

GraphVAE: Towards Generation of Small Graphs Using Variational Autoencoders

Martin Simonovsky Nikos Komodakis

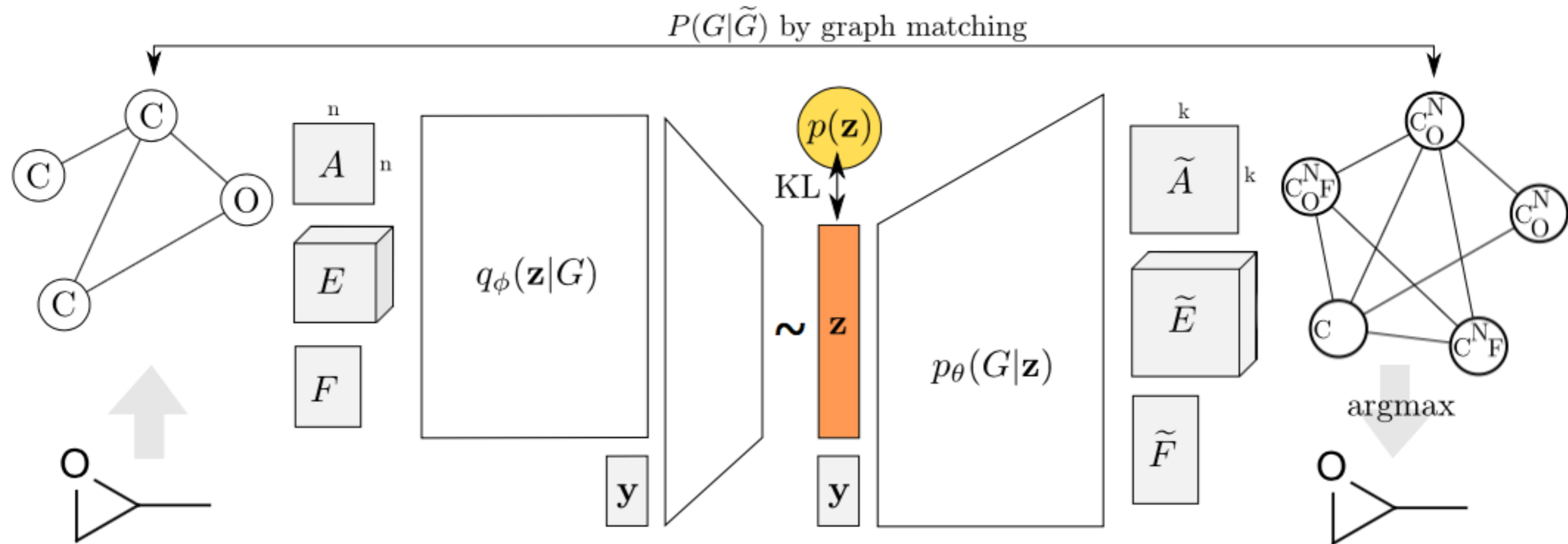
Presenter: Yevgeny Tkach

2019 Spring @
<https://qdata.github.io/deep2Read/>

Executive Summary

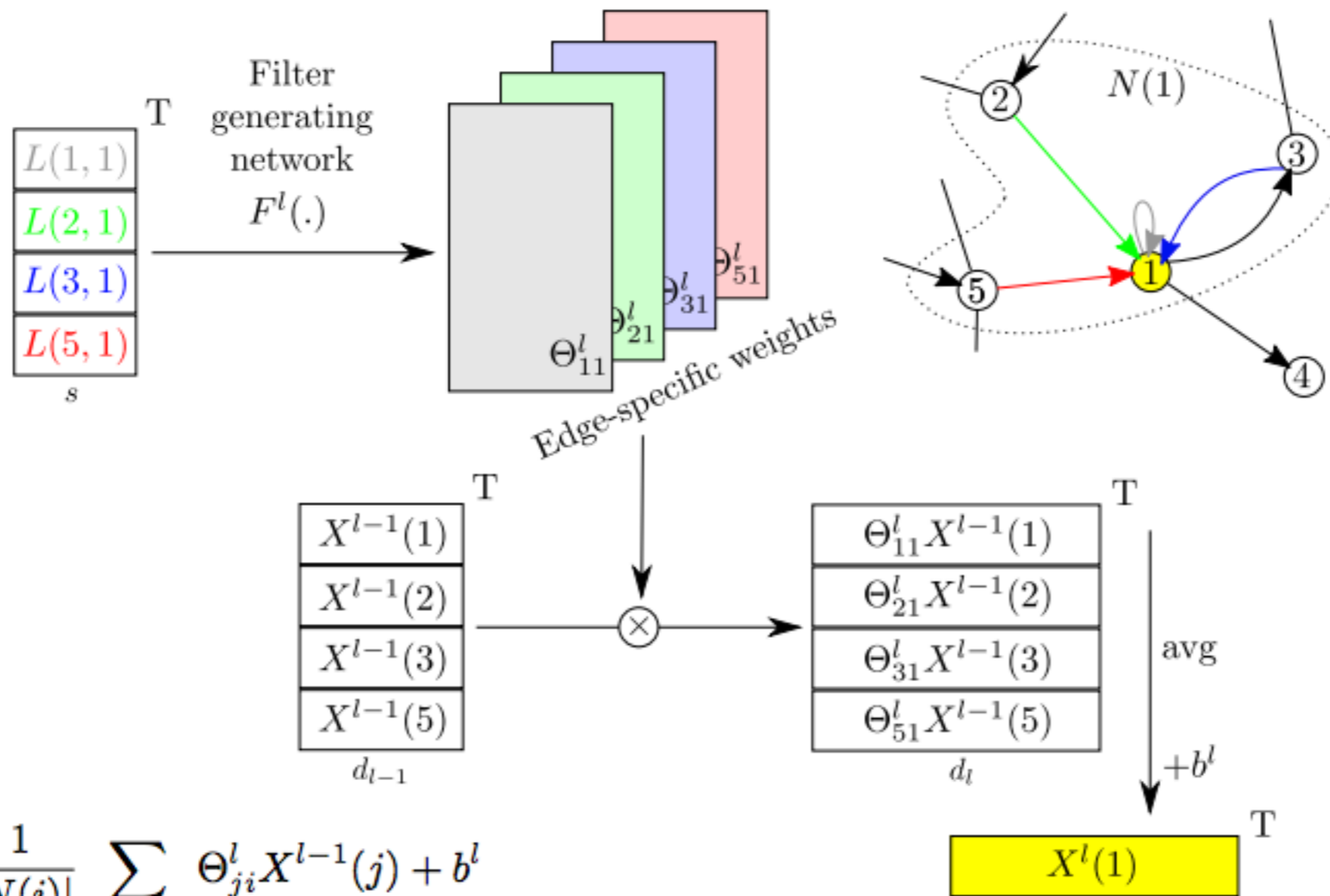
- Molecule generation using VAE
- Single step generation of the adjacency matrix and node and edge labels
- Reconstruction loss of VAE is multi target loss based on approximate graph matching solution
- The encoder is spatial graph method with edge conditioned convolutions, taken from another paper by the same authors
- The decoder architecture is a simple MLP with three outputs

GraphVAE



- A - adjacency matrix
- E - edge attribute tensor
- F - node attribute tensor

Graph Encoding



$$X^l(i) = \frac{1}{|N(i)|} \sum_{j \in N(i)} \Theta_{ji}^l X^{l-1}(j) + b^l$$

- F - a filter generating network given edge labels outputs edge specific weight matrix Θ

Graph Decoder

- MLP that outputs node probability, edge probability, i.e. the adjacency matrix, as well as node and edge label probability
- For a discrete estimation of the graph, nodes $\{a : \tilde{A}_{a,a} \geq 0.5\}$ are chosen
- Add edges between the chosen nodes if their probability is greater than 1/2
- Construct a maximum spanning tree between the chosen nodes and augment the chosen edges with the maximum spanning tree edges

Loss function

$$\begin{aligned}\mathcal{L}(\phi, \theta; G) &= \\ &= \mathbb{E}_{q_\phi(\mathbf{z}|G)}[-\log p_\theta(G|\mathbf{z})] + \text{KL}[q_\phi(\mathbf{z}|G)||p(\mathbf{z})] \quad (1) \\ p_\theta(G|\mathbf{z}) &= P(\tilde{G}|\tilde{G}).\end{aligned}$$

