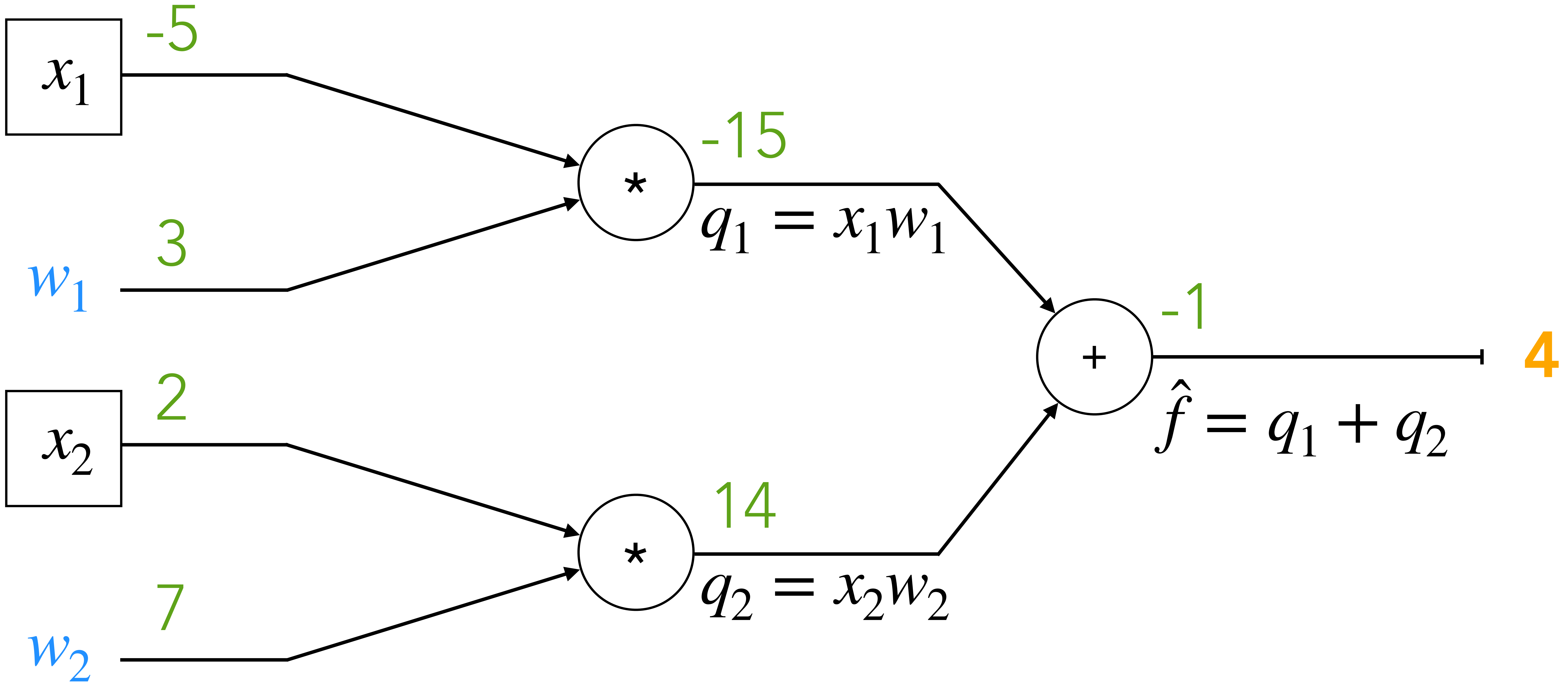
MACHINE LEARNING IN QCD THEORY

MICHELLE KUCHERA
DAVIDSON COLLEGE

FRIB THEORY SEMINAR
