

Winning the Lottery Twice

Author: Maximilian Burger
Supervisor: Pieter-Jan Hoedt

Intro

- Lottery tickets are highly sparse
- ‘unnecessary’ parameters are removed
- Since they are very reduced, their structure could provide insights

Question

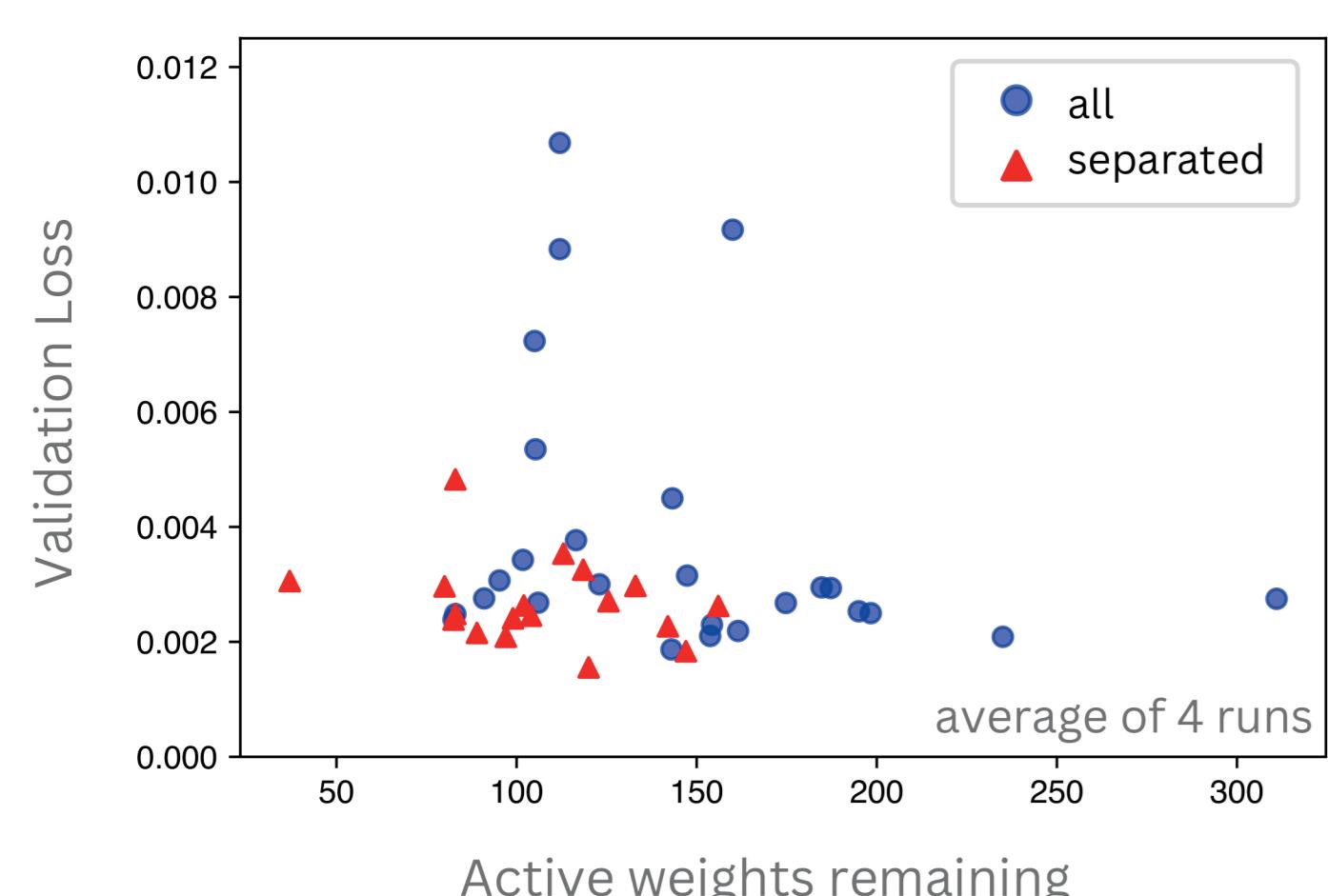
Do lottery tickets reveal the structure of the problem they were trained on?

