INTRO TO MACHINE LEARNING IN NUCLEAR PHYSICS

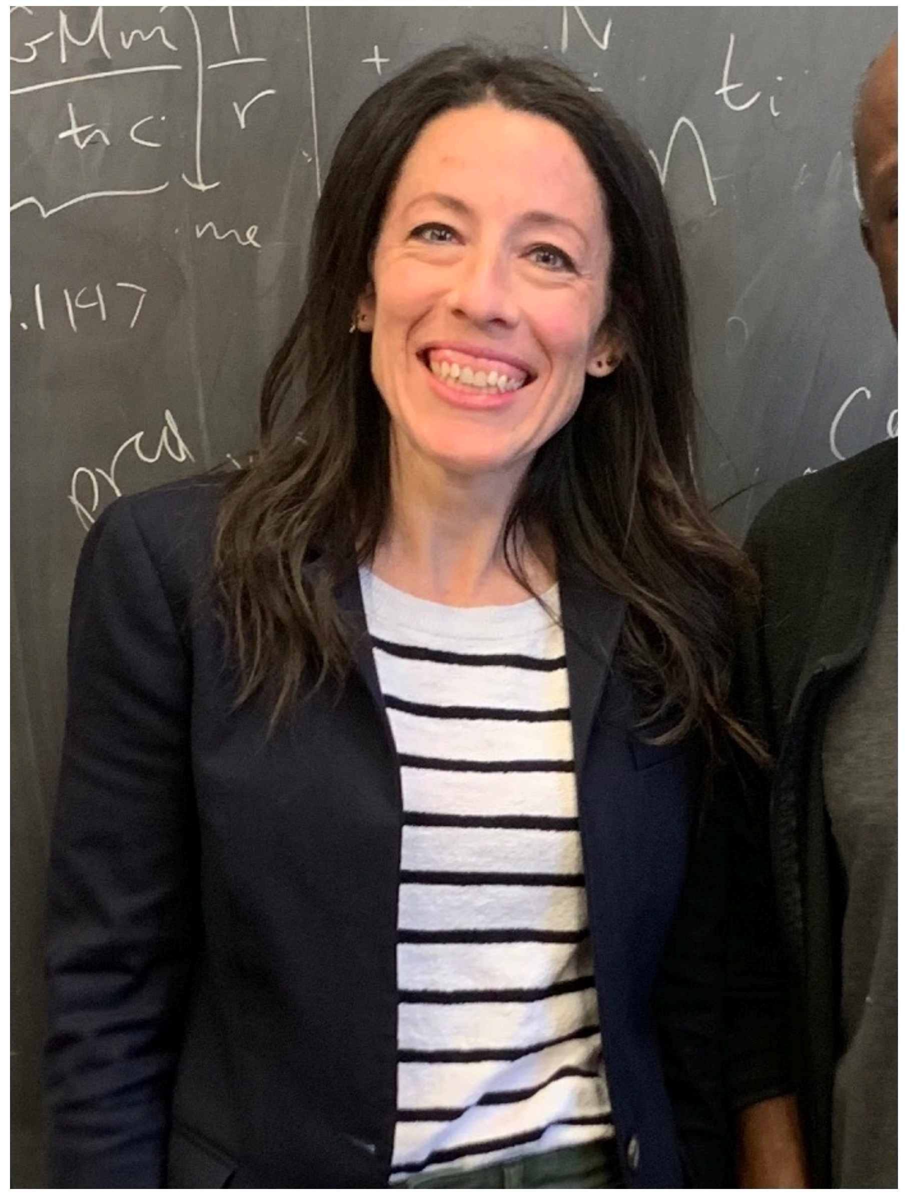
MICHELLE KUCHERA
DAVIDSON COLLEGE

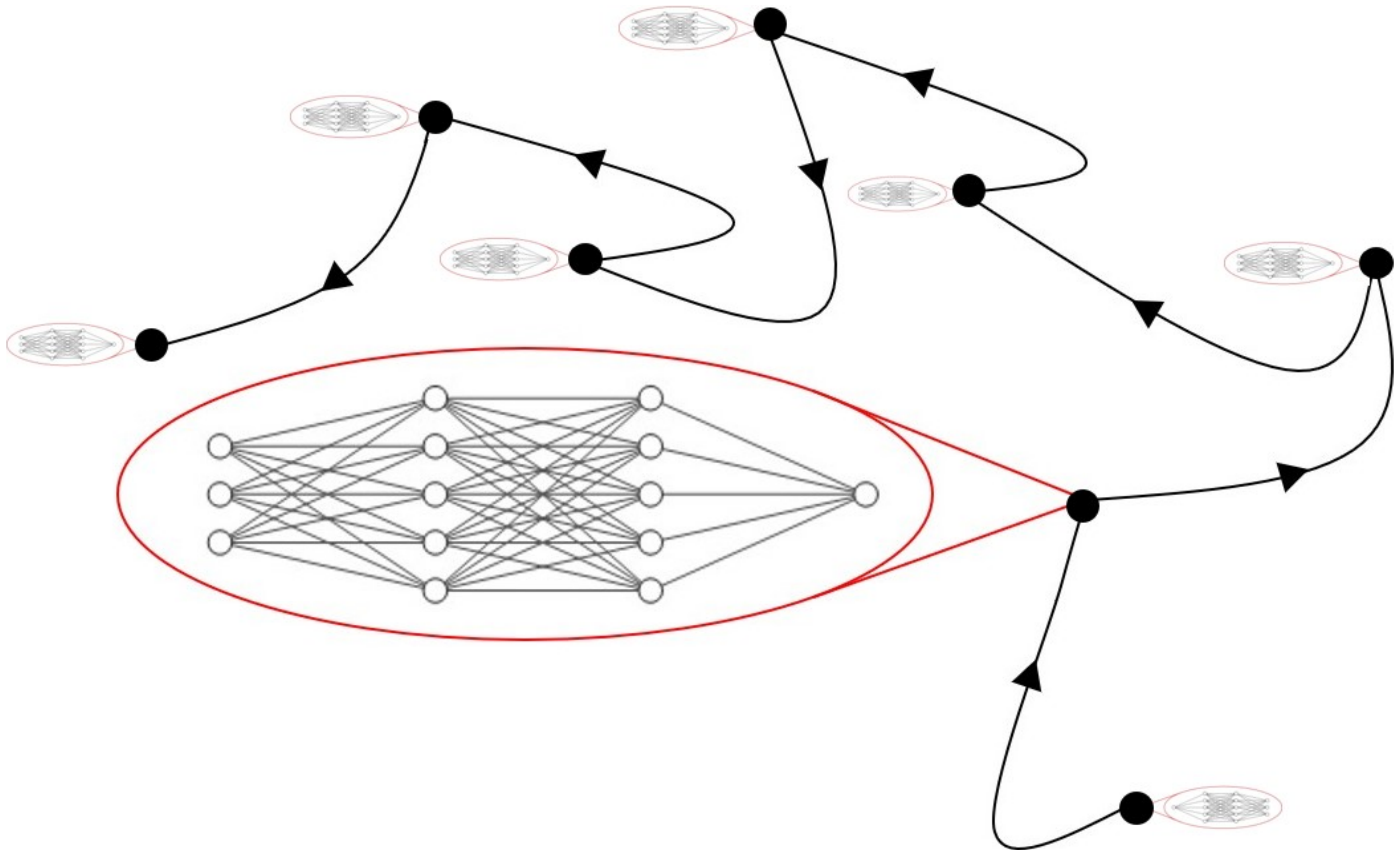
JOINT ICTP-IAEA SCHOOL ON AI FOR NUCLEAR, PLASMA, AND FUSION SCIENCE

ICTP
22 MAY 2023

MICHELLE KUCHERA

B.S., M.S. PHYSICS
M.S., PH.D. COMPUTATIONAL SCIENCE





SESSION 1 TOPICS

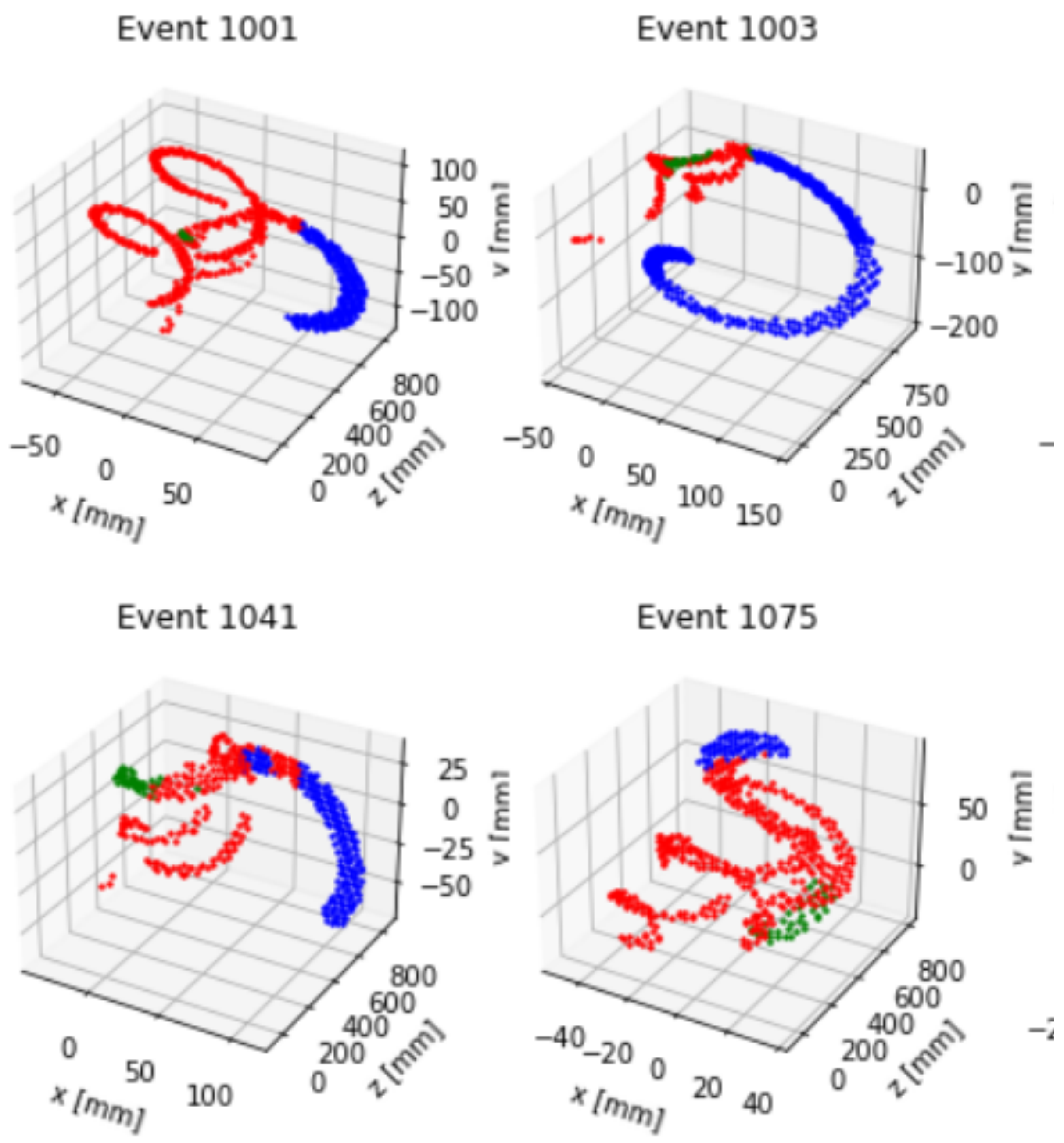
- Nuclear physics motivation
- Gradient descent on graphs
- Neural Networks

