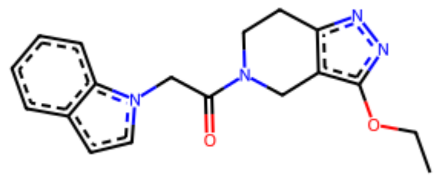
DIRECTO

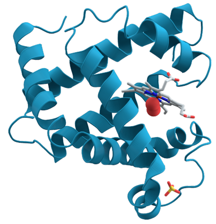
Generating Directed Graphs with Dual Attention and Asymmetric Encoding

Alba Carballo Castro, Manuel Madeira, Yiming Qin, Dorina Thanou, Pascal Frossard

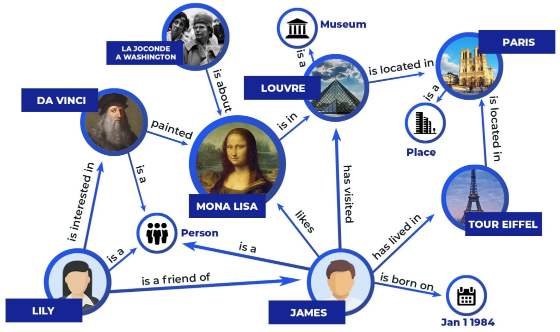
International Conference on Learning Representations (ICLR) 2026

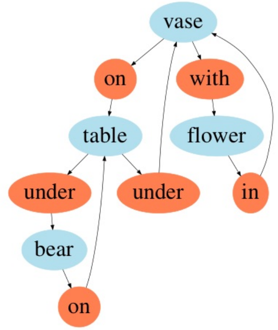
Generating *directed* graphs



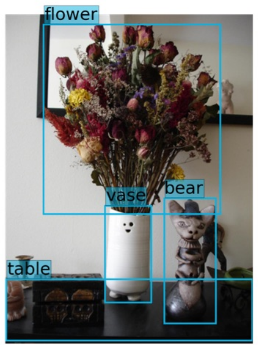


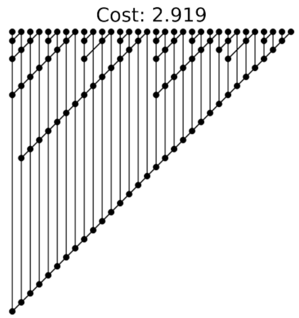
Undirected graph generation is seeing unprecedented success.





Directed graph generation is underexplored, despite its real-world applications.

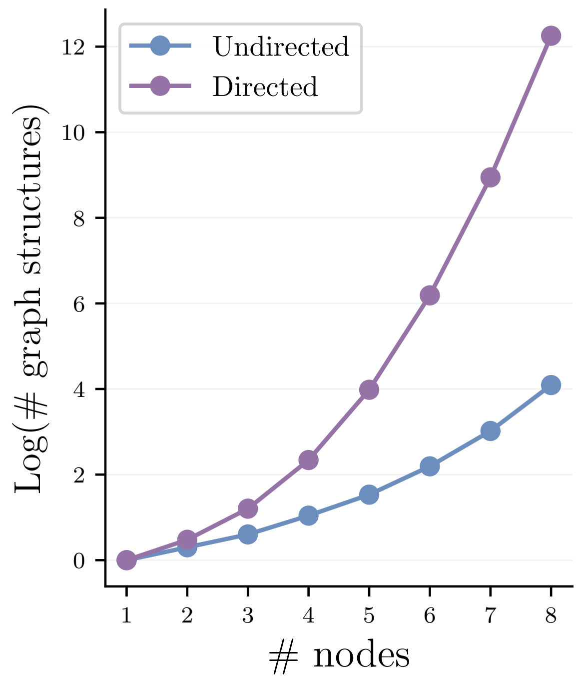




Generating *directed* graphs

Problem: the search space grows combinatorially in n^2 and current methods fail to capture directionality

