. Elsevier, 1969.

Kim, M., Yun, T., Bengio, E., Dinghuai Zhang, Y. B., Ahn, S., and Park, J. Local search gflownets. In *International Conference on Learning Representations (ICLR)*, 2024.

Kingma, D. P. and Gao, R. Understanding diffusion objectives as the ELBO with simple data augmentation. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

Kingma, D. P., Salimans, T., Poole, B., and Ho, J. On density estimation with diffusion models. In Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems*, 2021.

Lahlou, S. e. a. A theory of continuous generative flow networks. In *International Conference on Machine Learning (ICML)*, 2023.

Li, P., Wang, Y., Wang, H., and Leskovec, J. Distance encoding: Design provably more powerful neural networks for graph representation learning. *Advances in Neural Information Processing Systems*, 33, 2020.

Madan, K., Rector-Brooks, J., Korablyov, M., Bengio, E., Jain, M., Nica, A. C., Bosc, T., Bengio, Y., and Malkin, N. Learning gflownets from partial episodes for improved convergence and stability. In *International Conference on Machine Learning*, 2022.

Malkin, N., Jain, M., Bengio, E., Sun, C., and Bengio, Y. Trajectory balance: Improved credit assignment in GFlownets. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

Malkin, N., Lahlou, S., Deleu, T., Ji, X., Hu, E., Everett, K., Zhang, D., and Bengio, Y. GFlowNets and variational inference. *International Conference on Learning Representations (ICLR)*, 2023.

Nica, A. C., Jain, M., Bengio, E., Liu, C.-H., Korablyov, M., Bronstein, M. M., and Bengio, Y. Evaluating generalization in gflownets for molecule design. In *ICLR2022 Machine Learning for Drug Discovery*, 2022.

Pan, L., Malkin, N., Zhang, D., and Bengio, Y. Better training of GFlowNets with local credit and incomplete trajectories. In *International Conference on Machine Learning (ICML)*, 2023a.

Pan, L., Zhang, D., Courville, A., Huang, L., and Bengio, Y. Generative augmented flow networks. In *International Conference on Learning Representations (ICLR)*, 2023b.

Pan, L., Jain, M., Madan, K., and Bengio, Y. Pre-training and fine-tuning generative flow networks. In *The Twelfth International Conference on Learning Representations*, 2024.

Rector-Brooks, J., Madan, K., Jain, M., Korablyov, M., Liu, C.-H., Chandar, S., Malkin, N., and Bengio, Y. Thompson sampling for improved exploration in gflownets. *arXiv preprint arXiv:2306.17693*, 2023.

Shen, M. W., Bengio, E., Hajiramezanali, E., Loukas, A., Cho, K., and Biancalani, T. Towards understanding and improving gflownet training. In *International Conference on Machine Learning*, 2023.

Tiapkin, D., Morozov, N., Naumov, A., and Vetrov, D. P. Generative flow networks as entropy-regularized rl. In *AISTATS*, 2024.

Weisfeiler, B. and Lehman, A. A. A reduction of a graph to a canonical form and an algebra arising during this reduction. *Nauchno-Technicheskaya Informatsia*, 2(9): 12–16, 1968.

Xu, K., Hu, W., Leskovec, J., and Jegelka, S. How powerful are graph neural networks? In *International Conference on Learning Representations (ICLR)*, 2019a.

Xu, K., Hu, W., Leskovec, J., and Jegelka, S. How powerful are graph neural networks? *International Conference on Learning Representations (ICLR)*, 2019b.

Zaheer, M., Kottur, S., Ravanbakhsh, S., Poczos, B., Salakhutdinov, R., and Smola, A. Deep sets. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

Zeng, D., Liu, W., Chen, W., Zhou, L., Zhang, M., and Qu, H. Substructure aware graph neural networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023.

Zhang, D., Malkin, N., Liu, Z., Volokhova, A., Courville, A., and Bengio, Y. Generative flow networks for discrete probabilistic modeling. In *International Conference on Machine Learning (ICML)*, 2022.

Zhang, D., Chen, R. T. Q., Liu, C.-H., Courville, A., and Bengio, Y. Diffusion generative flow samplers: Improving learning signals through partial trajectory optimization, 2023a.

Zhang, D., Dai, H., Malkin, N., Courville, A., Bengio, Y., and Pan, L. Let the flows tell: Solving graph combinatorial optimization problems with gflownets. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023b.

Zhou, M. Y., Yan, Z., Layne, E., Malkin, N., Zhang, D., Jain, M., Blanchette, M., and Bengio, Y. PhyloGFN: Phylogenetic inference with generative flow networks. In *The Twelfth International Conference on Learning Representations*, 2024.

Zhu, Y., Wu, J., Hu, C., Yan, J., Hsieh, C.-Y., Hou, T., and Wu, J. Sample-efficient multi-objective molecular optimization with gflownets, 2023.

A. Proofs

This section contains self-contained and rigorous proofs for our theoretical results.