- Approximate graph matching can be used to obtain a binary assignment matrix $X \in \{0, 1\}^{k \times n}$ where $X_{a,i} = 1$ only if node $a \in \tilde{G}$ is assigned to $i \in G$ and $X_{a,i} = 0$ otherwise.
- Assuming we know X we can define $A' = XAX^T$, $\tilde{F}' = X^T \tilde{F}$, $\tilde{E}'_{\cdot, \cdot, l} = X^T \tilde{E}_{\cdot, \cdot, l} X$

$$\begin{aligned}\log p(A'|\mathbf{z}) &= \\ &= 1/k \sum_a A'_{a,a} \log \tilde{A}_{a,a} + (1 - A'_{a,a}) \log(1 - \tilde{A}_{a,a}) + \\ &+ 1/k(k-1) \sum_{a \neq b} A'_{a,b} \log \tilde{A}_{a,b} + (1 - A'_{a,b}) \log(1 - \tilde{A}_{a,b}) \\ \log p(F|\mathbf{z}) &= 1/n \sum_i \log F_{i,\cdot}^T \tilde{F}'_{i,\cdot} \\ \log p(E|\mathbf{z}) &= 1/(\|A\|_1 - n) \sum_{i \neq j} \log E_{i,j,\cdot}^T \tilde{E}'_{i,j,\cdot}, \quad (2)\end{aligned}$$

$$\begin{aligned}-\log p(G|\mathbf{z}) &= -\lambda_A \log p(A'|\mathbf{z}) - \lambda_F \log p(F|\mathbf{z}) - \\ &- \lambda_E \log p(E|\mathbf{z}) \quad (3)\end{aligned}$$

Loss function

- The matching algorithm solves the quadratic maximization problem:

$$\mathbf{x}^* = \arg \max_{\mathbf{x}} \mathbf{x}^T S \mathbf{x}$$

$$s.t. \sum_{i=1}^n \mathbf{x}_{ia} \leq 1, \sum_{a=1}^k \mathbf{x}_{ia} \leq 1, \mathbf{x} \in [0, 1]^{kn}$$

- The affinity matrix is defined:

$$\begin{aligned} S((i, j), (a, b)) &= \\ &= (E_{i,j}^T, \tilde{E}_{a,b,\cdot}) A_{i,j} \tilde{A}_{a,b} \tilde{A}_{a,a} \tilde{A}_{b,b} [i \neq j \wedge a \neq b] + \\ &+ (F_{i,\cdot}^T, \tilde{F}_{a,\cdot}) \tilde{A}_{a,a} [i = j \wedge a = b] \end{aligned} \quad (4)$$

- Max Pooling algorithm for graph matching:

repeat

for each candidate match (i, a) **do**

$$\mathbf{x}_{ia} \leftarrow \mathbf{x}_{ia} S_{ia:ia} + \sum_{j \in N_i} \max_{b \in N_a} \mathbf{x}_{jb} S_{ia;jb}$$