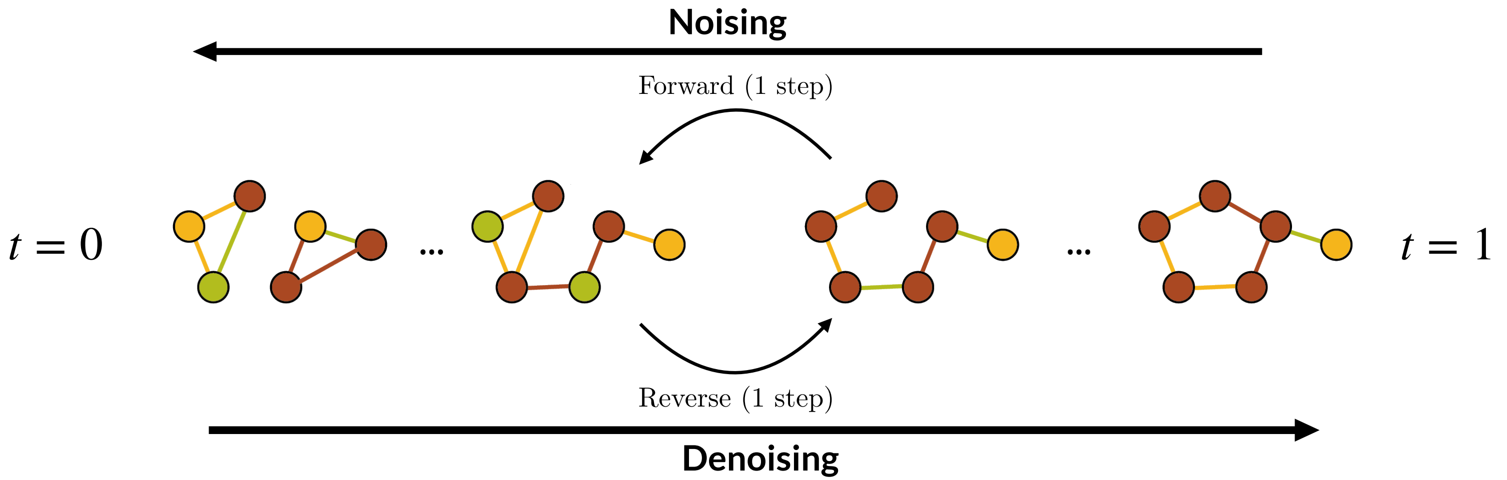
Solution: dual attention + directed Positional Encodings



Background: Discrete Flow Matching for Graph Generation

We model graphs as a tuple of nodes and edges: $G = \left(x^{(1:n:N)}, e^{(1:i\neq j:N)}\right)$

Noising process ($t = 1 \rightarrow t = 0$):
Linear interpolation between data distribution p_1 and initial distribution p_0.



Denoising process ($t = 0 \rightarrow t = 1$):
Reverts noising process based on CTMC formulation + model backbone: **graph transformer**

Generative Flows on Discrete State-Spaces: Enabling Multimodal Flows with Applications to Protein Co-Design, A. Campbell, J. Yim, R. Barzilay, T. Rainforth and T. Jaakkola, in ICML, 2024
DeFoG: Discrete Flow Matching for Graph Generation, Y. Qin, M. Madeira, D. Thanou and P. Frossard, in ICML, 2025

Improvement 1: Directed Positional Encodings

Positional encodings are added to enhance the **(directional) prior** in the network and improve performance:

- Magnetic Laplacian:

$$\boldsymbol{L}_U^{(q)} := \boxed{\boldsymbol{D}_s - \boldsymbol{A}_s} \odot \boxed{\exp\left(i\boldsymbol{\Theta}^{(q)}\right)}$$

Encodes edge direction
(complex valued)

Standard Laplacian
(real valued)