A.1. Proof of Theorem 1

The terminal states of the modified flow network will have two types of nodes, with flow $\frac{F}{g^h}$ and $\frac{F}{g^h} + \delta_i$, with $\delta_i \geq 0$ and $\sum_{i=1}^{g^{h-1}} \delta_i = \delta$. We normalize those probabilities to obtain the individual probabilities for each terminal state, which determines the density of each sample. From that, we can proceed to compute the total variation distance between $\tilde{p}_T$ and π.

$$\begin{aligned}
\|\tilde{p}_T - \pi\|_{TV} &= \frac{1}{2} \sum_{x \in \mathcal{X}} |\tilde{p}_T(x) - \pi(x)| \\
&= \frac{1}{2} \left[(g^h - g^{h-1}) \left| \frac{F}{g^h} \frac{1}{F+\delta} - \frac{1}{g^h} \right| + \sum_{i=1}^{g^{h-1}} \left| \frac{F + g^h \delta_i}{g^h} \frac{1}{F+\delta} - \frac{1}{g^h} \right| \right] \\
&= \frac{1}{2} \left[\frac{g^h \delta - g^{h-1}\delta + \sum_{i=1}^{g^{h-1}} |g^h \delta_i - \delta|}{g^h (F+\delta)} \right].
\end{aligned}$$

We can lower bound $\sum_{i=1}^{g^{h-1}} |g^h \delta_i - \delta|$, by considering that $\sum_{i=1}^{g^{h-1}} (g^h \delta_i - \delta) = g^h \delta - g^{h-1}\delta$, taking the absolute value of the result and each element of the sum to obtain $g^h \delta - g^{h-1}\delta \leq \sum_{i=1}^{g^{h-1}} |g^h \delta_i - \delta|$. Thus we obtain the lower bound

$$\begin{aligned}
\frac{1}{2} \left[\frac{g^h \delta - g^{h-1}\delta + g^h \delta - g^{h-1}\delta}{g^h (F+\delta)} \right] &\leq \frac{1}{2} \left[\frac{g^h \delta - g^{h-1}\delta + \sum_{i=1}^{g^{h-1}} |g^h \delta_i - \delta|}{g^h (F+\delta)} \right] \\
\left(1 - \frac{1}{g}\right) \frac{\delta}{F+\delta} &\leq ||\tilde{p}_T - \pi||_{TV}.
\end{aligned}$$

This lower bound is reached when all error terms in the terminal states have the same value $\delta_i = \frac{\delta}{g^h}$.

To upper bound $|g^h \delta_i - \delta|$ we apply the triangle inequality, obtaining $|g^h \delta_i - \delta| \leq g^h \delta_i + \delta$ and $\sum_{i=1}^{g^{h-1}} |g^h \delta_i - \delta| \leq g^h \delta + g^{h-1}\delta$, from which we obtain the upper bound

$$\begin{aligned}
\|\tilde{p}_T - \pi\|_{TV} &\leq \frac{1}{2} \left[\frac{g^h \delta - g^{h-1}\delta + g^h \delta + g^{h-1}\delta}{g^h (F+\delta)} \right] \\
&\leq \frac{\delta}{F+\delta}.
\end{aligned}$$

To obtain a tighter bound we break the sum $\sum_{i=1}^{g^{h-1}} |g^h \delta_i - \delta|$ by partitioning the sum into the first I terms $S_A = g^h \sum_{i=1}^{I} |\delta_i - \frac{\delta}{g^h}|$ with $\delta_i < \frac{\delta}{g^h}$ and subsequent $g^{h-1} - I$ terms $S_B = g^h \sum_{j=I+1}^{g^{h-1}} |\delta_j - \frac{\delta}{g^h}|$ with $\delta_j \geq \frac{\delta}{g^h}$. By construction, we know that $S_A + g^h \sum_{i=1}^{I} \delta_i + g^h \sum_{j=I+1}^{g^{h-1}} \delta_j - S_B = g^{h-1}\delta$, simplifying to $S_B - S_A = \delta(g^h - g^{h-1})$. We rewrite $S_A + S_B = S_B - S_A + 2S_A = \delta(g^h - g^{h-1}) + 2S_A$, and by triangle inequality on S_A, we obtain the upper bound $\sum_{i=1}^{g^{h-1}} |g^h \delta_i - \delta| = S_A + S_B \leq g^h \delta - g^{h-1}\delta + 2I\delta$. Setting $I = g^{h-1} - 1$ (the biggest value it can have without breaking the constraints on δ_i), it simplifies to $S_A + S_B \leq g^h \delta + g^{h-1}\delta - 2\delta$

$$\begin{aligned}\|\tilde{p}_T - \pi\|_{TV} &\leq \frac{1}{2}\left[\frac{g^h\delta - g^{h-1}\delta + \sum_{i=1}^{g^{h-1}} |g^h\delta_i - \delta|}{g^h(F+\delta)}\right] \\ &\leq \frac{1}{2}\left[\frac{g^h\delta - g^{h-1}\delta + g^h\delta + g^{h-1}\delta - 2\delta}{g^h(F+\delta)}\right] \\ &\leq \left[\frac{g^h\delta - \delta}{g^h(F+\delta)}\right] \\ &\leq \left(1 - \frac{1}{g^h}\right)\frac{\delta}{F+\delta}.\end{aligned}$$

A.2. Proof of Theorem 2

To demonstrate this result, we will need the following facts regarding the function $f(x)\colon x \in \mathbb{R}^n \mapsto \sum_{i=1}^n |x_i - a_i|$ for positive constants a_i.