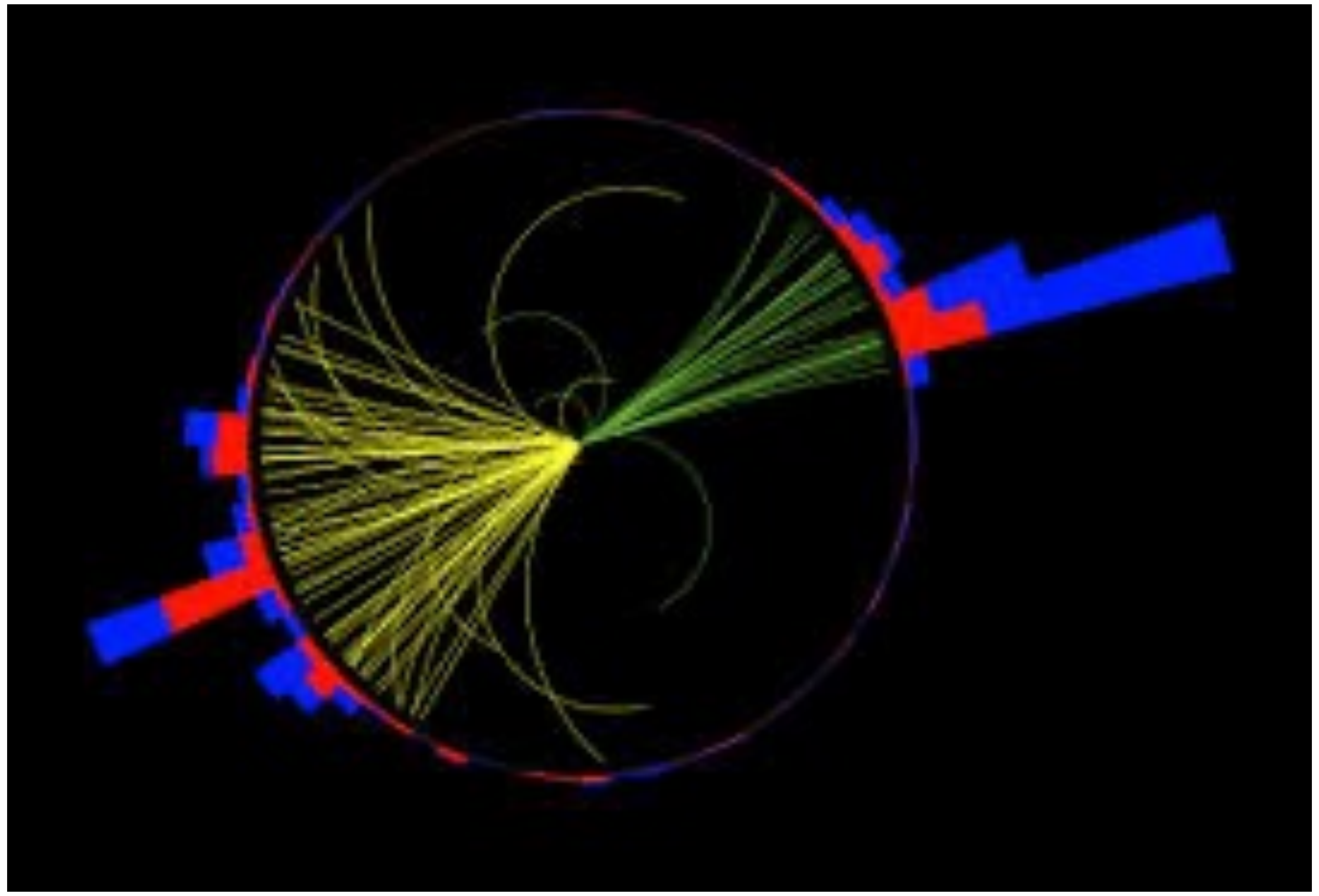
MICHELLE KUCHERA
DAVIDSON COLLEGE

JOINT ICTP-IAEA SCHOOL ON AI FOR NUCLEAR, PLASMA, AND FUSION SCIENCE

ICTP
22 MAY 2023

EXPERIMENTAL DATA

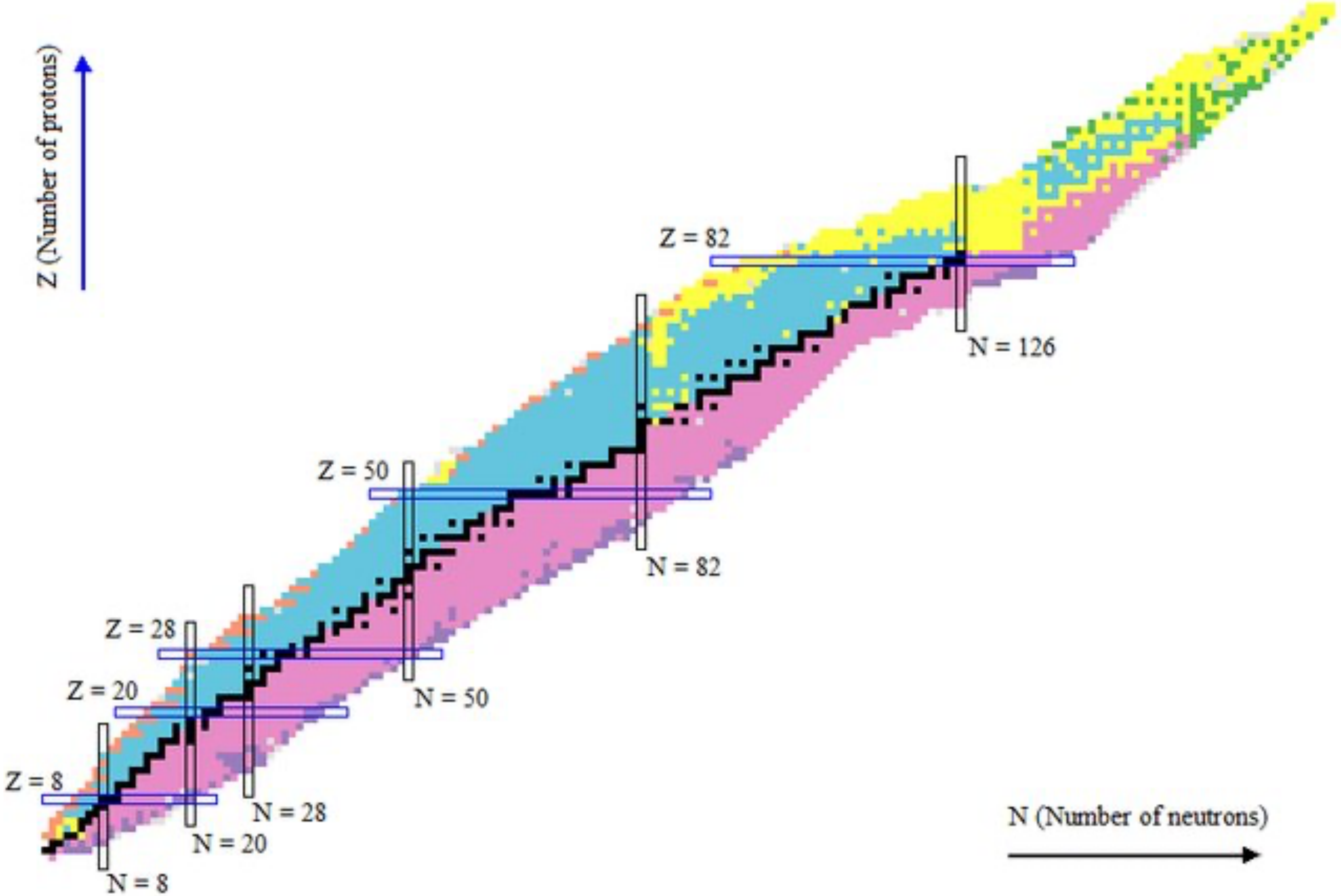




AT-TPC

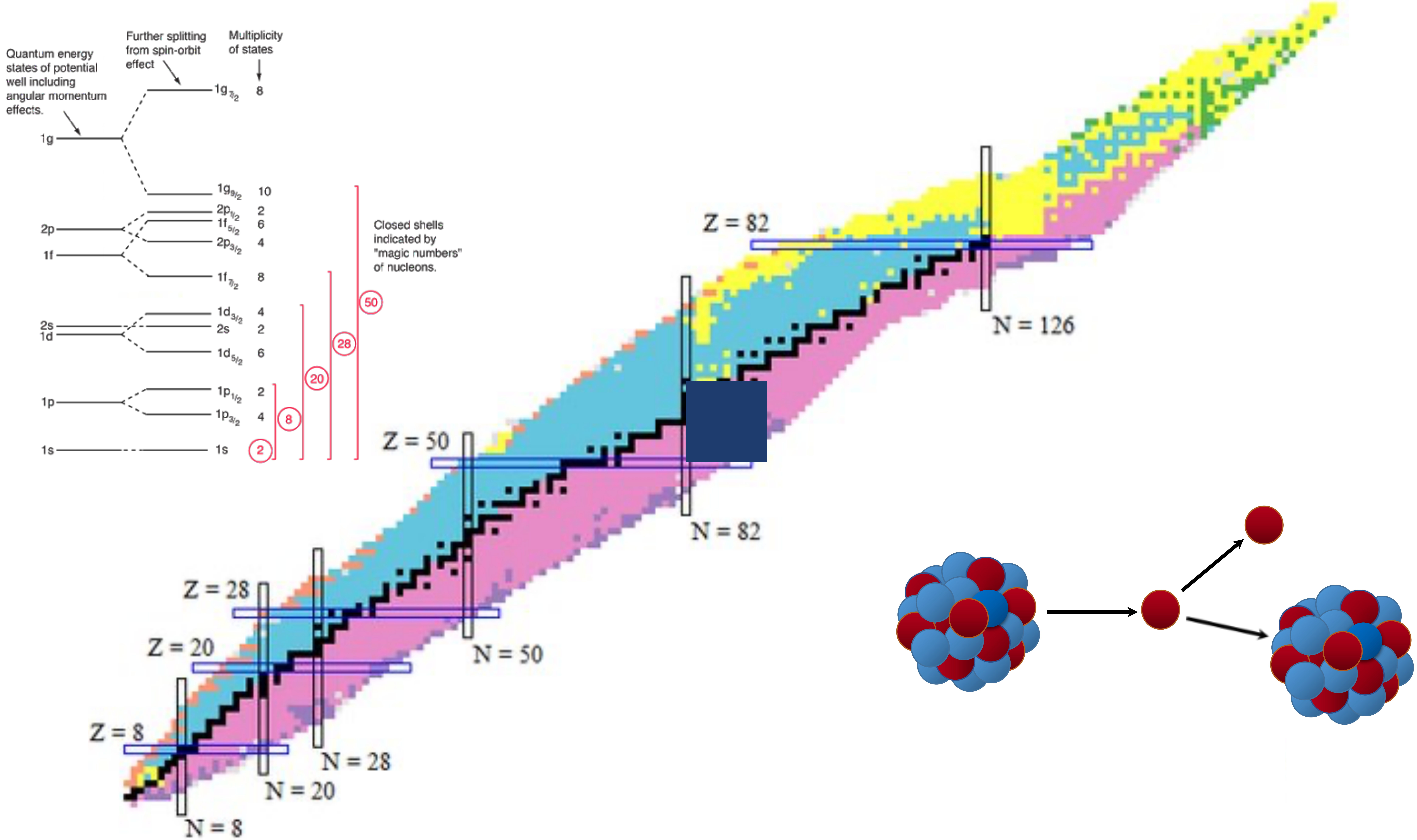
CLAS 12

CMS

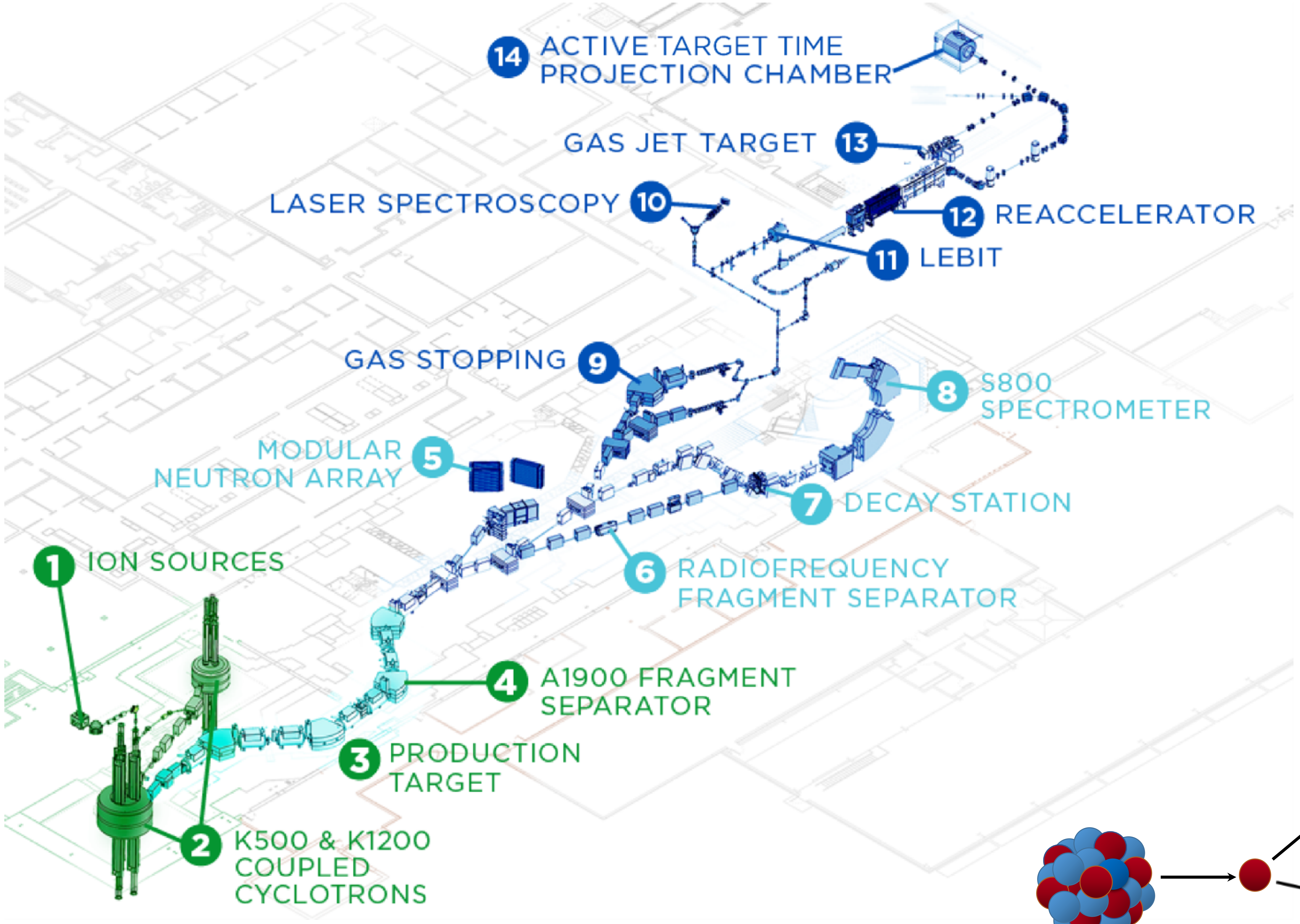


Z (Number of protons)
Z = 82
N = 126
Z = 50
N = 82
Z = 28
Z = 20
N = 50
Z = 8
N = 28
N = 20
N = 8
N (Number of neutrons)

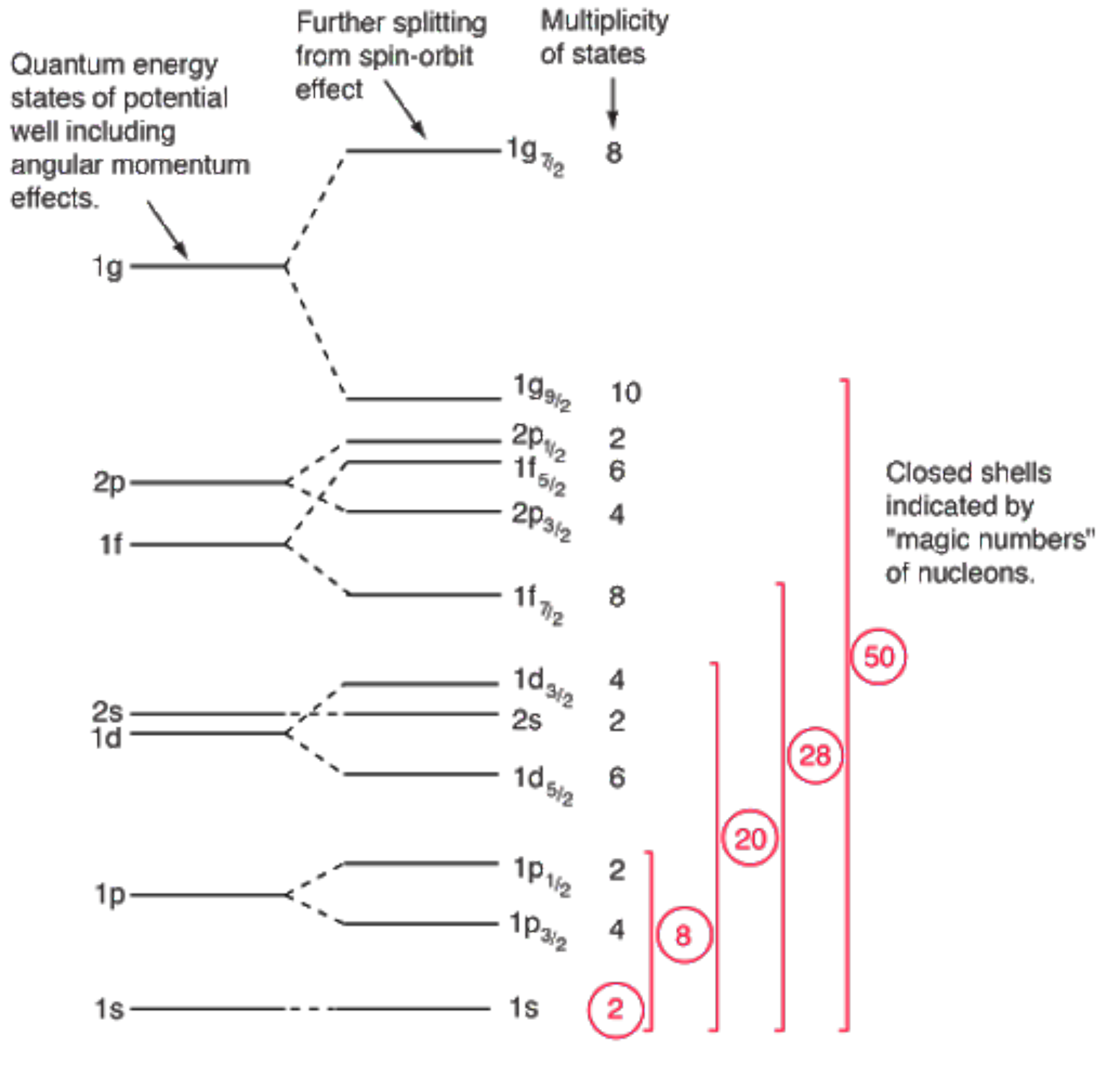
FRIB

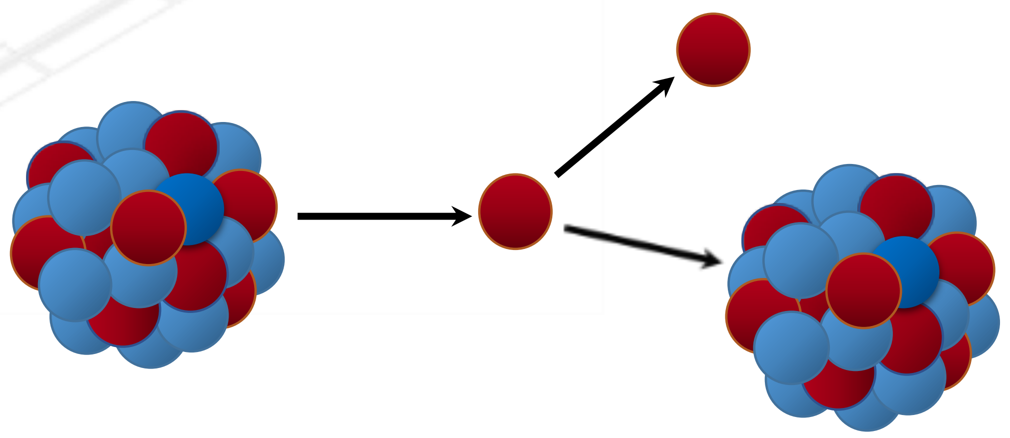


Quantum energy states of potential well including angular momentum effects.
Further splitting from spin-orbit effect
Multiplicity of states
1g
$1g_{7/2}$ 8
$1g_{9/2}$ 10
2p
$2p_{1/2}$ 2
1f
$1f_{5/2}$ 6
$2p_{3/2}$ 4
$1f_{7/2}$ 8
2s
1d
$1d_{3/2}$ 4
2s 2
$1d_{5/2}$ 6
1p
$1p_{1/2}$ 2
$1p_{3/2}$ 4
1s
1s
2
8
20
28
50
Closed shells indicated by "magic numbers" of nucleons.
Z = 82
N = 126
Z = 50
N = 82
Z = 28
Z = 20
N = 50
Z = 8
N = 28
N = 20
N = 8



ALPhA
DAVIDSON COLLEGE



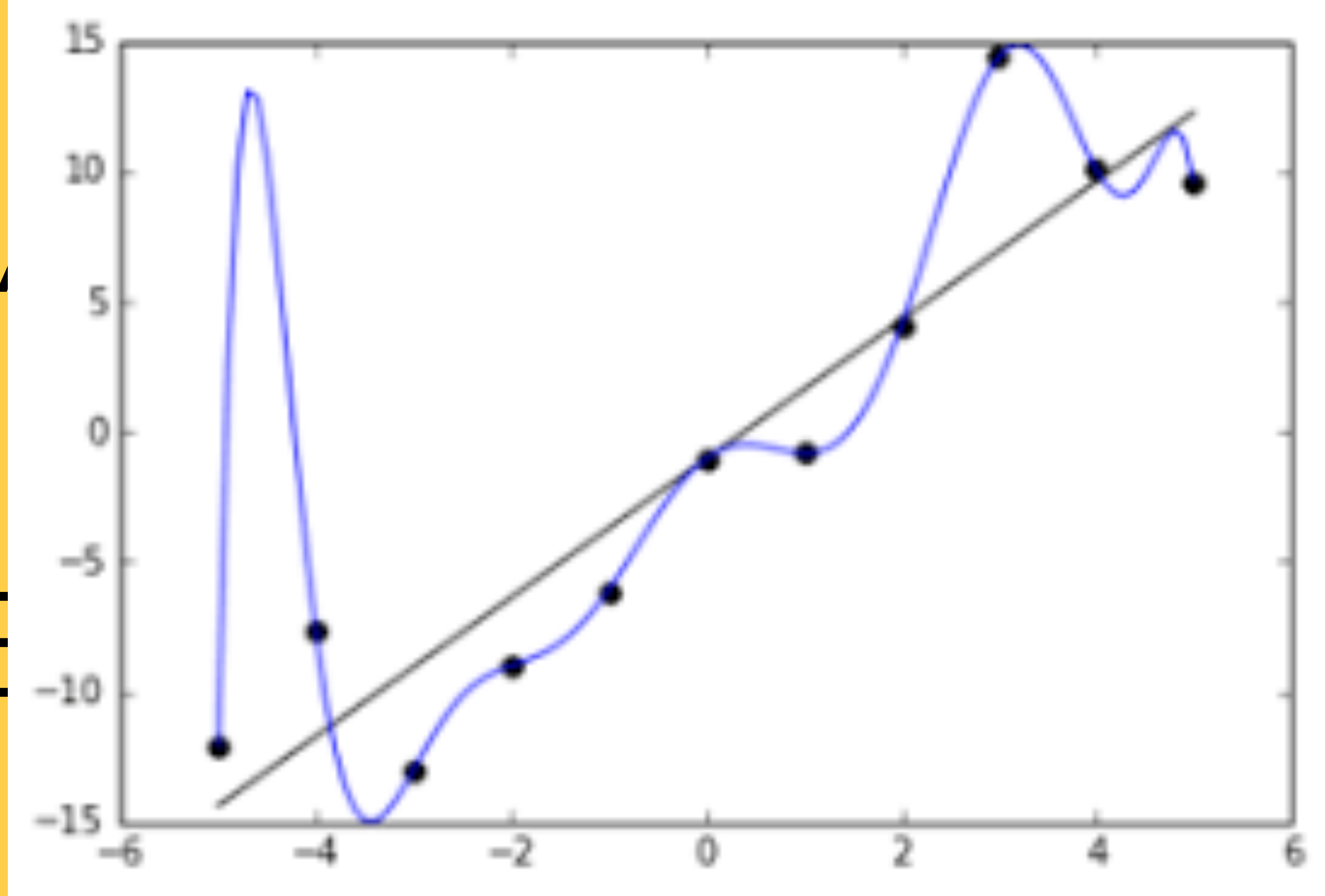


FRIB

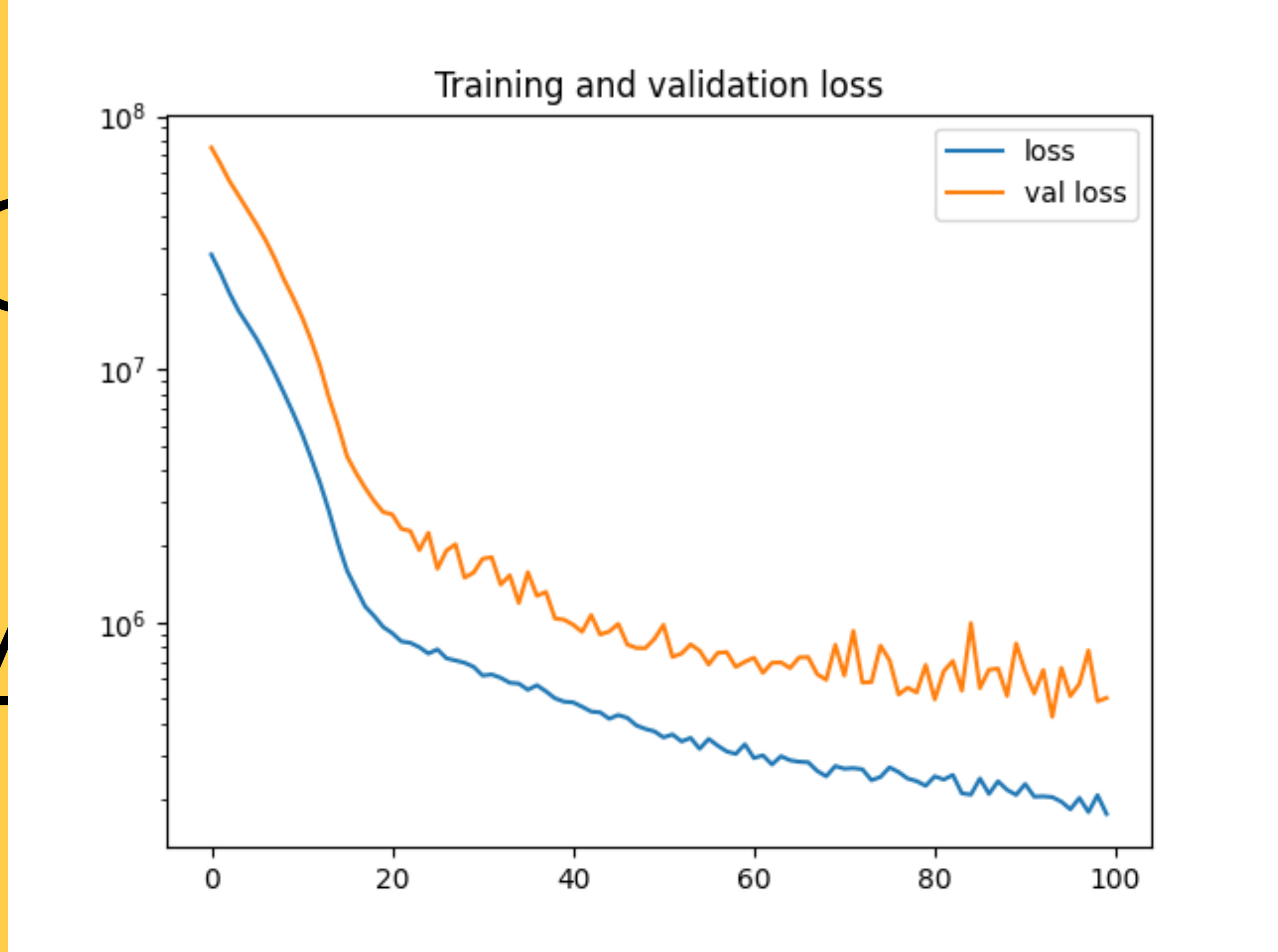
NSCL

MA ING

LE DA

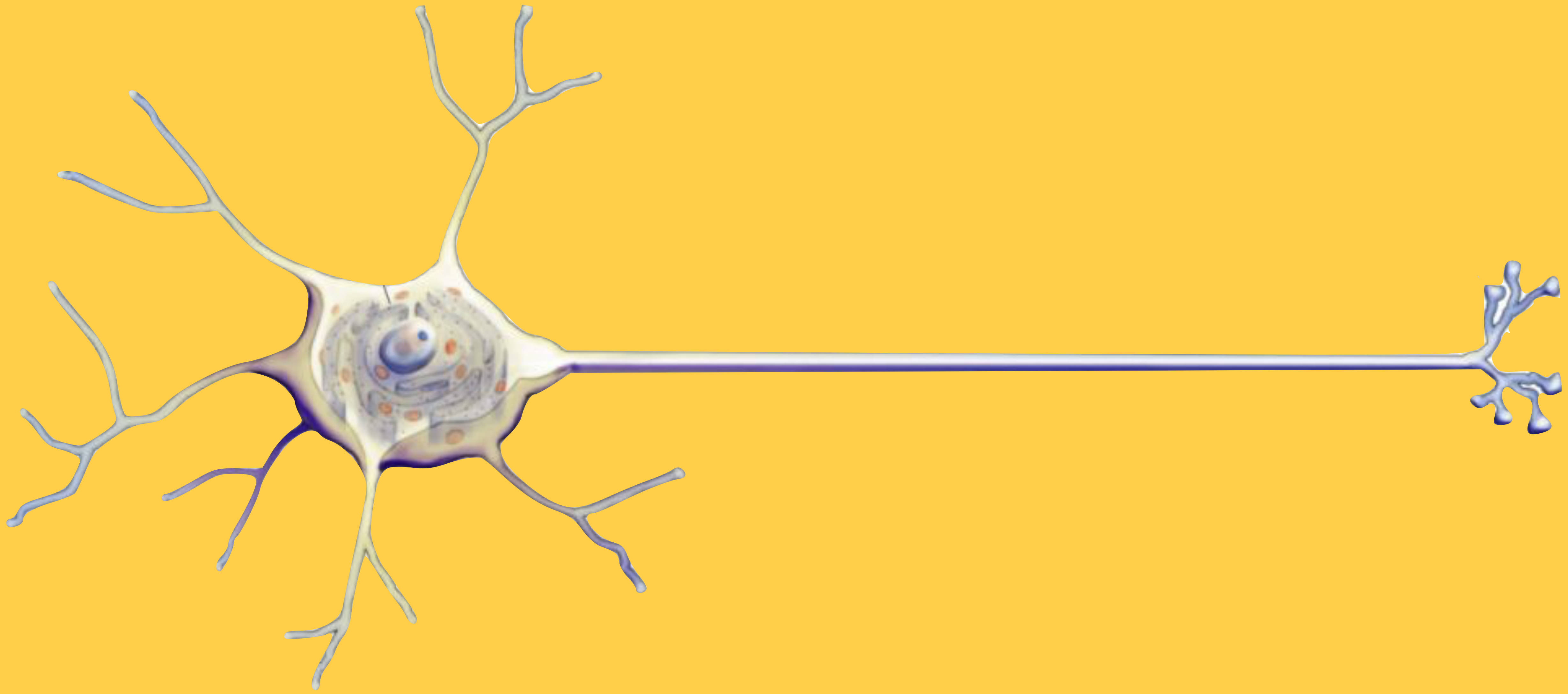


15
10
5
0
−5
−10
−15
−6
−4
−2
0
2
4
6



Training and validation loss
loss
val loss
10^8
10^7
10^6
0
20
40
60
80
100

NEURON



MATHEMATICS

ELSEVIER

Neural Networks
Volume 4, Issue 2, 1991, Pages 251-257

Approximation capabilities of multilayer feedforward networks

Kurt Hornik

Show more

Share Cite

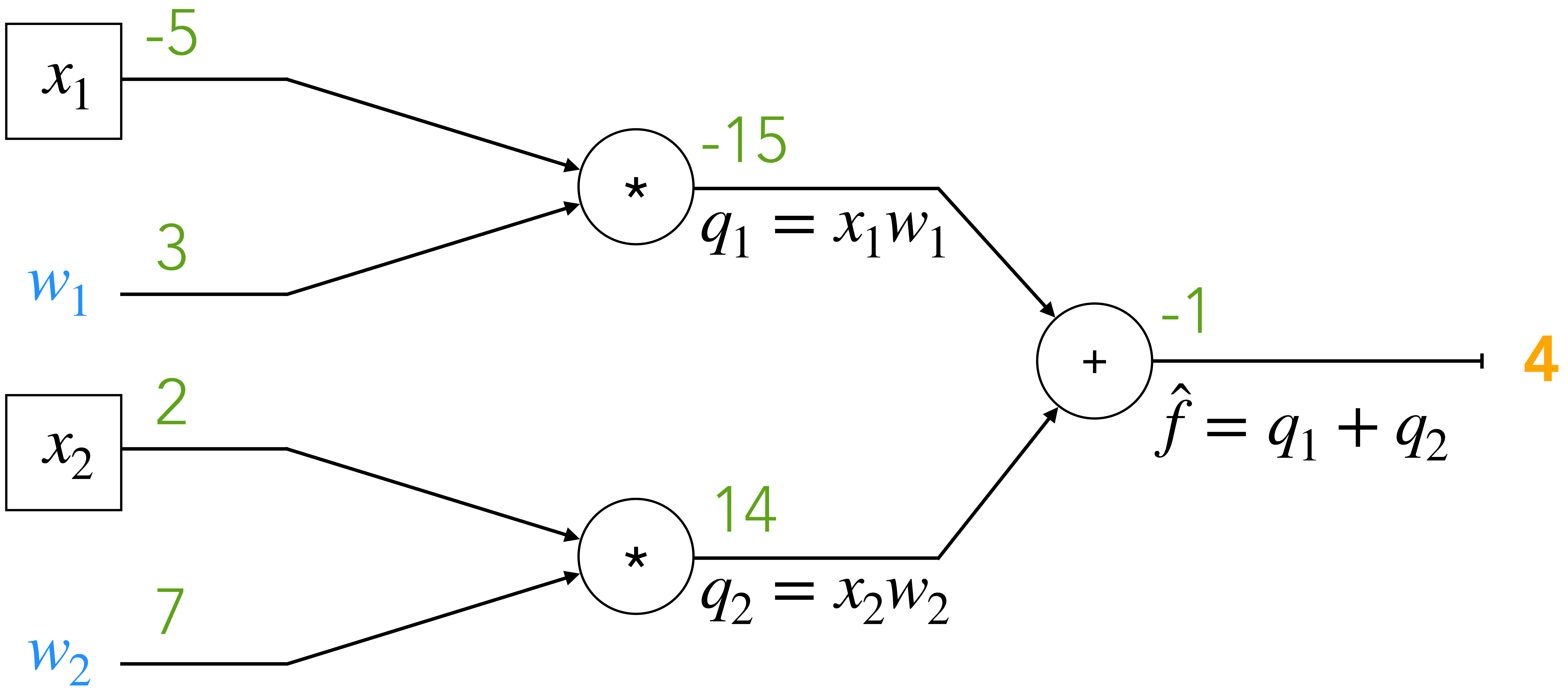
https://doi.org/10.1016/0893-6080(91)90009-T

Abstract

We show that standard multilayer feedforward networks with as few as a single hidden layer and arbitrary bounded and nonconstant activation function are universal approximators with respect to $L^p(\mu)$ performance criteria, for arbitrary finite input environment measures μ, provided only that sufficiently many hidden units are available. If the activation function is continuous, bounded and nonconstant, then continuous mappings can be learned uniformly over compact input sets. We also give very general conditions ensuring that networks with sufficiently smooth activation functions are capable of arbitrarily accurate approximation to a function and its derivatives.