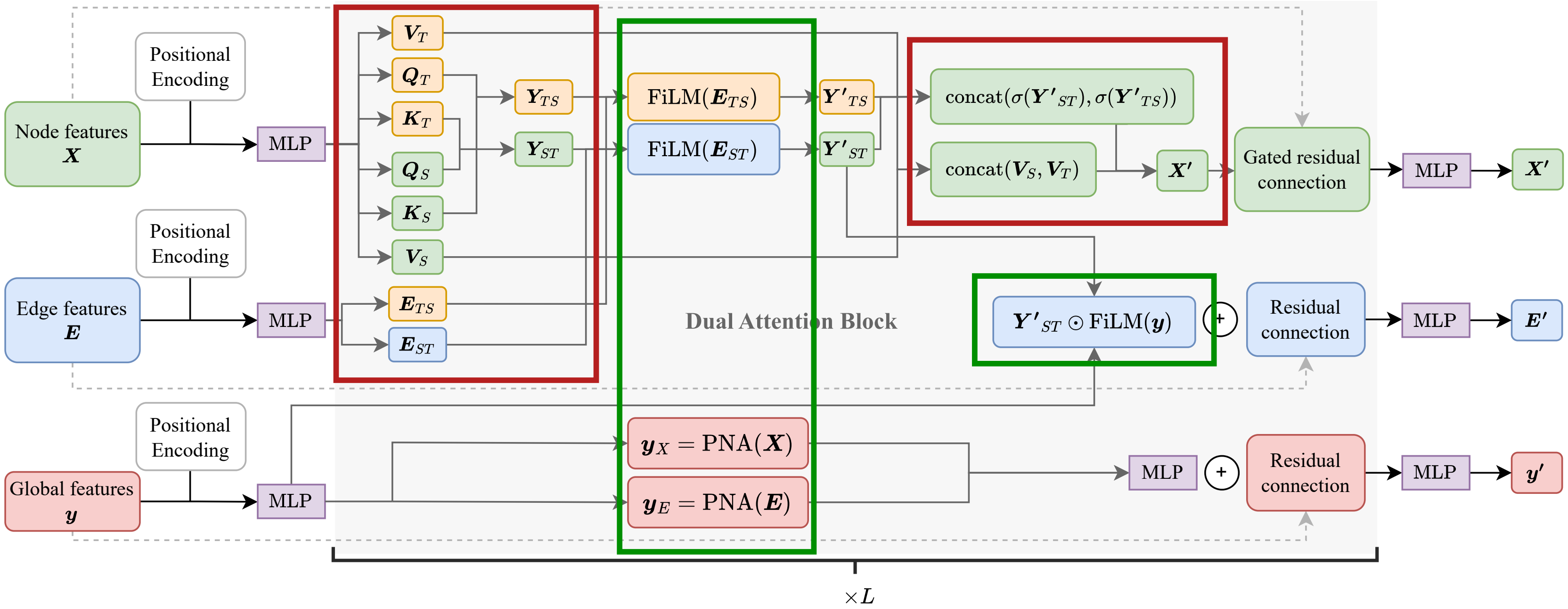
- **Multi-Q Magnetic Laplacian:** stacking multiple potentials together
- **Directed RRWP:** both $T = AD_{out}^{-1}$ and $R = A^\top D_{in}^{-1}$ are considered and stacked together

Transformers Meet Directed Graphs, S.Geisler, Y. Li, D. Mankowitz, A. T. Cemgil, S. Günnemann, C. Paduraru, in ICLR 2023.
What Are Good Positional Encodings for Directed Graphs?, Y. Huang, H. Wang, P. Li, in ICLR 2025.

Improvement 2: Dual Attention architecture

Enhancing the **directional awareness** of the Graph Transformer.

Attention-based aggregation to capture directionality

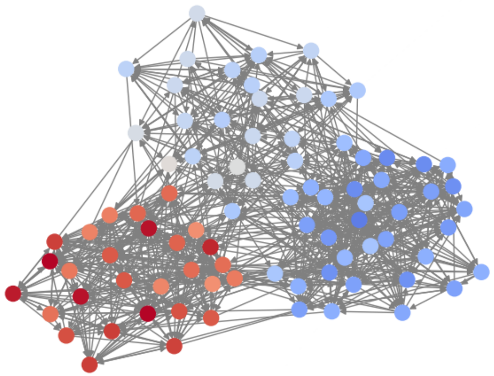


Modulation layers to incorporate information across graph components (nodes, edges, global features).

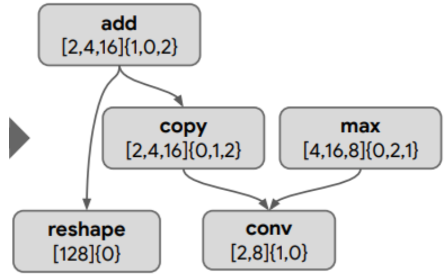
Benchmarking

Synthetic and real-world datasets with relevant metrics

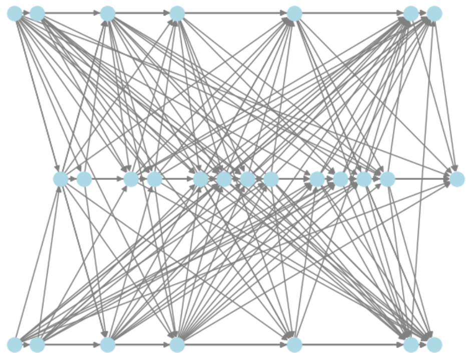
Directed Stochastic Block Model (SBM)



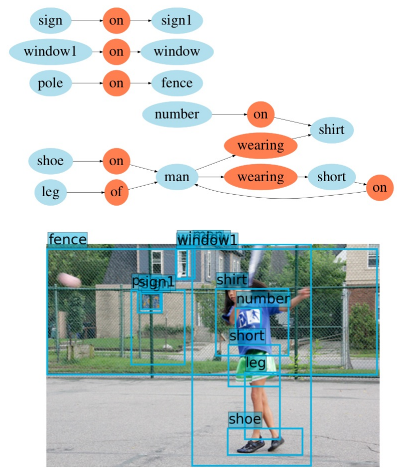
TPU Tiles (Neural Architecture Search)



