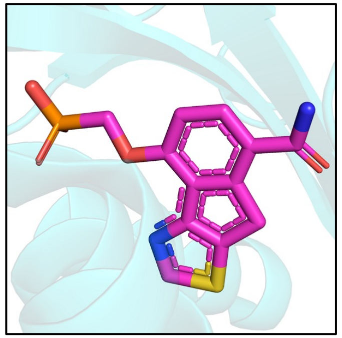
Learning Energy Decompositions for Partial Inference in GFlowNets

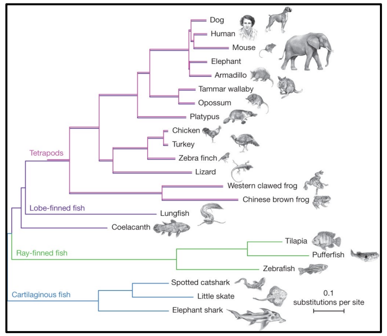
Hyosoon Jang[1] Minsu Kim[2] Sungsoo Ahn[1]

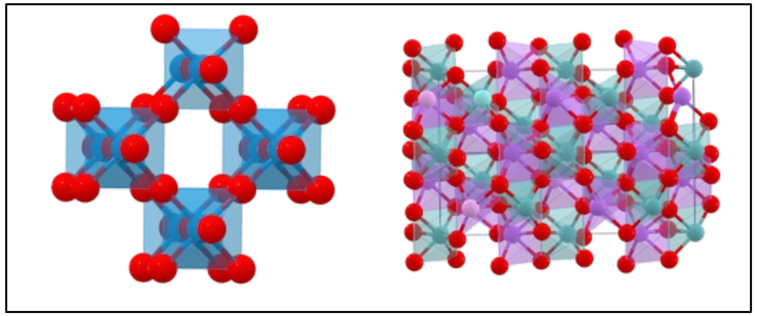
[1]POSTECH [2]KAIST

Probabilistic Modeling Tasks

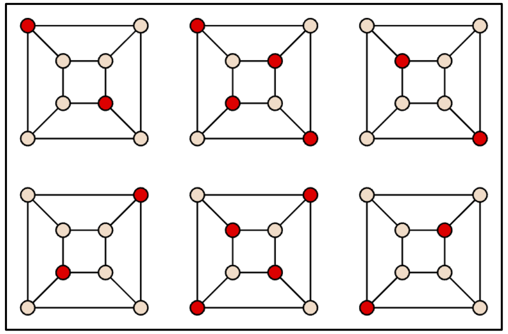
- Many real-world problems require diverse high-scoring solutions



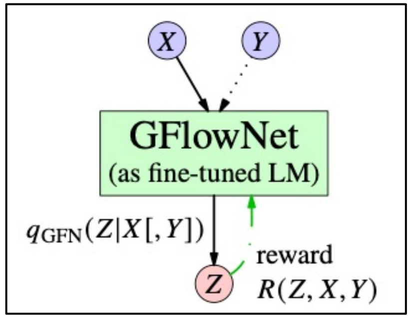


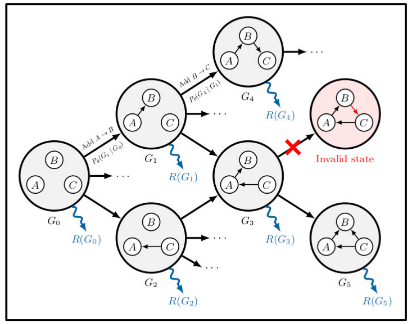


scientific discovery or generation



algorithmic problems



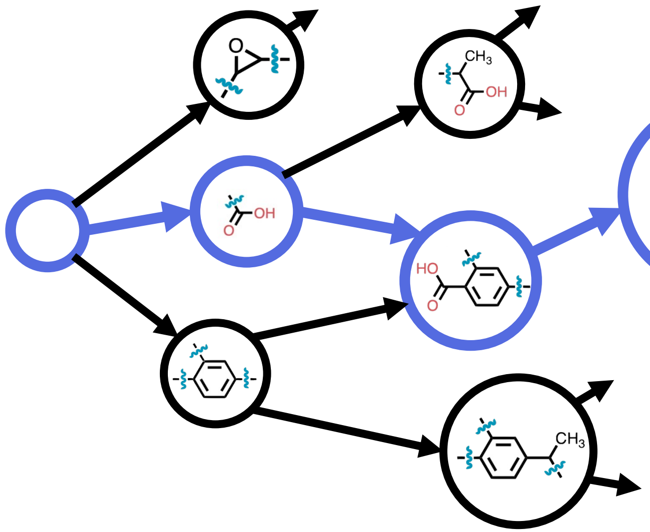


etc

Generative Flow Networks

- Generative Flow Networks (**GFlowNets**) are attractive models for these problems!

$x \sim \prod_i P(\text{action}_i|\text{state}_i)$



- generate with a sequence of actions

represent a rich multimodal distribution

$P(x) \propto \exp(-\mathcal{E}(x))$

energy function:
negative energy ↔ score

- sample from the Boltzmann distribution

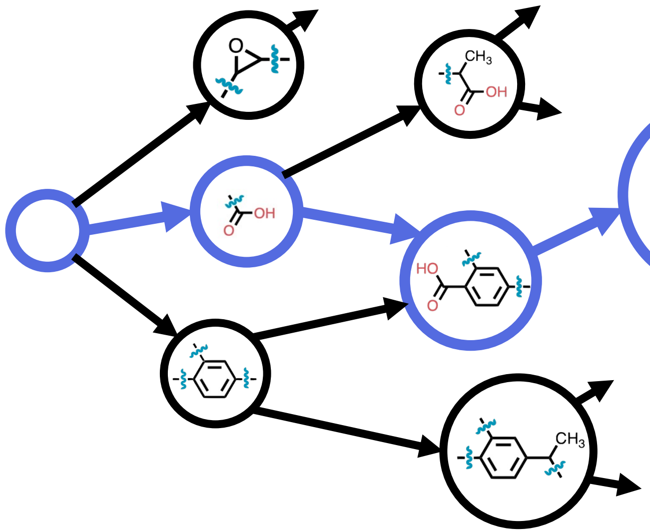
discover diverse high scoring solutions

[1] Jain et al., GFlowNets for AI-Driven Scientific Discovery, 2022

Generative Flow Networks

- Generative Flow Networks (**GFlowNets**) are attractive models for these problems!

$x \sim \prod_i P(\text{action}_i | \text{state}_i)$

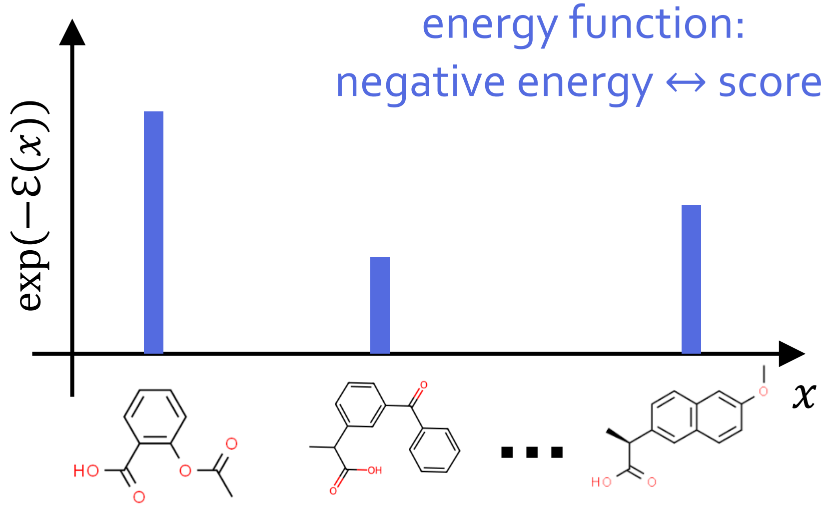


- generate with a sequence of actions

represent a rich multimodal distribution

$P(x) \propto \exp(-\mathcal{E}(x))$

energy function:
negative energy ↔ score



- sample from the Boltzmann distribution

discover diverse high scoring solutions

[1] Jain et al., GFlowNets for AI-Driven Scientific Discovery, 2022

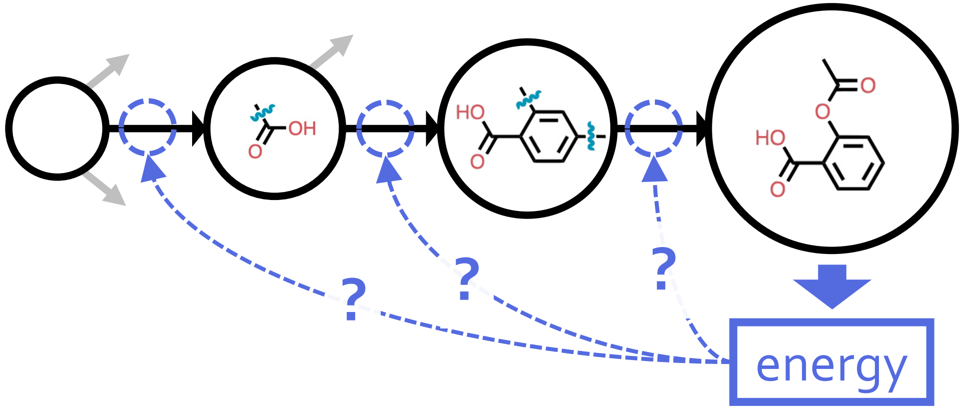
Limitations in Credit Assignment

(better modeling $P(x) \propto \exp(-\mathcal{E}(x))$)

- our goal: improving credit assignment for better training of GFlowNets!