10 MARCH 2020

COMPUTATIONAL GRAPH

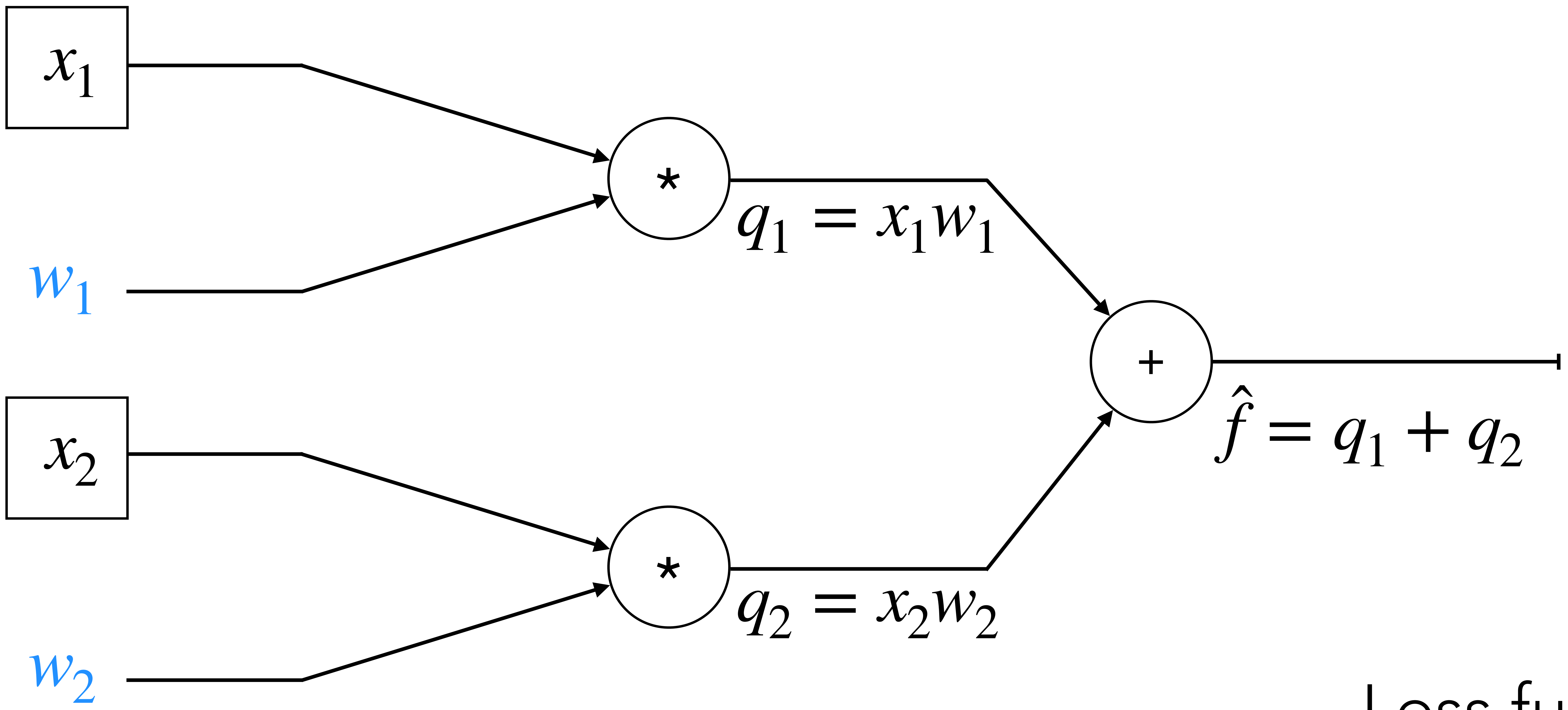


$$\hat{f} = x_1 w_1 + x_2 w_2$$

MACHINE LEARNING:

LEARNING FROM DATA

REGRESSION

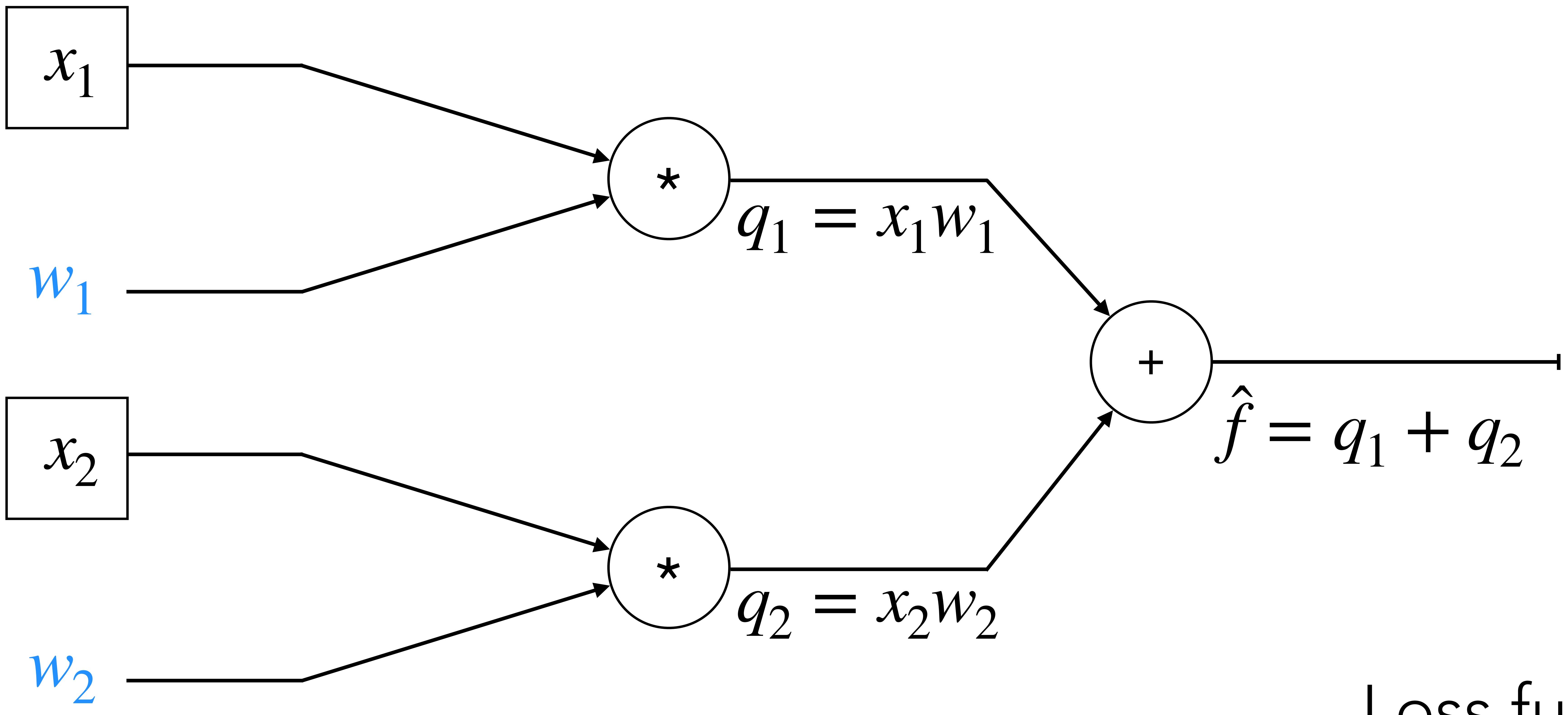


Loss function

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

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

SUPERVISED LEARNING



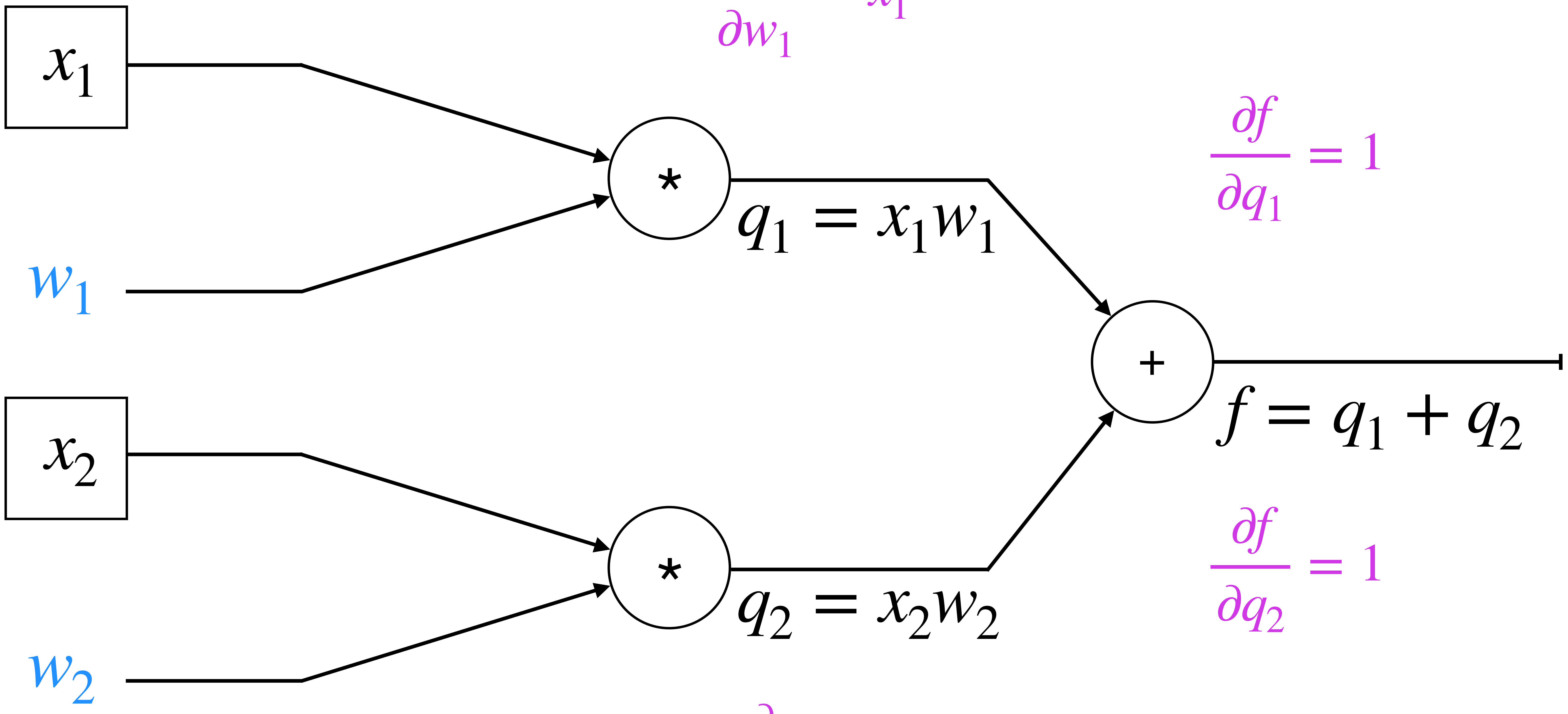
Loss function

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

BACKPROPAGATION

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