Lemma 1 (Convexity)**.** *Let* $\Delta_{n+1} = \{x \in \mathbb{R}^n \colon x_i \geq 0 \wedge \sum_{i=1}^n x_i = 1\}$ *and* $a \in \mathbb{R}^n$*. Then,* $f\colon \Delta_{n+1} \to \mathbb{R}$ *defined by* $f(x) = \sum_{i=1}^n |x_i - a_i|$ *is convex.*

Proof. It follows from $f(\alpha x + (1-\alpha)y) = \sum_{i=1}^n |\alpha x_i - \alpha a_i + (1-\alpha)y_i - (1-\alpha)a_i| \leq \alpha \sum_{i=1}^n |x_i - a_i| + (1-\alpha)\sum_{i=1}^n |y_i - a_i| = \alpha f(x) + (1-\alpha)f(y)$ for any $\alpha \in [0,1]$ and $x, y \in \Delta_{n+1}$. □

Lemma 2 (Maximality at edges)**.** *Let* $e_i \in \mathbb{R}^n$ *satisfy* $e_{ij} = 0$ *for* $j \neq i$ *and* $e_{ii} = 1$*. Then, the function* f *from Lemma 1 achieves its maximum at* $\arg\max_{1\leq i\leq n} f(e_i)$.

Proof. We will show that, for each $x \in \Delta_{n+1}$, there is a i for which $f(e_i) \geq f(x)$. In particular, f is maximized at one of the e_i's. For this, note that

$$f(x) = f\left(\sum_{i=1}^n x_i e_i\right) \leq \sum_{i=1}^n x_i f(e_i) \leq \max_{1\leq i\leq n} f(e_i) \tag{12}$$

due to the convexity of f. Thus, f is upper bounded by $\max_{1\leq i\leq n} f(e_i)$. Conversely, there is a e_i for which this upper bound is attained. Hence, $\arg\max_x f(x) \supseteq \arg\max_{1\leq i\leq n} f(e_i)$. □

Lemma 3 (Minimality)**.** *Let* f *be the function of Lemma 1 and assume that* $a \geq 0$ *and* $\sum_{i=1}^n a_i \leq 1$*. Then,* f *is minimized by* $1 - \sum_{i=1}^n a_i$.

Proof. Choose a $j \in \{1, \ldots, n\}$ arbitrarily. Since $x_j = 1 - \sum_{i=1, i\neq j}^n x_i$,

$$\sum_{i=1}^n |x_i - a_i| = \sum_{i=1, i\neq j}^n |a_i - x_i| + \left|a_j - 1 + \sum_{i=1, i\neq j}^n x_i\right| \geq \left|\sum_{i=1}^n a_i - 1\right|. \tag{13}$$

Correspondingly, the lower bound in Equation 13 is achieved when $x_i = a_i$ for $i \neq j$ and $x_j = 1 - \sum_{i=1, i\neq j}^n a_i \geq 0$. This ensures that f is minimized by $1 - \sum_{i=1}^n a_i$. □

In words, Lemma 1 and Lemma 2 ensure that the TV distance between finitely supported distributions is convex and attains its maximum at a Dirac delta.

Proof of Theorem 2. Initially, let δ_x be the amount of extra flow reaching $x \in \mathcal{X}$ and define $\beta_x = {}^{\delta_x}/_{\delta}$. Then,

$$\|p_T - \tilde{\pi}\|_{TV} = \frac{1}{2}\sum_{x\in\mathcal{X}} |p_T(x) - \pi(x)| = \frac{1}{2}\sum_{x\in\mathcal{D}_{s^\star}} |p_T(x) - \pi(x)| + \frac{1}{2}\sum_{x\in\mathcal{D}_{s^\star}^c} |p_T(x) - \pi(x)|. \tag{14}$$

Since $p_T(x) = {}^{\tilde{\pi}(x)+\delta_x}/_{F+\delta}$ for $x \in \mathcal{D}_{s^\star}$ and $p_T(x) = {}^{\tilde{\pi}(x)}/_{F+\delta}$ for $x \in \mathcal{D}_{s^\star}^c$,

$$\sum_{x \in \mathcal{D}_{s^\star}^c} |p_T(x) - \pi(x)| = \frac{\delta}{F+\delta} \sum_{x \in \mathcal{D}_{s^\star}^c} \pi(x). \tag{15}$$

On the other hand,

$$\sum_{x \in \mathcal{D}_{s^\star}} |p_T(x) - \pi(x)| = \sum_{x \in \mathcal{D}_{s^\star}} \left| \frac{\tilde{\pi}(x) + \delta_x}{F+\delta} - \frac{\tilde{\pi}(x)}{F} \right| = \frac{\delta}{F+\delta} \sum_{x \in \mathcal{D}_{s^\star}} \left| \beta_x - \frac{\tilde{\pi}(x)}{F} \right|. \tag{16}$$

By Lemma 2, the function $f\colon \beta \mapsto \sum_{x \in \mathcal{D}_{s^\star}} |\beta_x - \pi(x)|$ is maximized at

$$\begin{aligned} \max_{y \in \mathcal{D}_{s^\star}} f(e_y) &= \max_{y \in \mathcal{D}_{s^\star}} \sum_{x \in \mathcal{D}_{s^\star}} |e_{xy} - \pi(x)| \\ &= \max_{y \in \mathcal{D}_{s^\star}} \left(\sum_{x \in \mathcal{D}_{s^\star}, x \neq y} \pi(x) \right) + (1 - \pi(y)) \\ &= 1 + \sum_{x \in \mathcal{D}_{s^\star}} \pi(x) - 2 \min_{y \in \mathcal{D}_{s^\star}} \pi(y). \end{aligned} \tag{17}$$