  - credit assignment: identifying
    the contribution of the action to the energy
    e.g., high probability to the action responsible for the low energy

  - motivation: limitations in the credit assignment



assign $\prod_i P(\text{action}_i|\text{state}_i)$ from $\exp(-\mathcal{E}(x))$

assigning multiple probabilities solely relies on the energy;
associating an action with the observed energy is challenging

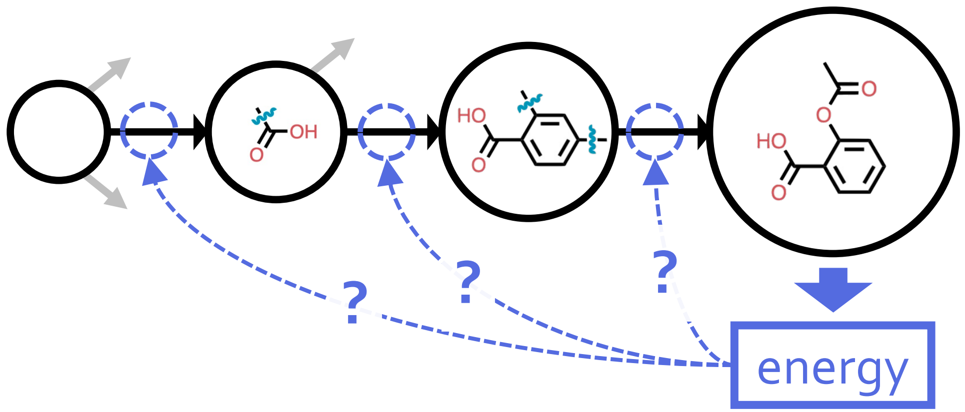
Limitations in Credit Assignment

(better modeling $P(x) \propto \exp(-\mathcal{E}(x))$)

- our goal: improving credit assignment for better training of GFlowNets!

  - credit assignment: identifying the contribution of the action to the energy
    e.g., high probability to the action responsible for the low energy

  - motivation: limitations in the credit assignment



assign $\prod_i P(\text{action}_i|\text{state}_i)$ from $\exp(-\mathcal{E}(x))$

assigning multiple probabilities solely relies on the energy;
associating an action with the observed energy is challenging

Key Contributions

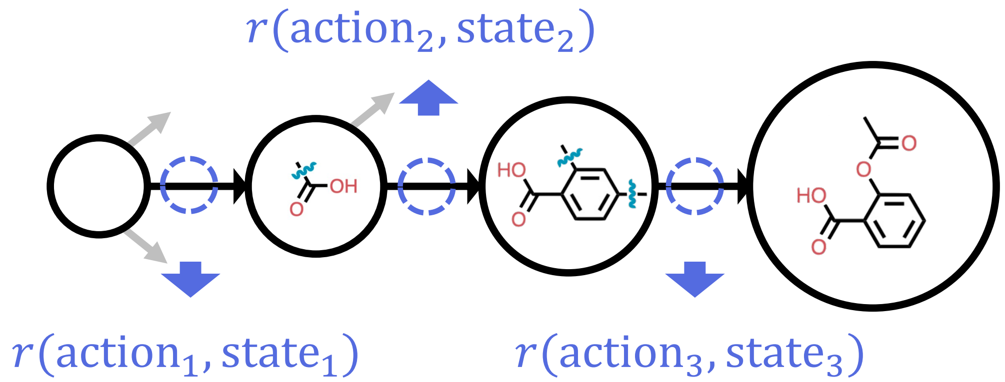
$$\mathcal{E}(x) = \sum_i r(\text{action}_i, \text{state}_i)$$

- We train GFlowNets with local credits r that decompose the energy $\mathcal{E}$

local credits enable partial inference:

learning from the evaluation of individual action

"before reaching the terminal state"



$$\text{assign} \prod_i P(\text{action}_i|\text{state}_i) \ \text{from} \ \exp(-\sum_i r(\text{action}_i, \text{state}_i))$$

specifying the contribution of an action;

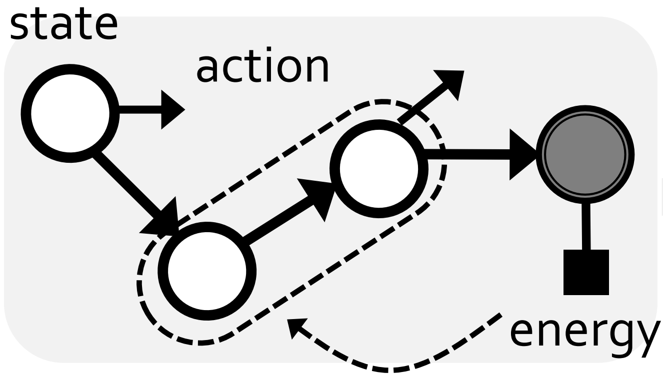
better associating an action with the observed energy

Learning Energy Decompositions for Partial Inference in GFlowNets

(method)

GFlowNet Training

- Given the energy of the terminal state $\mathcal{E}(s_T = x)$, GFlowNets train a policy P_F that makes a transition (, $s_t \rightarrow s_{t+1}$) with an action



$$\text{minimize} \left(\log \frac{F_\theta(s_t)}{F_\theta(s_{t+1})} \frac{P_F(s_{t+1}|s_t)}{P_B(s_t|s_{t+1})} \right)^2$$

$$\text{subject to } F_\theta(s_T = x) = \exp(-\mathcal{E}(x))$$

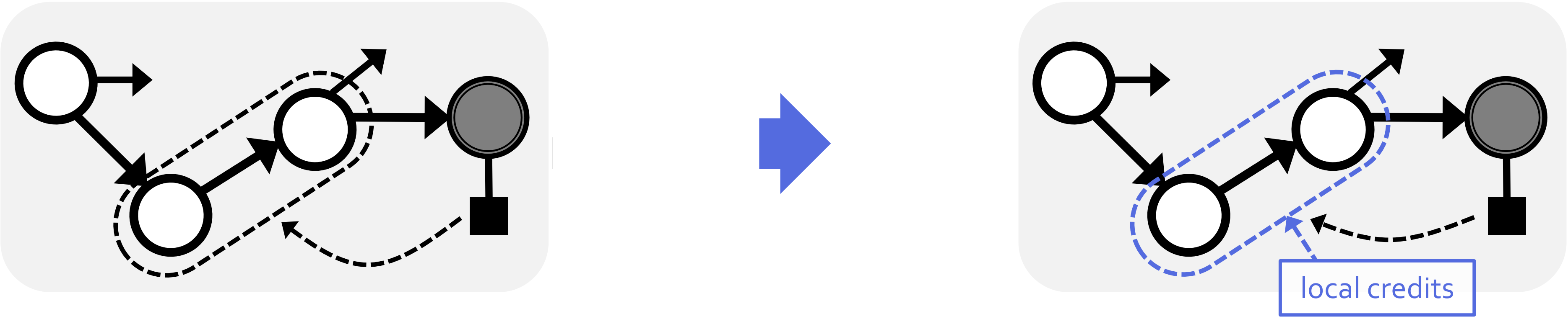
training with transitions $s_t \rightarrow s_{t+1}$ and energies of the terminal states

- $F_\theta(s)$ is the flow (unnormalized probability) estimation
- $P_B(\cdot \mid \cdot)$ is a backward policy

Partial Inference in GFlowNets

conventional training

partial inference



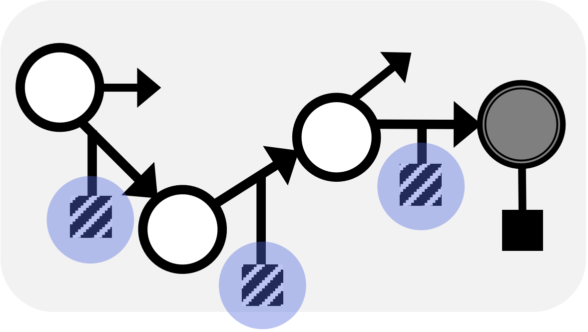
- We are interested in incorporating partial inference capabilities…

  (a) how to enable partial inference with local credits?