MATHEMATICS

COMPUTATIONAL GRAPH

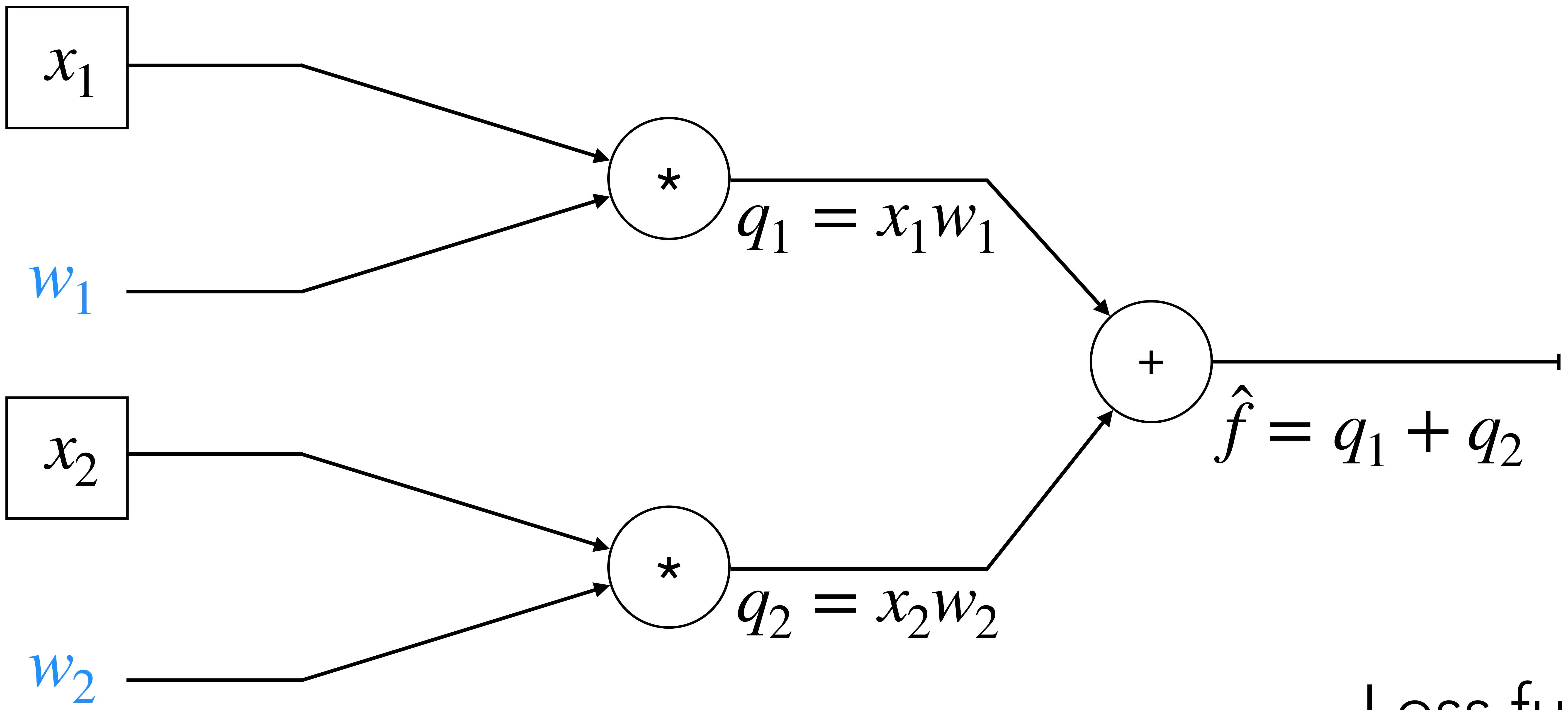


$$\hat{f} = x_1 w_1 + x_2 w_2$$

MACHINE LEARNING

SUPERVISED LEARNING

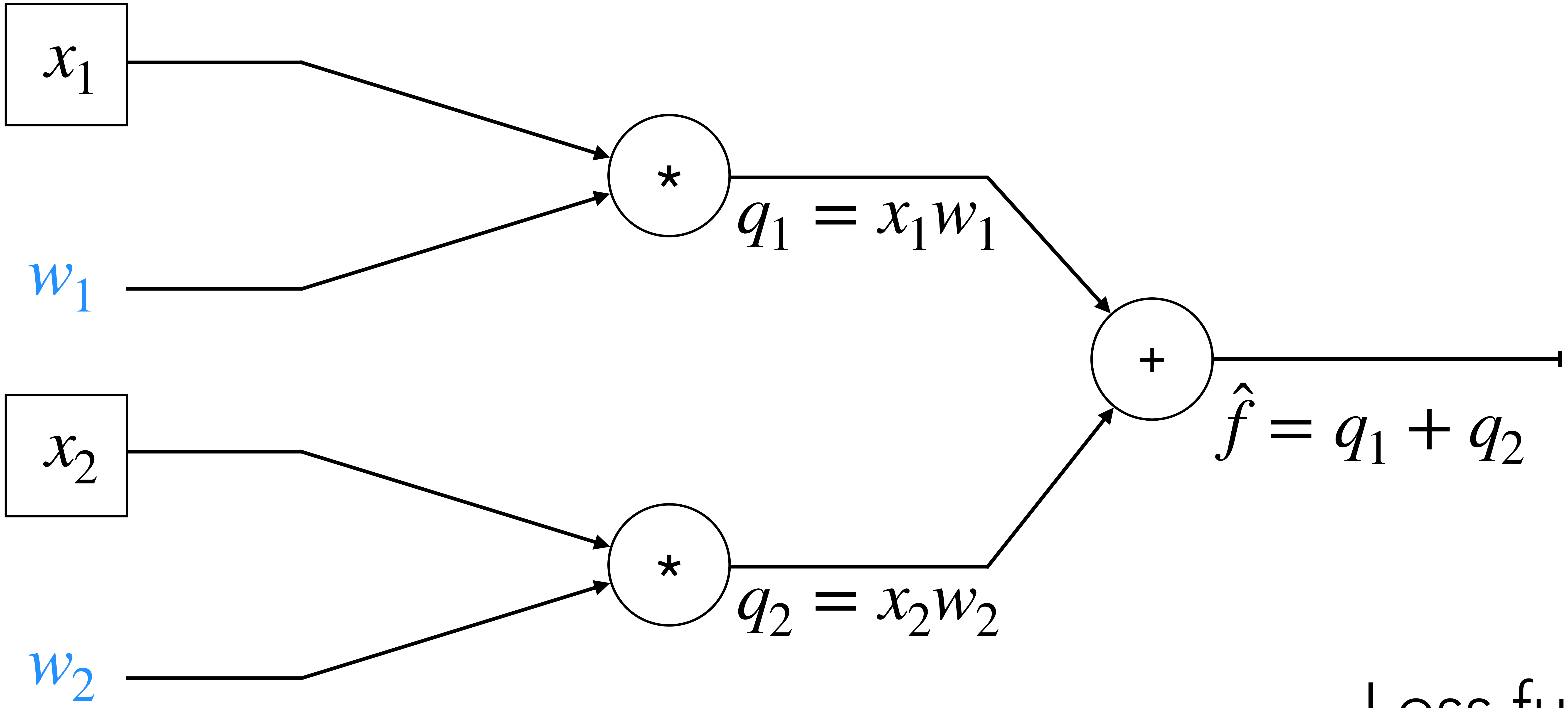
REGRESSION



Loss function

$$\hat{f} = x_1 w_1 + x_2 w_2 \qquad J(w) = \hat{f} - f$$

SUPERVISED LEARNING



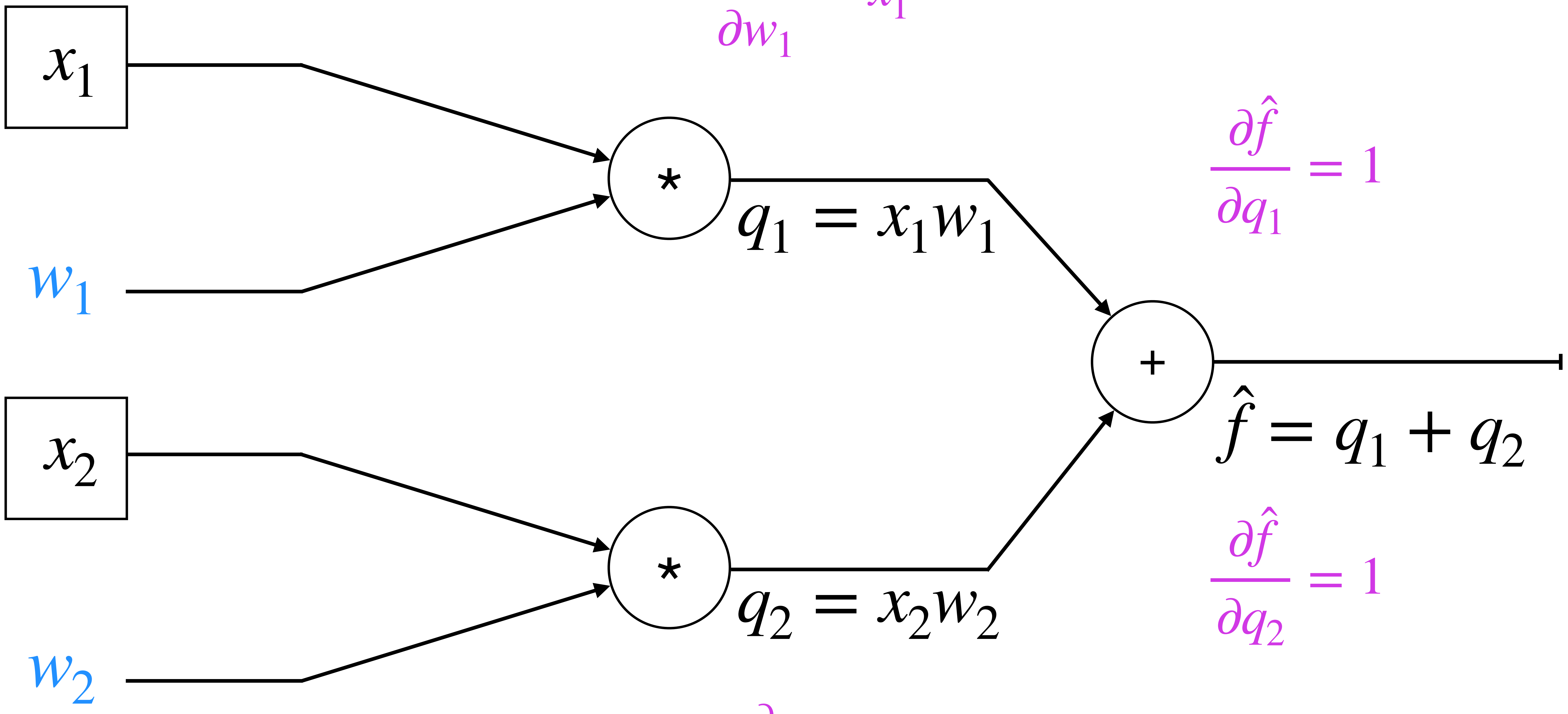
Loss function

$$\hat{f} = x_1 w_1 + x_2 w_2 \qquad J(w) = \hat{f} - f$$

BACKPROPAGATION

$$w_1 = w_1 - \eta * \frac{\partial J}{\partial \hat{f}} \frac{\partial \hat{f}}{\partial q_1} \frac{\partial q_1}{\partial w_1}$$

$$\frac{\partial q_1}{\partial w_1} = x_1$$



$$\frac{\partial \hat{f}}{\partial q_1} = 1$$

$$\frac{\partial J}{\partial \hat{f}} = 1$$

$$\frac{\partial \hat{f}}{\partial q_2} = 1$$

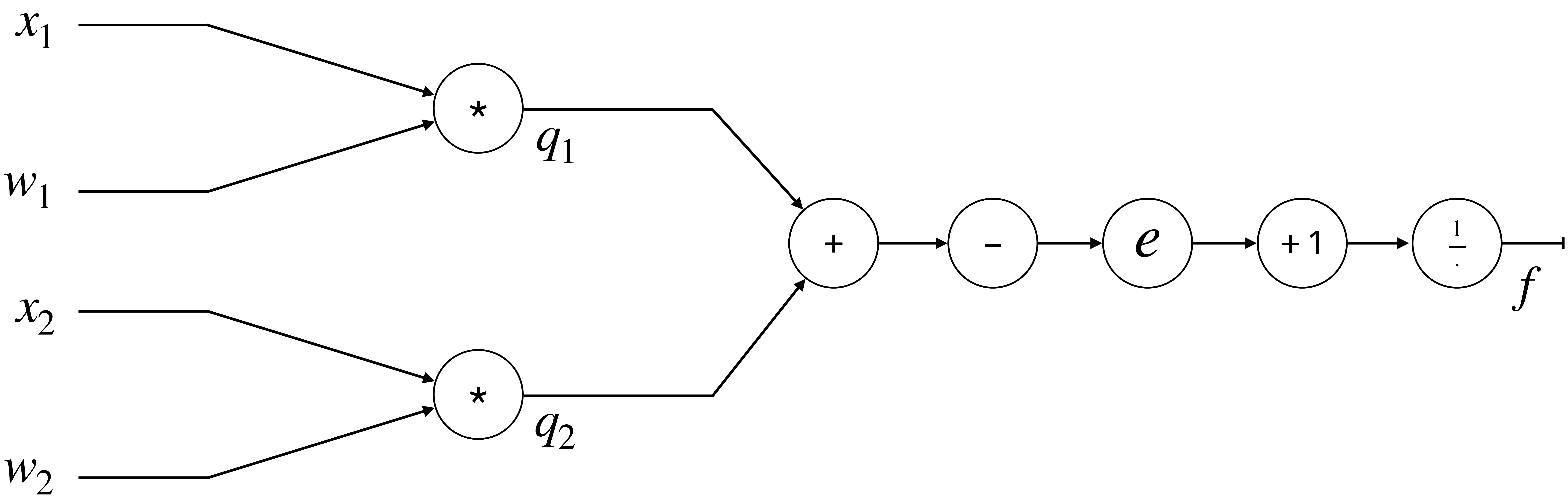
$$\frac{\partial q_2}{\partial w_2} = x_2$$

$$w_2 = w_2 - \eta * \frac{\partial J}{\partial \hat{f}} \frac{\partial \hat{f}}{\partial q_2} \frac{\partial q_2}{\partial w_2}$$

