Diffusion Twigs with Loop Guidance for Conditional Graph Generation

The 38th Annual Conference on Neural Information Processing Systems (NeurIPS 2024)

Giangiacomo Mercatali[*,†]
HES-SO Genève

Yogesh Verma[*]
Aalto University

Andre Freitas
University of Manchester
Idiap Research Institute

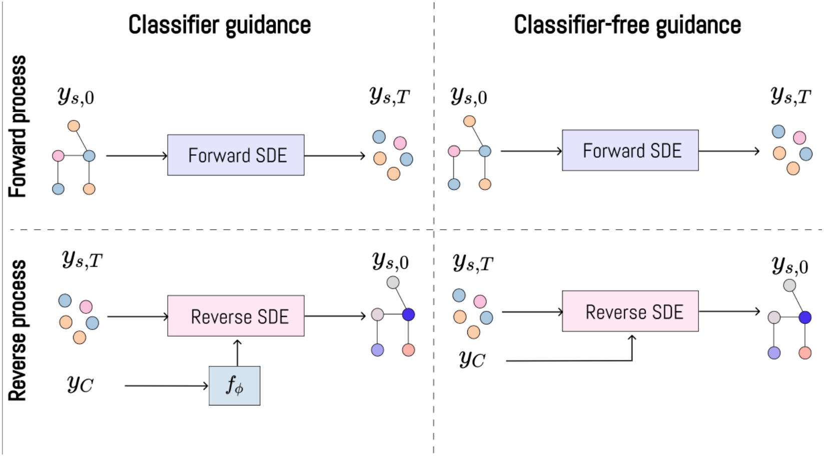
Vikas Garg
Aalto University
YaiYai Ltd

[*]Equal contribution
[†]Work done at the University of Manchester

Problem: Conditional graph generation

- Fundamental problem for: de-novo drug design and material design.
- Task: Generate a graph with certain desired properties.

Existing Guiding diffusion procedures

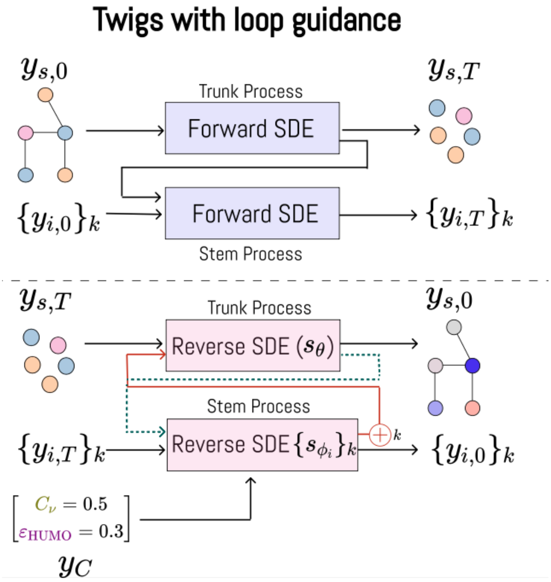


Related works

Table 1: Comparison of related diffusion models

Method	Conditional	Asymmetric	Multiple flows	Continuous (SDEs)
GDSS	✗	✗	✓	✓
EEGSDE	✓	✗	✗	✓
MOOD	✓	✗	✗	✓
JODO	✓	✗	✗	✓
EDGE	✗	✗	✓	✗
GraphMaker	✓	✗	✓	✗
Klarner et al	✓	✓	✗	✗
Twigs (ours)	✓	✓	✓	✓

Twigs Diffusion models



Twigs Forward process

$$\underbrace{\mathrm{d}\mathbf{y}_s = \mathbf{f}_s(\mathbf{y}_{s,t}, t)\mathrm{d}t + g_s(t)\mathrm{d}\mathbf{w}}_{\text{graph structure}} \tag{1}$$

$$\underbrace{\mathrm{d}\mathbf{y}(t) = \begin{pmatrix} \mathrm{d}\mathbf{y}_1(t) \\ \vdots \\ \mathrm{d}\mathbf{y}_k(t) \end{pmatrix} = \begin{pmatrix} \mathbf{f}_p(\mathbf{y}_{1,t}, \mathbf{y}_{s,t}, t)\mathrm{d}t + g_p(t)\mathrm{d}\mathbf{w} \\ \vdots \\ \mathbf{f}_p(\mathbf{y}_{k,t}, \mathbf{y}_{s,t}, t)\mathrm{d}t + g_p(t)\mathrm{d}\mathbf{w} \end{pmatrix}}_{\text{graph properties}} \tag{2}$$

Reverse process (Unconditional Twigs)

$$\mathrm{d}\mathbf{y}_t = [\mathbf{f}(\mathbf{y}_t, t) - g_t^2(\underbrace{\nabla_{\mathbf{y}_t} \log p_t(\mathbf{y}_{s,t})}_{\text{Trunk process}} + \underbrace{\textstyle\sum_i^k \nabla_{\mathbf{y}_t} \log p_t(\mathbf{y}_{i,t} \mid \mathbf{y}_{s,t})}_{\text{Stem processes}})]\mathrm{d}t + g_t \mathrm{d}\bar{\mathbf{w}} \tag{3}$$

