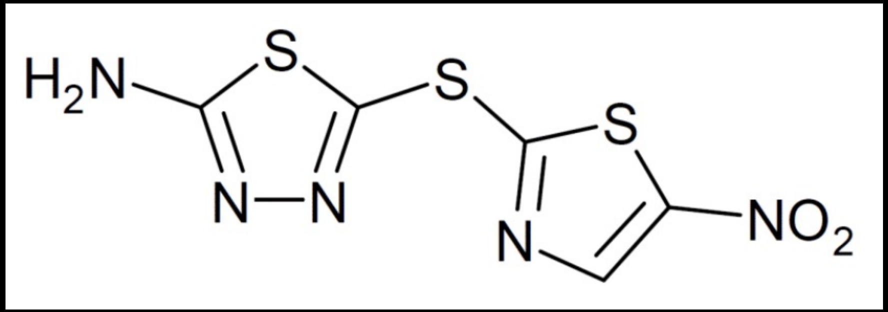
Contractive Systems Inspired GNNs

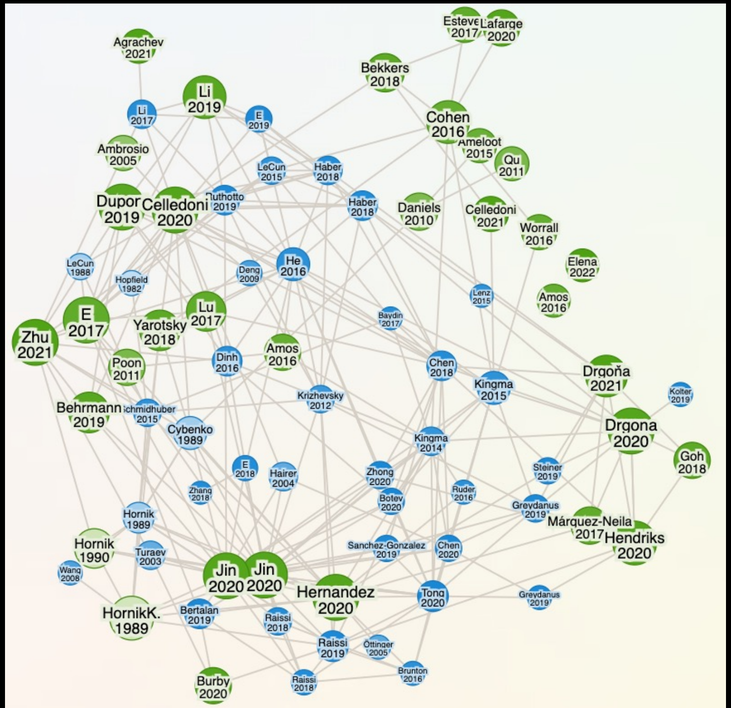
Davide Murari (NTNU)
davide.murari@ntnu.no

In collaboration with Moshe Eliasof, Ferdia Sherry and Carola Schönlieb (Cambridge University)