Loss function

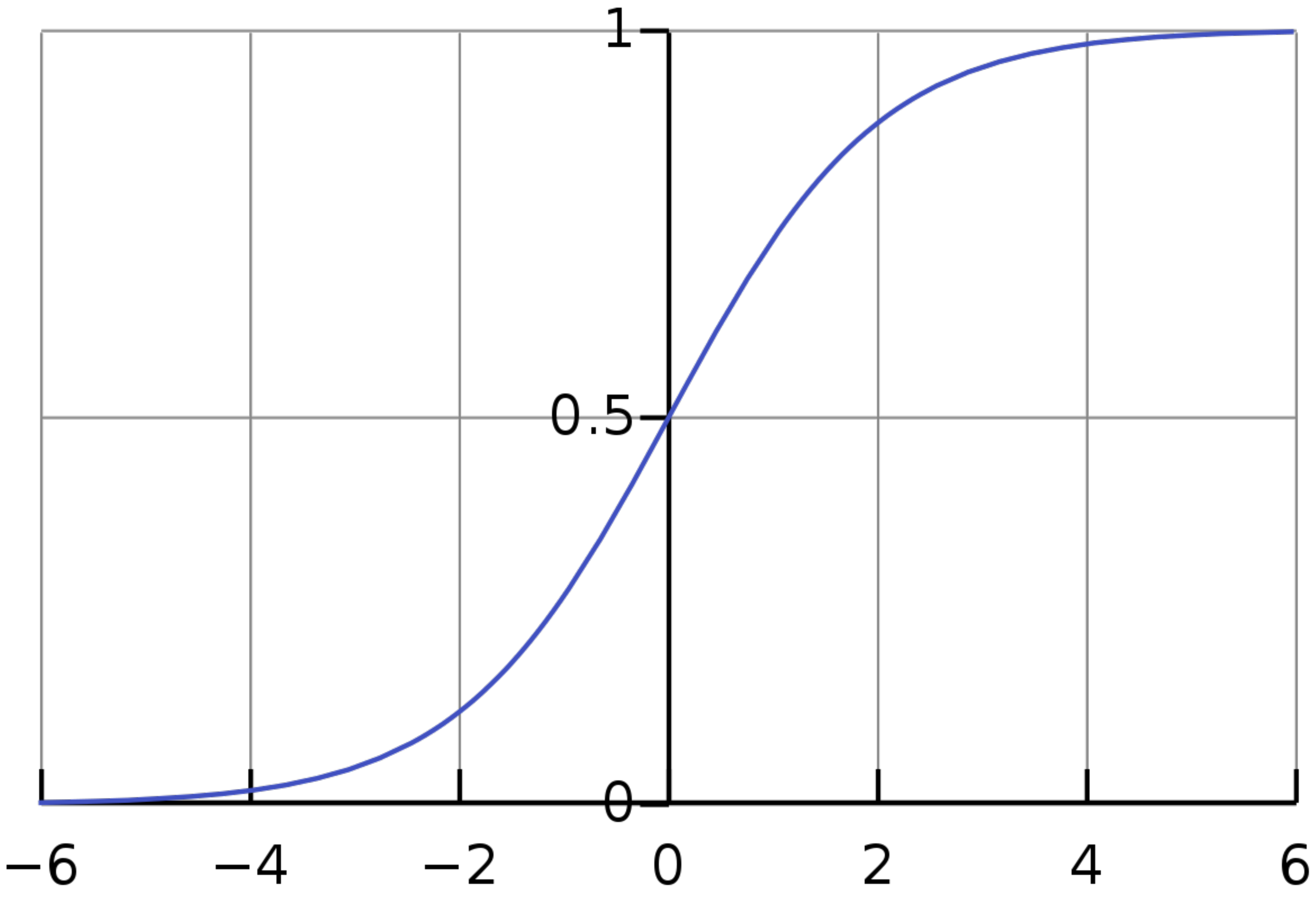
$$J(w) = \hat{f} - f$$

LOGISTIC REGRESSION

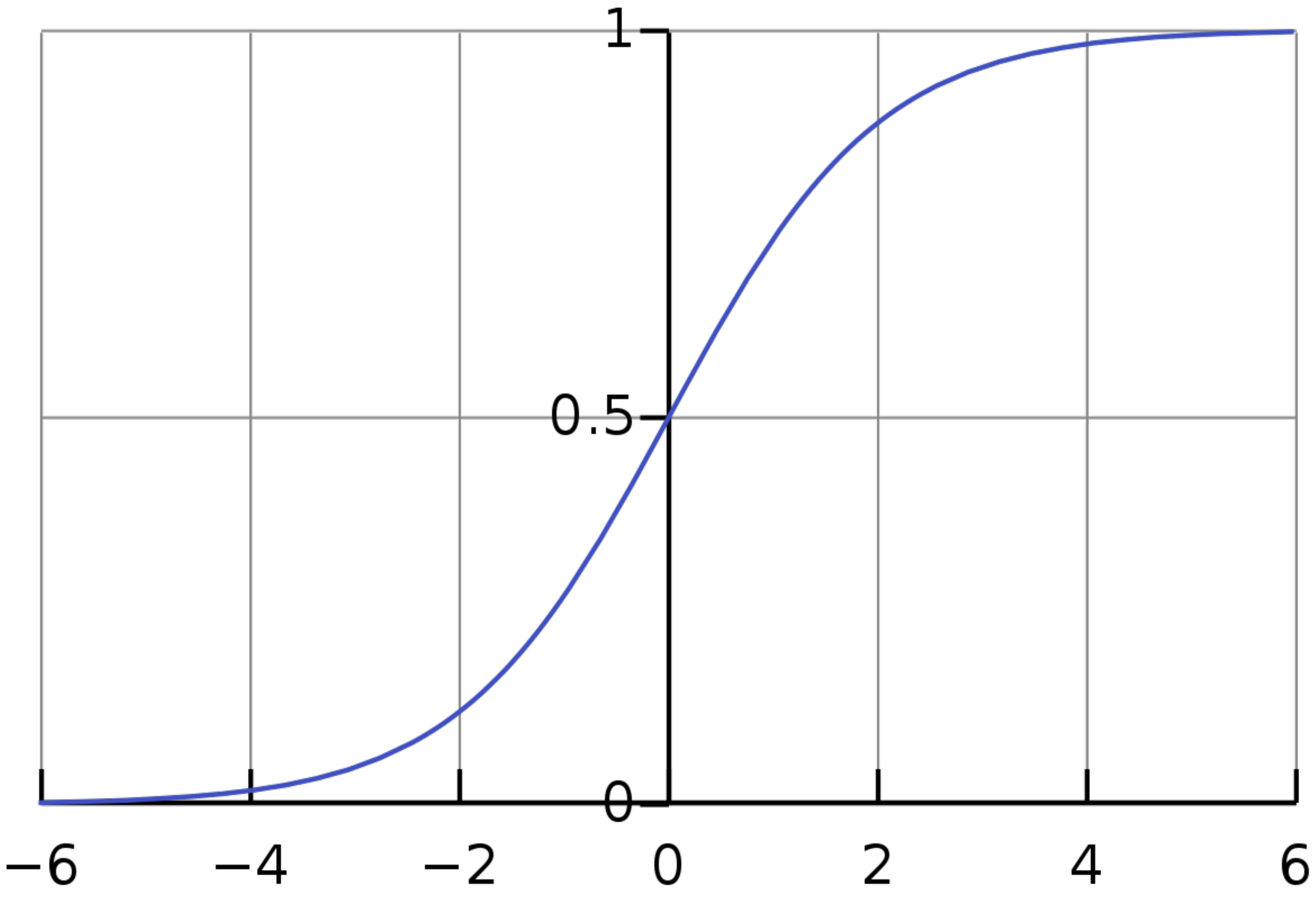


$$f = \frac{1}{1 + e^{-(x_1 w_1 + x_2 w_2)}}$$

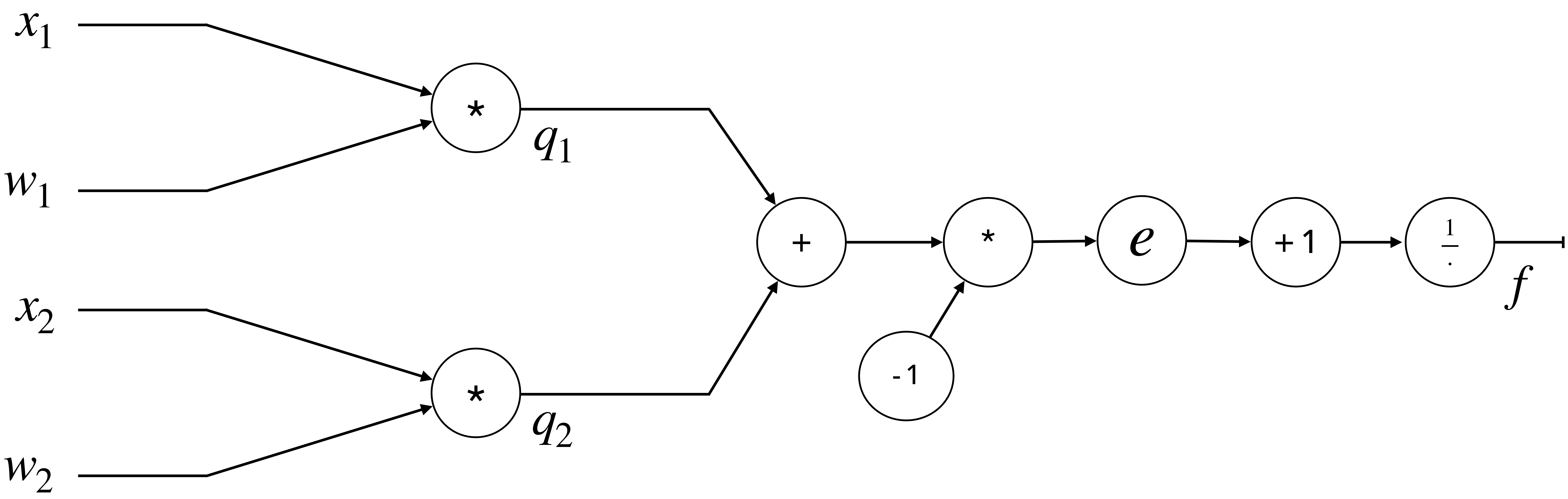
LOGISTIC REGRESSION



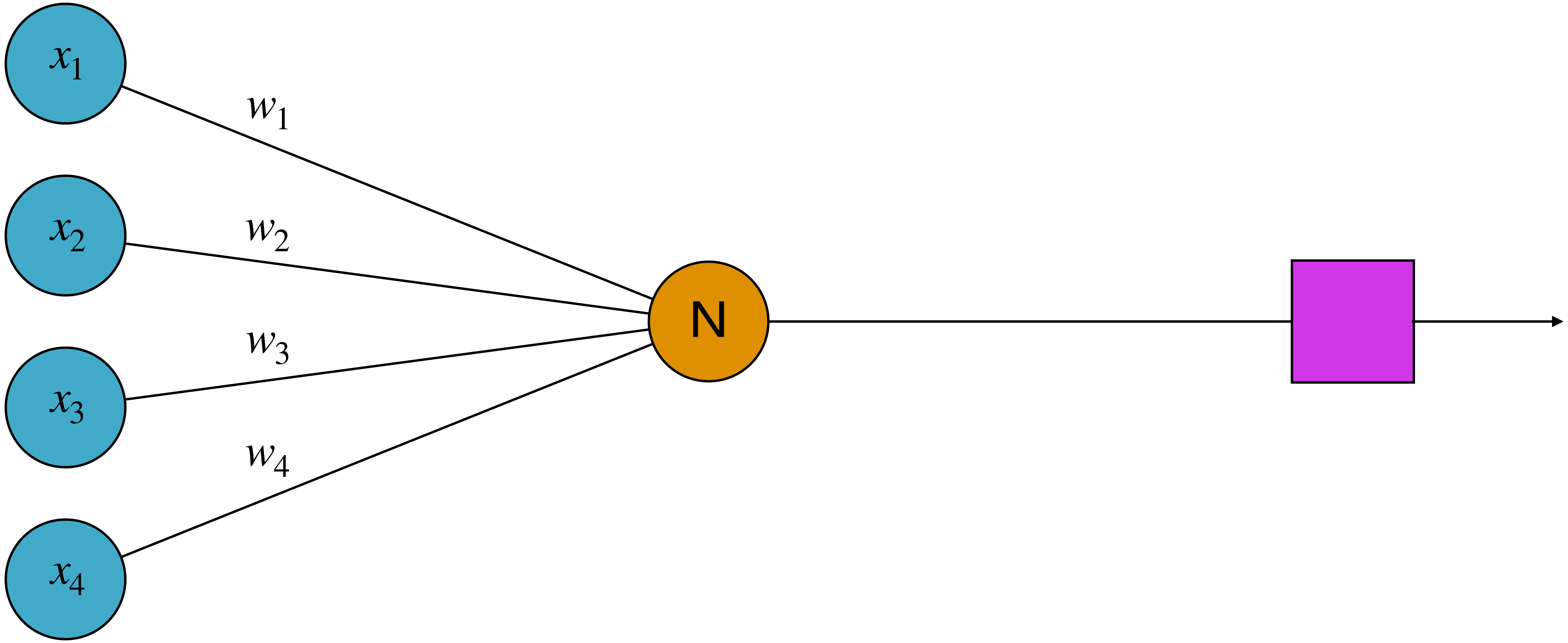
BINARY CLASSIFICATION



LOGISTIC REGRESSION



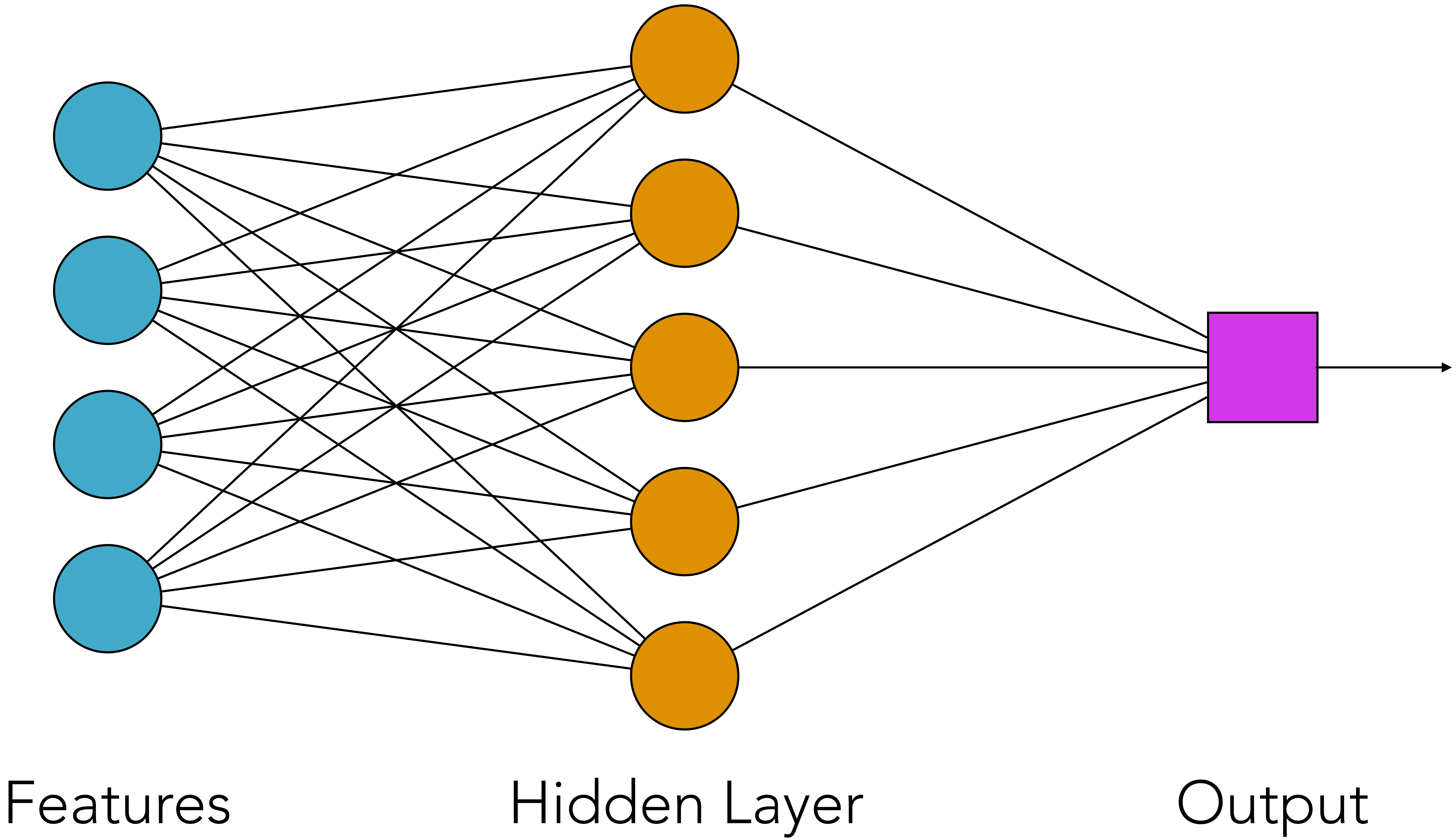
$$f = \frac{1}{1 + e^{-(x_1 w_1 + x_2 w_2)}}$$



Features

Summation
+ Nonlinearity

Output

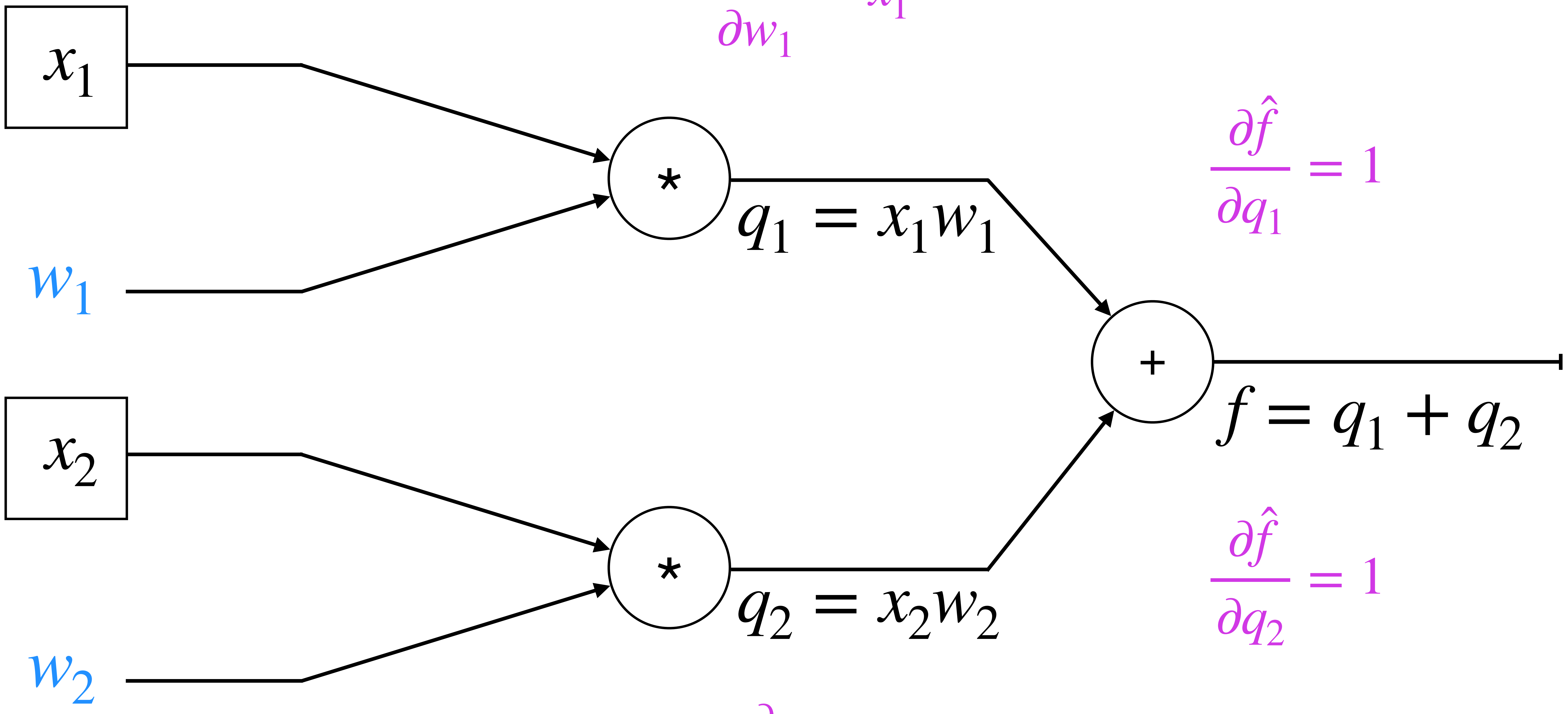


Features
Hidden Layer
Output

BACKPROPAGATION

$$w_1 = w_1 + \eta * \frac{\partial J}{\partial \hat{f}} \frac{\partial \hat{f}}{\partial q_1} \frac{\partial q_1}{\partial w_1}$$

$$\frac{\partial q_1}{\partial w_1} = x_1$$



$$\frac{\partial \hat{f}}{\partial q_1} = 1$$

$$\frac{\partial J}{\partial \hat{f}} = 1$$

$$\frac{\partial \hat{f}}{\partial q_2} = 1$$

$$\frac{\partial q_2}{\partial w_2} = x_2$$

Loss function

$$J(w) = \hat{f} - f$$

$$w_2 = w_2 + \eta * \frac{\partial J}{\partial \hat{f}} \frac{\partial \hat{f}}{\partial q_2} \frac{\partial q_2}{\partial w_2}$$

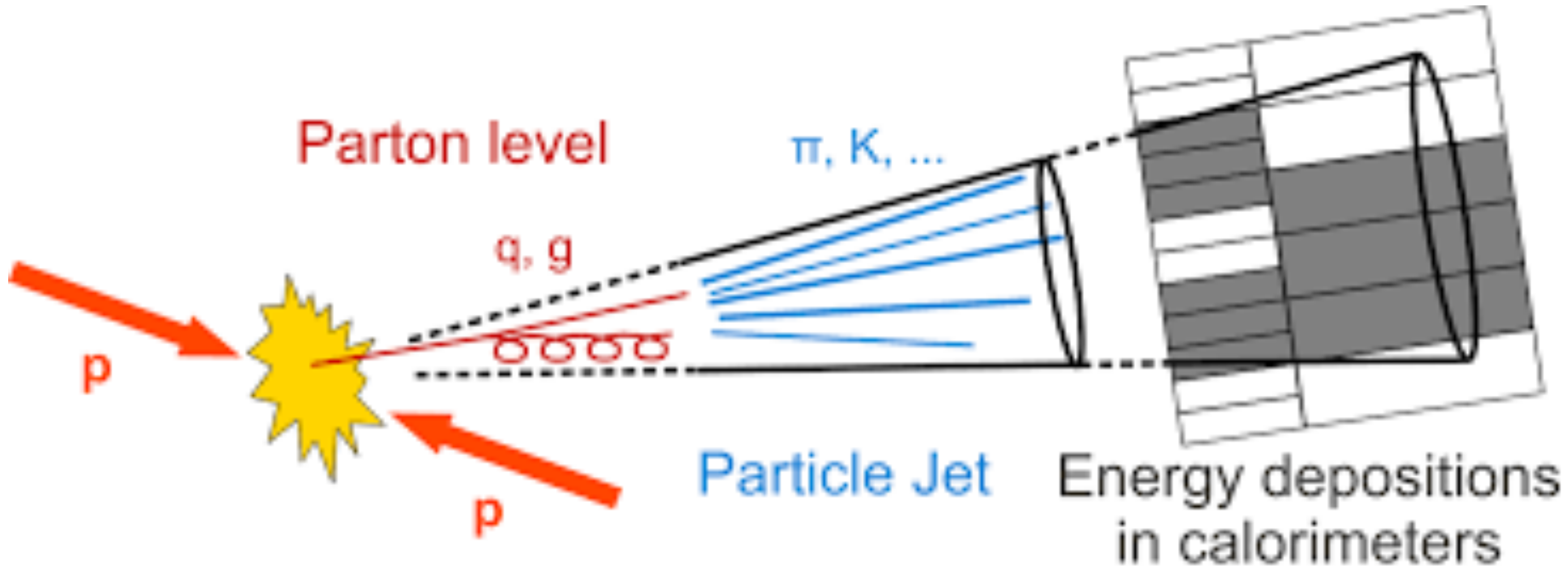
AUTOMATIC DIFFERENTIATION

TensorFlow

PyTorch

Application 1: Can we model stochastic processes?