Experiments

- Classify of 2 independent problems, combined in one dataset
- Find lottery tickets with IMP
- Check if separate networks formed

Results

- The network reliably splits into two
- at similar number of active weights
- over model sizes and pruning levels



dictionary

active parameters
parameters connected to input and output

separate / split networks
networks are not connected by any weight, considering only active parameters

degraded networks
at least one input or output is not connected

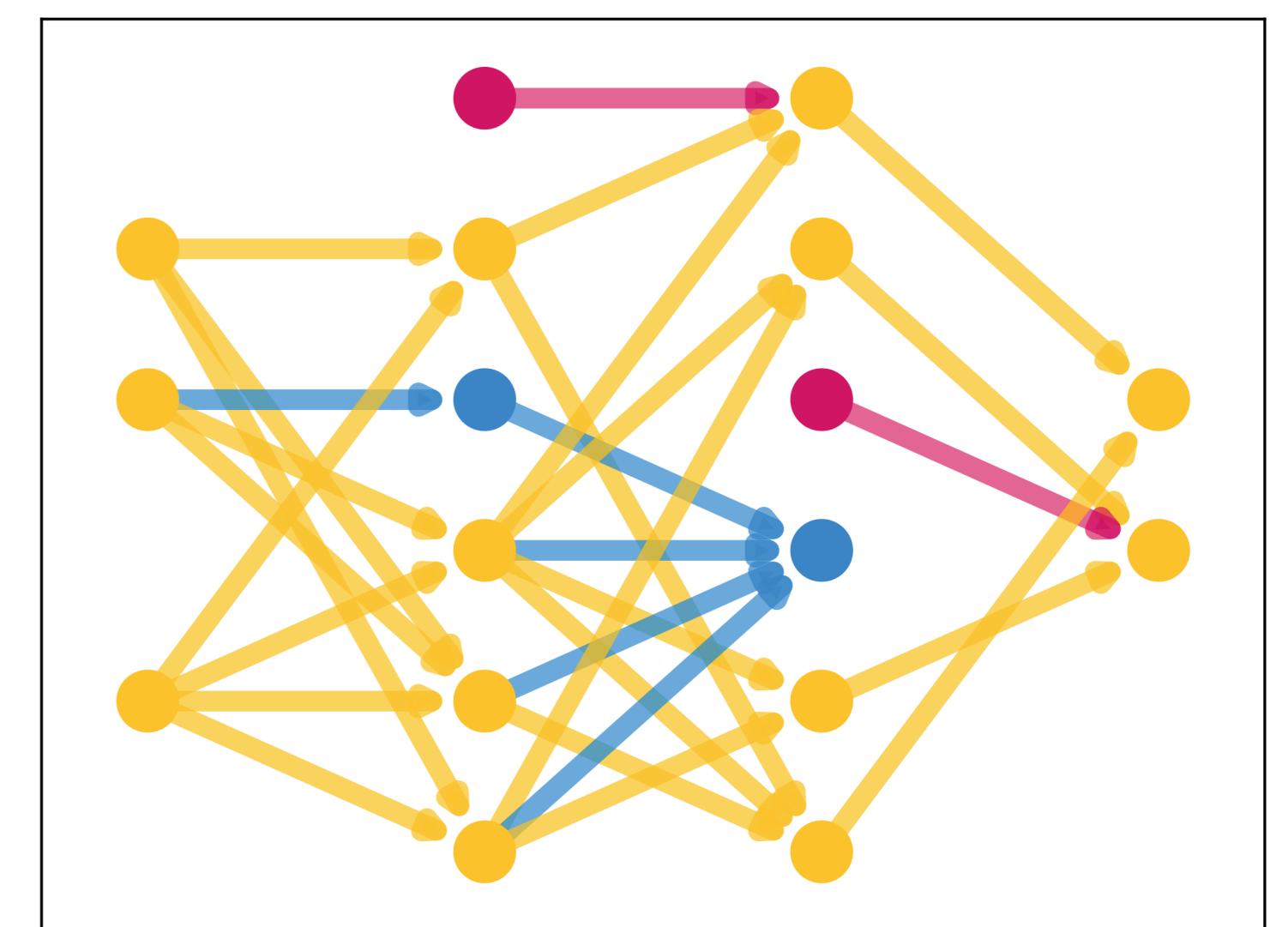
pruning levels
number of iterations for IMP