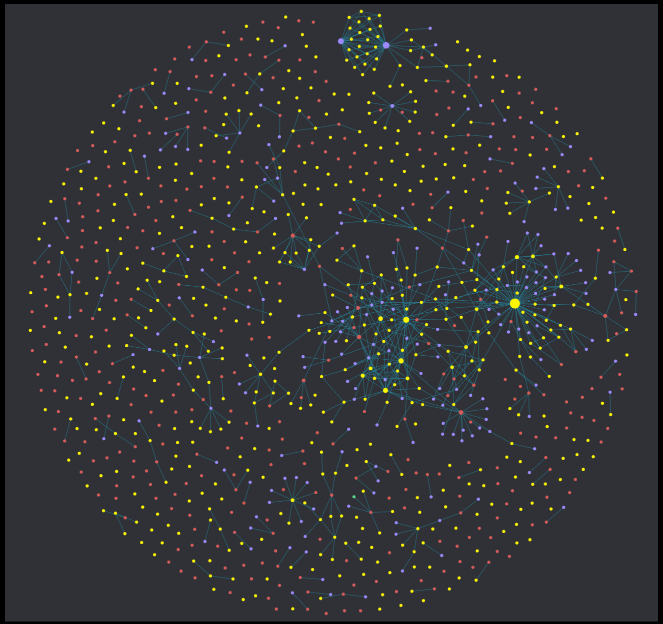
Graphs are everywhere



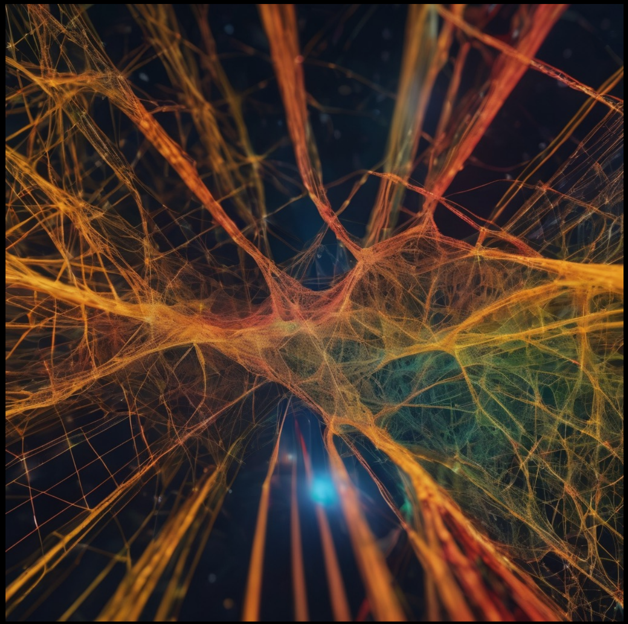
Molecule structure



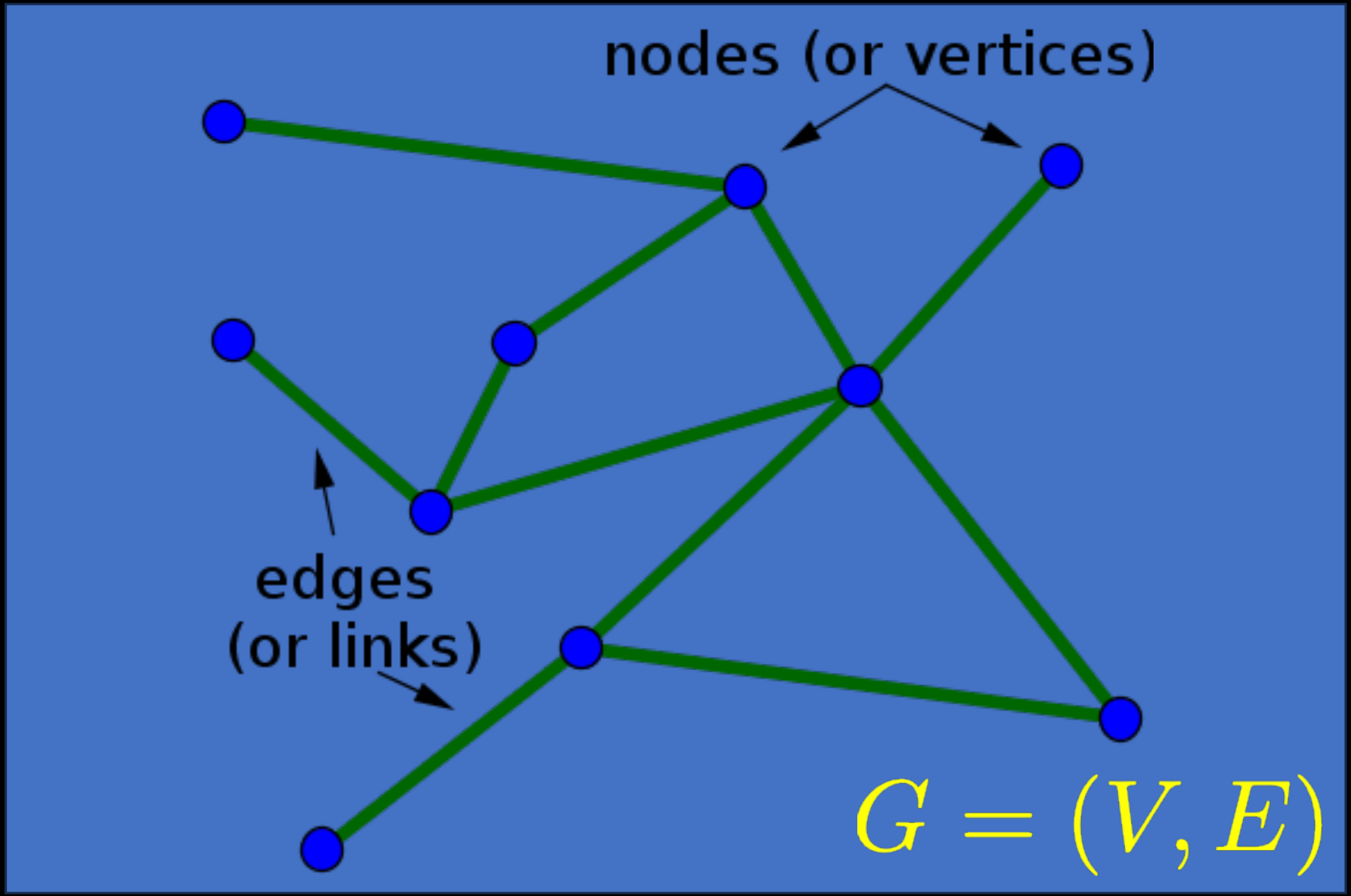
Citation graph of some papers I saved