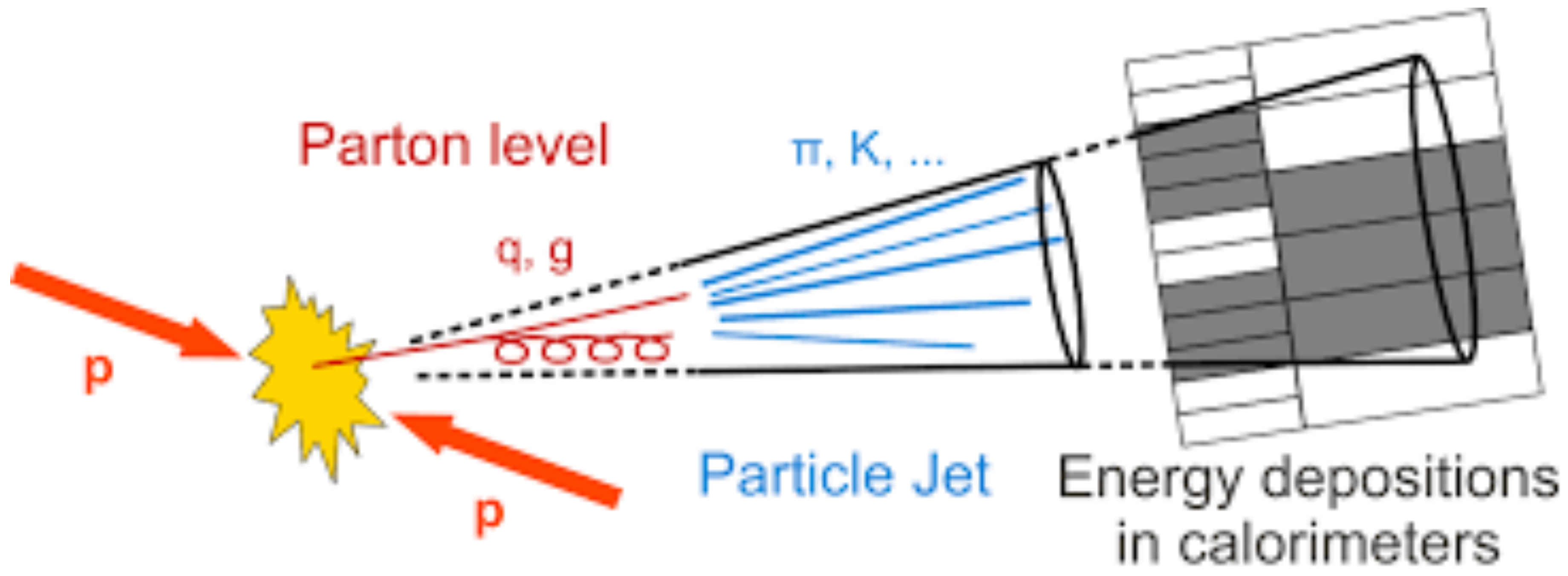
JET SIMULATION AND CORRECTION

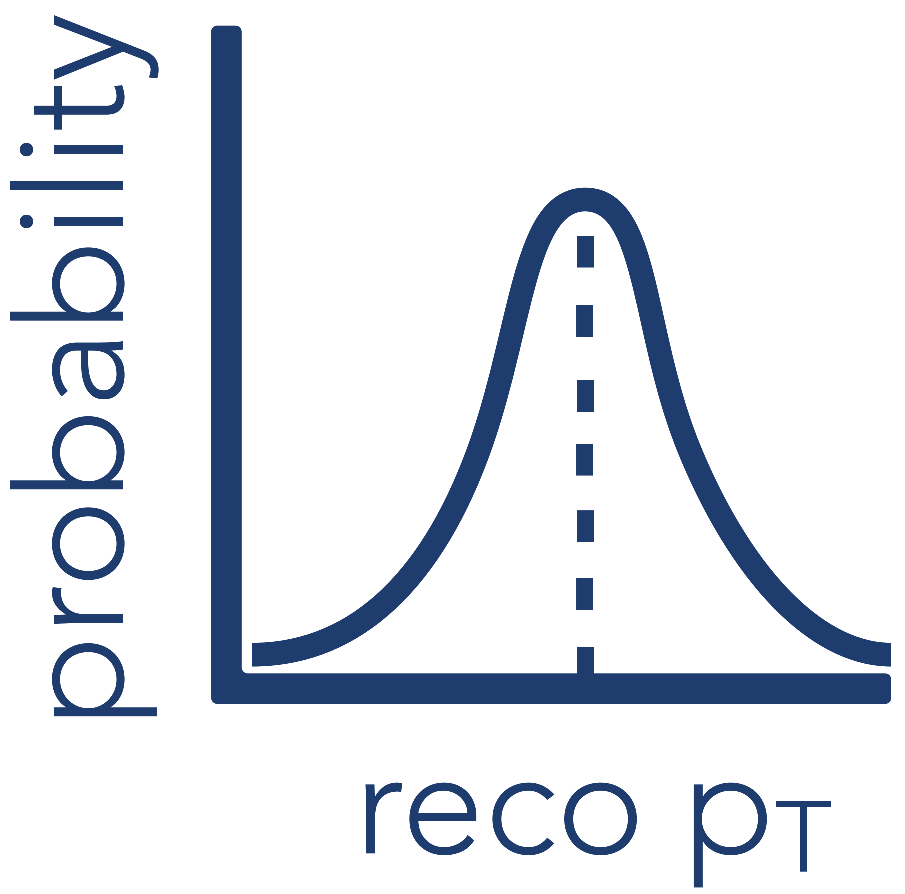


EVENT GENERATION → PARTON JETS → HADRONIZATION CLUSTERING → GEN JETS → GEANT4 PUPPI CLUSTERING → RAW RECO → JET CORRECTIONS → RECO JETS

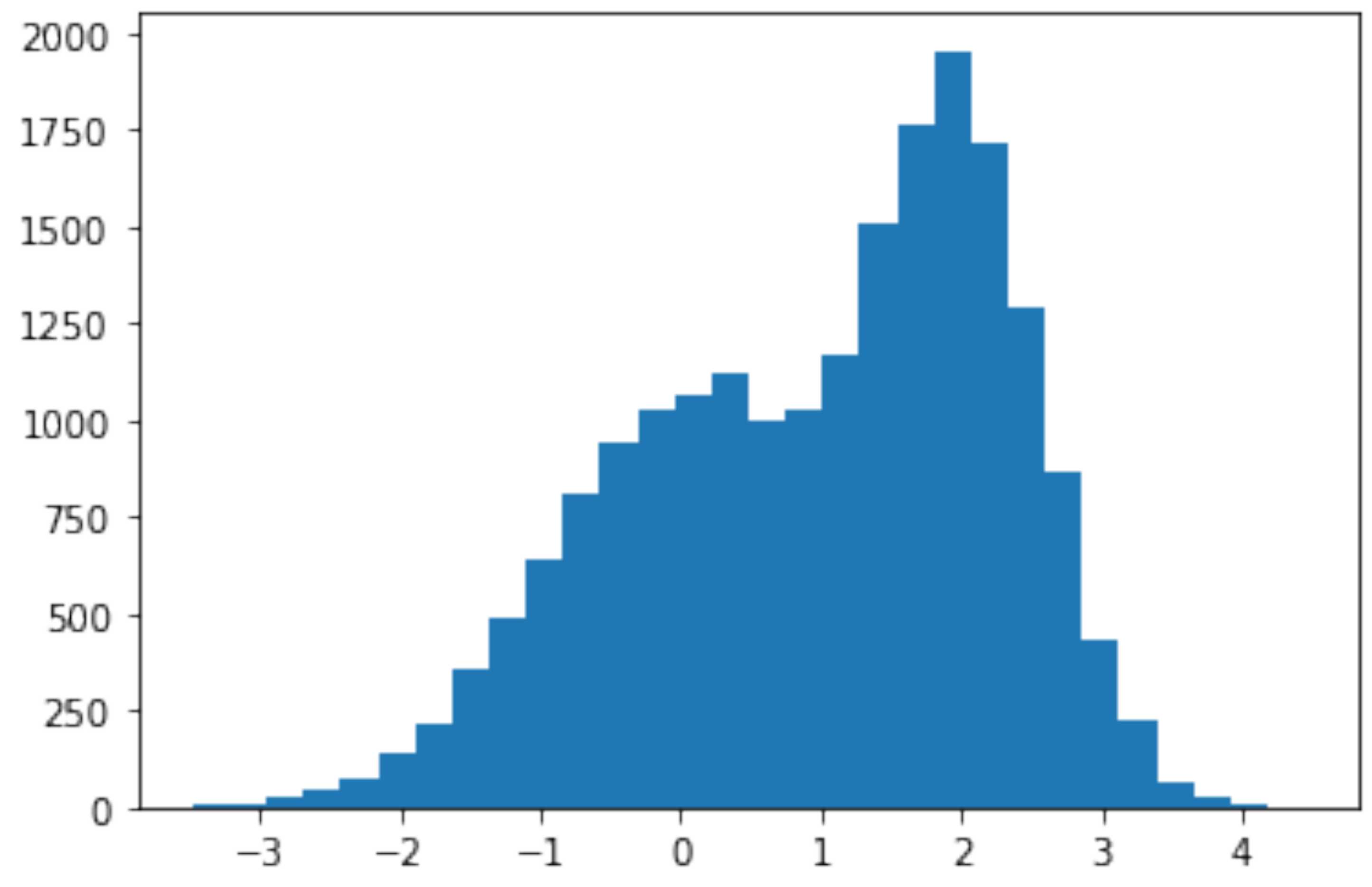
DATASET: Generated via Pythia8 + Delphes + anti-K_T + FastJet