  (b) how to evaluate local credits?

Training with Local Credits

- **key component (a):** incorporating local credits $r(s_t \to s_{t+1})$ into the objective (instead of the terminal energy)



$$\text{minimize}\left(r(s_t \to s_{t+1}) + \log\frac{F_\theta(s_t)}{F_\theta(s_{t+1})}\frac{P_F(s_{t+1}|s_t)}{P_B(s_t|s_{t+1})}\right)^2$$

➔ promotes $p_F(s_{t+1}|s_t)$ for high $r(s_t \to s_{t+1})$

still enabling $P_F(x) \propto \exp(-\mathcal{E}(x))$ when $\sum_t r(s_t \to s_{t+1}) = \mathcal{E}(x)$ [1]

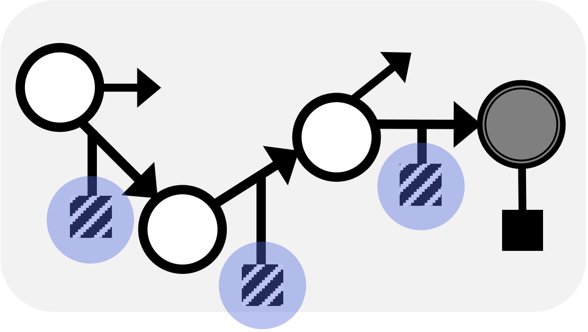
a necessary condition for local credits

[1] Pan et al., Better Training of GFlowNets with Local Credit and Incomplete Trajectories, ICML 2023

Training with Local Credits

- **key component (a):** incorporating local credits $r(s_t \to s_{t+1})$ into the objective

(instead of the terminal energy)



$$\text{minimize}\left(r(s_t \to s_{t+1}) + \log \frac{F_\theta(s_t)}{F_\theta(s_{t+1})} \frac{P_F(s_{t+1}|s_t)}{P_B(s_t|s_{t+1})} \right)^2$$

still enabling $P_F(x) \propto \exp(-\mathcal{E}(x))$ when $\sum_t r(s_t \to s_{t+1}) = \mathcal{E}(x)$ [1]

a necessary condition for local credits

[1] Pan et al., Better Training of GFlowNets with Local Credit and Incomplete Trajectories, ICML 2023

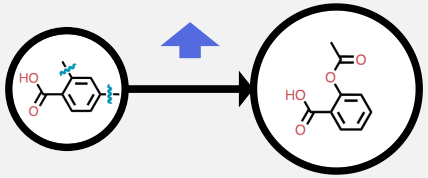
How to Evaluate Local Credits?

- prior approach: the heuristic evaluation $\hat{r}(s_t \rightarrow s_{t+1})$ based on the energy [1]

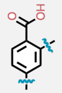
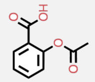
$$\boxed{\hat{r}(s_t \rightarrow s_{t+1}) = \mathcal{E}(s_t) - \mathcal{E}(s_{t+1})}$$

  - e.g., heuristic evaluation in molecular generation

$\hat{r}(s_t \rightarrow s_{t+1})$



s_t $\qquad$ s_{t+1}

$\hat{r}(s_t \rightarrow s_{t+1}) = \mathcal{E}($  $) - \mathcal{E}($  $)$

e.g., changes in the molecular property by adding a fragment

[1] Pan et al., Better Training of GFlowNets with Local Credit and Incomplete Trajectories, ICML 2023

Pitfalls of Heuristic Evaluation

$$\hat{r}(s_t \to s_{t+1}) = \mathcal{E}(s_{t+1}) - \mathcal{E}(s_t)$$

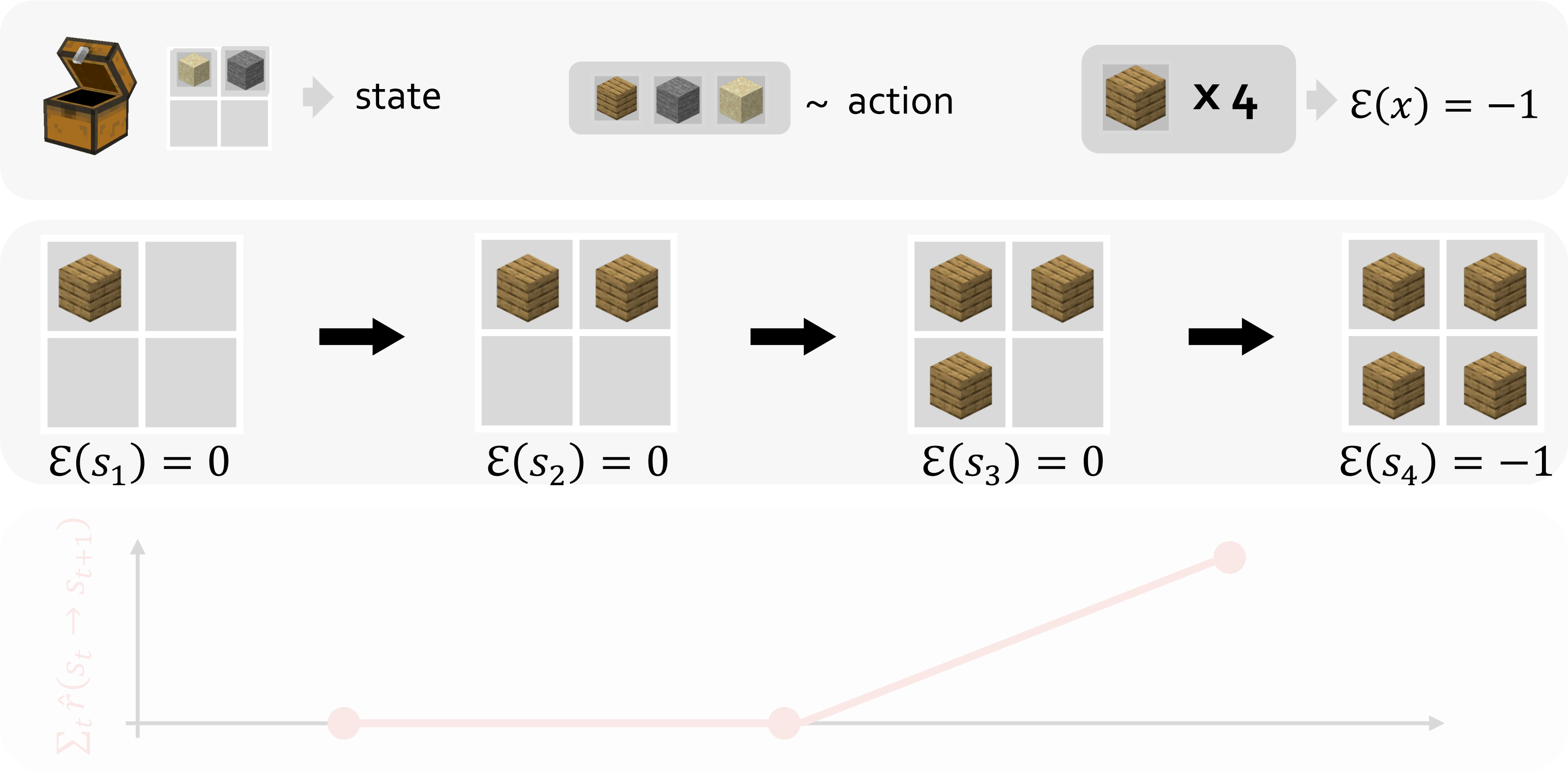
- The heuristic local credits may not be informative
  - may noy provide useful hints to enhance credit assignments

Is the "benefit of repeating the same item" identifiable?