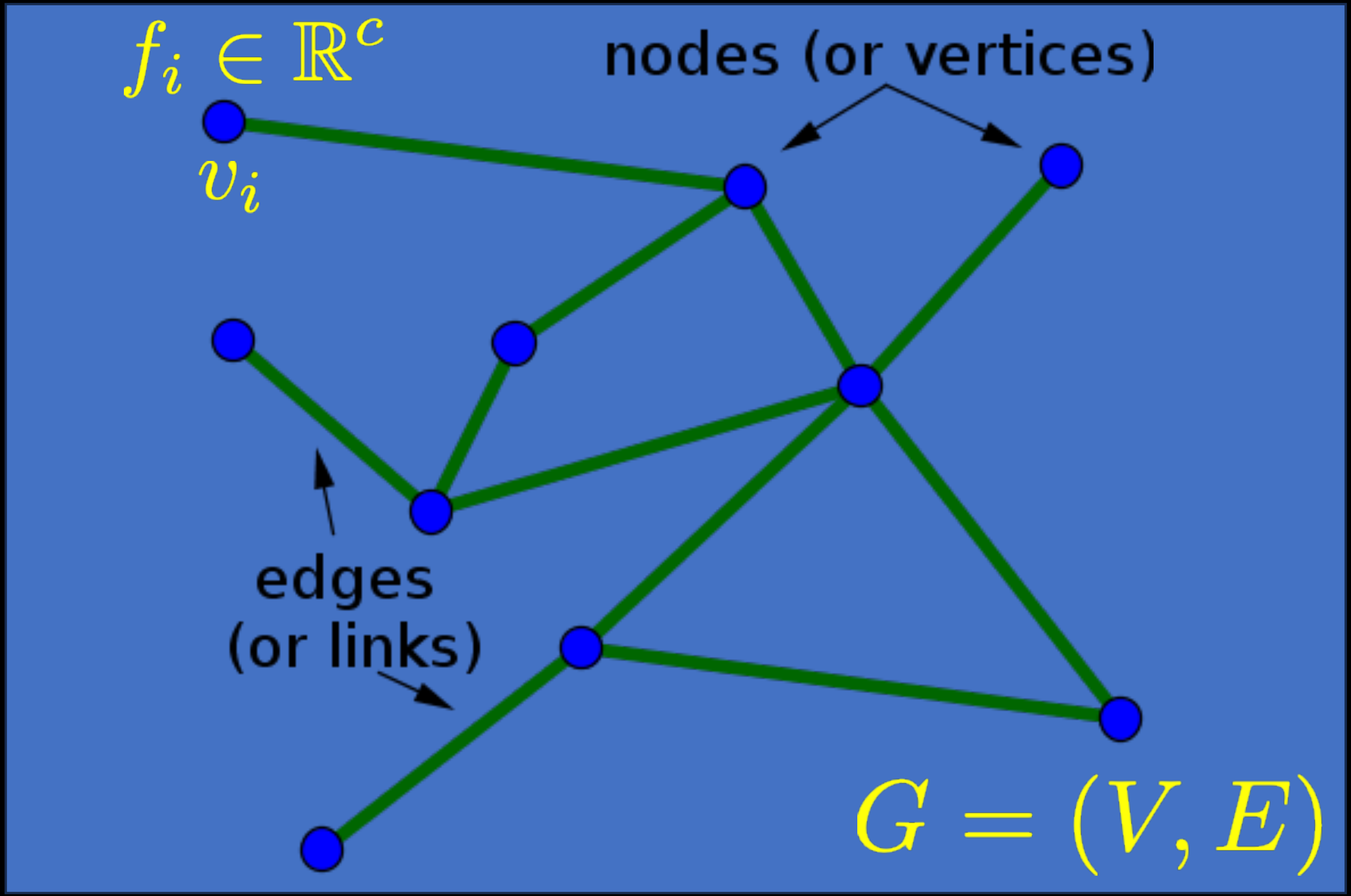
Reference graph some notes I wrote



Connectivity of neurons in the brain



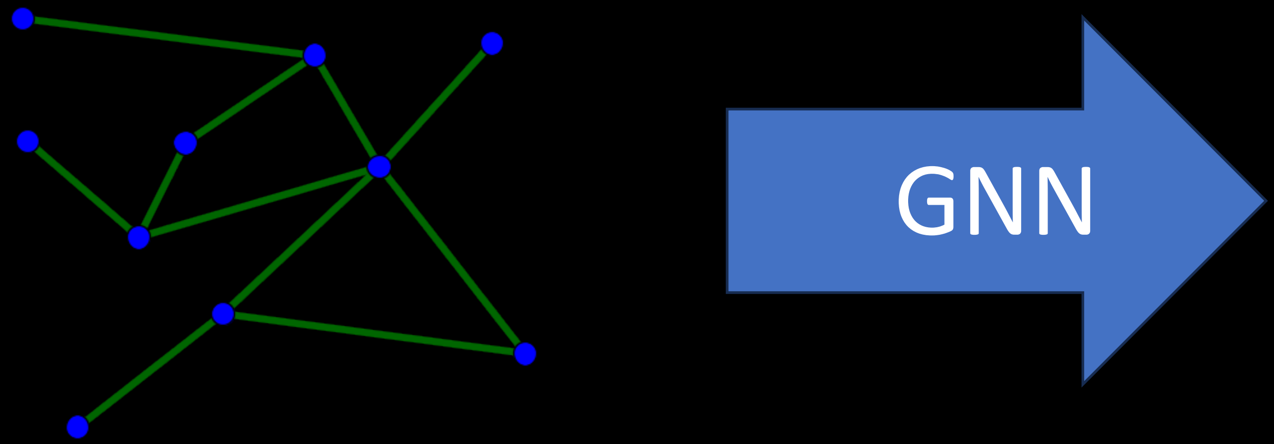
$$A \in \{0,1\}^{n\times n},\ \ A^T = A$$