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



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

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

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

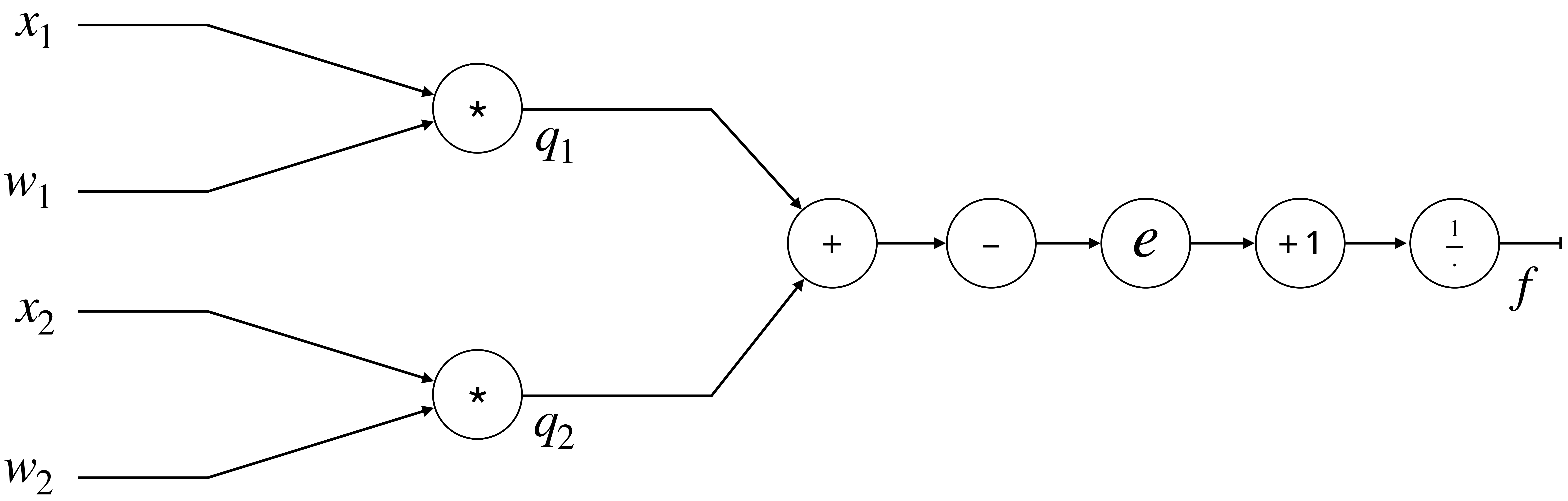
$$\frac{\partial q_2}{\partial w_2} = x_2$$

$$w_2 = w_2 + \eta * \frac{\partial f}{\partial q_1}\frac{\partial q_1}{\partial w_2}$$