Similarly, Lemma 3 ensures that

$$\min_{\beta \in \Delta_{|\mathcal{D}_{s^\star}|+1}} f(\beta) = 1 - \sum_{x \in \mathcal{D}_{s^\star}} \pi(x). \tag{18}$$

Thus, since $\sum_{x \in \mathcal{D}_{s^\star}} \pi(x) = 1 - \sum_{x \in \mathcal{D}_{s^\star}^c} \pi(x)$,

$$\frac{\delta}{F+\delta} \left(1 - \sum_{x \in \mathcal{D}_{s^\star}} \pi(x) \right) \leq \|p_T - \pi\|_{TV} \leq \frac{\delta}{F+\delta} \left(1 - \min_{y \in \mathcal{D}_{s^\star}} \pi(y) \right). \tag{19}$$

□

A.3. Proof of Theorem 3

As stepping stones towards proving Theorem 3, we first lay down Lemma 4 and Lemma 5.

Lemma 4. *Let $G = (V, E)$ and $G' = (V', E')$ be two non-isomorphic trees of size at most n. Let ϕ be the node embedding map of a 1-WL GNN with at least $2n - 1$ layers. Then, $\phi_v \neq \phi_{v'}$ for all $v \in V$ and $v' \in V'$.*

Proof. Recall 1-WL GNNs can distinguish any pair of non-isomorphic trees. Let $\mathcal{T}_n$ and $\mathcal{T}'_n$ denote the sets of computation trees (CTs) for each node in G and G' after n layers, respectively. Likewise, let $\mathcal{T}_{2n-1}$ and $\mathcal{T}'_{2n-1}$ denote the sets of CTs after $2n+1$ layers. Since both graphs are non-isomorphic, 1-WL has already converged with n steps — the maximum diameter of a tree is $n-1$. Without loss of generality, $\mathcal{T}_n - \mathcal{T}'_n \neq \varnothing$, i.e., there is at least one CT in $\mathcal{T}_n$ that is not isomorphic to any tree in $\mathcal{T}'_n$. The same holds for $2n-1$ layers, i.e., $\mathcal{T}_{2n-1} - \mathcal{T}'_{2n-1} \neq \varnothing$. Note that a CT $T_n \in \mathcal{T}_n - \mathcal{T}'_n$ is also a subtree of any $T_{2n-1} \in \mathcal{T}_{2n-1}$. Since $T_n \notin \mathcal{T}'_n$, T_n is not a subtree of any CT in $\mathcal{T}'_{2n-1}$ — otherwise it would be in $\mathcal{T}'_n$ too. In other words, $\mathcal{T}_{2n-1} \cap \mathcal{T}'_{2n-1} = \varnothing$, implying directly our claim. □

Lemma 5. *Let $G = (V, E)$ and $G' = (V', E')$ be any two trees of size at most n, i.e., $|V|$ and $|V'| \leq n$. Also, let $I = (U, \emptyset)$ and $I' = (U', \emptyset)$ be graphs comprising isolated nodes, and ϕ be the node embedding map of a 1-WL GNN with at least $2n - 1$ layers. If $\{\phi_v, \phi_u\} = \{\phi_{v'}, \phi_{u'}\}$ for any $(v, u) \in V \times U$ and $(v', u') \in V' \times U'$, then the trees $(V \cup \{u\}, E \cup \{(v, u)\})$ and $(V' \cup \{u'\}, E' \cup \{(v', u')\})$ are isomorphic.*

Proof. If $\{\phi_v, \phi_u\} = \{\phi_{v'}, \phi_{u'}\}$, then we either have that *i)* $\phi_v = \phi_{v'}$ and $\phi_u = \phi_{u'}$ or *ii)* $\phi_v = \phi_{u'}$ and $\phi_{v'} = \phi_u$. In the first case, we can apply Lemma 4 to conclude that $G \cong G'$ (with associated bijection g_1). Since $\phi_u = \phi_{u'}$, we know that $x_u = x_{u'}$ and the corresponding singleton graphs are trivially isomorphic as well (with bijection g_2). Finally, we can build a bijection g between the vertices of the merged graphs by making $g(v) = g_1(v)$ if $v \in V$ and $g(u) = g_2(u) = u'$. For the second case, Lemma 4 implies G and G' are singletons with $x_u = x_{v'}$ and $x_v = x_{u'}$. The result is a totally disconnected graph, except for an edge linking nodes with identical features in both graphs. □

Armed with the previous lemmata, Theorem 3 is straightforward assuming GNN depth $2n-1$. From Lemma 5, we know that the action embeddings for any two nodes have an empty intersection. Likewise, two actions have the same embedding only if they leave from the same state and arrive at the same state. Therefore, all edges in the SG receive different embeddings. Recall that GNN embeddings are fed to MLP layers, which are universal approximators given enough width. Therefore, a 1-WL GNN followed by MLP can approximate any policy forward p_F. The same applies to the backward policy p_B. We can use the same combination to get state embeddings, which allow approximating any node flow function F. Therefore, we can choose the triplet (p_F, p_B, F) respecting the DB conditions, for instance.