Reverse SDE (Conditional Twigs)

$$\mathrm{d}\mathbf{y}_t = [\mathbf{f}(\mathbf{y}_t, t) - g_t^2 \nabla_{\mathbf{y}_t} \log p_t(\mathbf{y}_t, \mathbf{y}_C)]\mathrm{d}t + g_t \mathrm{d}\bar{\mathbf{w}} \tag{4}$$

$$\begin{aligned} &\underbrace{\nabla_{\mathbf{y}_t} \log p_t(\mathbf{y}_{s,t}, \mathbf{y}_C)}_{\text{graph structure}} + \underbrace{\sum_{i \notin C}^{k} \nabla_{\mathbf{y}_t} \log p_t(\mathbf{y}_{i,t} \mid \mathbf{y}_{s,t})}_{\text{non-conditional properties}} \\ &\qquad + \underbrace{\sum_{c}^{C} \sum_{i}^{k} \delta_{i=c} \nabla_{\mathbf{y}_t} \log p_t(\mathbf{y}_{i,t} \mid \mathbf{y}_{s,t}, \mathbf{y}_c)}_{\text{conditional properties}} \end{aligned} \tag{5}$$

Training (Conditional Twigs)

$$\min_{\theta,\phi_i} \mathbb{E}_t \left\{\lambda_{\mathbf{y}_t}(t)\mathbb{E}_{\mathbf{y}_0}\mathbb{E}_{\mathbf{y}_t|\mathbf{y}_0}\|\text{Twigs}_{\theta,\phi} - \underbrace{\nabla_{\mathbf{y}_t} \log p_{0t}(\mathbf{y}_t, \mathbf{y}_C)}_{\text{Score function}}\|_2^2\right\} \tag{6}$$

$$\text{where } \text{Twigs}_{\theta,\phi} = \underbrace{\boldsymbol{s}_{\theta,t}(\mathbf{y}_{s,t}, \mathbf{y}_c)}_{\text{Trunk process}} + \underbrace{\sum_i^k \boldsymbol{s}_{\phi_i,t}(\mathbf{y}_{i,t}, \mathbf{y}_{s,t}, \mathbf{y}_c)}_{\text{Stem processes}}$$

$$\text{and} \quad \mathbb{E}_{\mathbf{y}_0} = \mathbb{E}_{\mathbf{y}_{s,0},\mathbf{y}_{i,0}}, \mathbb{E}_{\mathbf{y}_t} = \mathbb{E}_{\mathbf{y}_{s,t},\mathbf{y}_{i,t}}$$

Experiments

Conditional graph generation for

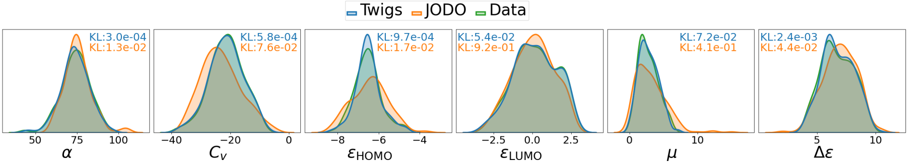
1. Single / multiple properties QM9
2. Molecule optimization ZINC250K
3. Single / multiple properties Network data

Conditioning on a single QM9 property

Table 2: MAE↓ results.

Method	C_v	μ	α	$\Delta\epsilon$	ϵ_{HOMO}	ϵ_{LUMO}
EDM	1.065 (± 0.010)	1.123 (± 0.013)	2.78 (± 0.04)	671 (± 5)	371 (± 2)	601 (± 7)
GeoLDM	1.025 (± na)	1.108 (± na)	2.37 (± na)	587 (± na)	340 (± na)	522 (± na)
EEGSDE	0.941 (± 0.005)	0.777 (± 0.007)	2.50 (± 0.02)	487 (± 3)	302 (± 2)	447 (± 6)
EquiFM	1.033 (± na)	1.106 (± na)	2.41 (± na)	591 (± na)	337 (± na)	530(± na)
TEDMol	0.847 (± na)	0.840 (± na)	2.24 (± na)	443 (± na)	279 (± na)	412 (± na)
JODO	0.581 (± 0.001)	0.628 (± 0.003)	1.42 (± 0.01)	335 (± 3)	226 (± 1)	256 (± 1)
`Twigs`	**0.559** (± 0.002)	**0.627** (± 0.001)	**1.36** (± 0.01)	**323** (± 2)	**225** (± 1)	**244** (± 3)

Figure 1: Top: samples, bottom KDE plots.



Conditioning on two QM9 properties

Table 3: MAE (↓) for conditional generation on QM9 with two properties.