pruning rate
% of parameters removed at each pruning level

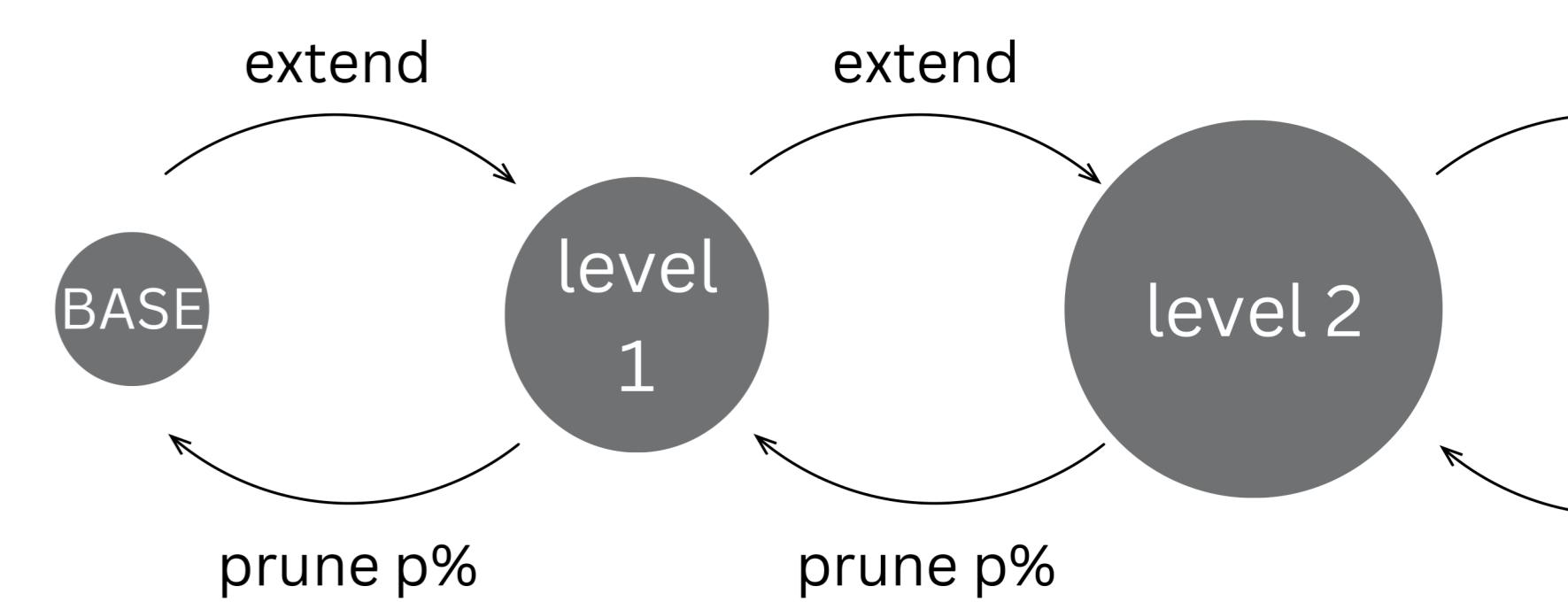
pruning target
the final number of parameters after pruning

Lottery Tickets contain independent subnetworks for independent tasks.