Loss function

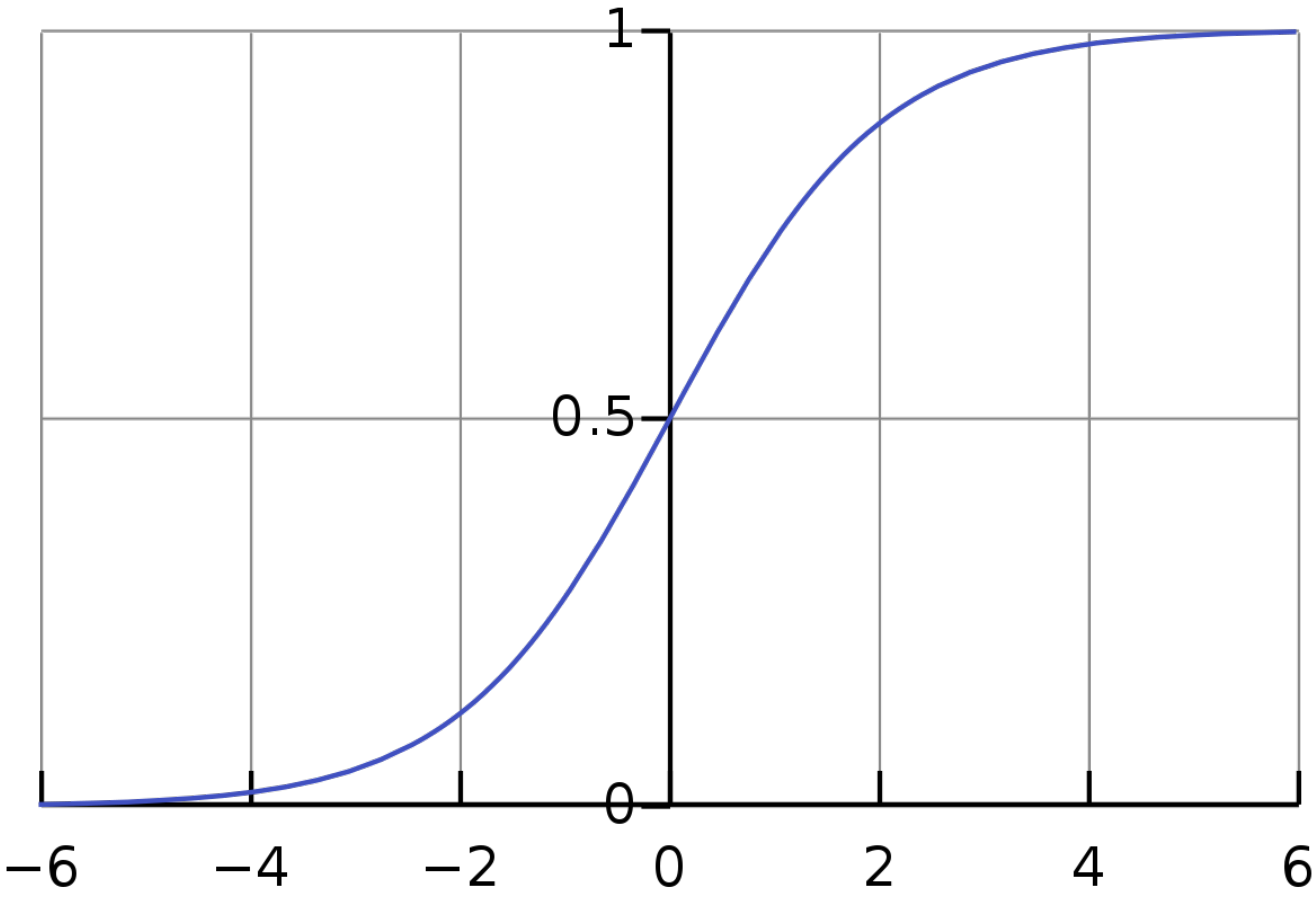
$$J(w) = f - \hat{f}$$

LOGISTIC REGRESSION

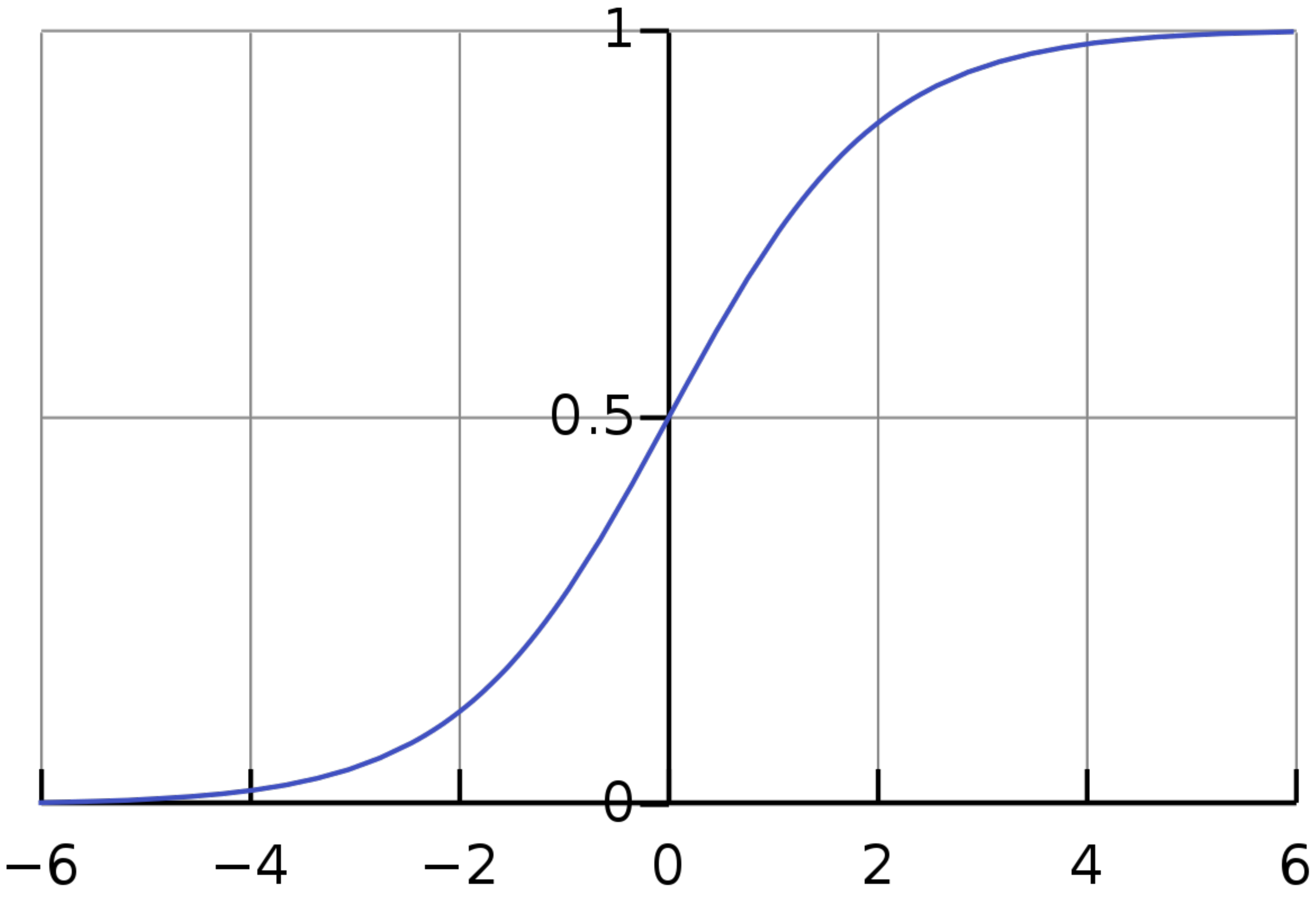