A.4. Proof of Theorem 4

Assume there is a 1-WL GFlowNet sampling from π. Since $\mathcal{G}$ is tree-structured, the mass arriving at $T(s_1) \cup T(s_2)$ must arrive through s — i.e., all paths from s_0 to some $x \in T(s_1) \cup T(s_2)$ traverse s. Furthermore, there is no directed path from s' to any terminal in $T(s'')$ or vice-versa, otherwise the skeleton (i.e., undirected structure) of $\mathcal{G}$ would contain a cycle. Then, $F(s, s') = \sum_{x \in T(s')} R(x)$ and $F(s, s') = \sum_{x \in T(s')} R(x)$, implying $F(s, s') \neq F(s, s'')$.

A.5. Proof of Theorem 5

Since child embeddings are included as additional inputs to LA-GFlowNets, it follows directly that LA-GFlowNets are at least as expressive as 1-WL GFlowNets. We are left with showing the converse does not hold. In Figure 3, we provide a construction for which 1-WL GFlowNets fail but LA-GFlowNets do not.

A.6. Proof of Theorem 6

Henceforth, let $e(p) = \mathbb{E}_{S \sim p}[e(S, \theta)]$. Also, we abuse notation by defining $q(S) = \sum_{x \in S} q(x)$ for a probability mass function q. Then, we first show that $e(p) = 0$ when $d_{TV} = 0$. For this, note that $d_{TV} = 0$ implies $p_T(x) = \pi(x)$ for every x and, in particular, $p_T(S) = \pi(S)$ for every $S \subseteq \mathcal{X}$. Thus,

$$e(p) \coloneqq \mathbb{E}_{S \sim p}\left[\frac{1}{2}\sum_{x \in S} |p_T^{(S)}(x; \theta) - \pi^{(S)}(x)|\right] = 0. \tag{20}$$

On the other hand, assume that $e(p) = 0$. Recall that p is a distribution of full support over $\{S \subseteq \mathcal{X} \colon |S| = B\}$ and that $B \geq 2$. In particular, $e(p)$ ensures that

$$e(S, \theta) \coloneqq \frac{1}{2}\sum_{x \in S}\left|\frac{p_T(x; \theta)}{p_T(S; \theta)} - \frac{\pi(x)}{\pi(S)}\right| = 0. \tag{21}$$

Hence, $p_T(S)\pi(x) = \pi(S)p_T(x)$ for each S and $x \in S$. Write then $S = S' \cup \{x\}$ and conclude that $p_T(S')\pi(x) = \pi(S')p_T(x)$ for every S' and $x \notin S'$. Thus, by summing both members of this equality across $x' \notin S'$, we notice that

$$p_T(S')(1 - \pi(S')) = \pi(S')(1 - p_T(S')), \tag{22}$$

i.e., $p_T(S') = \pi(S')$. Thus, $p_T(x) = \pi(x)$ for all S' and $x \notin S'$. Since S' and x were chosen arbitrarily, $p_T(x) = \pi(x)$ for every $x \in \mathcal{X}$. Thus, $d_{TV} = 0$. This ensures the equivalence between $e(p)$ and d_{TV}.

A.7. Proof of Corollary 1

Let $\hat{e} = \mathbb{E}_{S \sim p}[e(S, \theta)]$, $\mathcal{P}_\beta = \{S \subseteq \mathcal{X} \colon |S| = \beta\}$, and $\Delta = \frac{n}{2\beta} \max_{S \in \mathcal{P}_\beta} |p_T(S) - \pi(S)|$. We will first show that

$$d_{TV} - \hat{e} \leq \Delta. \tag{23}$$

Then, we will verify that $d_{TV} - \hat{e} \geq -\Delta$. These inequalities will jointly imply Corollary 1. In this scenario, note that there are $\binom{n-1}{\beta-1}$ subsets of $\mathcal{X}$ with B elements containing a fixed $x \in \mathcal{X}$. Thus,

$$d_{TV} = \frac{1}{2}\sum_{S \in \mathcal{P}_\beta}\sum_{x \in S}\binom{n-1}{\beta-1}^{-1}|p_T(x) - \pi(x)|. \tag{24}$$

As a consequence,

$$\begin{aligned} d_{TV} - \hat{e} &= \frac{1}{2} \sum_{S \in \mathcal{P}_\beta} \sum_{x \in S} \binom{n-1}{\beta-1}^{-1} |p_T(x) - \pi(x)| - P_S(S) \left| \frac{p_T(x)}{p_T(S)} - \frac{\pi(x)}{\pi(S)} \right| \\ &\leq \frac{1}{2} \sum_{S \in \mathcal{P}_\beta} \sum_{x \in S} \binom{n-1}{\beta-1}^{-1} \left(\left| \pi(x) - \frac{\pi(S)}{p_T(S)} p_T(x) \right| + p_T(x) \left| 1 - \frac{\pi(S)}{p_T(S)} \right| \right) \\ &\qquad - \frac{P_S(S)}{p_T(S)} \left| p_T(x) - \frac{\pi(S)}{p_T(S)} \pi(x) \right| \\ &= \frac{1}{2} \sum_{S \in \mathcal{P}_\beta} \sum_{x \in S} \binom{n-1}{\beta-1}^{-1} p_T(x) \left| 1 - \frac{\pi(S)}{p_T(S)} \right| \\ &= \frac{1}{2} \binom{n-1}{\beta-1}^{-1} \sum_{S \in \mathcal{P}_\beta} |p_T(S) - \pi(S)| \\ &\leq \frac{1}{2} \binom{n-1}{\beta-1}^{-1} \binom{n}{\beta} \max_{S \in \mathcal{P}_\beta} |p_T(S) - \pi(S)| = \frac{n}{2\beta} \Delta \end{aligned} \tag{25}$$