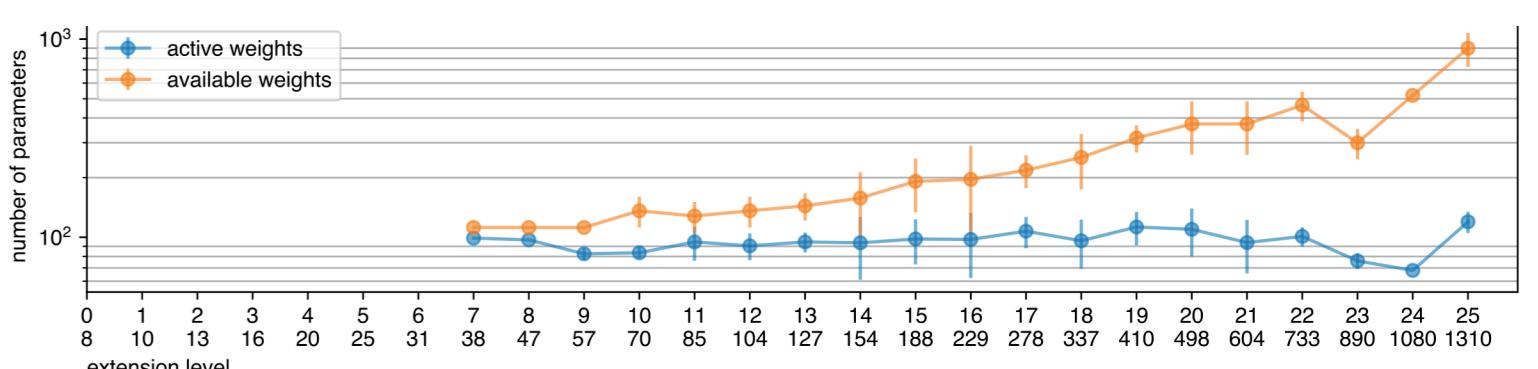
active, inactive and zombie parameters



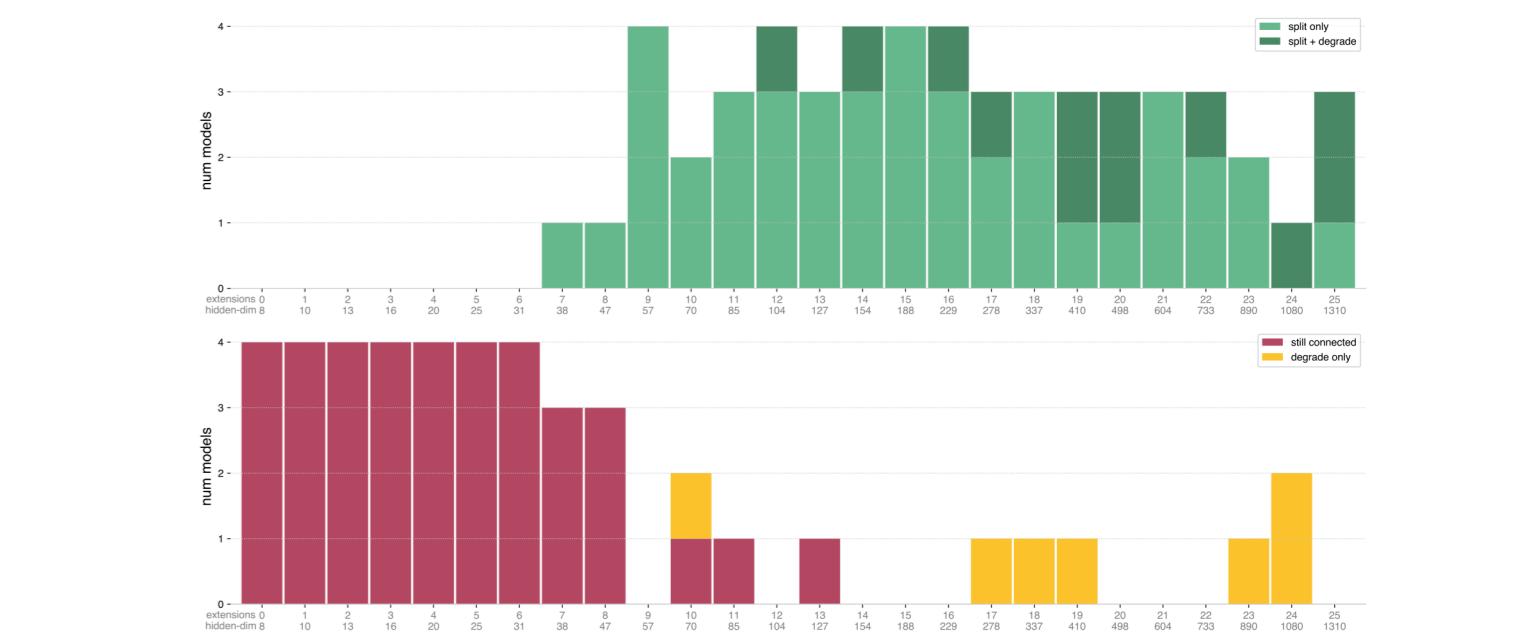
extending model with fixed pruning rate