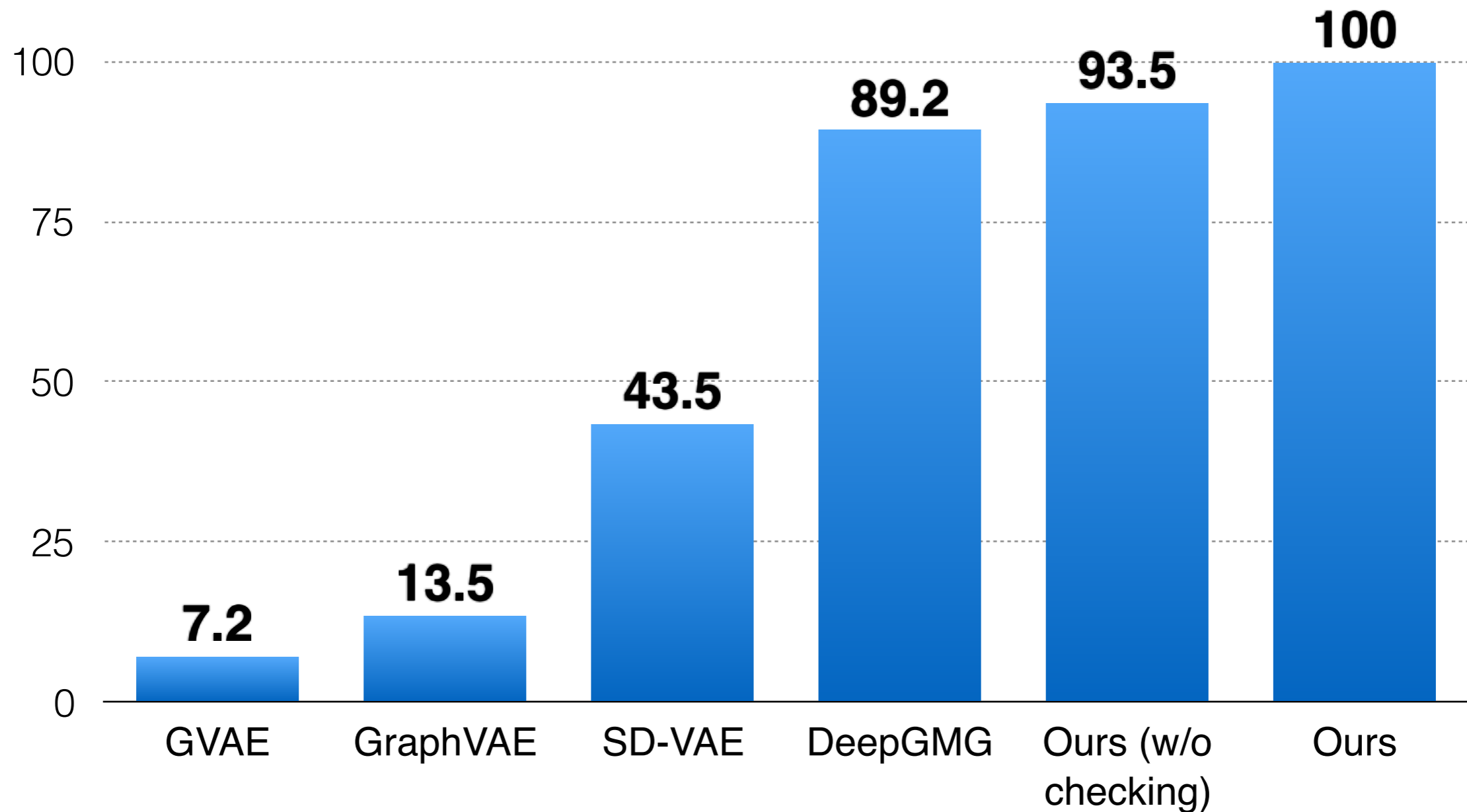
$$\mathbf{x} \leftarrow \frac{1}{\|\mathbf{x}\|_2} \mathbf{x};$$

until \mathbf{x} converges;

Experiment results - QM9

		$\log p_{\theta}(G \mathbf{z})$	ELBO	Valid	Accurate	Unique	Novel
Cond.	Ours $c = 20$	-0.578	-0.722	0.565	0.467	0.314	0.598
	Ours $c = 40$	-0.504	-0.617	0.511	0.416	0.484	0.635
	Ours $c = 60$	-0.492	-0.585	0.520	0.406	0.583	0.613
	Ours $c = 80$	-0.475	-0.557	0.458	0.353	0.666	0.661
Unconditional	Ours $c = 20$	-0.660	-0.916	0.485	0.485	0.457	0.575
	Ours $c = 40$	-0.537	-0.744	0.542	0.542	0.618	0.617
	Ours $c = 60$	-0.486	-0.656	0.517	0.517	0.695	0.570
	Ours $c = 80$	-0.482	-0.628	0.557	0.557	0.760	0.616
	NoGM $c = 80$	-2.388	-2.553	0.810	0.810	0.241	0.610
	CVAE $c = 60$	–	–	0.103	0.103	0.675	0.900
	GVAE $c = 20$	–	–	0.602	0.602	0.093	0.809

Experiment results - Zync



Discussion

- Using a matching algorithm to overcome the limitation imposed by lack of natural order is an interesting approach
- sequence generation approaches outperform single step generation for complex correlated data structures
- Single step generation is not scalable
- The work is not accompanied with code so it is hard to fill in the gaps about the actual algorithm that was used