$$f = \frac{1}{1 + e^{-(x_1 w_1 + x_2 w_2)}}$$

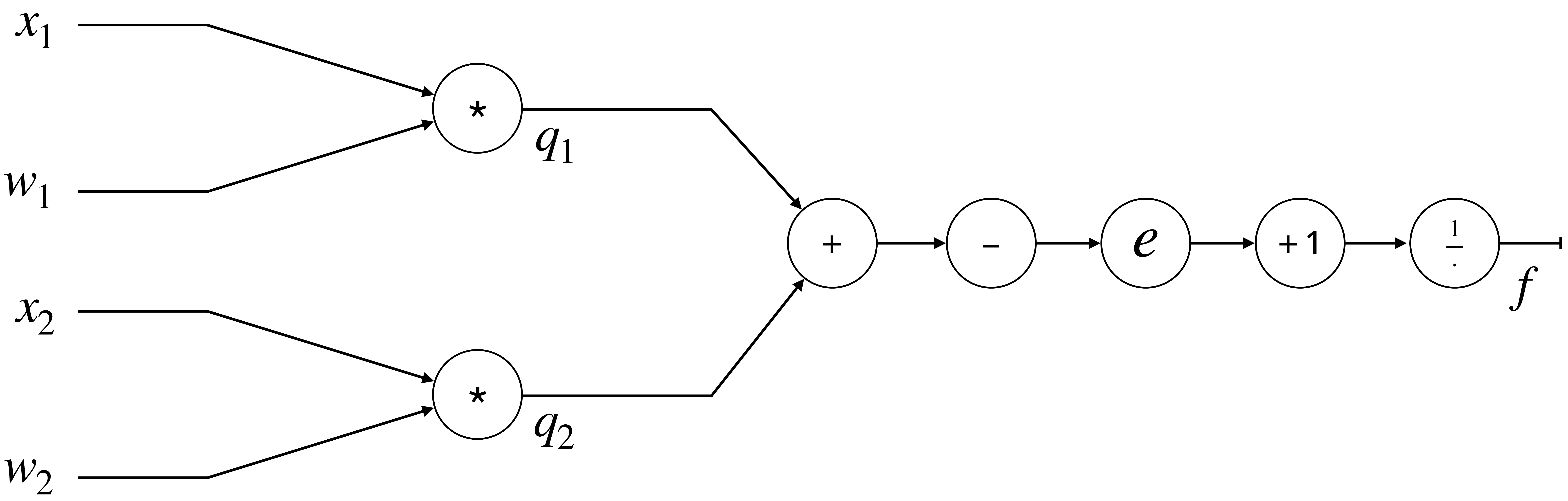
LOGISTIC REGRESSION



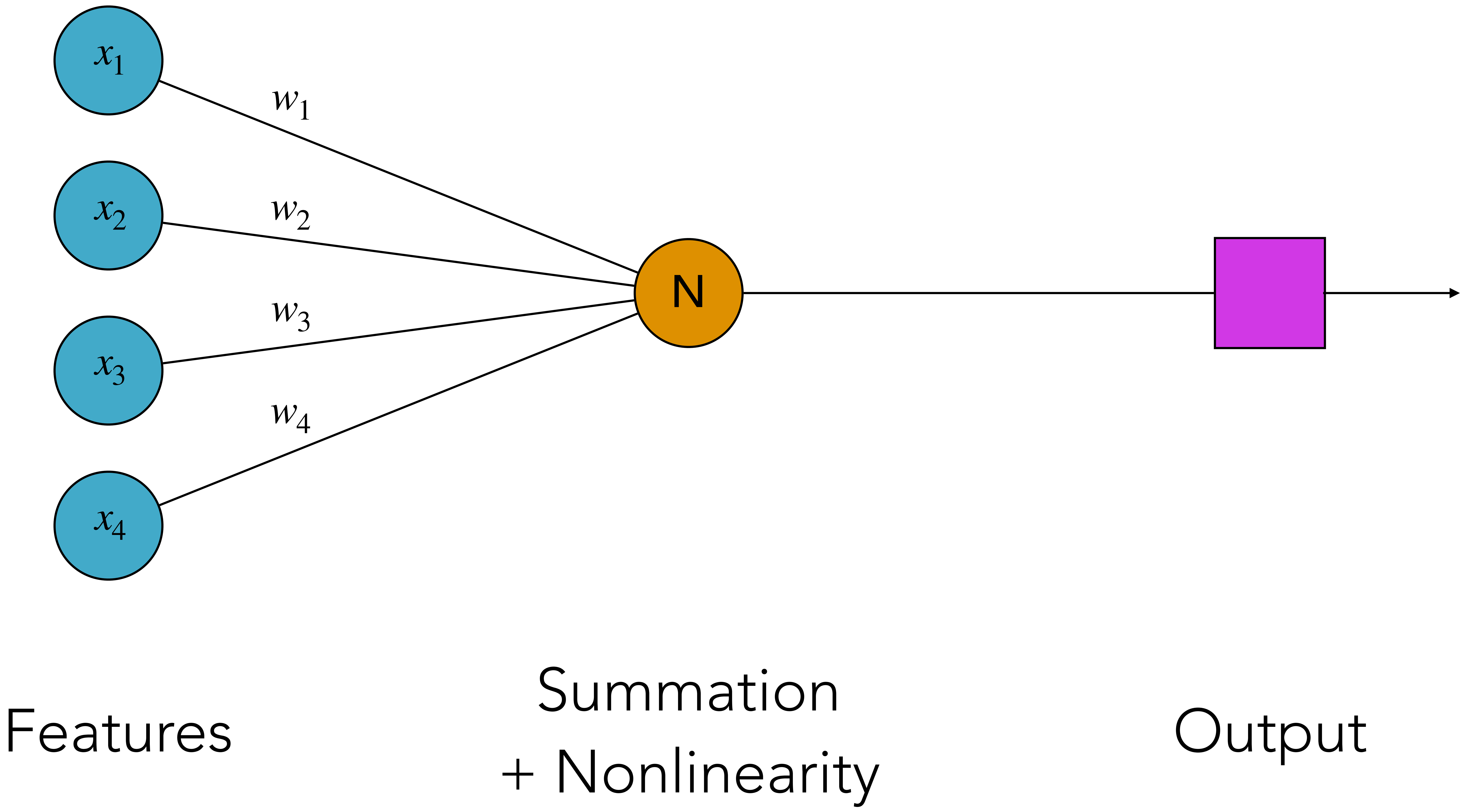
CLASSIFICATION