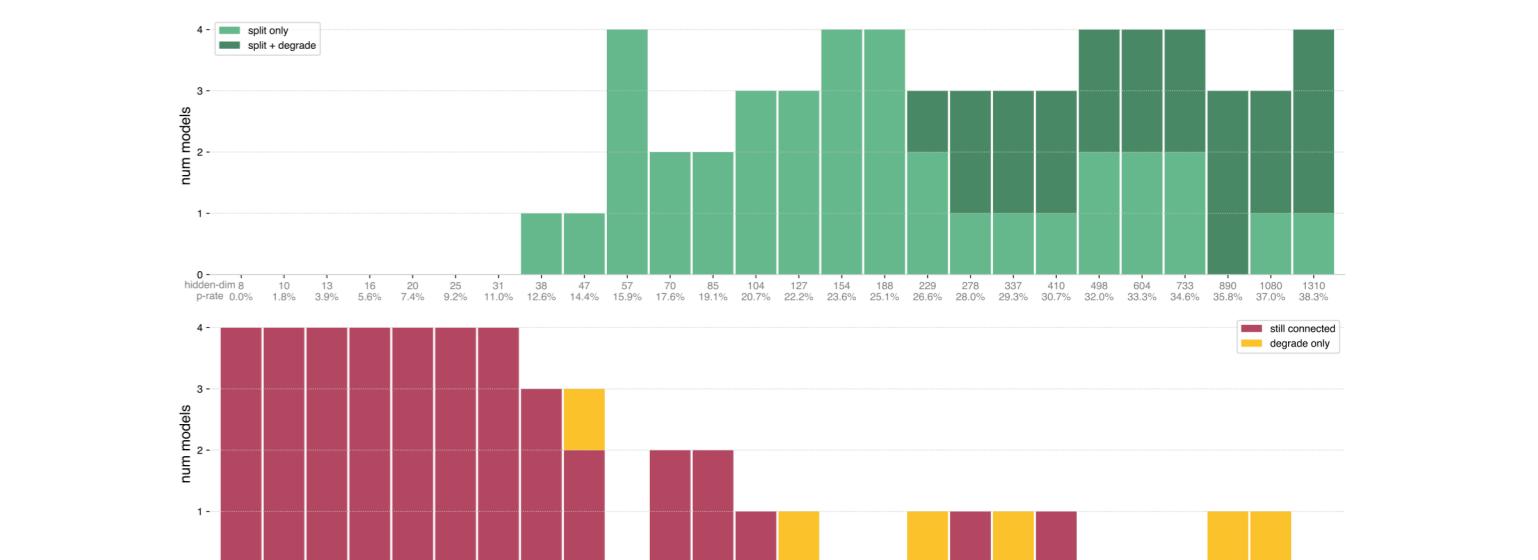
Experiment: Extending model with pruning rate of 0.32 from 8-1310 hidden neurons (25 extension levels)