Erdős-Renyi DAGs



Visual Genome (Scene Graphs)



TpuGraphs: A Performance Prediction Dataset on Large Tensor Computational Graphs, P. M. Phothilimthana, S. Abu-El-Haija, K. Cao, B. Fatemi, M. Burrows, C. Mendis, B. Perozzi, in NeurIPS Datasets and Benchmarks, 2025
Visual Genome: Connecting language and vision using crowdsourced dense image annotations, R. Krishna, Y. Zhu, O. Groth, J. Johnson, K. Hata, J. Kravitz, S. Chen, Y. Kalantidis, et al., in IJCV, 2017

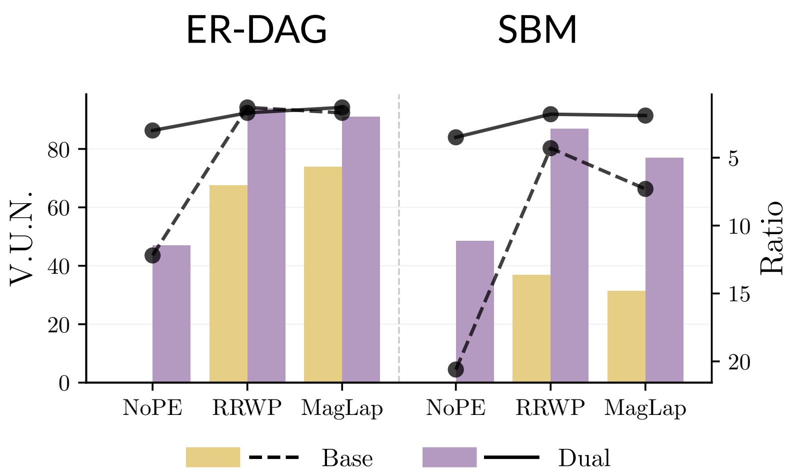
Results

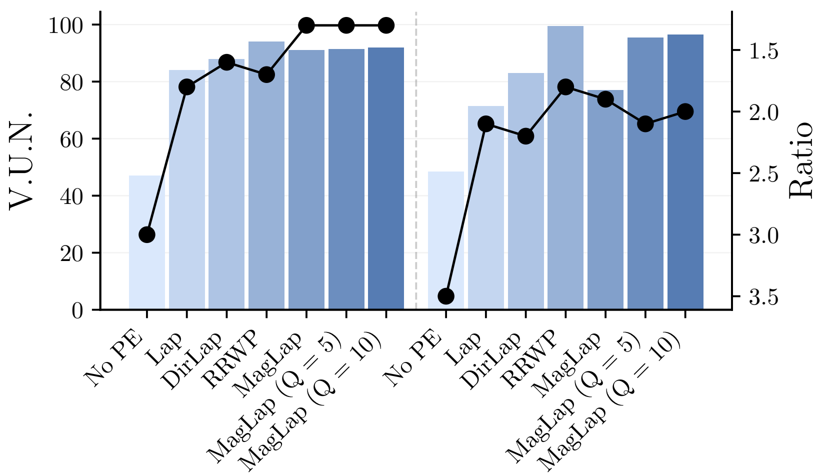
DIRECTO: original Discrete Flow Matching implementation