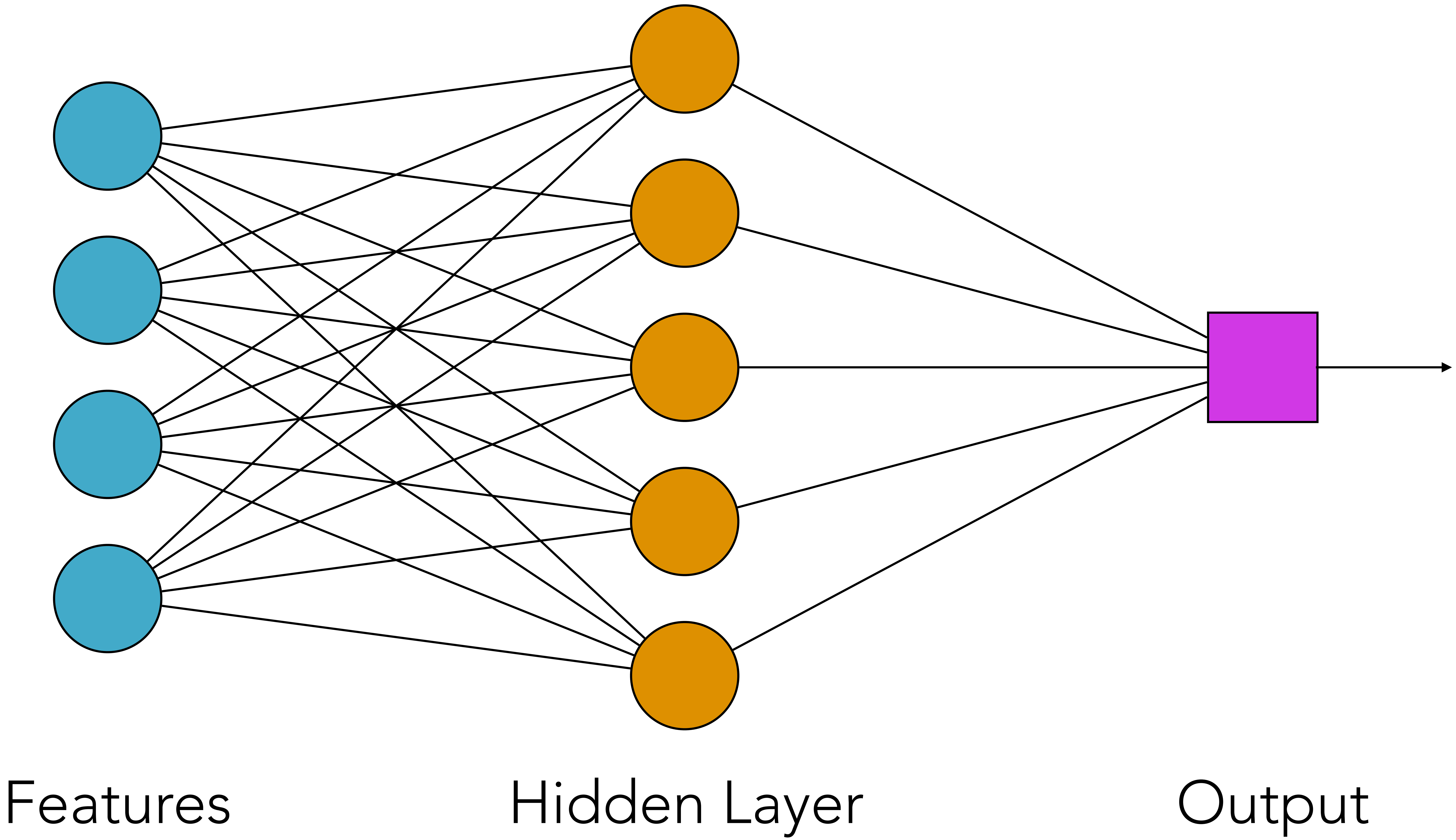


LOGISTIC REGRESSION



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


x_1
x_2
x_3
x_4
w_1
w_2
w_3
w_4
N
Features
Summation
+ Nonlinearity