$$F \in \mathbb{R}^{n \times c}, e_i^T F = f_i$$

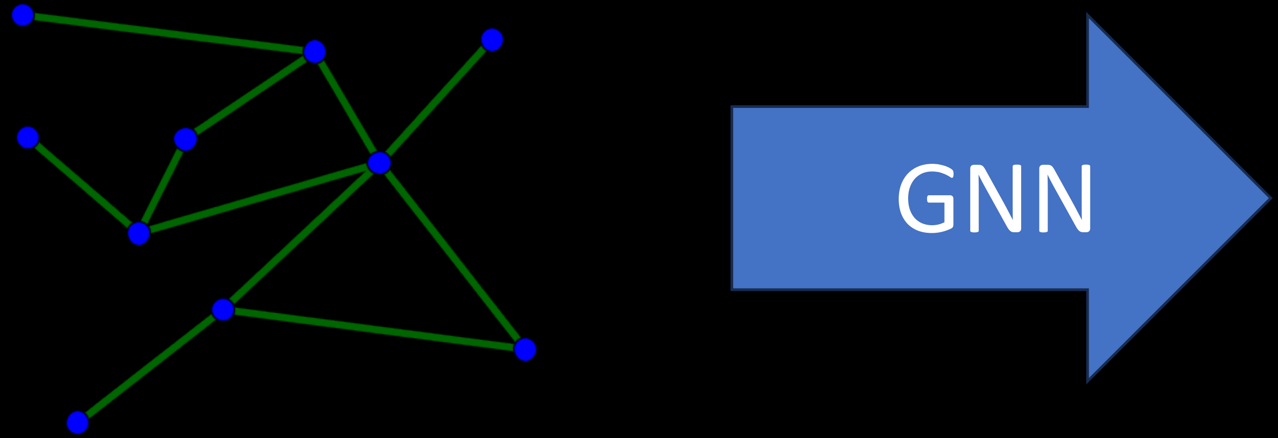
Classical tasks solved with GNNs

Graph classification



It is a protein

Node classification



Red
Blue
Blue
…
Green

Usual structure of GNNs

$$F^{(0)} = F$$

$$F^{(l+1)} = T_l\left(F^{(l)}, A\right), l = 0, \ldots L - 1$$

$$R = \mathrm{MLP}\left(\mathrm{F^{(L)}}\right) =: \mathrm{GNN(F, A)}$$

Invariant

$$\mathrm{GNN}(F, A) = \mathrm{GNN}(PF, PAP^T)$$

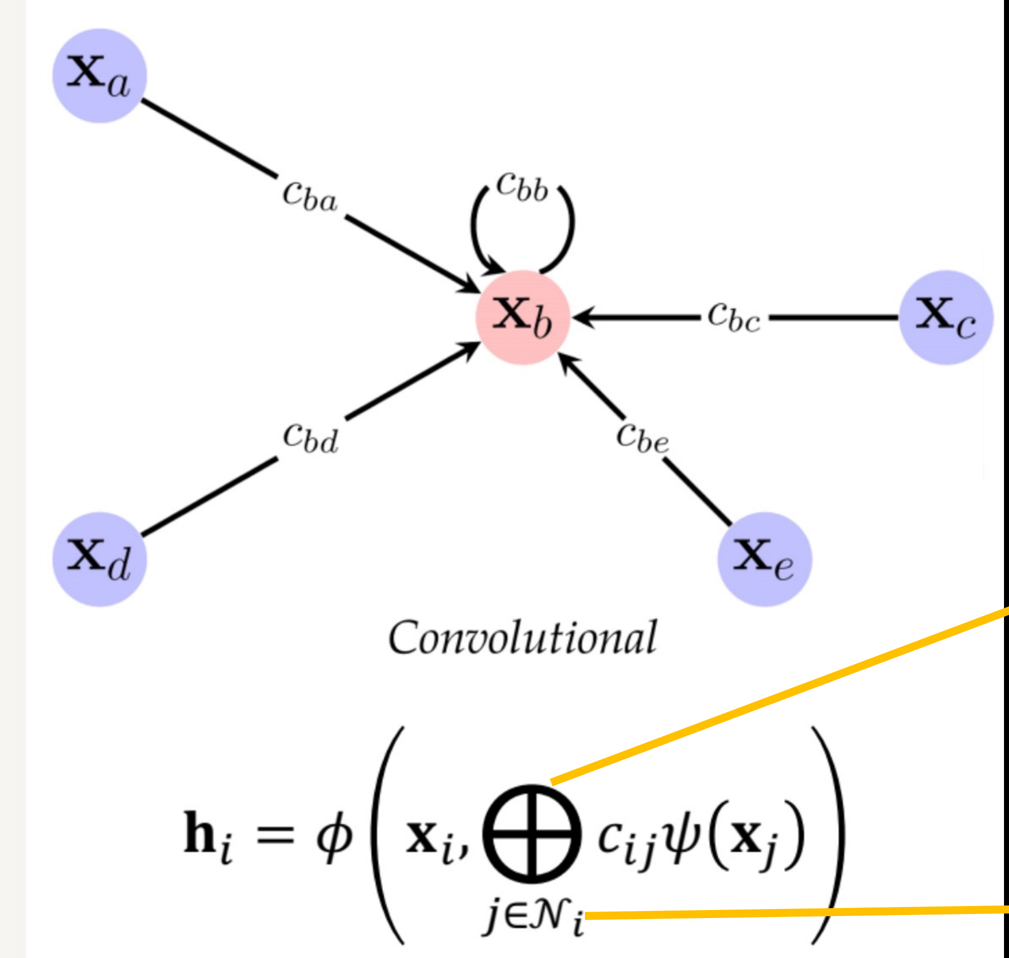
$$P\,\mathrm{GNN}(F, A) = \mathrm{GNN}(PF, PAP^T)$$