	C_v	μ	$\Delta\epsilon$	μ	α	μ
EDM	1.097(± 0.007)	1.156(± 0.011)	683(± 1)	1.130(± 0.007)	2.76(± 0.01)	1.158(± 0.002)
EEGSDE	0.981(± 0.008)	0.912(± 0.006)	563(± 3)	0.866(± 0.003)	2.61(± 0.01)	0.855(± 0.007)
TEDMol	0.645(± n/a)	0.836(± n/a)	489(± n/a)	0.843(± n/a)	2.27(± n/a)	0.809(± n/a)
JODO	0.634(± 0.002)	0.716(± 0.006)	350(± 4)	0.752(± 0.006)	1.52(± 0.01)	0.717(± 0.006)
Twigs	**0.602**(± 0.001)	**0.708**(± 0.002)	**343**(± 2)	**0.740**(± 0.003)	**1.46**(± 0.01)	**0.712**(± 0.002)

Molecule optimization on ZINC250K

Table 4: Novel top 5% docking score on ZINC250K.

Model	*parp1*	*fa7*	*5ht1b*	*braf*	*jak2*
REINVENT	8.702(± 0.523)	7.205(± 0.264)	8.770(± 0.316)	8.392(± 0.400)	8.165(± 0.277)
MORLD	7.532(± 0.260)	6.263(± 0.165)	7.869(± 0.650)	8.040(± 0.337)	7.816(± 0.133)
HierVAE	9.487(± 0.278)	6.812(± 0.274)	8.081(± 0.252)	8.978(± 0.525)	8.285(± 0.370)
GDSS	9.967(± 0.028)	7.775(± 0.039)	9.459(± 0.101)	9.224(± 0.068)	8.926(± 0.089)
MOOD	10.409(± 0.030)	7.947(± 0.034)	10.487(± 0.069)	**10.421**(± 0.050)	9.575(± 0.075)
Twigs	**10.449**(± 0.009)	**8.182**(± 0.012)	**10.542**(± 0.025)	10.343(± 0.024)	**9.678**(± 0.032)

Table 5: Novel hit ratio (↑) results on ZINC250K.

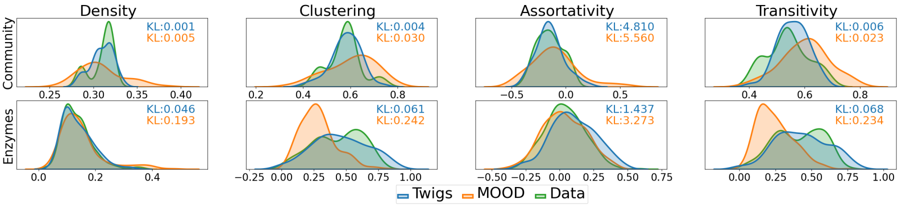
Model	*parp1*	*fa7*	*5ht1b*	*braf*	*jak2*
REINVENT	0.480(± 0.344)	0.213(± 0.081)	2.453(± 0.561)	0.127(± 0.088)	0.613(± 0.167)
MORLD	0.047(± 0.050)	0.007(± 0.013)	0.880(± 0.735)	0.047(± 0.040)	0.227(± 0.118)
HierVAE	0.553(± 0.214)	0.007(± 0.013)	0.507(± 0.278)	0.207(± 0.220)	0.227(± 0.127)
GDSS	1.933(± 0.208)	0.368(± 0.103)	4.667(± 0.306)	0.167(± 0.134)	1.167(± 0.281)
MOOD	3.400(± 0.117)	0.433(± 0.063)	11.873(± 0.521)	**2.207**(± 0.165)	3.953(± 0.383)
Twigs	**3.733**(± 0.081)	**0.900**(± 0.012)	**16.366**(± 0.029)	1.933(± 0.023)	**5.100**(± 0.312)

Network graphs with desired single properties

Table 6: MAE (↓) values.

	Community Small				Enzymes			
Model	Density	Clustering	Assortativity	Transitivity	Density	Clustering	Assortativity	Transitivity
GDSS	2.95	12.1	19.6	11.4	8.04	2.53	1.98	2.55
GDSS-T	2.30	11.5	19.2	10.1	9.25	3.27	2.03	2.68
Digress	2.34	10.6	17.8	9.42	8.04	2.39	1.95	2.55
MOOD-1	2.35	11.1	18.8	10.5	7.94	2.34	1.83	2.12
MOOD-4	2.12	11.3	16.7	8.76	7.98	2.44	1.99	2.43
`Twigs`	**2.07**	**9.67**	**15.2**	**8.35**	**7.35**	**2.23**	**1.72**	**2.03**

Figure 2: KDE plots



Conclusion and discussion

Summary:

- Conditional diffusion method
- Forks the diffusion into multiple stem diffusion flows
- Re-integrates them into the trunk flow
- Resembles a loop guidance

Future work:

1. Investigating efficacy in other domains (e.g., image, text, audio)
2. Grouping together potentially correlated properties into a joint stem process.

More details.

Poster Session 5

Fri 13 Dec 11 a.m. PST — 2 p.m. PST