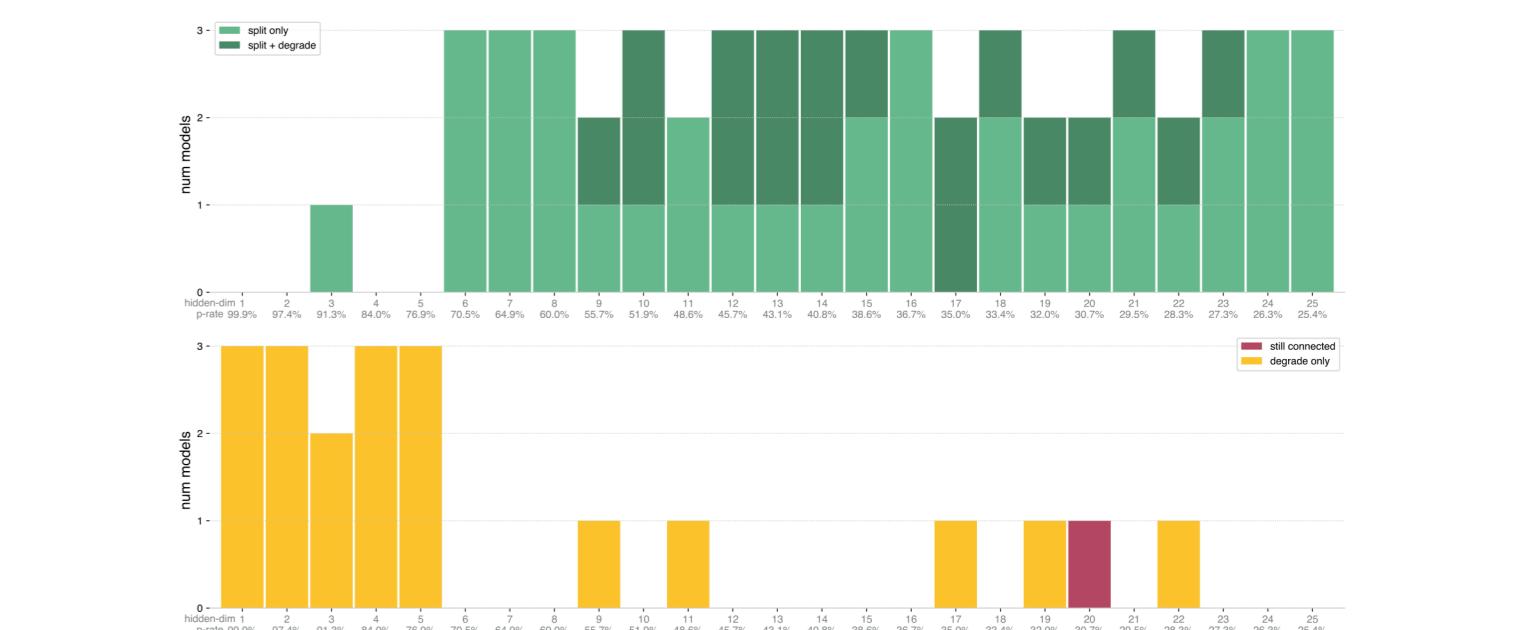
count how many of the networks split (multiple seeds)



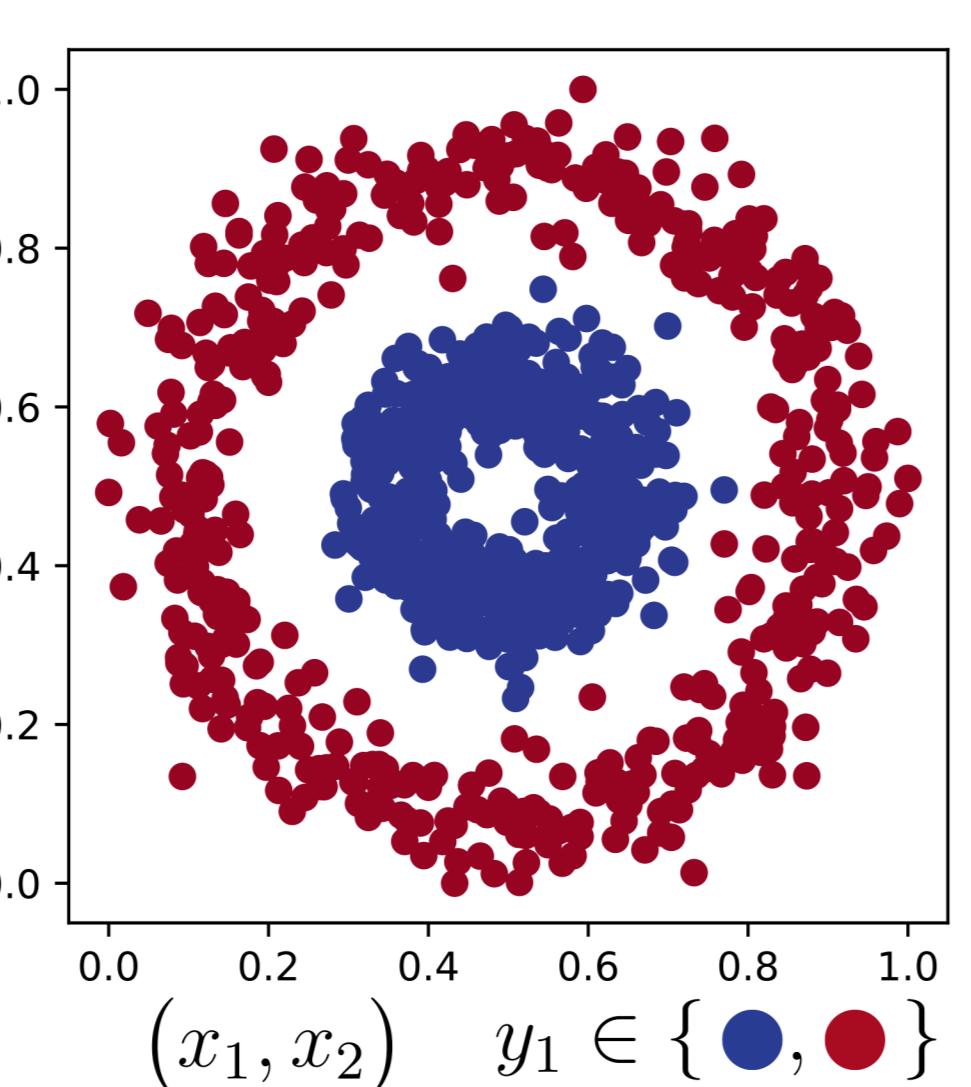
Experiment: same model sizes but 20 pruning levels for each