Pitfalls of Heuristic Evaluation

$$\hat{r}(s_t \to s_{t+1}) = \mathcal{E}(s_{t+1}) - \mathcal{E}(s_t)$$

- The heuristic local credits may not be informative
  - may noy provide useful hints to enhance credit assignments

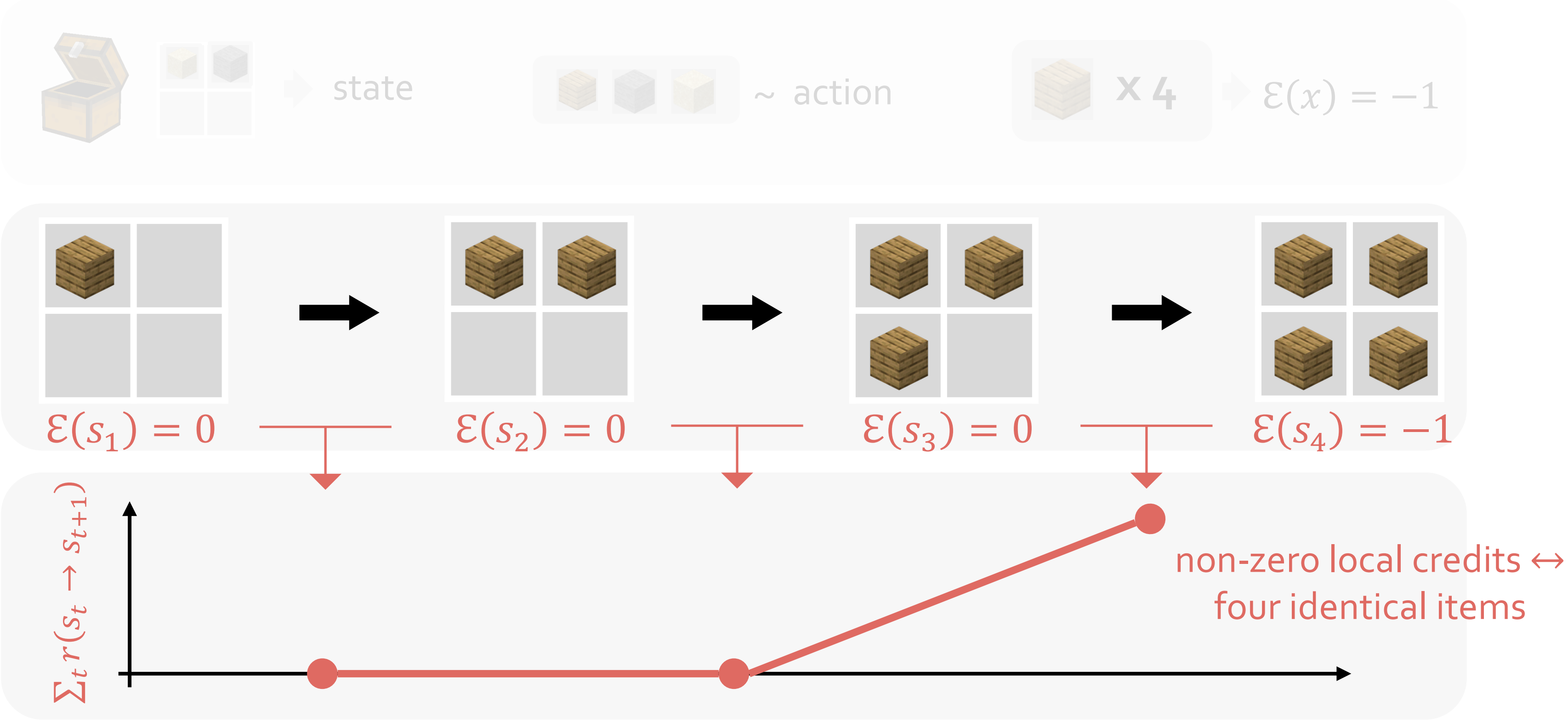


Is the "benefit of repeating the same item" identifiable?

Pitfalls of Heuristic Evaluation

$$\hat{r}(s_t \rightarrow s_{t+1}) = \mathcal{E}(s_{t+1}) - \mathcal{E}(s_t)$$

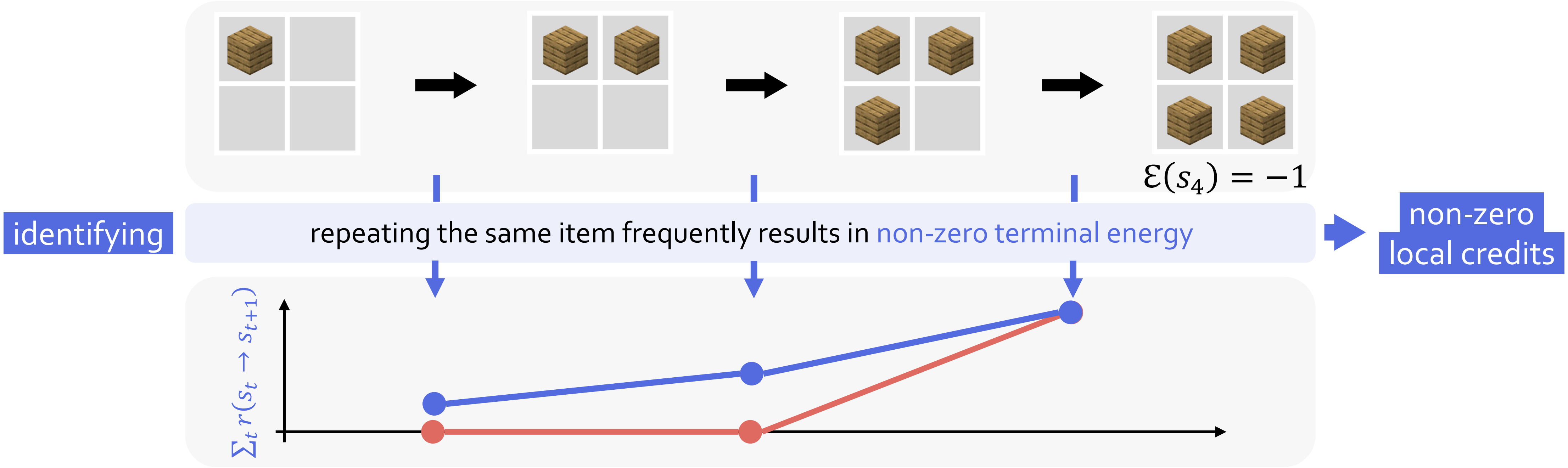
- The heuristic local credits may not be informative
  - may noy provide useful hints to enhance credit assignments



Is the "benefit of repeating the same item" identifiable?

Learning Regularized Local Credits

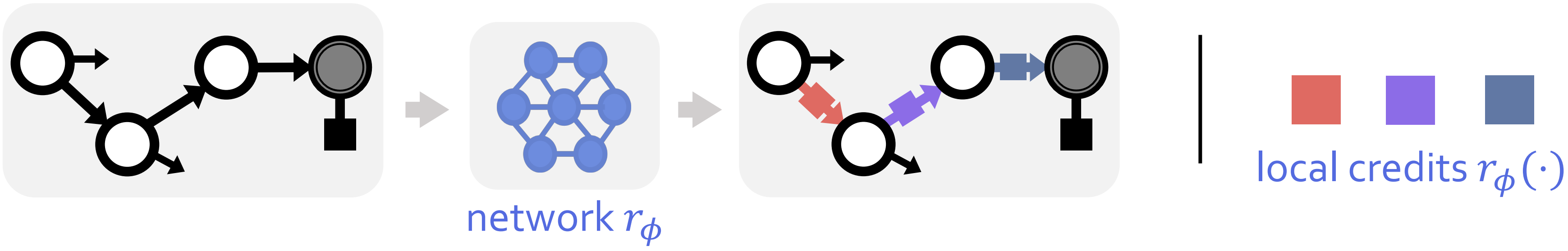
- **key component (b)**: learning informative local credits that are regularized to identify the terminal state energies



identifying underlying contributions of actions to the energy

Learning Energy Decompositions

- We train a local credit network r_ϕ to induce informative local credits



- To replace the energy (key component (a)),
  local credits learn to decompose the energy: + + ≈