NEED FOR DISTRIBUTION PREDICTIONS

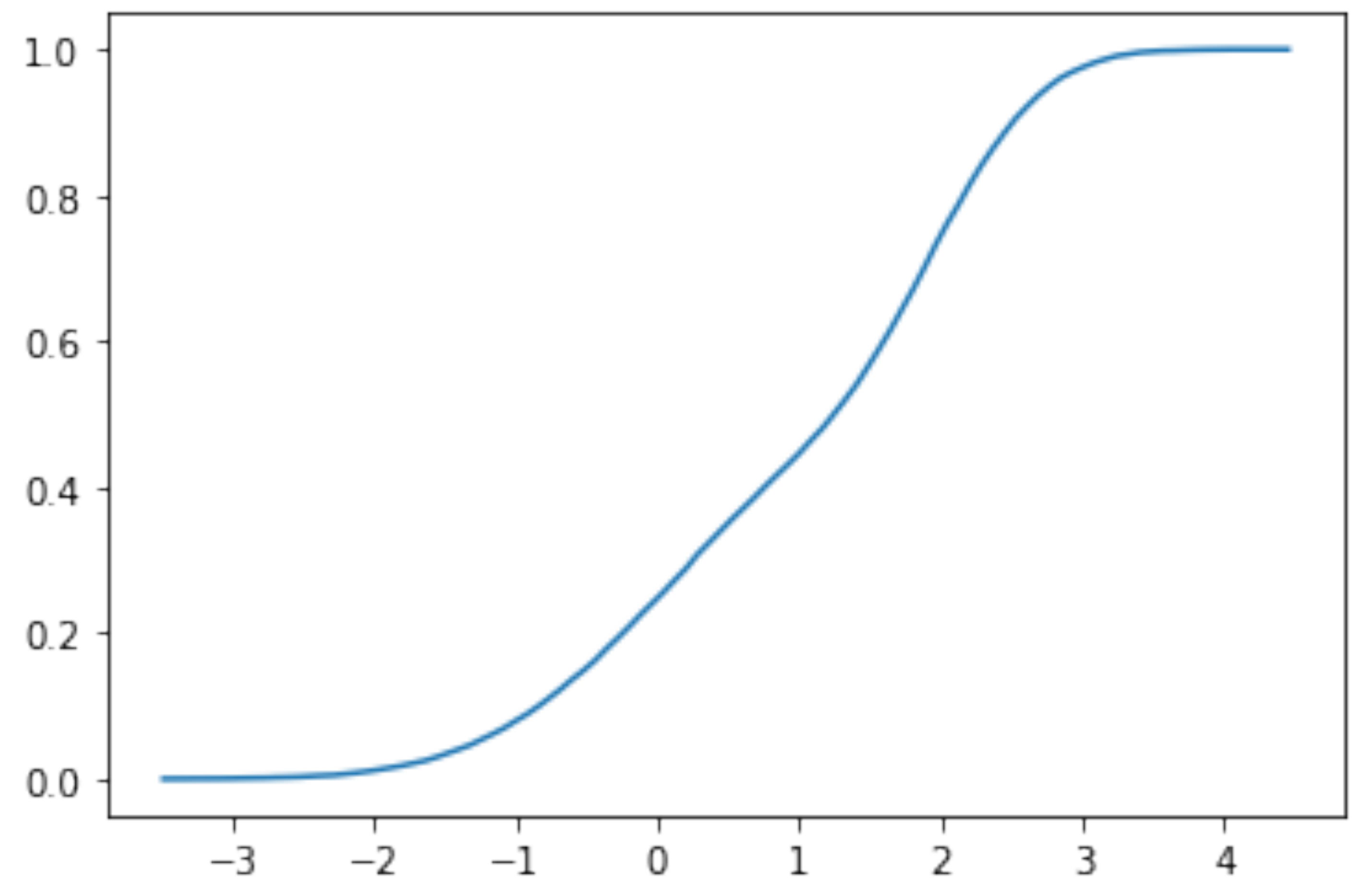




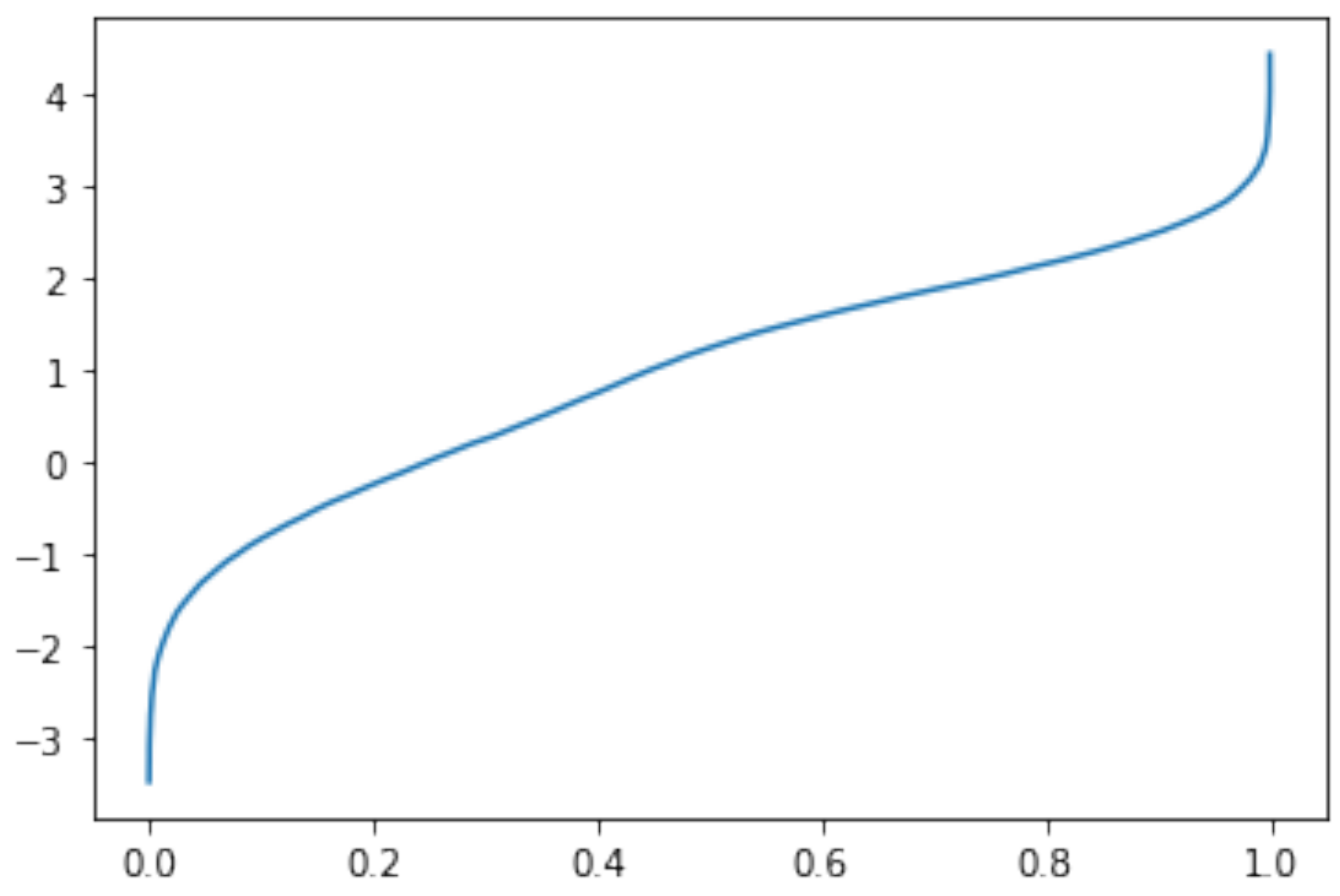
IMPLICIT QUANTILE NETWORKS STATS REVIEW



data

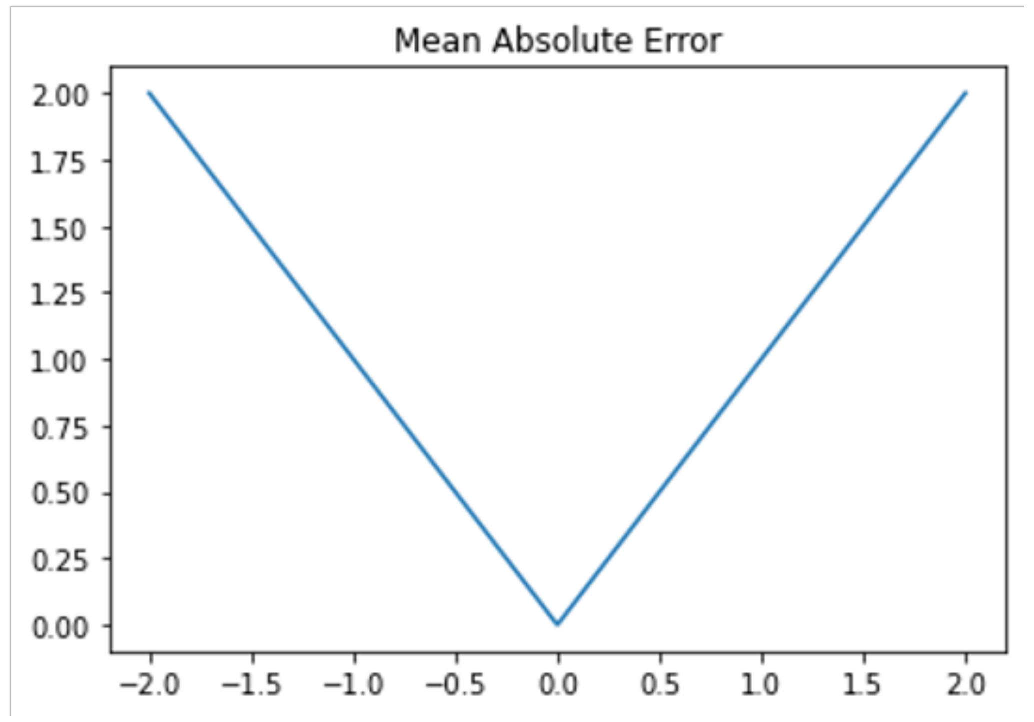


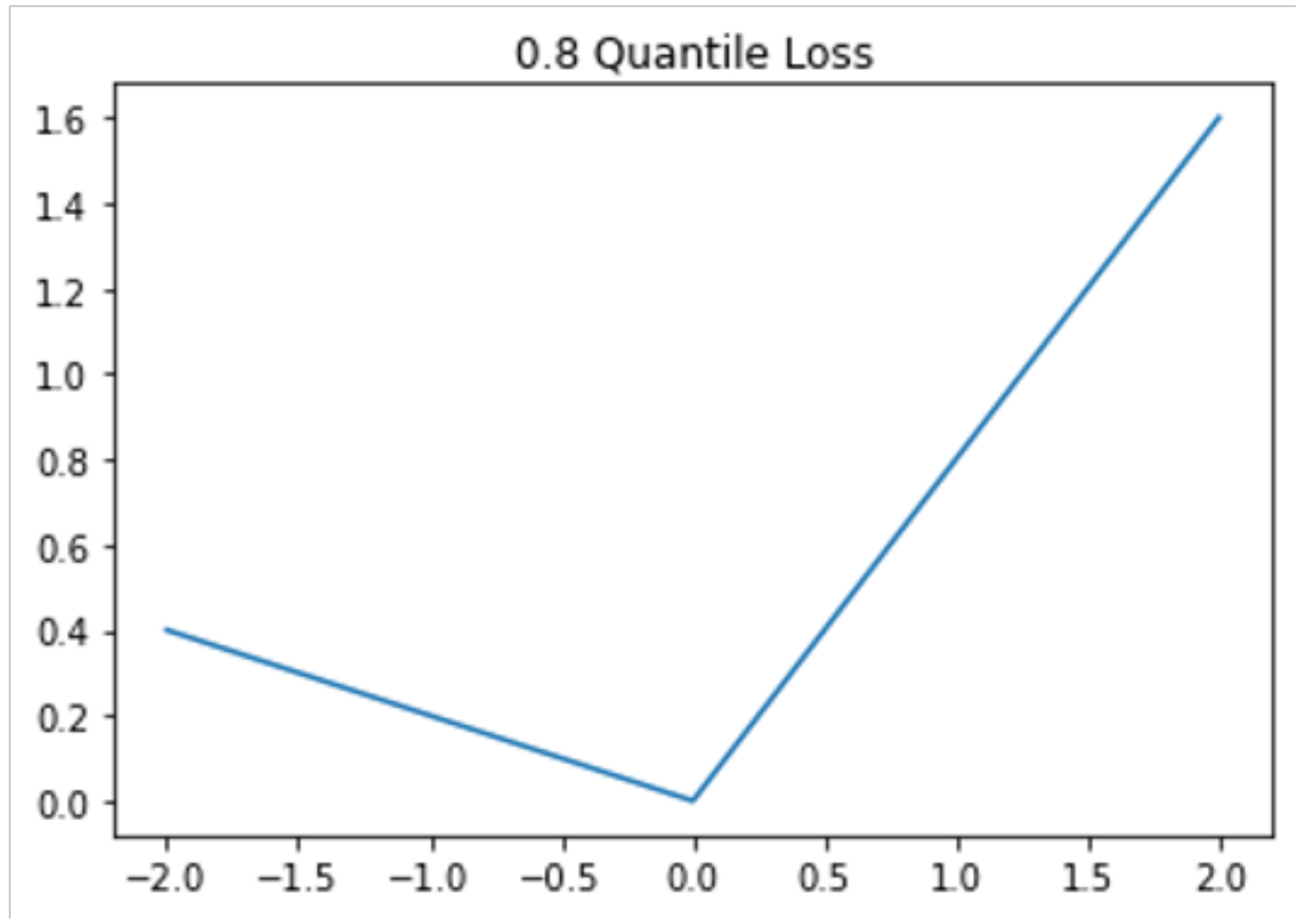
CDF

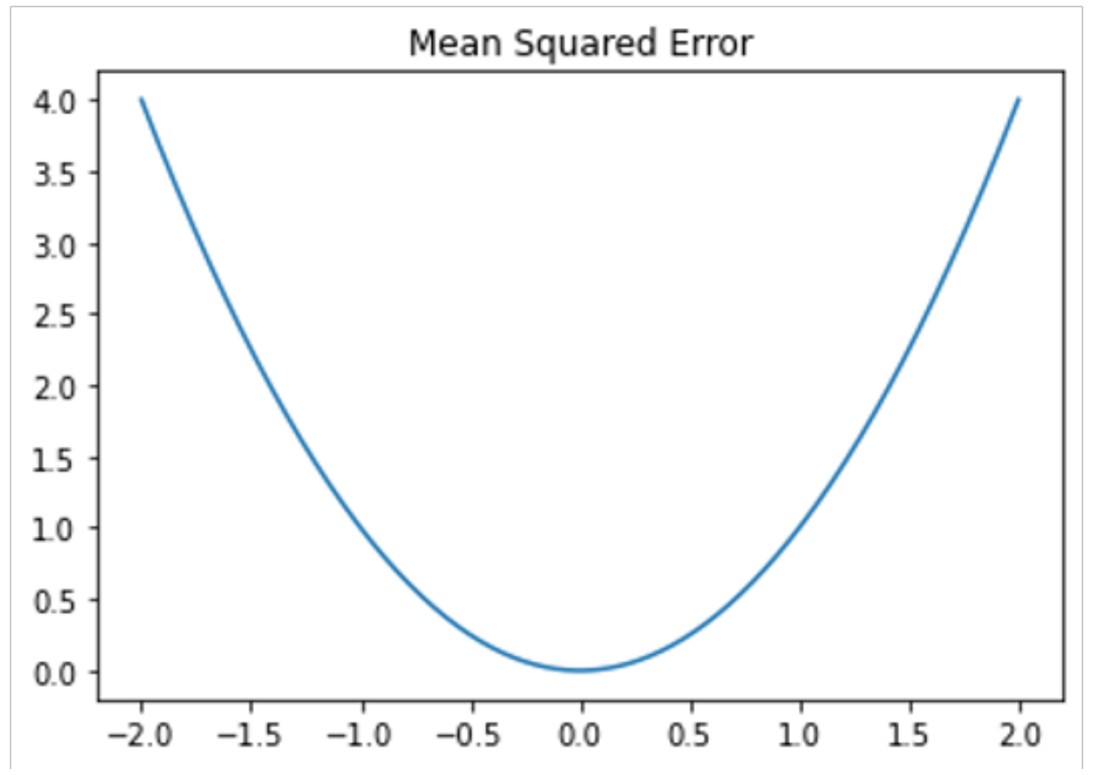


quantile function

IMPLICIT QUANTILE NETWORKS LOSS FUNCTION

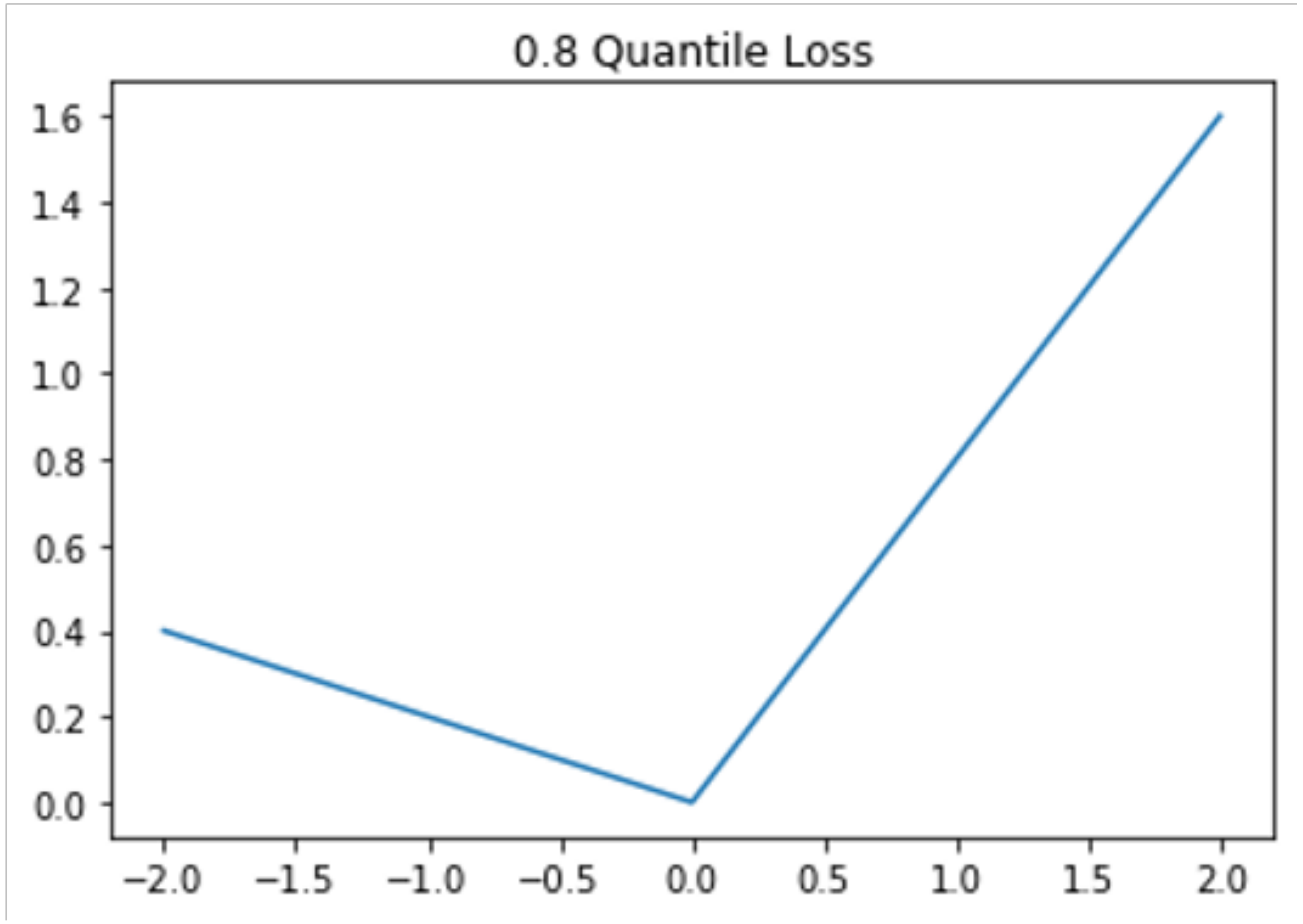






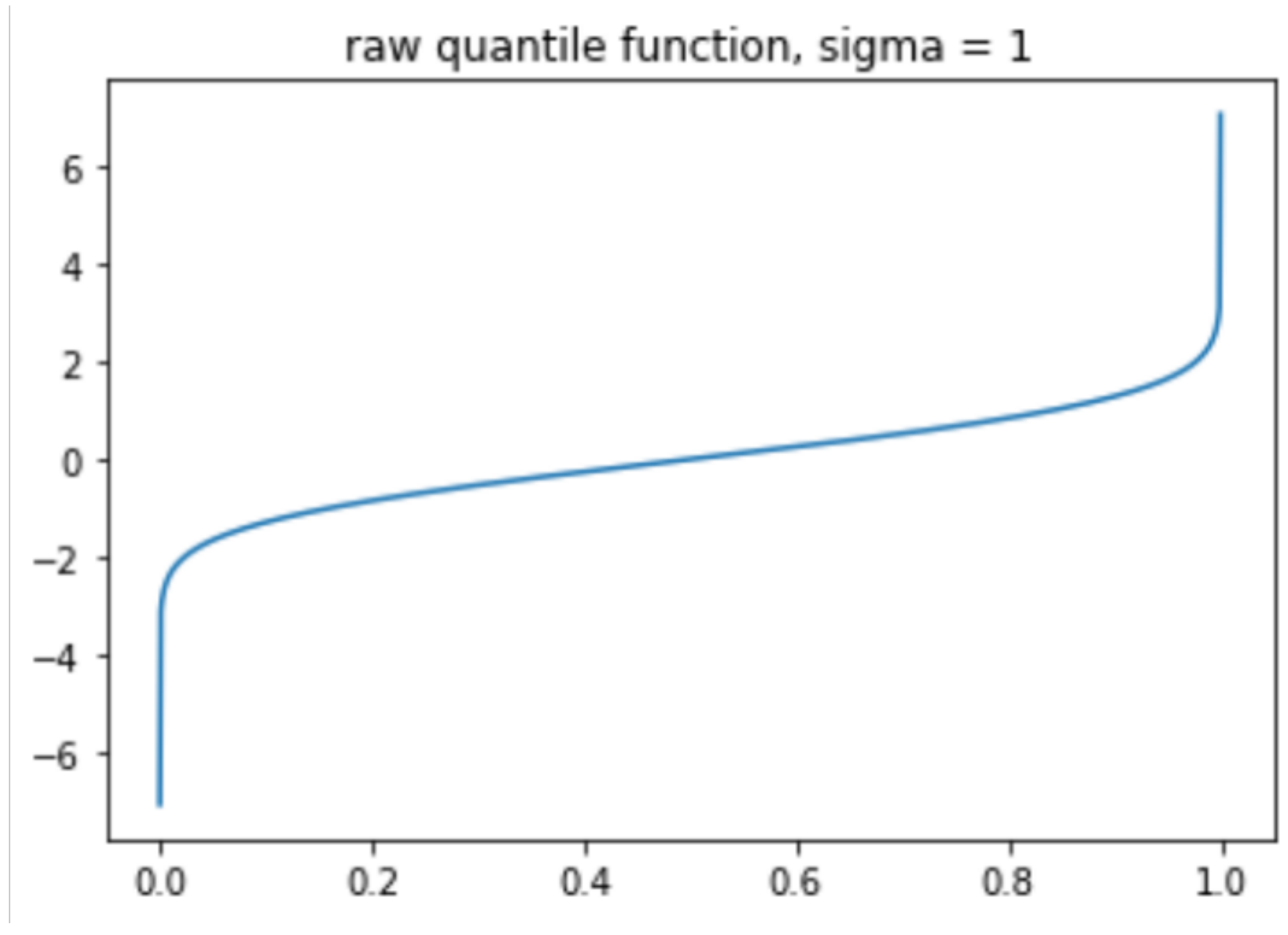
$$\mathcal{L}(f, x, y, \tau) = \begin{cases} \tau(y - f(x, \tau)) & y \geq f(x, \tau) \\ (\tau - 1)(y - f(x, \tau)) & y < f(x, \tau) \end{cases}$$

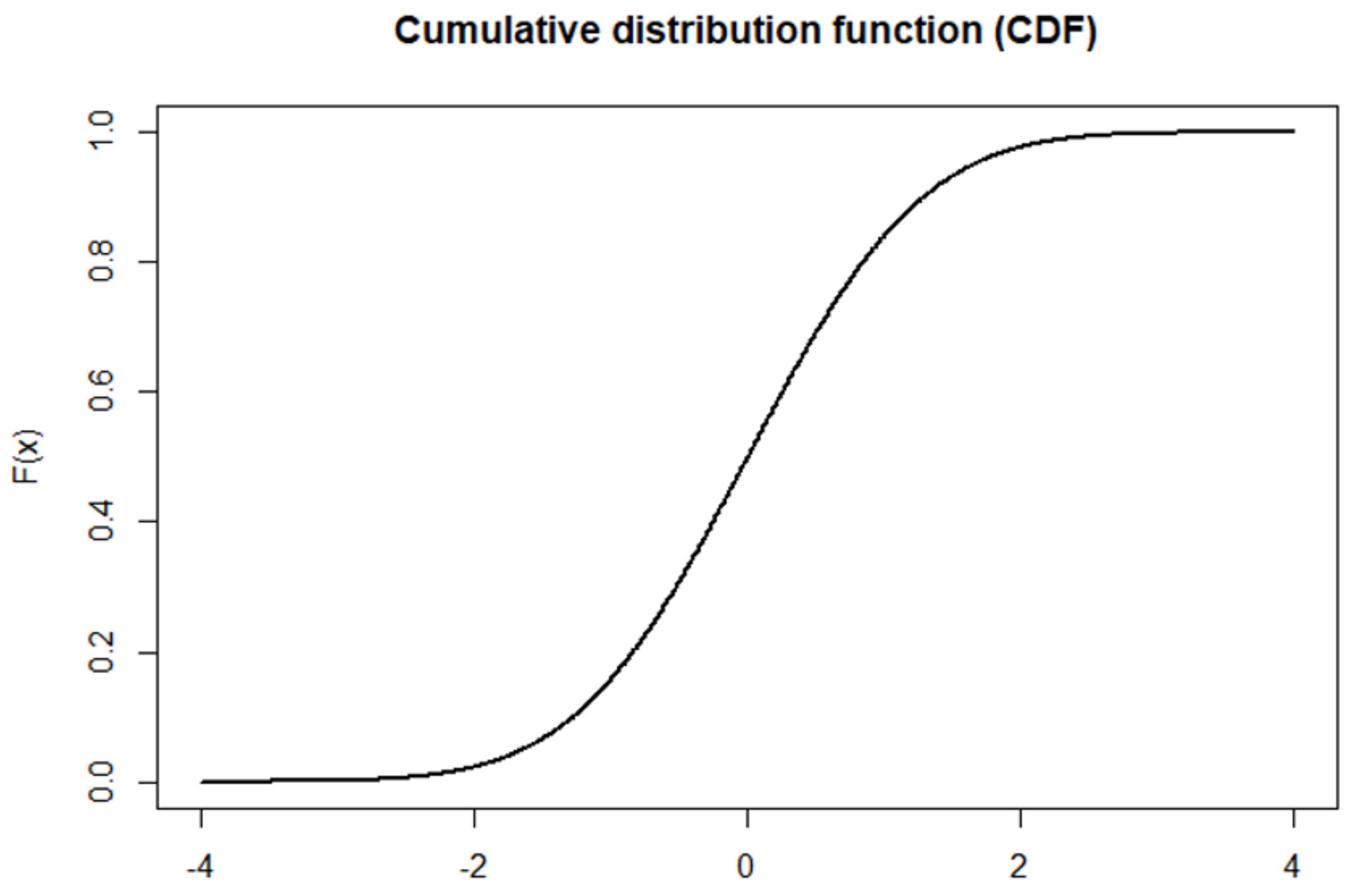
IMPLICIT QUANTILE NETWORKS LOSS FUNCTION

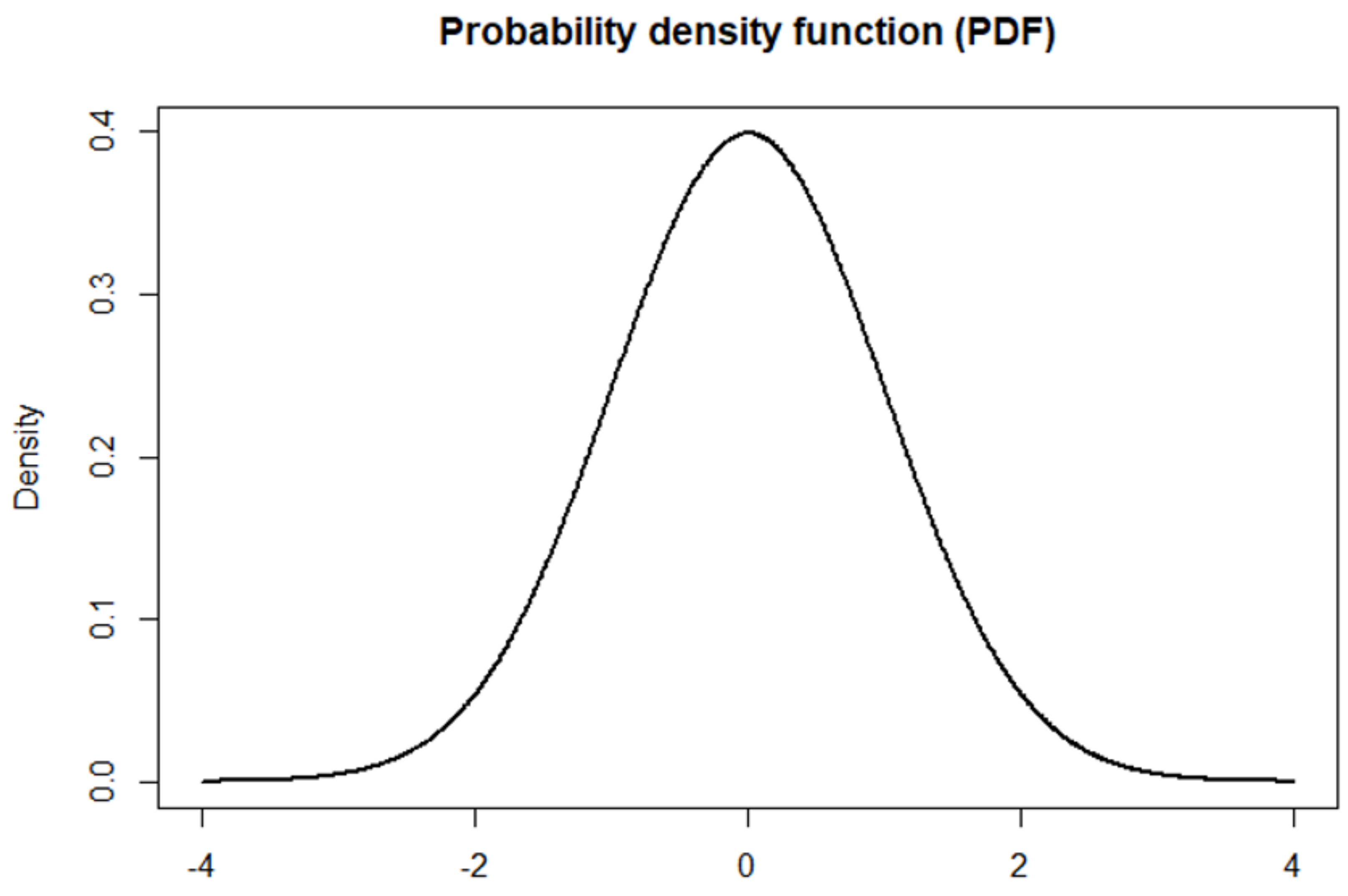


$$\mathcal{L}(f, x, y, \tau) = \begin{cases} \tau(y - f(x,\tau)) & y \geq f(x,\tau) \\ (\tau - 1)(y - f(x,\tau)) & y < f(x,\tau) \end{cases}$$

IMPLICIT QUANTILE NETWORKS LOSS FUNCTION

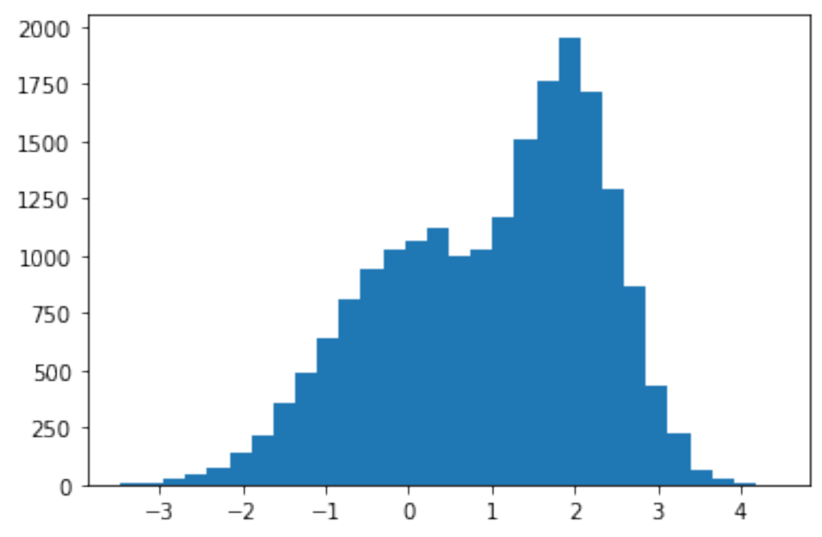




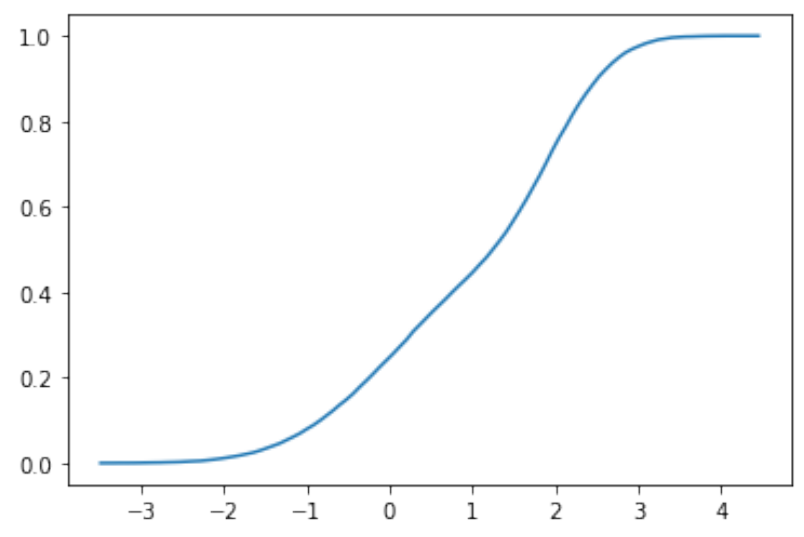


$$\mathcal{L}(f, x, y, \tau) = \begin{cases} \tau(y - f(x, \tau)) & y \geq f(x, \tau) \\ (\tau - 1)(y - f(x, \tau)) & y < f(x, \tau) \end{cases}$$

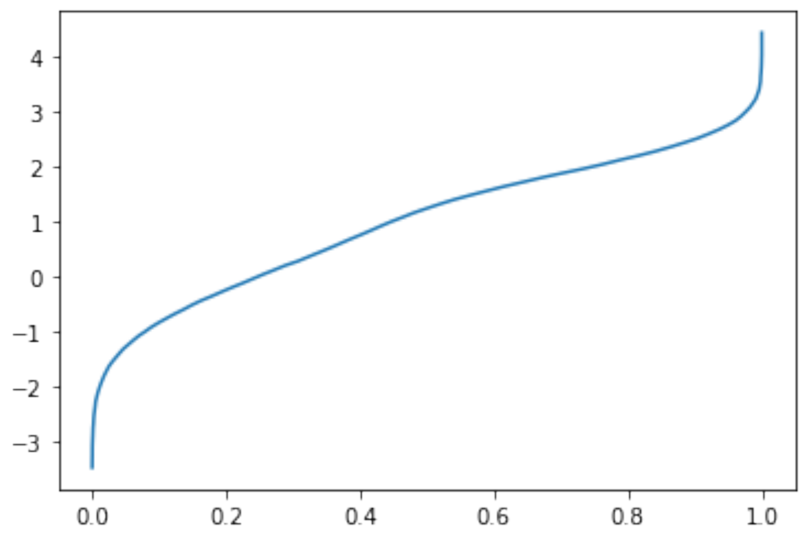
IMPLICIT QUANTILE NETWORKS STATS REVIEW



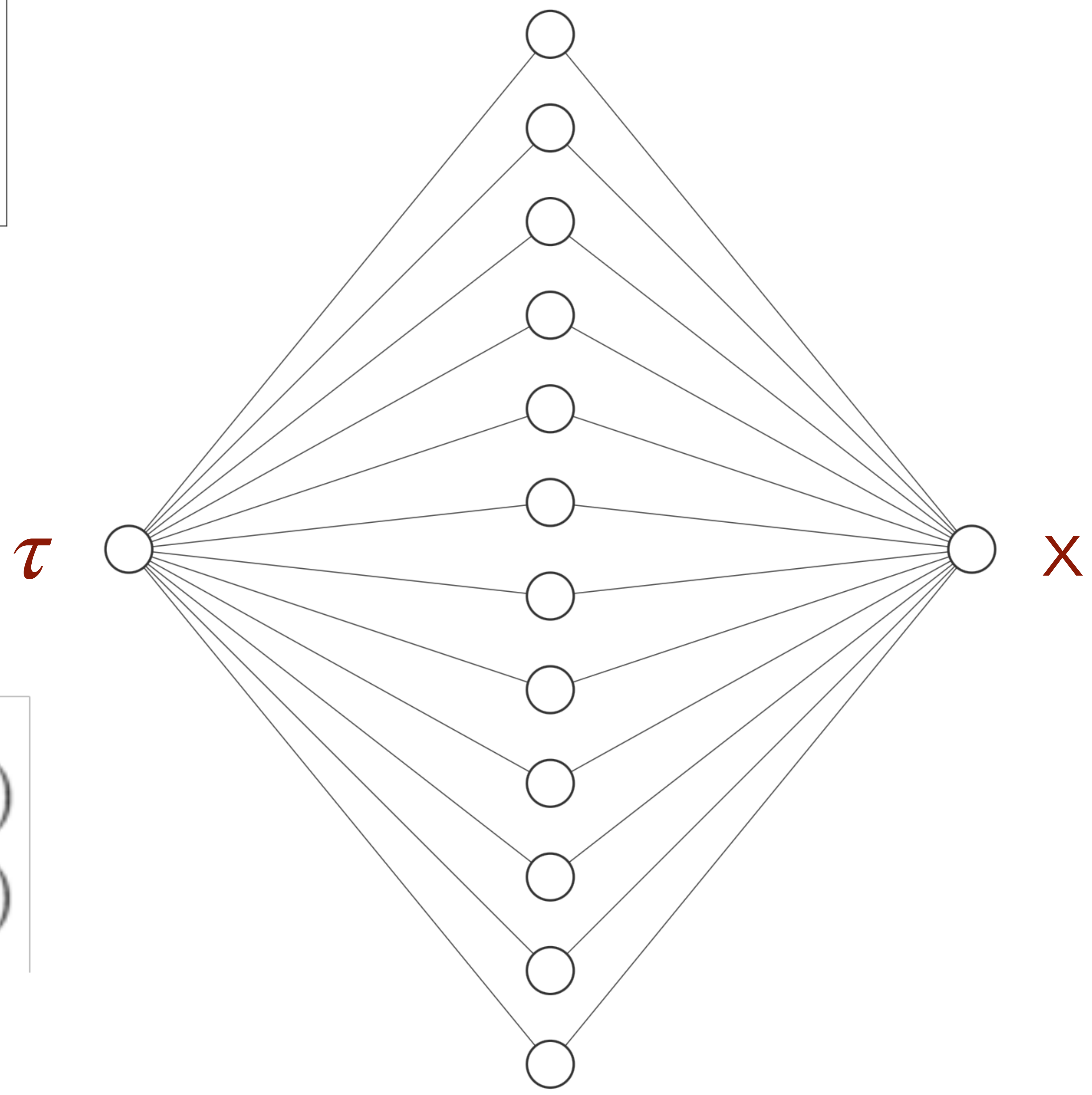
data (x)