DIRECTO-DD: implementation with Discrete Diffusion

	ER-DAG		SBM		TPU Tiles			Visual Genome		
Model	**Ratio** ↓	**V.U.N.** ↑	**Ratio** ↓	**V.U.N.** ↑	**Ratio** ↓	**V.U.N.** ↑	**RBF** ↓	**Ratio** ↓	**V.U.N.** ↑	**RBF** ↓
Training set	1.0	0.0	1.0	0.0	1.0	0.0	0.002	1.0	0.0	0.021
MLE	15.1 ± 0.2	0.0 ± 0.0	11.6 ± 0.2	0.0 ± 0.0	149.8 ± 0.7	24.7 ± 0.0	1.039 ± 0.033	17.0 ± 0.6	0.0 ± 0.0	0.618 ± 0.025
D-VAE	106.6 ± 5.4	0.0 ± 0.0	-	-	OOT	OOT	OOT	-	-	-
LayerDAG	4.2 ± 3.2	21.5 ± 2.7	-	-	413.6 ± 70.1	**98.5** ± 3.0	1.021 ± 0.023	-	-	-
DiGress	1.9 ± 0.3	34.0 ± 4.1	3.9 ± 0.9	41.5 ± 5.1	<u>57.5</u> ± 1.7	70.9 ± 3.4	0.097 ± 0.033	17.0 ± 0.6	0.3 ± 0.6	0.232 ± 0.028
DeFoG	1.6 ± 0.2	75.0 ± 2.2	4.3 ± 0.8	37.0 ± 6.6	63.7 ± 2.6	72.0 ± 2.4	0.059 ± 0.015	10.8 ± 0.7	50.8 ± 8.4	0.085 ± 0.023
DIRECTO-DD RRWP	<u>1.4</u> ± 0.3	79.0 ± 3.7	<u>1.7</u> ± 0.4	81.5 ± 3.2	61.0 ± 2.9	76.8 ± 1.9	0.058 ± 0.023	15.3 ± 0.8	<u>72.7</u> ± 3.9	<u>0.039</u> ± 0.004
DIRECTO-DD MagLap	1.5 ± 0.2	85.0 ± 9.2	**1.5** ± 0.4	95.5 ± 3.7	64.3 ± 5.3	77.0 ± 7.0	0.079 ± 0.027	7.6 ± 0.7	61.9 ± 4.4	0.042 ± 0.006
DIRECTO RRWP	1.7 ± 0.1	**94.0** ± 1.0	1.8 ± 0.5	**99.5** ± 1.0	75.4 ± 8.1	77.0 ± 2.9	<u>0.044</u> ± 0.018	12.8 ± 0.6	**83.8** ± 4.3	**0.038** ± 0.005
DIRECTO MagLap	**1.3** ± 0.2	<u>92.0</u> ± 3.7	2.0 ± 0.3	<u>96.5</u> ± 2.5	**44.0** ± 7.1	<u>80.5</u> ± 4.6	**0.042** ± 0.001	**6.2** ± 0.5	67.0 ± 4.3	0.051 ± 0.012

Results: ablations



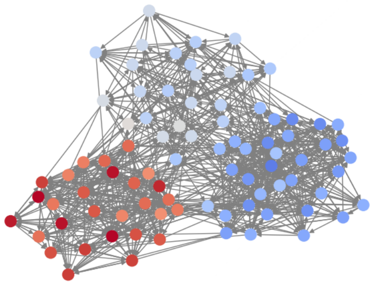


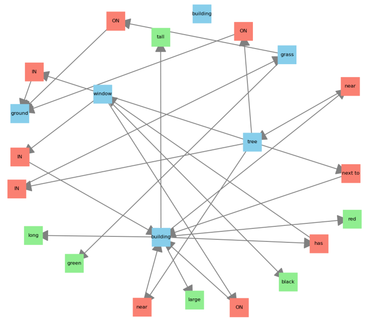
- Dual Attention is an essential component to capture directionality.
- PEs are also important for performance, and directed PEs help further.
- Our method is also efficient: for the same number of parameters, much better results!

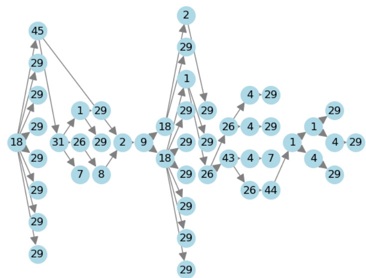
Dataset	**Model**	**Ratio** ↓	**V.U.N.** ↑
ER-DAG	RRWP - Double	4.9 ± 0.2	72.0 ± 9.8
	RRWP - Dual	$\mathbf{1.7 \pm 0.1}$	$\mathbf{94.0 \pm 1.0}$
	MagLap - Double	4.3 ± 0.7	80.0 ± 6.3
	MagLap - Dual	$\underline{1.3 \pm 0.2}$	$\underline{91.0 \pm 2.5}$
SBM	RRWP - Double	26.3 ± 2.5	0.0 ± 0.0
	RRWP - Dual	$\mathbf{1.8 \pm 0.5}$	$\mathbf{99.5 \pm 1.0}$
	MagLap - Double	14.2 ± 1.5	8.0 ± 7.5
	MagLap - Dual	$\underline{1.9 \pm 0.3}$	$\underline{77.0 \pm 7.6}$

Takeaways

- **DIRECTO** allows to generate directed graphs in a principled way.
- Dual attention is the most important component.
- Directed PEs further enhance generations.
- Experiments on a new benchmarking suite with synthetic and real-world datasets as well as tailored metrics.







International Conference On
Learning Representations

DIRECTO

Generating Directed Graphs with Dual Attention and Asymmetric Encoding

Alba Carballo Castro, Manuel Madeira, Yiming Qin, Dorina Thanou, Pascal Frossard

International Conference on Learning Representations (ICLR) 2026