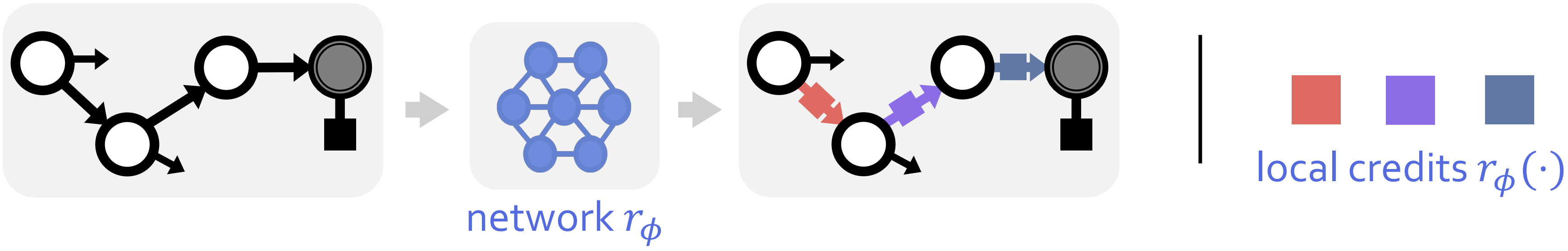
$$\sum_{t=0}^{T-1} r_\phi(s_t \to s_{t+1}) \approx \mathcal{E}(s_T)$$

$$\text{minimize}_{r_\phi} \sum_{\tau} \mathbb{E}_{z\sim\text{Bern}} \left[\left(\frac{1}{C} \sum_{t=0}^{T-1} z_t \, r_\phi(s_t \to s_{t+1}) - \frac{1}{T} \mathcal{E}(s_T) \right)^2 \right]$$

training objective for valid decompositions

Learning Energy Decompositions

- We train a local credit network r_ϕ to induce informative local credits



- To replace the energy (key component (a)),
  local credits learn to decompose the energy: $\blacksquare + \blacksquare + \blacksquare \approx \blacksquare$

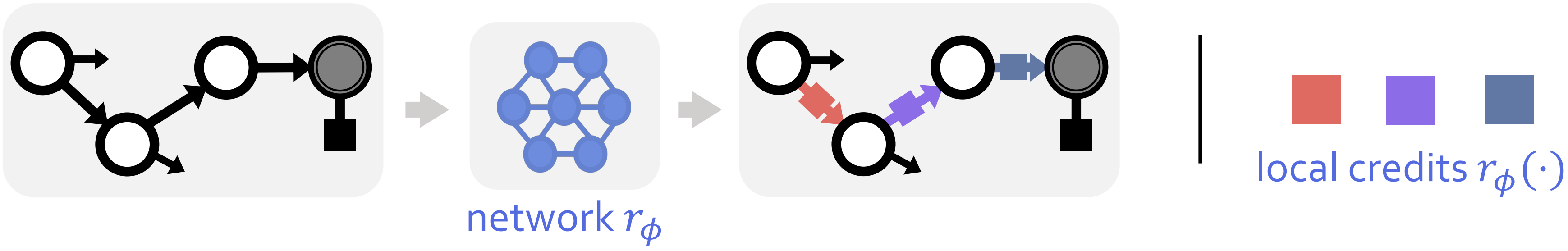
$$\boxed{\sum_{t=0}^{T-1} r_\phi(s_t \to s_{t+1}) \approx \mathcal{E}(s_T)}$$

$$\text{minimize}_{r_\phi} \sum_{\tau} \mathbb{E}_{z\sim\text{Bern}} \left[\left(\frac{1}{C} \sum_{t=0}^{T-1} z_t r_\phi(s_t \to s_{t+1}) - \frac{1}{T} \mathcal{E}(s_T) \right)^2 \right]$$

training objective for valid decompositions

Learning Energy Decompositions

- We train a local credit network r_ϕ to induce informative local credits



- To replace the energy (key component (a)),
  local credits learn to decompose the energy: $\blacksquare + \blacksquare + \blacksquare \approx \blacksquare$

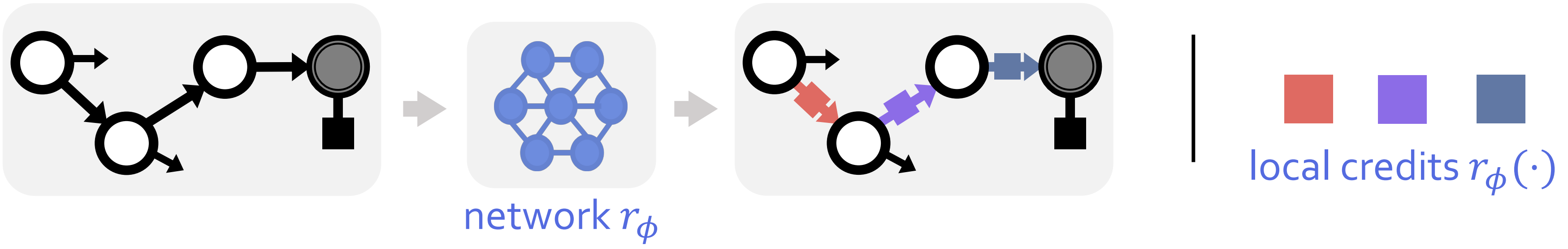
$$\boxed{\sum_{t=0}^{T-1} r_\phi(s_t \to s_{t+1}) \approx \mathcal{E}(s_T)}$$

$$\text{minimize}_{r_\phi} \sum_{\tau} \left[\left(\sum_{t=0}^{T-1} r_\phi(s_t \to s_{t+1}) - \mathcal{E}(s_T) \right)^2 \right]$$

training objective for valid decompositions

Learning Energy Decompositions

- We train a local credit network r_ϕ to induce informative local credits



- The local credits are regularized (key component (b)), to be correlated with the terminal energy : $\blacksquare \approx \blacksquare \approx \blacksquare \approx \frac{\blacksquare}{3}$

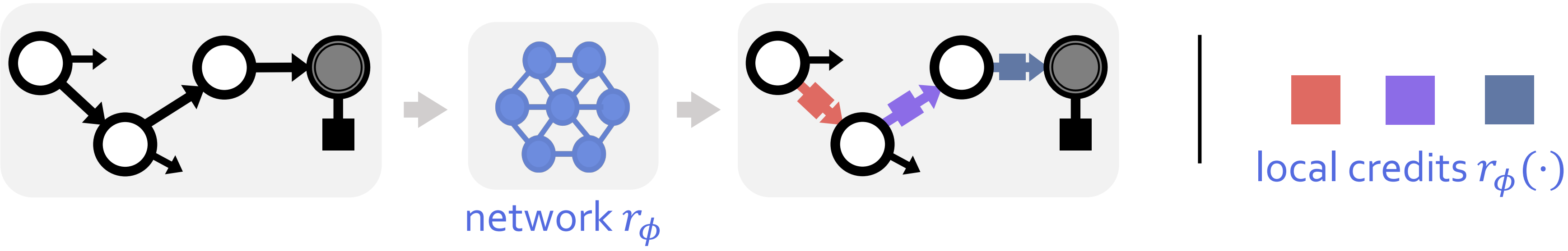
$$r_\phi(s_t \rightarrow s_{t+1}) \approx \mathbb{E}\left[\frac{\mathcal{E}(s_T)}{T}\right]$$

$$\text{minimize}_{r_\phi} \sum_{\tau} \boxed{?} \left[\left(\boxed{?} \sum_{t=0}^{T-1} \boxed{?}\, r_\phi(s_t \rightarrow s_{t+1}) - \boxed{?}\, \mathcal{E}(s_T)\right)^2\right]$$

incorporating a regularization in energy decompositions

Learning Energy Decompositions

- We train a local credit network r_ϕ to induce informative local credits



- The local credits are regularized (key component (b)), to be correlated with the terminal energy : $\blacksquare \approx \blacksquare \approx \blacksquare \approx \frac{\blacksquare}{3}$

$$r_\phi(s_t \to s_{t+1}) \approx \mathbb{E}\left[\frac{\mathcal{E}(s_T)}{T}\right]$$

$$\text{minimize}_{r_\phi} \sum_{\tau} \mathbb{E}_{z\sim\text{Bern}}\left[\left(\frac{1}{C}\sum_{t=0}^{T-1} z_t r_\phi(s_t \to s_{t+1}) - \frac{1}{T}\mathcal{E}(s_T)\right)^2\right]$$