CDF



quantile function (τ)



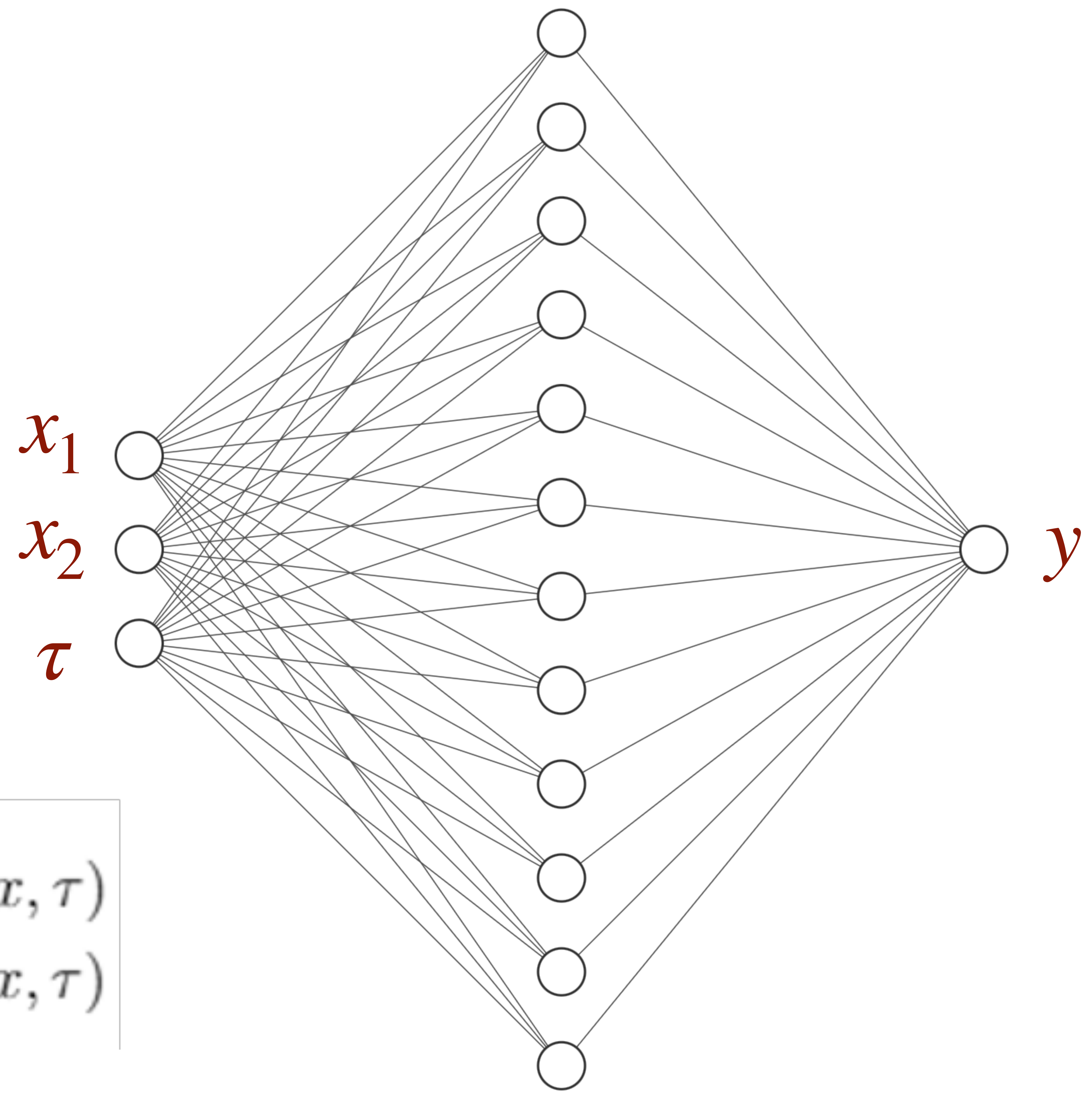
$$\mathcal{L}(f, x, y, \tau) = \begin{cases} \tau(y - f(x,\tau)) & y \geq f(x,\tau) \\ (\tau - 1)(y - f(x,\tau)) & y < f(x,\tau) \end{cases}$$

IMPLICIT QUANTILE NETWORKS ARCHITECTURE

$p(y \mid x_1, x_2)$

$(x_1, x_2) \rightarrow y$

$$\mathcal{L}(f, x, y, \tau) = \begin{cases} \tau(y - f(x, \tau)) & y \geq f(x, \tau) \\ (\tau - 1)(y - f(x, \tau)) & y < f(x, \tau) \end{cases}$$



IMPLICIT QUANTILE NETWORKS ARCHITECTURE

$$p(p_T', \eta', \phi', m' \,|\, p_T, \eta, \phi, m)$$

$$(p_T, \eta, \phi, m) \rightarrow (p_T', \eta', \phi', m')$$

$$p(y^{(1)}, y^{(2)}, \ldots, y^{(n)} \,|\, \mathbf{x}) = p(y^{(1)} \,|\, \mathbf{x}) p(y^{(2)} \,|\, \mathbf{x}, y^{(1)}) \prod_{i=3}^{n} p(y^{(i)} \,|\, \mathbf{x}, y^{(1)}, \ldots, y^{(i-1)})$$

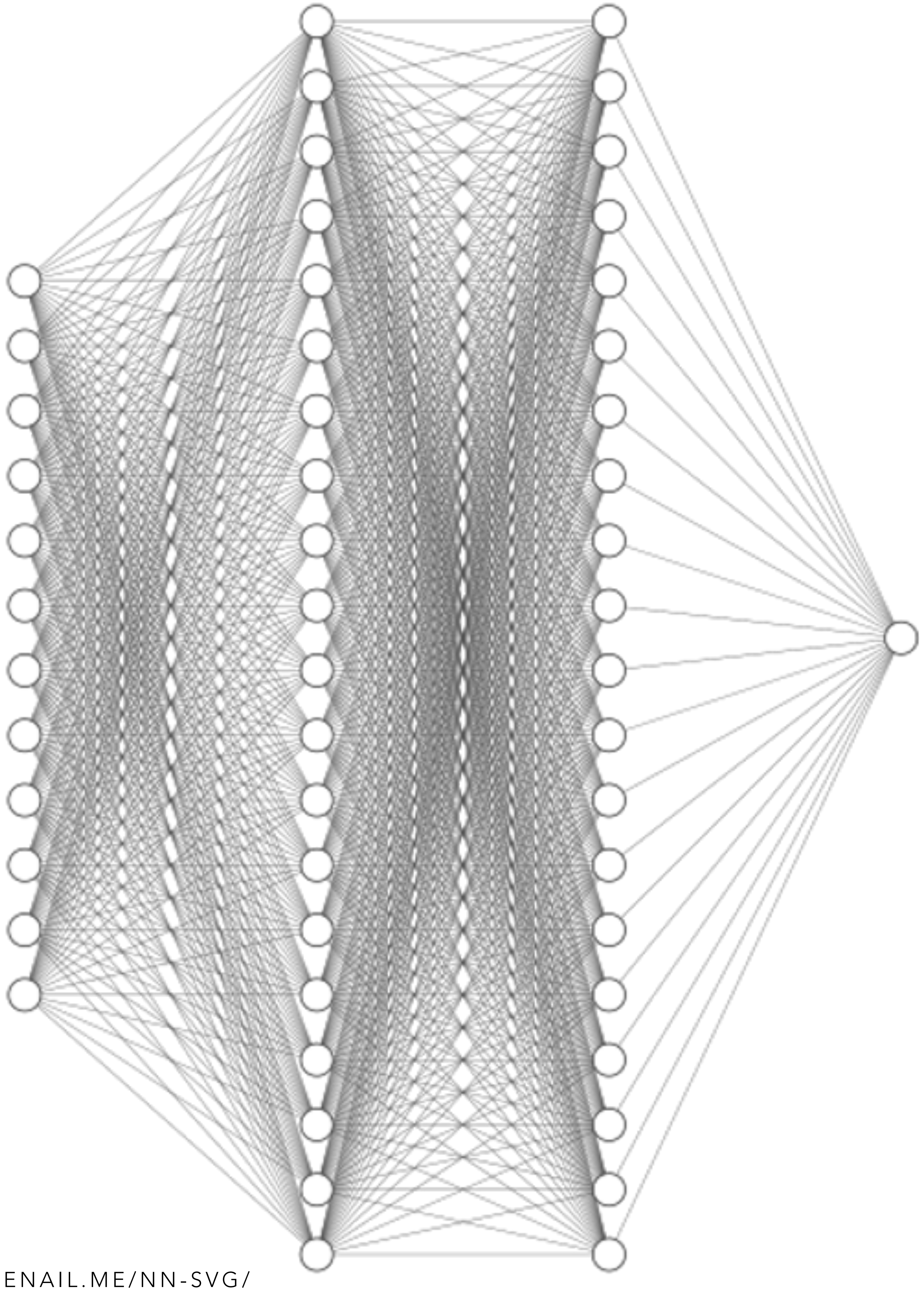
IMPLICIT QUANTILE NETWORKS ARCHITECTURE

$$p(p_T', \eta', \phi', m' \,|\, p_T, \eta, \phi, m)$$

$$(p_T, \eta, \phi, m) \rightarrow (p_T', \eta', \phi', m')$$

$$p(y^{(1)}, y^{(2)}, \ldots, y^{(n)} \,|\, \mathbf{x}) = \boxed{p(y^{(1)} \,|\, \mathbf{x})}\boxed{p(y^{(2)} \,|\, \mathbf{x}, y^{(1)})}\prod_{i=3}^{n}\boxed{p(y^{(i)} \,|\, \mathbf{x}, y^{(1)}, \ldots, y^{(i-1)})}$$

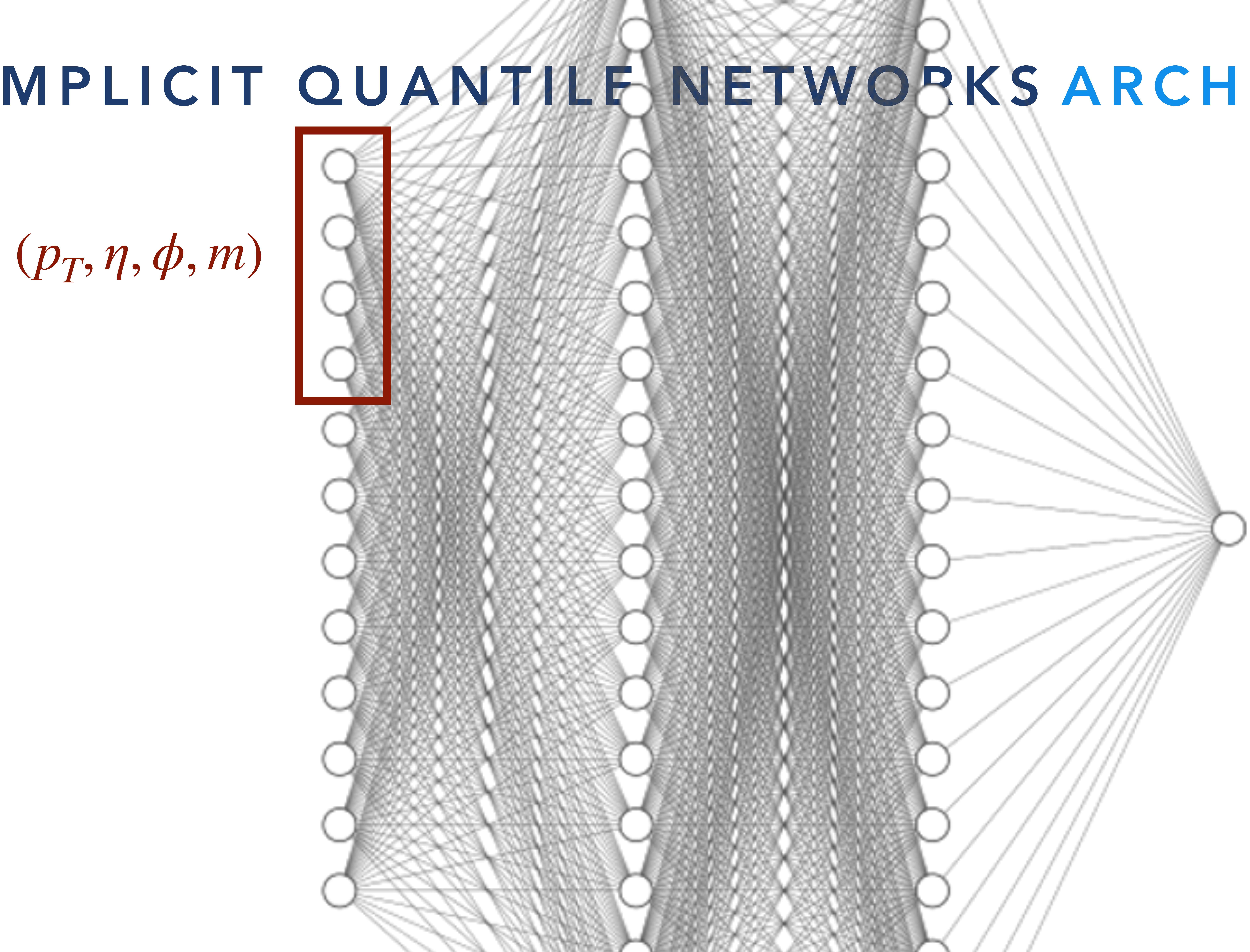
IMPLICIT QUANTILE NETWORKS ARCHITECTURE



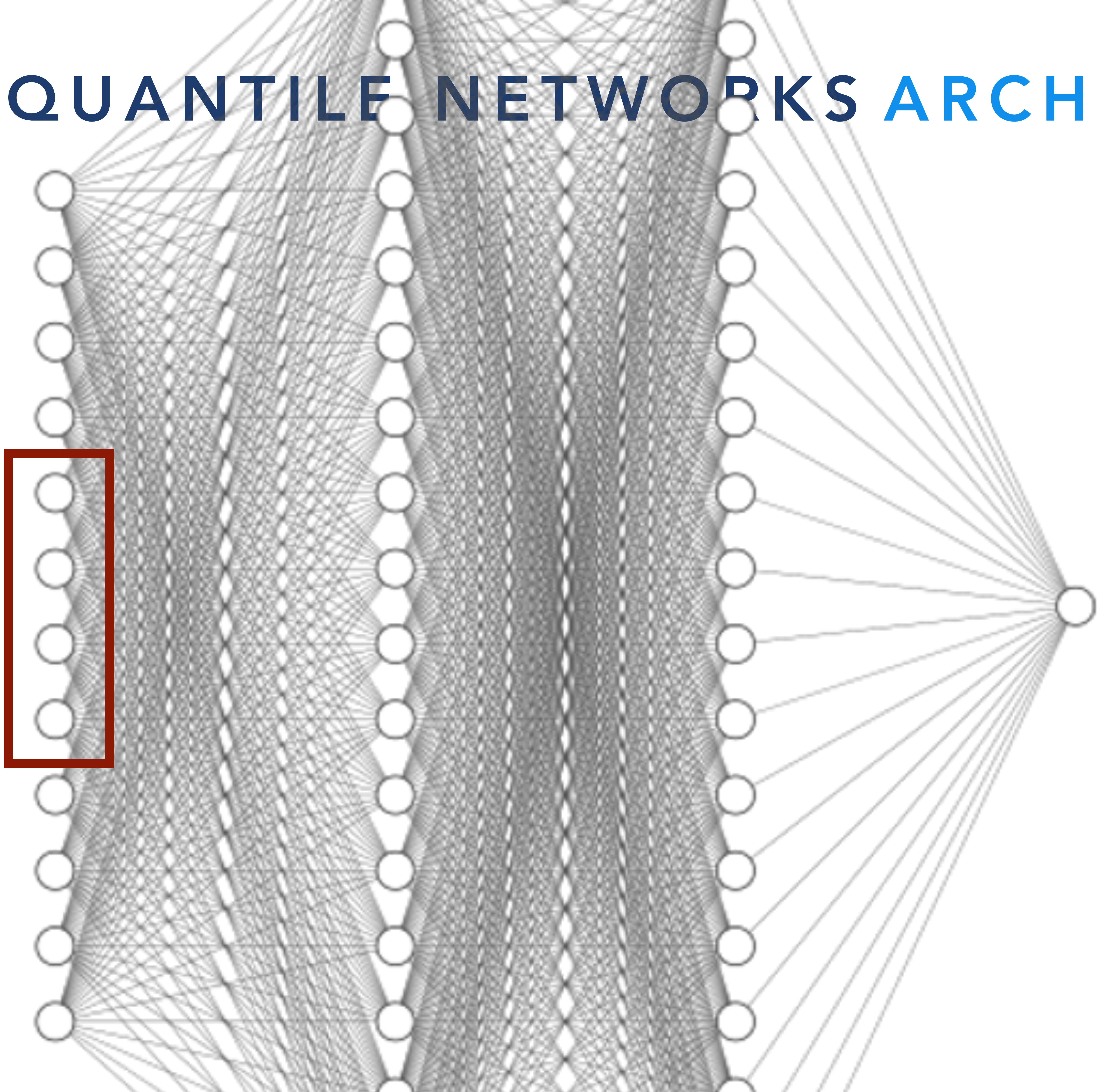
$$(p_T, \eta, \phi, m) \rightarrow (p'_T, \eta', \phi', m')$$

IMAGE CREDIT: HTTPS://ALEXLENAIL.ME/NN-SVG/

IMPLICIT QUANTILE NETWORKS ARCHITECTURE



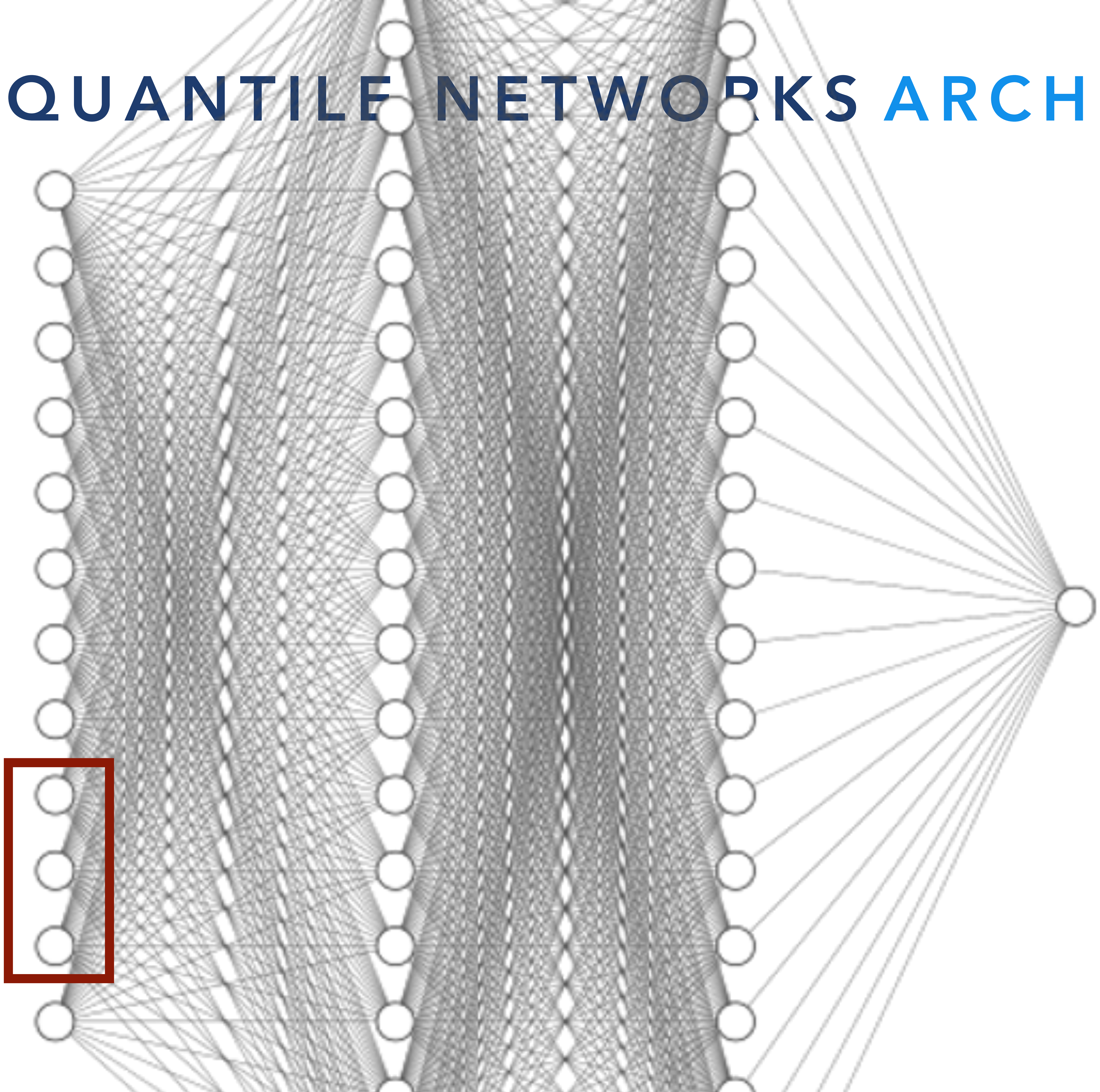
IMPLICIT QUANTILE NETWORKS ARCHITECTURE



(p_T', η', ϕ', m')