Equivariant

Usual structure of GNNs

ϕ, ψ

Are (learnable) functions

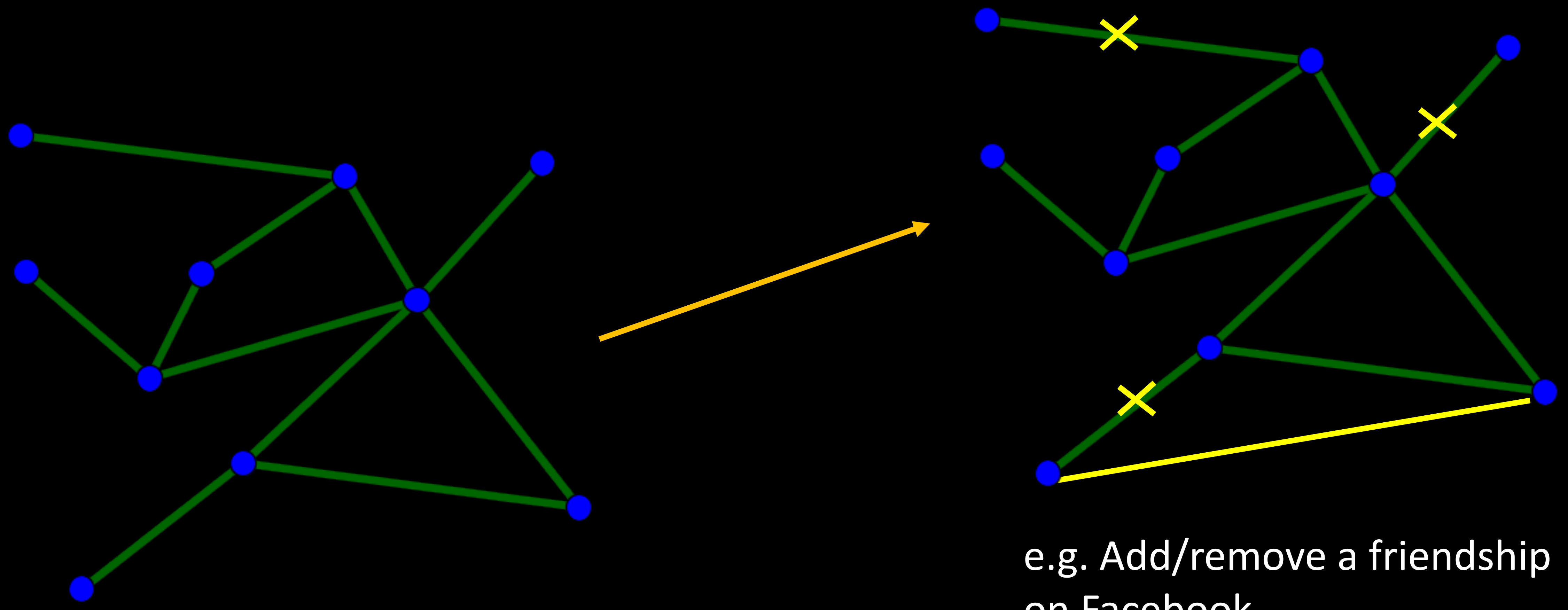


Permutation invariant aggregator, like a sum

Neighbourhood of the i-th node

Source : https://geometricdeeplearning.com/lectures/

Adversarial attacks



e.g. Add/remove a friendship
on Facebook

Adversarial attacks

$$F_* = F + \delta F, \quad \|\delta F\|_F \leq \varepsilon_1$$

$$A_* = A + \delta A, \quad \|\delta A\|_0 \leq \varepsilon_2$$

Attacks do not break the properties of symmetry generally

Goal:

$$\mathrm{GNN(F, A) \approx GNN(F_*, A_*)}$$

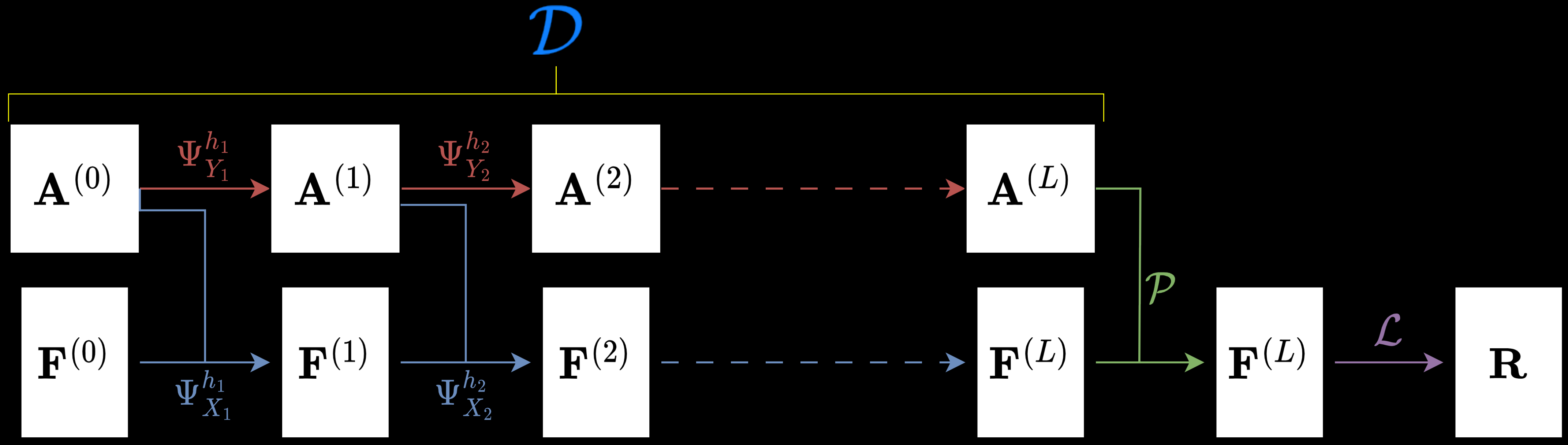
Remark on Nuclear Norm

$$A \in \{0,1\}^{n\times n} \implies \|A\|_0 = \#\{i,j \in \{1,\dots,n\} : A_{ij} \neq 0\} = \|\mathrm{vec(A)}\|_{\ell^1}$$

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \mathrm{vec}(A) = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{12} \\ a_{22} \end{bmatrix}$$

The 1-norm of the vectorisation is better suited for what we do, and we will use such norm instead of the nuclear norm.

Our proposed architecture: CSGNN



$$\left(F^{(0)}, A^{(0)}\right) := \left(\mathcal{K}\left(F_*\right), A_*\right)$$

$$\Psi_{X_i}^{h_i}(F, A) = F - h_i G(A)^T \sigma\left(G(A) F W_i\right) W_i^T$$

$$\Psi_{Y_i}^{h_i}(A) = A + h_i \sigma\left(M_i(A)\right)$$

Linear equivariant vector field

$$M(A) = k_1 A + k_2 \mathrm{diag}(\mathrm{diag}(A)) + \frac{k_3}{2n}(A\mathbf{1}_n\mathbf{1}_n^\top + \mathbf{1}_n\mathbf{1}_n^\top A) + k_4 \mathrm{diag}(A\mathbf{1}_n)$$
$$+ \frac{k_5}{n^2}(\mathbf{1}_n^\top A\mathbf{1}_n)\mathbf{1}_n\mathbf{1}_n^\top + \frac{k_6}{n}(\mathbf{1}_n^\top A\mathbf{1}_n)I_n + \frac{k_7}{n^2}(\mathbf{1}_n^\top \mathrm{diag}(A))\mathbf{1}_n\mathbf{1}_n^\top$$
$$+ \frac{k_8}{n}(\mathbf{1}_n^\top \mathrm{diag}(A))I_n + \frac{k_9}{2n}(\mathrm{diag}(A)\mathbf{1}_n^\top + \mathbf{1}_n(\mathrm{diag}(A))^\top)$$

$$M(PAP^T) = PM(A)P^T, \quad (M(A))^T = M(A)$$

Contractivity of feature updates

If $\sigma : \mathbb{R} \to \mathbb{R}$ is a non-decreasing 1-Lipschitz function, then the explicit Euler update is contractive in the F-norm when $h_i \leq 2/\|W_i\|_2^2$:

$$\left\|\Psi_{X_i}^{h_i}(\mathbf{F} + \delta\mathbf{F}, \mathbf{A}) - \Psi_{X_i}^{h_i}(\mathbf{F}, \mathbf{A})\right\|_F \leq \|\delta\mathbf{F}\|_F,$$

$$\delta\mathbf{F} \in \mathbb{R}^{n \times c}$$

Contractivity of adjacency updates

If $\sigma : \mathbb{R} \to \mathbb{R}$ is a non-decreasing 1-Lipschitz function, then the explicit Euler update is contractive in the vectorized 1-norm when

$$h_i \leq \frac{2}{\left(2\sum_{i=2}^{9}|k_i|\right) - \alpha}, \quad k_1 = \left(\alpha - \sum_{i=2}^{9}|k_i|\right), \; \alpha \leq 0.$$