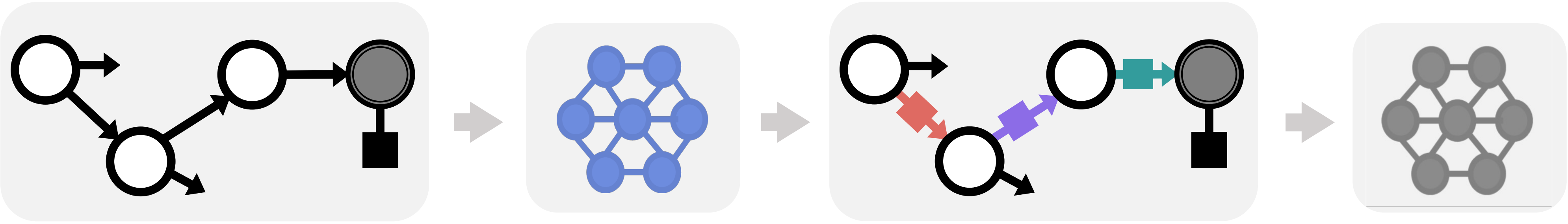
regularizing heavily relying on specific local credits

[1] Pan et al., Learning Long-Term Reward Redistribution via Randomized Return Decomposition, ICLR 2022

Overall Algorithm

- We alternatively train the local credit network and the policy:

$$\mathrm{minimize}_{r_\phi} \sum_{\tau} \mathbb{E}_{\boldsymbol{z}\sim\mathrm{Bern}} \left(\frac{1}{C} \sum_{t=0}^{T-1} z_t r_\phi(s_t \to s_{t+1}) - \frac{1}{T} \mathcal{E}(s_T) \right)^2$$



local credit network r_ϕ

policy P_F

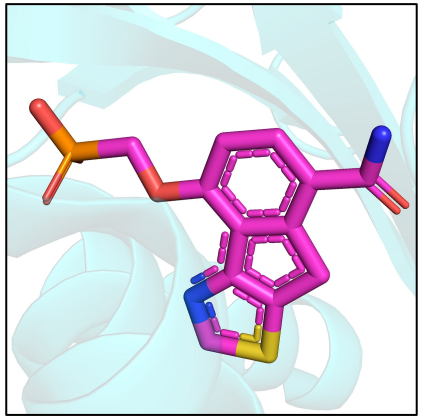
$$\mathrm{minimize}_{F_\theta, P_F, P_B} \left(r_\phi(s_t \to s_{t+1}) + \log \frac{F_\theta(s_t)}{F_\theta(s_{t+1})} \frac{P_F(s_{t+1}|s_t)}{P_B(s_t|s_{t+1})} \right)^2$$

Experiments

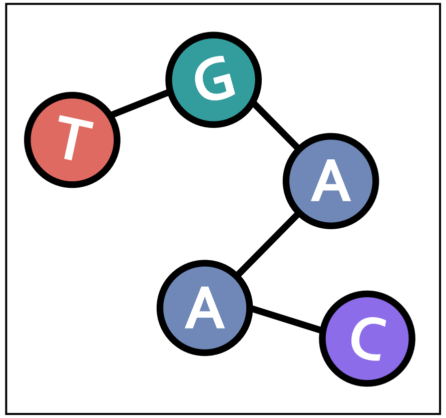
- our algorithm: Learning Energy Decomposition for GFlowNets (LED-GFN)
  - extensively validate LED-GFN on various tasks

    metric: the number of discovered modes and the performance of top-100 sampled objects

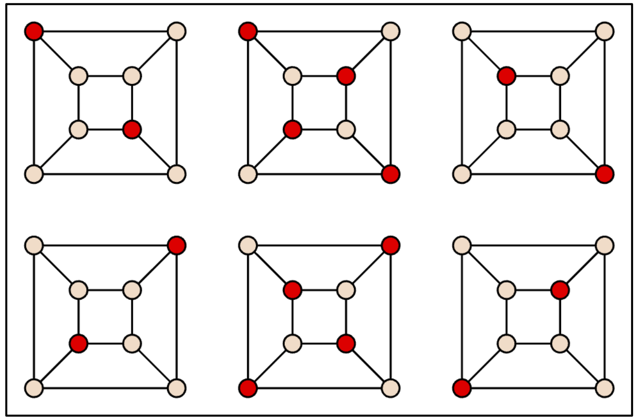
bag
(Shen et al.)



molecule
(Bengio et al.)



rna sequence
(Jain et al.)



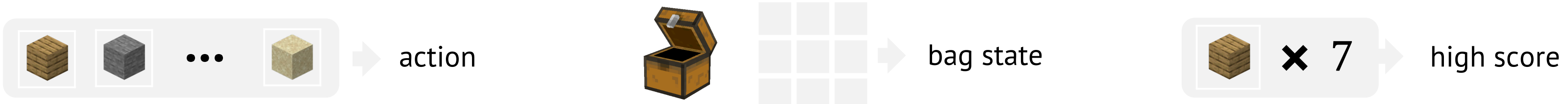
set-types
(Pan et al., Zhang et al.)

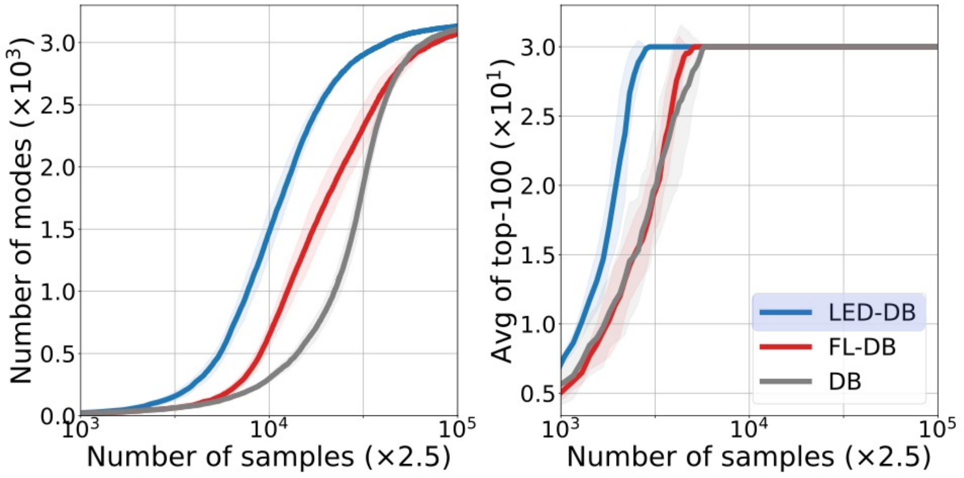
sub trajectory-based implementations (subTB)

$$\text{minimize}\ \left(\sum_{t=U}^{V-1}\phi(s_t \to s_{t+1}) + \log\frac{F_\theta(s_U)}{F_\theta(s_V)}\prod_{t=U}^{V-1}\frac{P_F(s_{t+1}|s_t)}{P_B(s_t|s_{t+1})}\right)^2$$

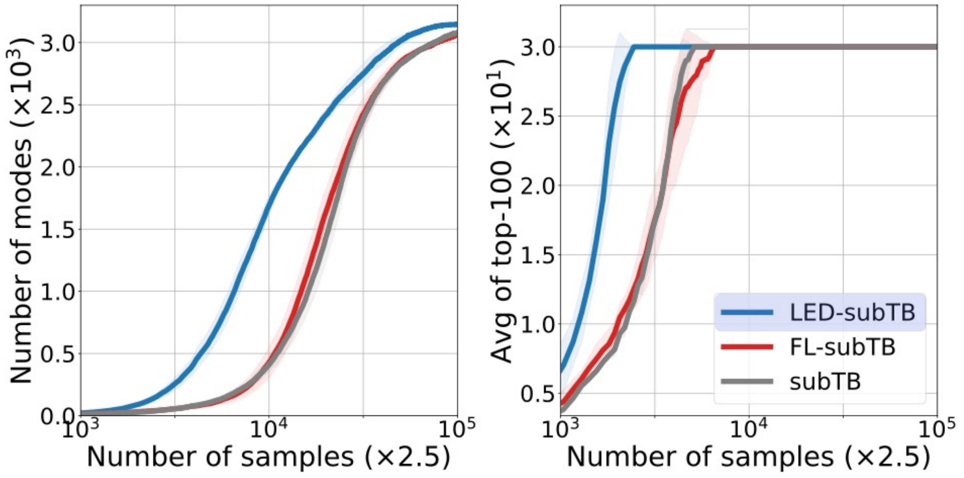
Experiments

- our algorithm: Learning Energy Decomposition for GFlowNets (LED-GFN)