Experiment: how many pruning levels are enough pruning levels:[0,20], same target, model (4,410,410,2)



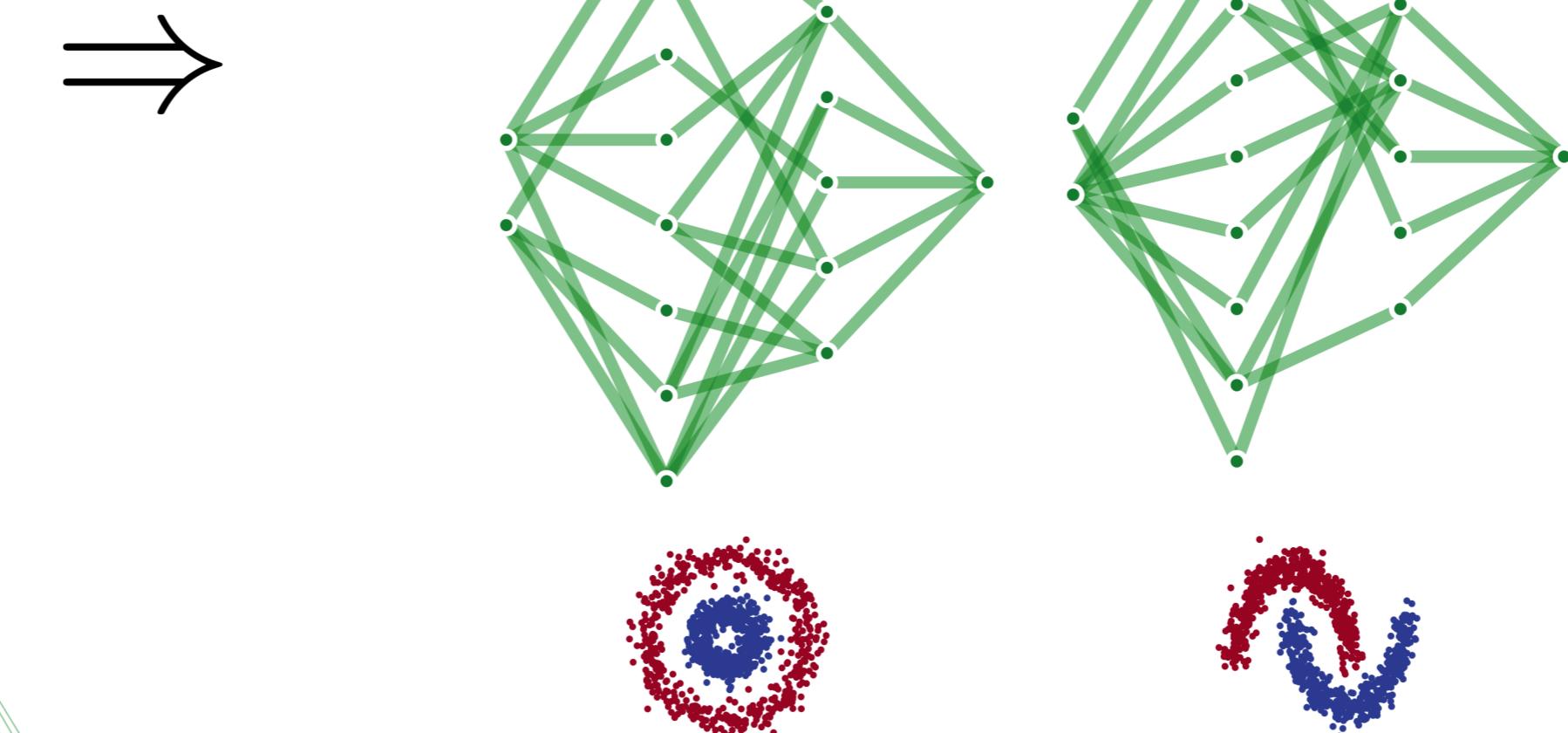
independent tasks



iterative magnitude pruning

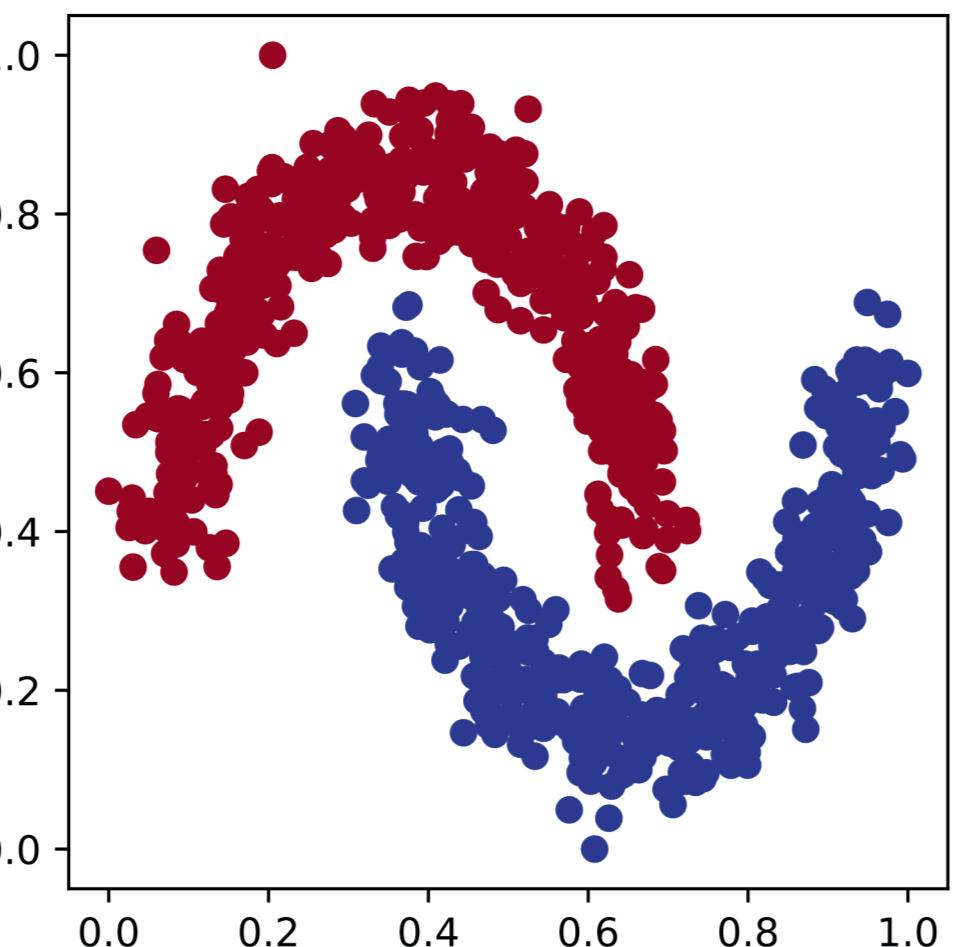
1. train
2. prune
3. reset
4. repeat

independent lottery tickets



concatenate \Rightarrow

(x_3, x_4) $y_2 \in \{\bullet, \bullet\}$



$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$

$\begin{pmatrix} p_{y_1} = \bullet \\ p_{y_2} = \bullet \end{pmatrix}$