since $P_S(S)/p_T(S) = \binom{n-1}{\beta-1}^{-1}$ and there are $\binom{n}{\beta}$ β-sized subsets of $\mathcal{X}$. For the reverse inequality, notice that

$$\begin{aligned} d_{TV} - \hat{e} &= \frac{1}{2} \sum_{S \in \mathcal{P}_\beta} \sum_{x \in S} \binom{n-1}{\beta-1}^{-1} |p_T(x) - \pi(x)| - P_S(S) \left| \frac{p_T(x)}{p_T(S)} - \frac{\pi(x)}{\pi(S)} \right| \\ &\geq \frac{1}{2} \sum_{S \in \mathcal{P}_\beta} \sum_{x \in S} \binom{n-1}{\beta-1}^{-1} |p_T(x) - \pi(x)| \\ &\qquad - P_S(S) \left(\left| \frac{p_T(x)}{p_T(S)} - \frac{\pi(x)}{p_T(S)} \right| + \left| \frac{\pi(x)}{p_T(S)} - \frac{\pi(x)}{\pi(S)} \right| \right) \\ &= -\frac{1}{2} \sum_{S \in \mathcal{P}_\beta} \binom{n-1}{\beta-1}^{-1} p_T(S) \sum_{x \in S} \pi(x) \left| \frac{1}{p_T(S)} - \frac{1}{\pi(S)} \right| \\ &= -\frac{1}{2} \binom{n-1}{\beta-1}^{-1} \sum_{S \in \mathcal{P}_\beta} |p_T(S) - \pi(S)| \geq -\frac{n}{2\beta} \max_{S \in \mathcal{P}_\beta} |p_T(S) - \pi(S)| . \end{aligned} \tag{26}$$

A.8. Proof of Corollary 2

We provide a self-contained proof of Corollary 2, which follows from Corollary 1 and Hoeffding's inequality (Alquier, 2021). Firstly, let $\hat{e} = \mathbb{E}_{S \sim p}[e(S, \theta)]$ and $e_i = e(S_i, \theta)$. Since $\hat{e} - e_i \in [-1, 1]$, Hoeffding's inequality yields

$$\mathbb{E}\left[\exp\left\{ \lambda \left(\hat{e} - \frac{1}{m} \sum_{1 \leq i \leq m} e_i \right) \right\} \right] \leq \exp\left\{ \frac{\lambda^2}{2m} \right\}. \tag{27}$$

Then, Chernoff's bound implies

$$\Pr_{S_1, \ldots, S_m} \left[\hat{e} \geq \frac{1}{m} \sum_{1 \leq i \leq m} e_i + s \right] \leq \mathbb{E}\left[\exp\left\{ \lambda \left(\hat{e} - \frac{1}{m} \sum_{1 \leq i \leq m} e_i \right) \right\} \right] e^{-\lambda s} \leq \exp\left\{ \frac{\lambda^2}{2m} - \lambda s \right\}$$

due to Equation 27. This upper bound is minimized when $\lambda = sm$. In this case, $\lambda^2/2m - \lambda s = -s^2 m/2$. By letting $s = -2 \log \delta / m$, we verify that

$$\Pr_{S_1, \ldots, S_m} \left[\hat{e} \geq \frac{1}{m} \sum_{1 \leq i \leq m} e_i + \sqrt{\frac{2 \log \frac{1}{\delta}}{m}} \right] \leq \delta. \tag{28}$$

Then, Corollary 1 and the complementary of the preceding inequality imply

$$\Pr_{S_1,\ldots,S_m}\left[d_{TV} \leq \frac{1}{m}\sum_{1\leq i\leq m} e_i + \max_{S\subseteq\mathcal{X},|S|=B} |p_T(S)-\pi(S)| + \sqrt{\frac{2\log\frac{1}{\delta}}{m}}\right] \geq 1-\delta. \tag{29}$$

A.9. Proof of Theorem 7

To prove Theorem 7, first note that $\mathcal{L}_{ls}=0$ for the FL formulation of GFlowNets. Then, assume that $\mathcal{L}_{LED}(s,s')=0$ and $\mathcal{L}_{ls}(\tau)=0$ for every transition $s\to s'$ and every trajectory τ. Thus,

$$F(s)\exp\{\phi_\theta(s,s')\}p_F(s,s') = p_B(s',s)F(s') \text{ and } \sum_{s\to s'\in\tau}\phi_\theta(s,s') = -\log R(x)$$

for every trajectory finishing at x. Therefore, for every trajectory $\tau \rightsquigarrow x$,

$$\begin{aligned} p_F(\tau) &= p_B(\tau|x)\frac{F(x)}{F(s_o)}\prod_{s,s'}\exp\{-\phi_\theta(s,s')\} \\ &= p_B(\tau|x)\frac{F(x)}{F(s_o)}\exp\left\{-\sum_{s\to s'\in\tau}\phi_\theta(s,s')\right\} \\ &= p_B(\tau|x)\frac{F(x)}{F(s_o)}R(x). \end{aligned}$$

Hence,

$$p_T(x) = \sum_{\tau\rightsquigarrow x} p_F(\tau) = \sum_{\tau\rightsquigarrow x}\frac{F(x)R(x)}{F(s_o)}p_B(\tau|x) \propto F(x)R(x)\sum_{\tau\rightsquigarrow x}p_B(\tau|x) = F(x)R(x), \tag{30}$$

ensuring that the marginal distribution learned by FL- and LED-GFlowNets do not necessarily match the target.