(a) DB-based objectives

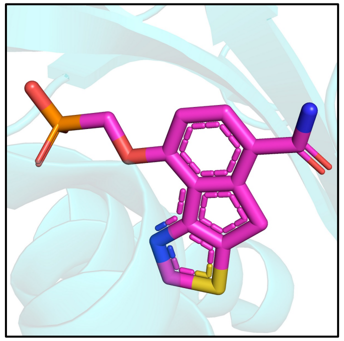


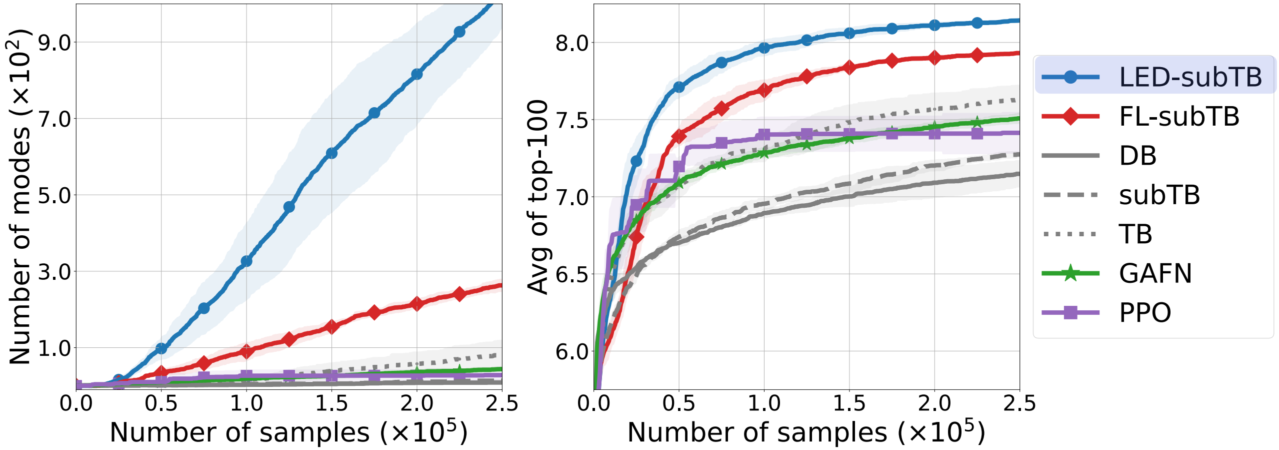
(b) SubTB-based objectives

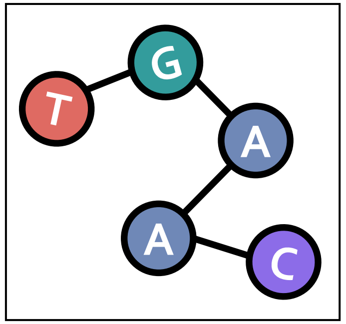
LED-GFN outperforms the baselines defined with the heuristic local credits

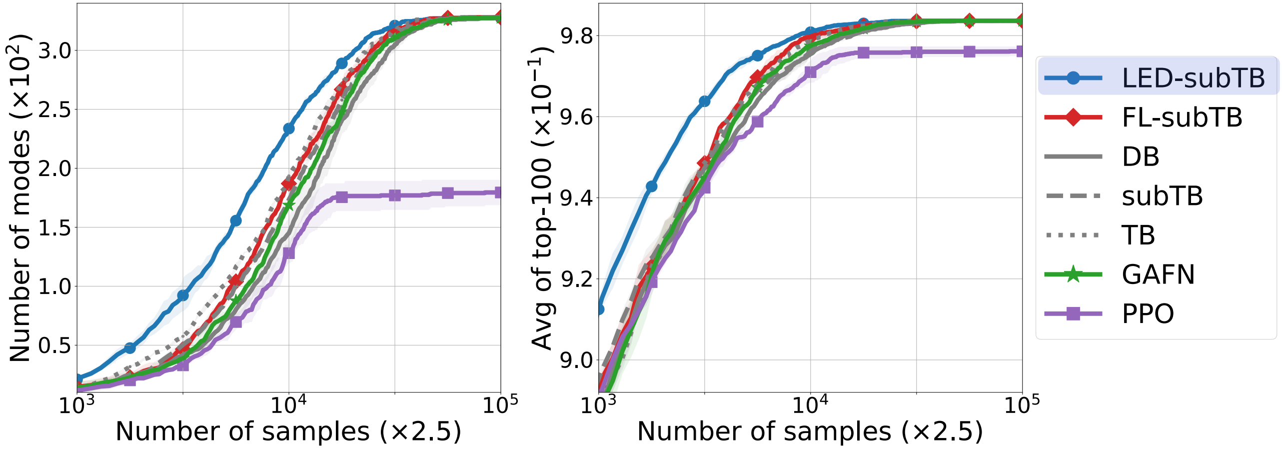
Experiments

- our algorithm: Learning Energy Decomposition for GFlowNets (LED-GFN)





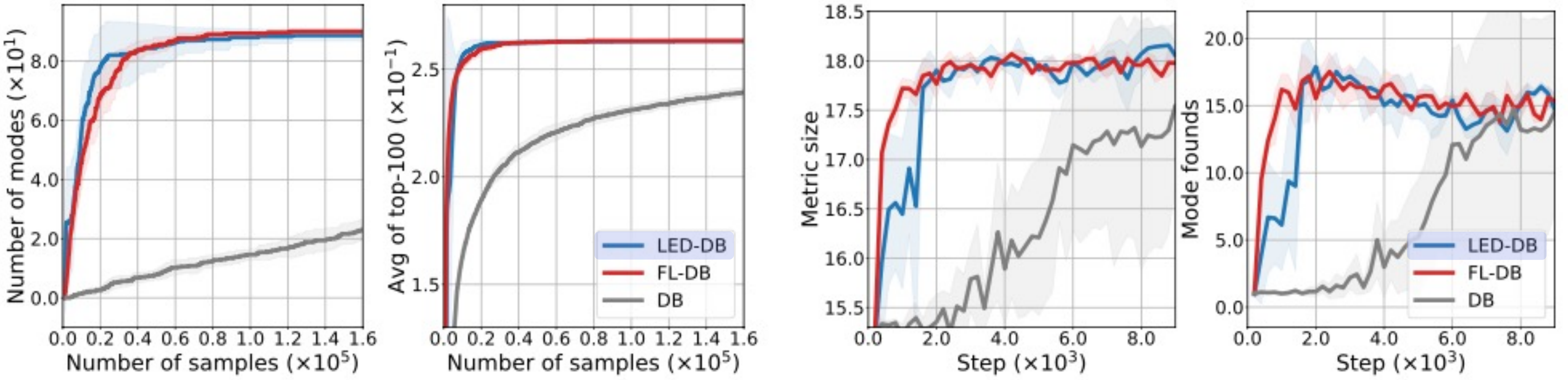




LED-GFN excels on the generation of molecules RNA sequences

Experiments

- our algorithm: Learning Energy Decomposition for GFlowNets (LED-GFN)



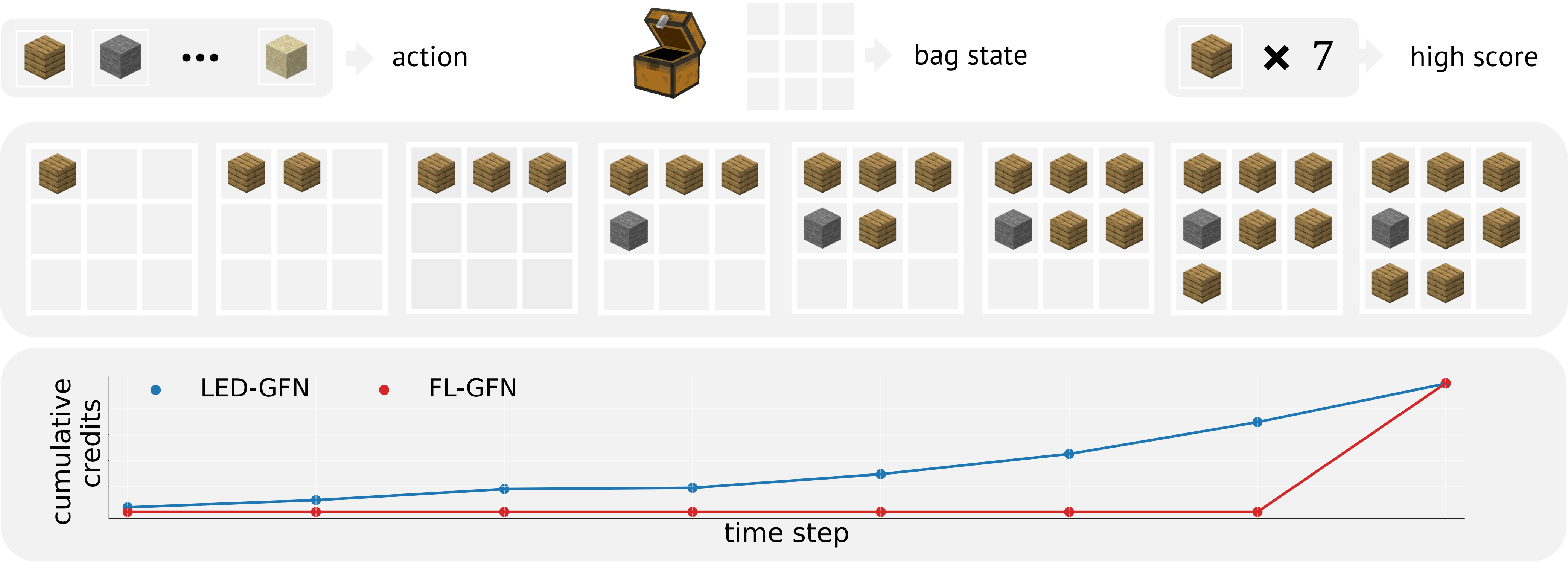
(a) Set generation (b) Maximum independent set problem