This means that:

$$\left\|\mathrm{vec}(\Psi_{Y_i}^{h_i}(\boldsymbol{A} + \delta\boldsymbol{A})) - \mathrm{vec}(\Psi_{Y_i}^{h_i}(\boldsymbol{A}))\right\|_1 \leq \|\mathrm{vec}(\delta\boldsymbol{A})\|_1,$$

$$\delta\boldsymbol{A} \in \mathbb{R}^{n\times n}$$

Robustness of the network

If the assumptions of the two previous theorems hold, and

$$\boldsymbol{A}_*^{(0)} = \boldsymbol{A}^{(0)} + \delta \boldsymbol{A}, \ \ \boldsymbol{F}_*^{(0)} = \boldsymbol{F}^{(0)} + \delta \boldsymbol{F}$$

$$\|\delta \boldsymbol{F}\|_F \leq \varepsilon_1, \ \ \|\mathrm{vec}(\delta \boldsymbol{A})\|_1 \leq \varepsilon_2,$$

it follows

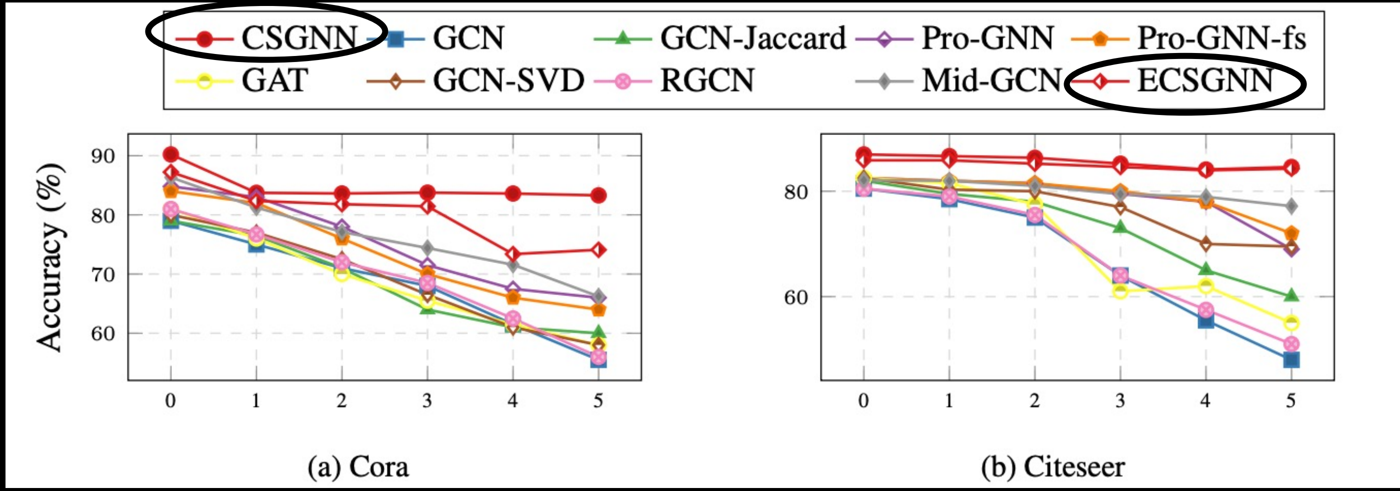
$$\begin{aligned} d\left(\mathcal{D}\left(\mathbf{F}^{(0)}, \mathbf{A}^{(0)}\right), \mathcal{D}\left(\mathbf{F}_*^{(0)}, \mathbf{A}_*^{(0)}\right)\right) &:= \left\|\mathrm{vec}\left(\mathbf{A}^{(L)}\right) - \mathrm{vec}\left(\mathbf{A}_*^{(L)}\right)\right\|_1 + \left\|\mathbf{F}^{(L)} - \mathbf{F}_*^{(L)}\right\|_F \\ &\leq \varepsilon_1 + \varepsilon_2 \left(1 + \sum_{i=1}^{L} \mathrm{Lip}\left(X_{i,\mathbf{F}^{(i-1)}}\right) h_i\right) \\ &=: \varepsilon_1 + c\left(h_1, \ldots, h_L\right) \varepsilon_2. \end{aligned}$$

Experimental setup

Hyperparameter	Range	Distribution
input/output embedding learning rate	$[10^{-5}, 10^{-2}]$	uniform
node dynamics learning rate	$[10^{-5}, 10^{-2}]$	uniform
adjacency dynamics learning rate	$[10^{-5}, 10^{-2}]$	uniform
input/output embedding weight decay	$[5 \cdot 10^{-8}, 5 \cdot 10^{-2}]$	log uniform
node dynamics weight decay	$[5 \cdot 10^{-8}, 5 \cdot 10^{-2}]$	log uniform
adjacency dynamics weight decay	$[5 \cdot 10^{-8}, 5 \cdot 10^{-2}]$	log uniform
input/output embedding dropout	$[0, 0.6]$	uniform
node dynamics dropout	$[0, 0.6]$	uniform
share weights between time steps	$\{\text{yes}, \text{no}\}$	discrete uniform
step size h	$[10^{-2}, 1]$	log uniform
adjacency contractivity parameter α	$[-2, 0]$	uniform
#layers L	$\{2, 3, 4, 5\}$	discrete uniform
#channels c	$\{8, 16, 32, 64, 128\}$	discrete uniform