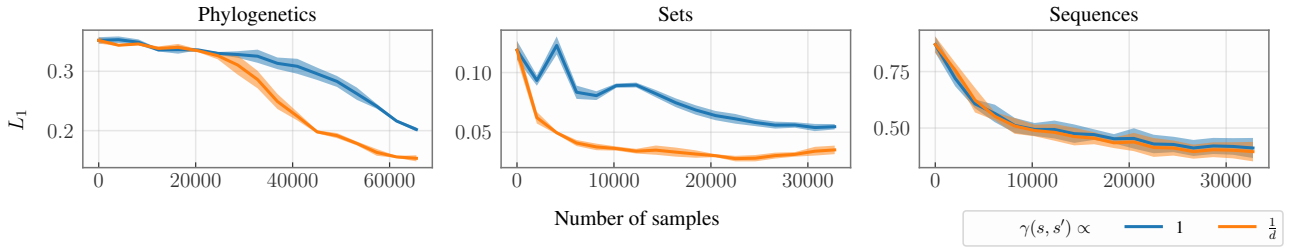

Figure 7: $\mathcal{L}_{WDB}$ **vs.** $\mathcal{L}_{DB}$**.** By weighting each transition (s, s') in inverse proportion to the number of terminal descendants $d(s')$ of s' (i.e., $\gamma(s, s') = 1/d(s')$), we often achieve a faster convergence in terms of L_1 wrt a standard detailed balance penalization (with $\gamma(s, s') = 1$) when training GFlowNets.

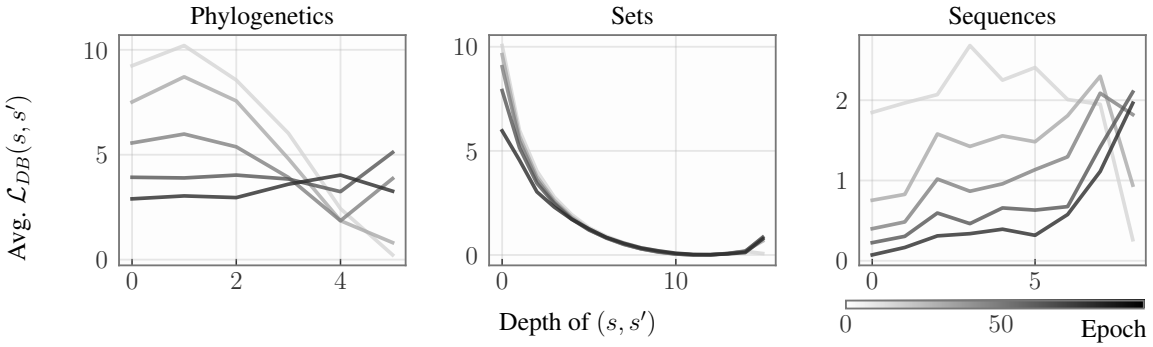

Figure 8: **Average** $\mathcal{L}_{DB}(s, s') = (\log(F(s)p_F(s, s') - \log(F(s')p_B(s, s')))^2$ **along randomly sampled trajectories.** As suggested by our analysis, the DB loss is dominated by the log-squared violation of early transitions near the initial state s_o, obfuscating the terminating reward's signal.

B. Additional experiments

Evaluation of $\mathcal{L}_{WDB}$**.** Figure 7 shows that the minimization of $\mathcal{L}_{WDB}$ often leads to faster training convergence when compared to $\mathcal{L}_{DB}$, underlining the effectiveness of our weighting scheme. Please refer to the main text in Section 3 for a thorough discussion regarding this experiment and to Appendix C for further technical details.

Dominance of early-state transitions in $\mathcal{L}_{DB}$**.** Our theoretical analysis suggested that the detailed balance loss may be dominated by early-state transitions, obfuscating the reward signal at the trajectory's end. Figure 8, which shows the average log-squared violation to the DB condition at each transition within randomly sampled trajectories, empirically corroborates this hypothesis. We provide a more extensive discussion of this result in Section 3. Appendix C describes the technical details for these experiments.

C. Experimental details

We provide further details regarding the experimental setup for each section below, along with computer code for reproducing them (attached to the manuscript). Experiments were run in a cluster equipped with A100 and V100 GPUs, using a single GPU per run.

C.1. Experiments for Section 3

Set generation. The support $\mathcal{X}$ is defined as the collection of sets with 16 elements sampled from a deposit $\mathcal{D} = \{1, \dots, 32\}$. To define the reward function, we let $f\colon \mathcal{D} \to \mathbb{R}$ with $f(d) \sim \mathcal{U}[0, 1]$ and let $\log R(x) = \sum_{d \in x} f(d)$. We implemented an MLP with 2 256-dimensional hidden layers to parameterize both the forward policy and the flow function. For the weighting function γ, we note that $d(s') = \binom{32-|s'|}{16-|s'|}$, in which $|s'|$ is the current state's size.