In both tasks, the local credits for baselines are hand-crafted;
but LED-GFN even shows similar performance

Summary

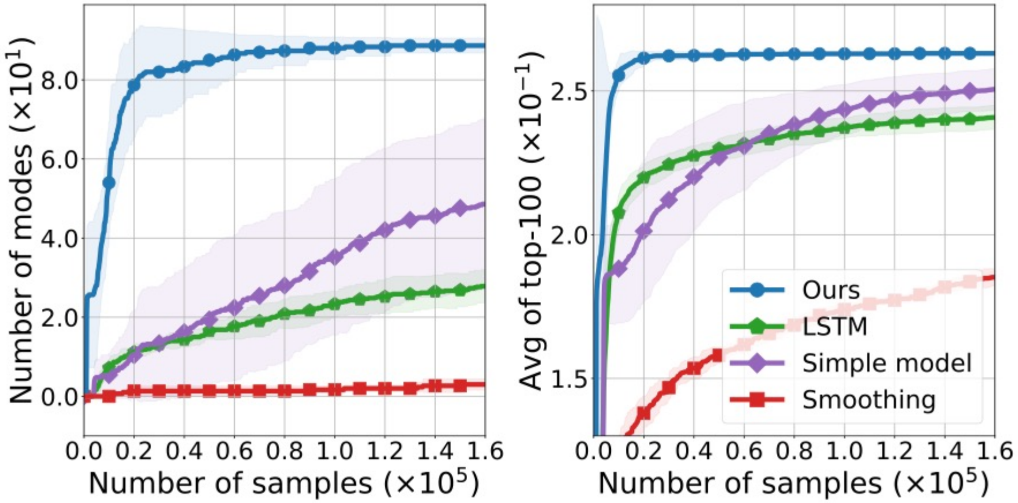
- We propose learning energy decomposition for GFlowNets (LED-GFN)
  - a simple and effective approach for improving GFlowNets
  - local credits identifying contribution of actions via learning
  - informative local credits acting as an important inductive bias

Appendix

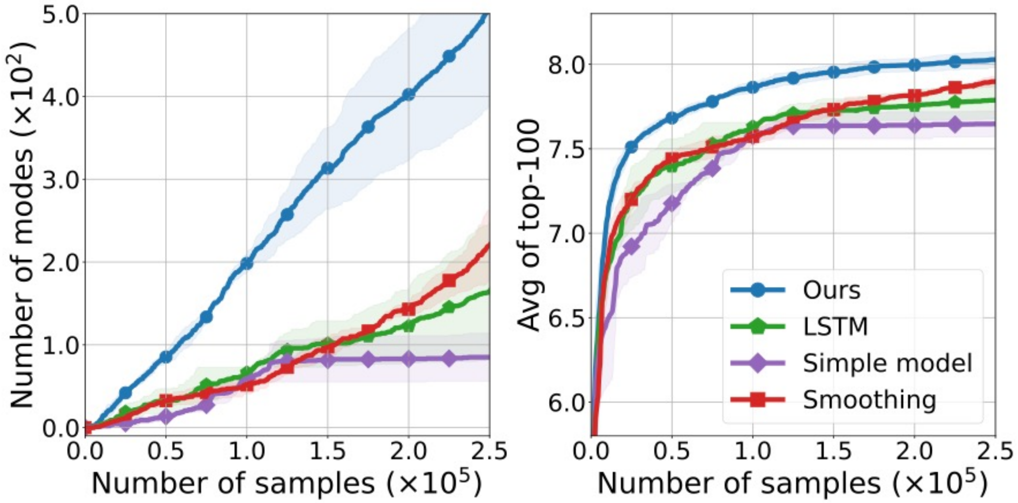


action
bag state
× 7
high score
LED-GFN
FL-GFN
cumulative credits
time step

Appendix

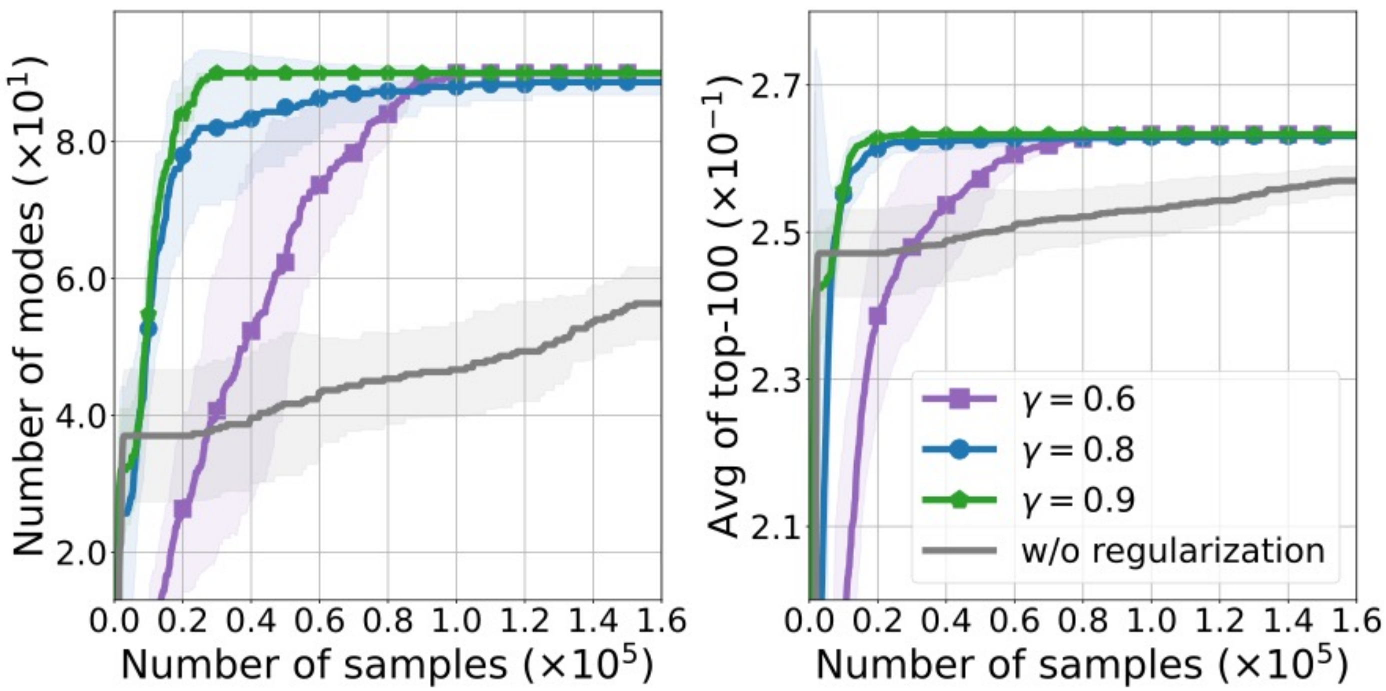


(a) Set generation



(b) Molecule generation

Appendix



(a) Set generation

Appendix

Table 1: **Time costs (sec) analysis.** The LED-GFN does not incur significant overheads.

Method	Time cost
subTB	5.80 (–)
FL-subTB	11.43 (↑ 5.63)
LED-subTB	6.61 (↑ 0.81)