Some experimental results

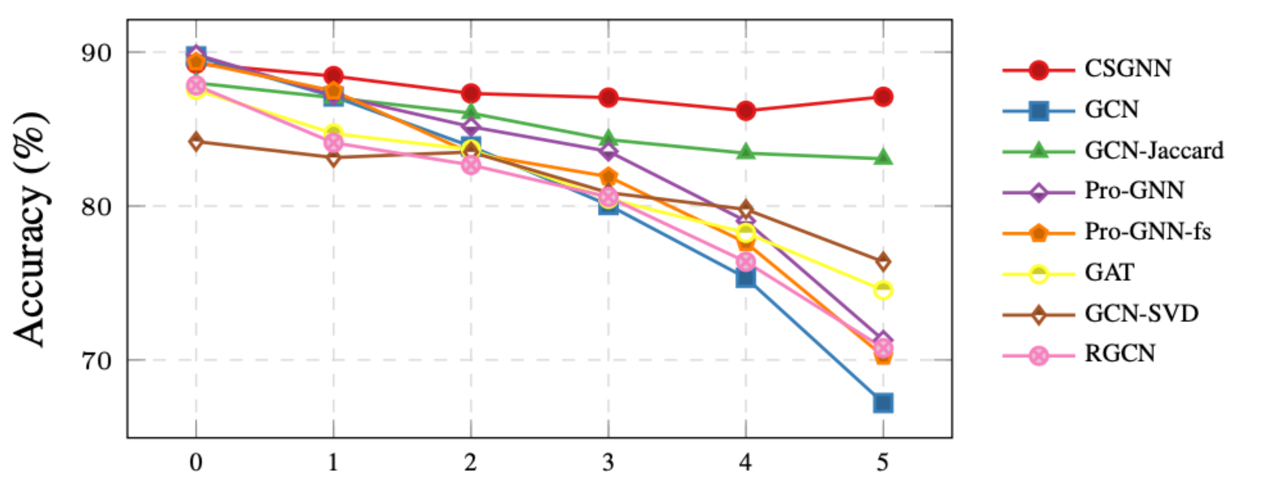
Method	Cora nettack	Cora metattack	Cora random	Citeseer nettack	Citeseer metattack	Citeseer random
$\text{CSGNN}_{\text{noAdj}}$	81.90	70.25	77.19	82.20	70.17	71.28
CSGNN	83.29	74.46	78.38	84.60	72.94	72.70



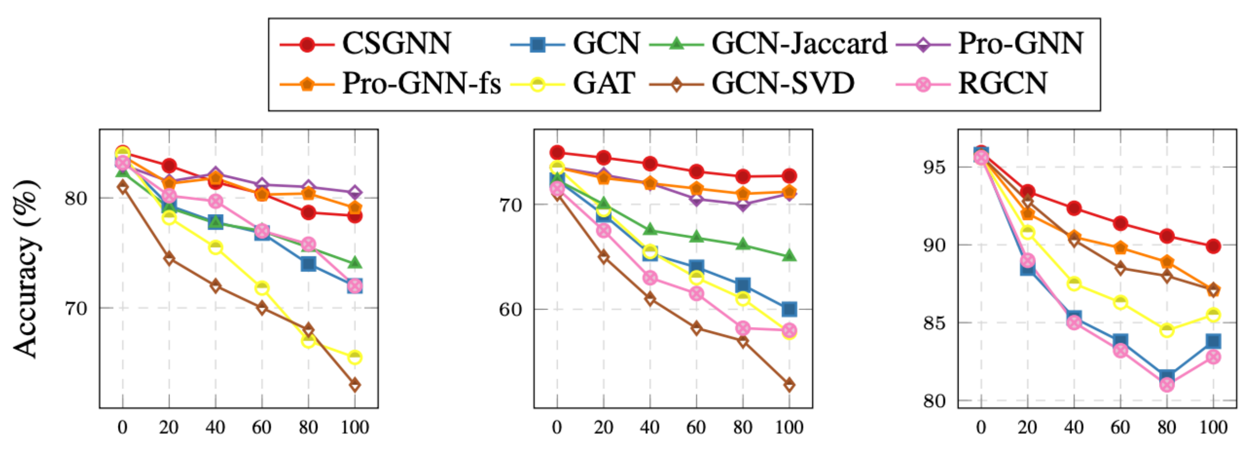
Node classification accuracy (%) of ECSGNN and other baselines, under a targeted attack generated by nettack. The horizontal axis describes the number of perturbations per node.

We target the nodes with degree at least 10 and flip few of their incident edges

Some experimental results



Classification accuracy for The Pubmed dataset using Nettack as attack method.



(a) Cora

(b) Citeseer

(c) Polblogs

The adjacency matrix is attacked by adding random fake edges, from 0% to 100% of the number of edges in the true one.

Thank you for the attention

Scan for the preprint

Eliasof, M., M., D., Sherry, F., & Schönlieb, C. B. (2023). Contractive Systems Improve Graph Neural Networks Against Adversarial Attacks. *arXiv preprint*.