Sequence design. The support $\mathcal{X}$ is defined as the collection of sequences of size up to 12 with elements extracted from a deposit $\mathcal{D} = \{1, \dots, 4\}$. We implemented an MLP with 2 256-dimensional hidden layers for both the forward policy and flow functions, both of which received as input a sequence of length 12 padded with 0s. Then, the reward function of a $\mathbf{x} \in \mathbb{R}^8$ is defined by $f\colon \mathcal{D} \to \mathbb{R}$ and $g\colon [[1, 12]] \to \mathbb{R}$,, with $f(d), g(i) \sim U[-1, 1]$ for $d \in \mathcal{D}$ and $i \in [[1, 12]]$, through $\log R(x) = \sum_i f(x_i)g(x_i)$. For the weighting function γ, we note that $d(s') = 1 + 4 + \dots + 4^{12-|s'|}$ is the number of s''s terminal descendants.

Phylogenetic inference. A *phylogenetic tree* is defined by a complete binary tree $\mathcal{G}$ with labeled leaves corresponding to observed biological species and anonymous internal nodes corresponding to their evolutionary ancestors. Also, we consider a set $\mathbf{Y} \in \mathbb{R}^{32\times 7}$ of DNA sequences of size 32 associated to the 7 observed species; the likelihood of $\mathbf{Y}$ is defined by the J&C69 (Jukes & Cantor, 1969)'s mutation model and computed by Felsenstein's algorithm (Felsenstein, 1981), and the reward function is the unnormalized posterior induced by an uniform prior distribution over trees. We adopt the iterative process proposed by Zhou et al. (2024) to sample phylogenetic trees with GFlowNets, and use a Graph Isomorphism Network (Xu et al., 2019b) to parameterize p_F. For the weighting function γ, $d(s') = (2 \cdot (7 - |s'|) - 1)!!$ is the number of terminal descendants of s', with $|s'|$ as the amount of connected components in s'.

Details on the experiments for Figure 8. To further understand the consequences of Theorem 2 to the training of GFlowNets, we show in Figure 8 the average log-squared balance violation along trajectories for the generative tasks considered in Section 3. As expected, the DB loss' magnitude is mostly dominated by early-transitions of the IGP. Also, this dominance is more notorious for the problems of set generation and phylogenetic inference and less noticeable for the problem of sequence design, consistently with the results observed in Figure 7 concerning the improved performance of minimizing our weighted loss in Equation 7 wrt the traditional approach. In this regard, we note that the design of sequences is the only task in Figures 7 and 8 with variable-length trajectories. Hence, in this case, our assumption of uniformly distributed extra flows is potentially inaccurate, rendering the noted inefficacy (yet harmlessness) of the corresponding weighted estimator relatively to the unweighted one. Under these circumstances, even though one may often achieve faster convergence when implementing the scheme we propose in Equation 7 and in Figure 7, our experiments suggest the design of an optimal γ for the DB loss should be chosen in a problem-by-problem basis.

C.2. Experiments for Section 4

Setup for Figure 4. This experiment is built upon simple 3-state SGs with the form $L \leftarrow P \rightarrow R$, in which P is a 3-regular graph of 8 nodes and L and R are P's non-isomorphic children obtained by the addition of a single edge. In particular, we choose four different tuples (L_i, P_i, R_i) for the four plots of Figure 4. See Figure 9 for an illustration of 2 of the implemented SGs. To parameterize the policies of both LA- and the standard GFlowNets, we use a 3-layer GIN (Xu et al., 2019b) having 32-dimensional layer, followed by an MLP of 2 32-dimensional layers. For LA-GFlowNet, the MLP's input size is twice as large as the one for the standard model.

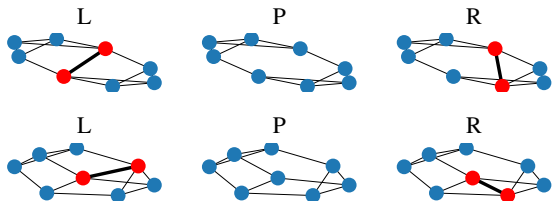

Figure 9: Examples of tuples (L_i, P_i, R_i). Added edges and their nodes are highlighted.

C.3. Experiments for Section 5

FL- and LED-GFlowNets. We rigorously followed the experimental setup of Pan et al. (2023a) and Jang et al. (2024) when implementing both FL- and LED-GFlowNets. To avoid implementation bias, we reproduced our experiments using

Pan et al. (2023a)'s publicly released code[1] and obtained similar results. In particular, both p_F and ϕ were parameterized with MLPs. For LED-GFlowNet, we carried out 8 stochastic gradient steps for learning ϕ for each epoch during training. For the standard GFlowNet trained by minimizing the TB loss, we followed Malkin et al. (2022)'s instructions.

Set generation. The experimental setup is identical to the one described at Section 3. To compute Shen et al. (2023)'s accuracy, we pre-computed the average of $R(x)$ under the target distribution by extensively enumerating the SG's terminal states. For FCS, we randomly sampled 16 batches of terminal states of size up to 32 for the Monte Carlo estimator.

Bag generation. The experimental setup is mostly the same we used for set generation. However, due to the space of bags being significantly larger than the space of sets, we fix $\mathcal{D} = \{1, \ldots, 16\}$ and consider sets of size up to 8.

[1] Available online at github.com/ling-pan/FL-GFN.