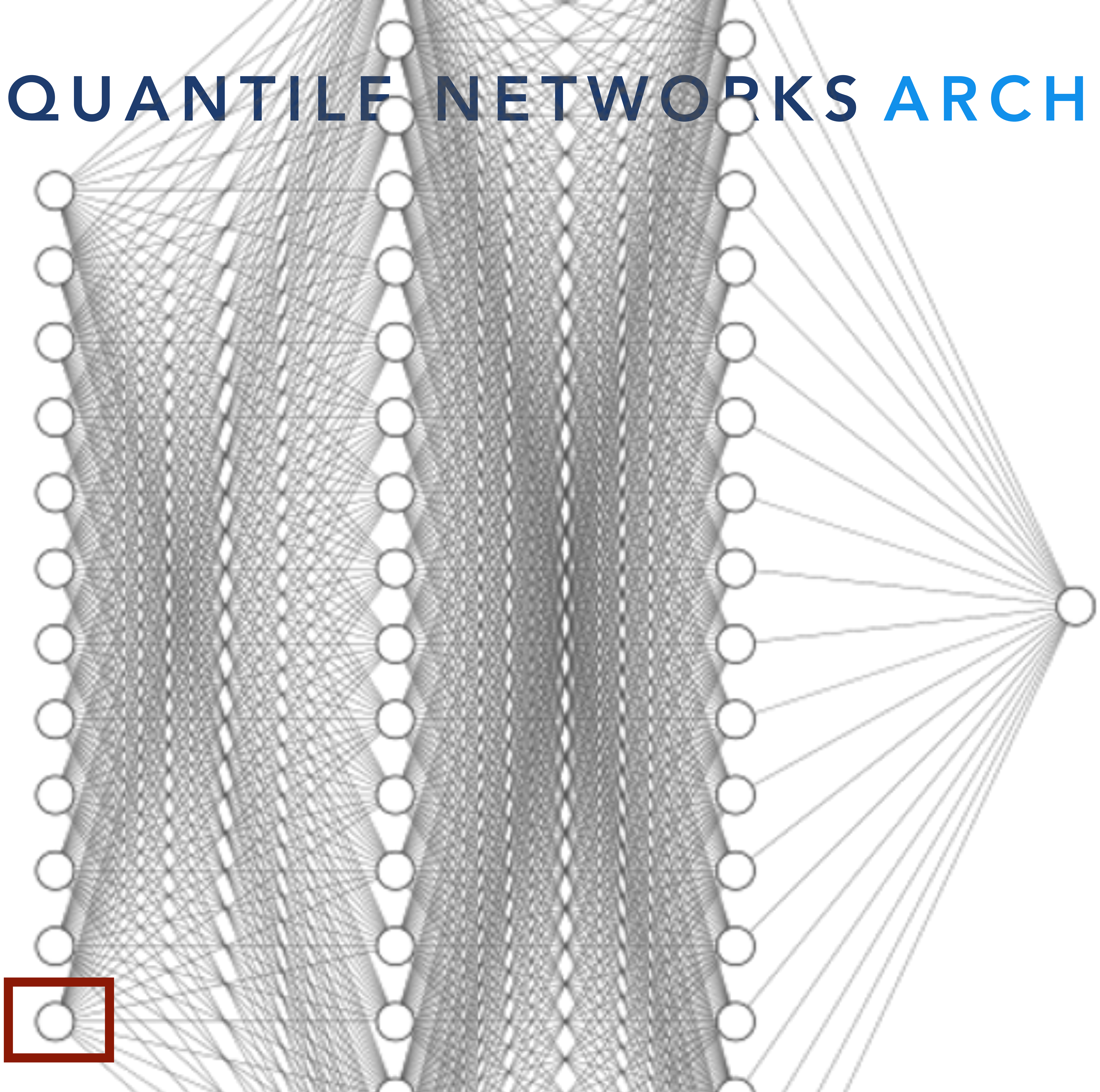
[0,0,1,0]

IMPLICIT QUANTILE NETWORKS ARCHITECTURE



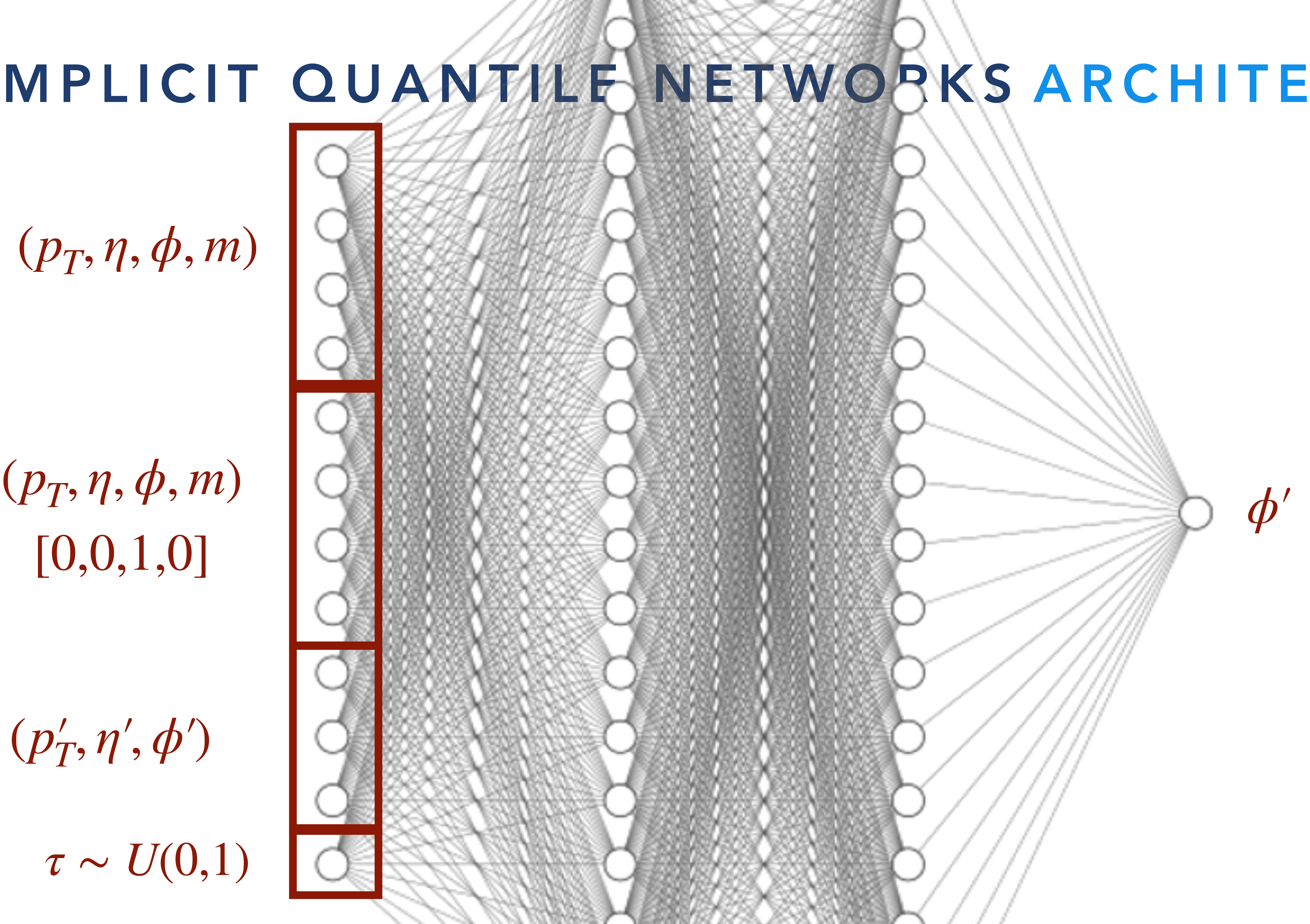
(p_T', η', ϕ')

IMPLICIT QUANTILE NETWORKS ARCHITECTURE

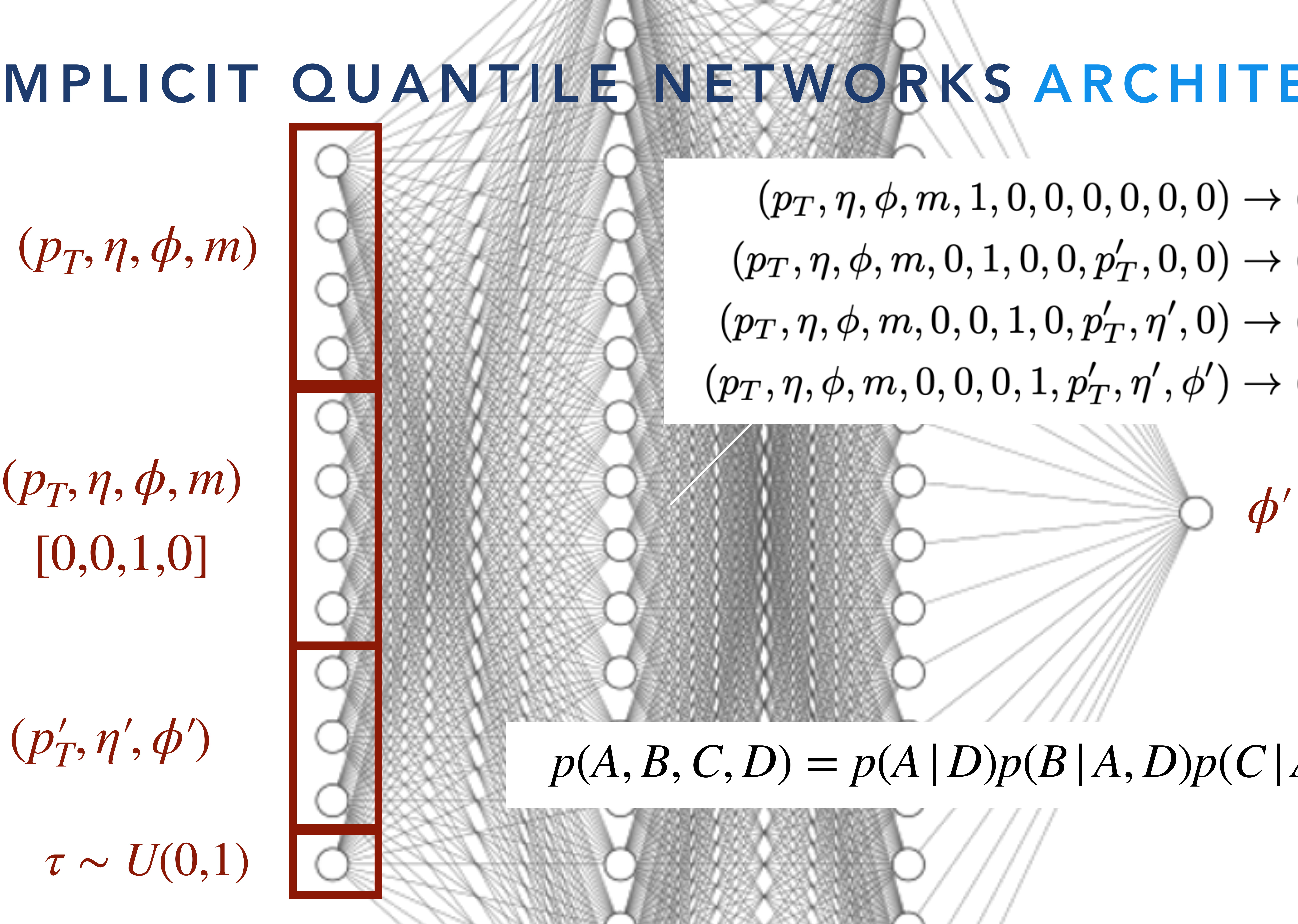


$\tau \sim U(0,1)$

IMPLICIT QUANTILE NETWORKS ARCHITECTURE



IMPLICIT QUANTILE NETWORKS ARCHITECTURE

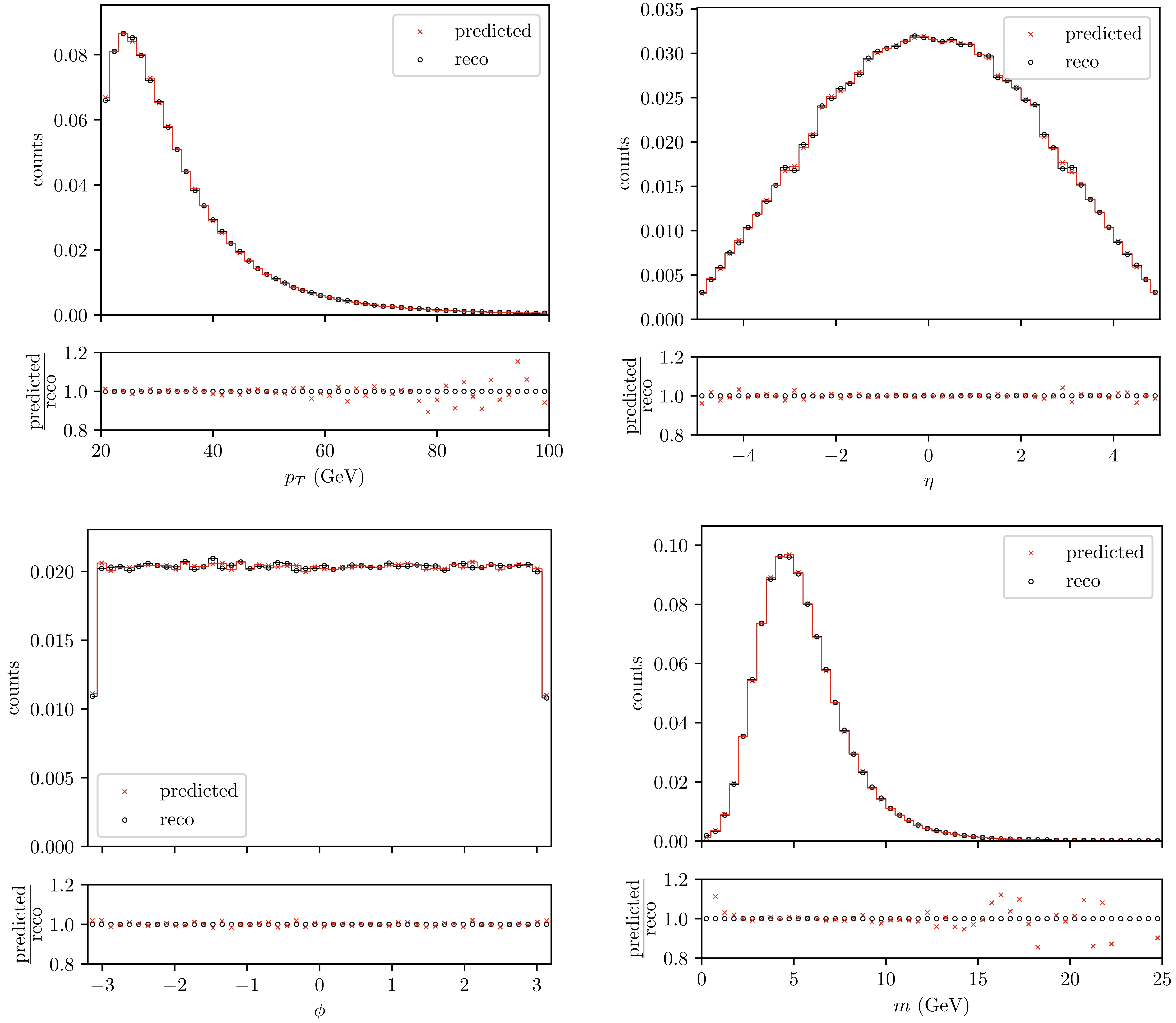


$$
\begin{aligned}
(p_T, \eta, \phi, m, 1, 0, 0, 0, 0, 0, 0) &\rightarrow (p_T'),\\
(p_T, \eta, \phi, m, 0, 1, 0, 0, p_T', 0, 0) &\rightarrow (\eta'),\\
(p_T, \eta, \phi, m, 0, 0, 1, 0, p_T', \eta', 0) &\rightarrow (\phi'),\\
(p_T, \eta, \phi, m, 0, 0, 0, 1, p_T', \eta', \phi') &\rightarrow (m'),
\end{aligned}
$$

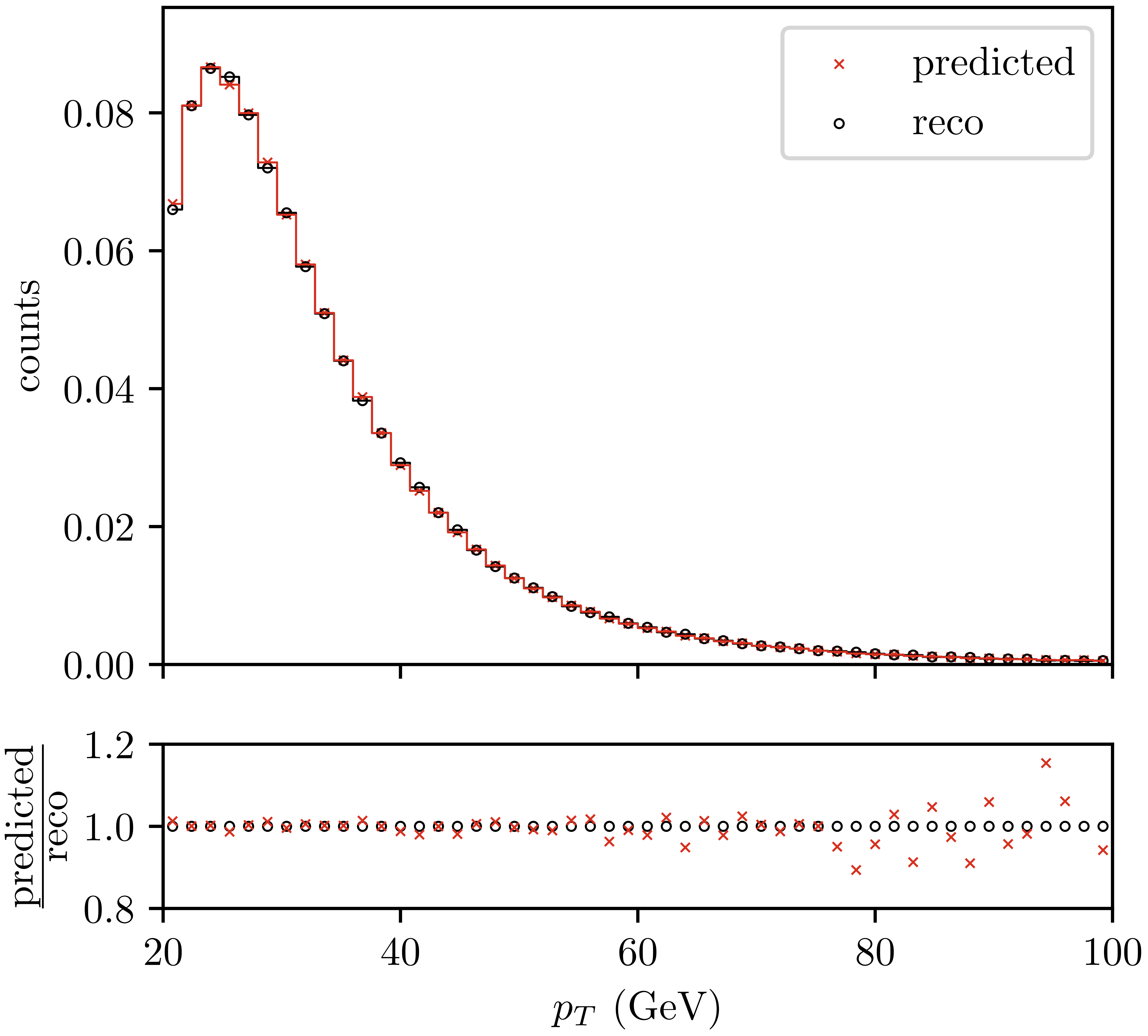
$$p(A, B, C, D) = p(A \,|\, D)\, p(B \,|\, A, D)\, p(C \,|\, A, B, D)$$

RESULTS JET SIMULATION

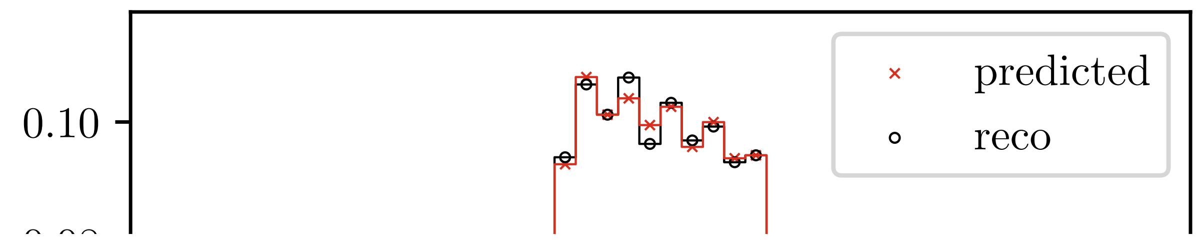
RESULTS JET SIMULATION



RESULTS JET SIMULATION



RESULTS JET SIMULATION: SUBSPACE



Implicit Quantile Neural Networks for Jet Simulation and Correction

Braden Kronheim
Department of Physics
University of Maryland
College Park, MD 20742
bkronhei@umd.edu

Michelle P. Kuchera
Department of Physics
Davidson College
Davidson, NC 28035
mikuchera@davidson.edu

Harrison B. Prosper
Department of Physics
Florida State University
Tallahassee, FL 32306
hprosper@fsu.edu

Raghuram Ramanujan
Department of Mathematics & Computer Science
Davidson College
Davidson, NC 28035
raramanujan@davidson.edu