Output


Features
Hidden Layer
Output

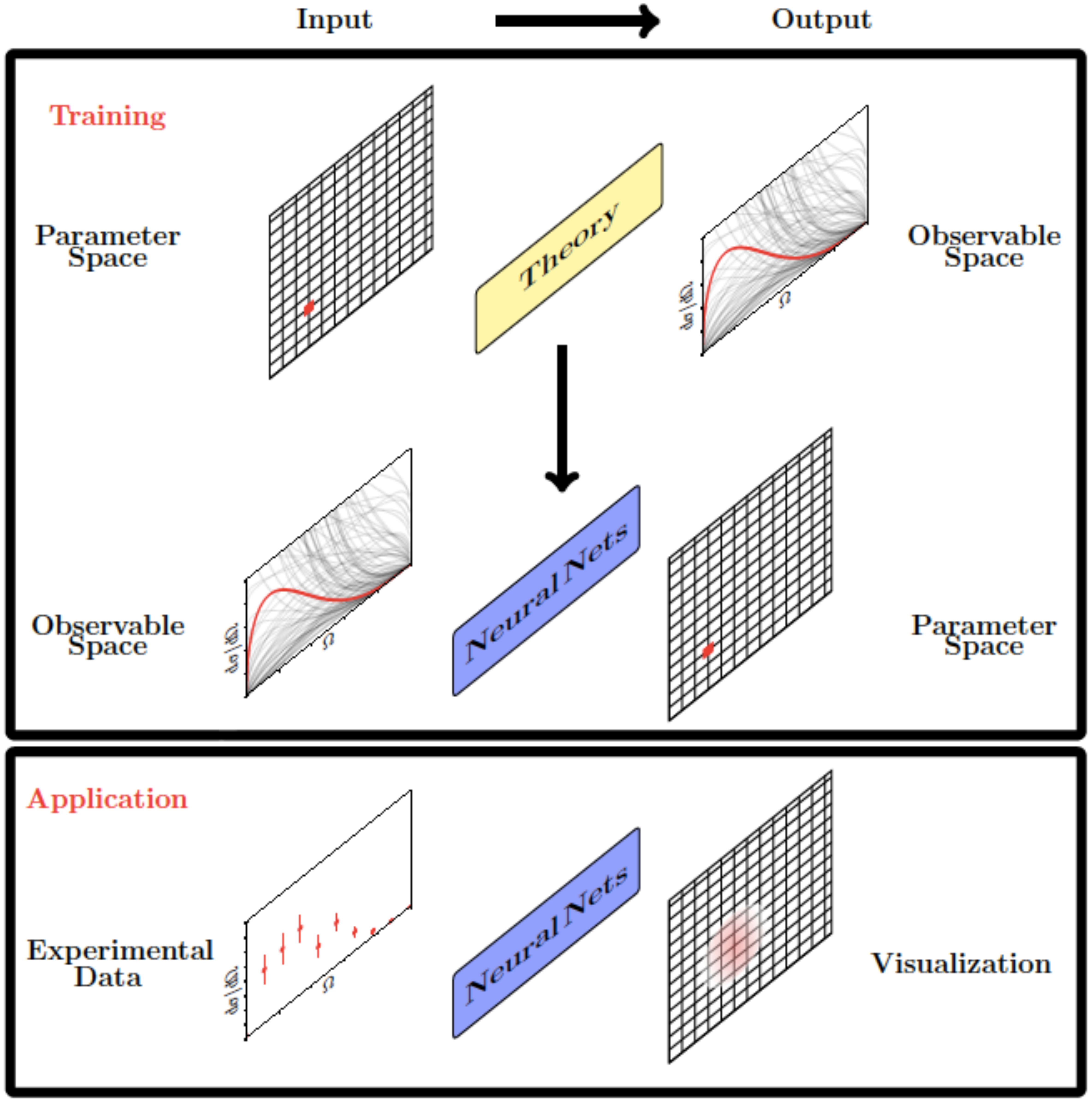
Application 1: How can experimental observables constrain theoretical models?

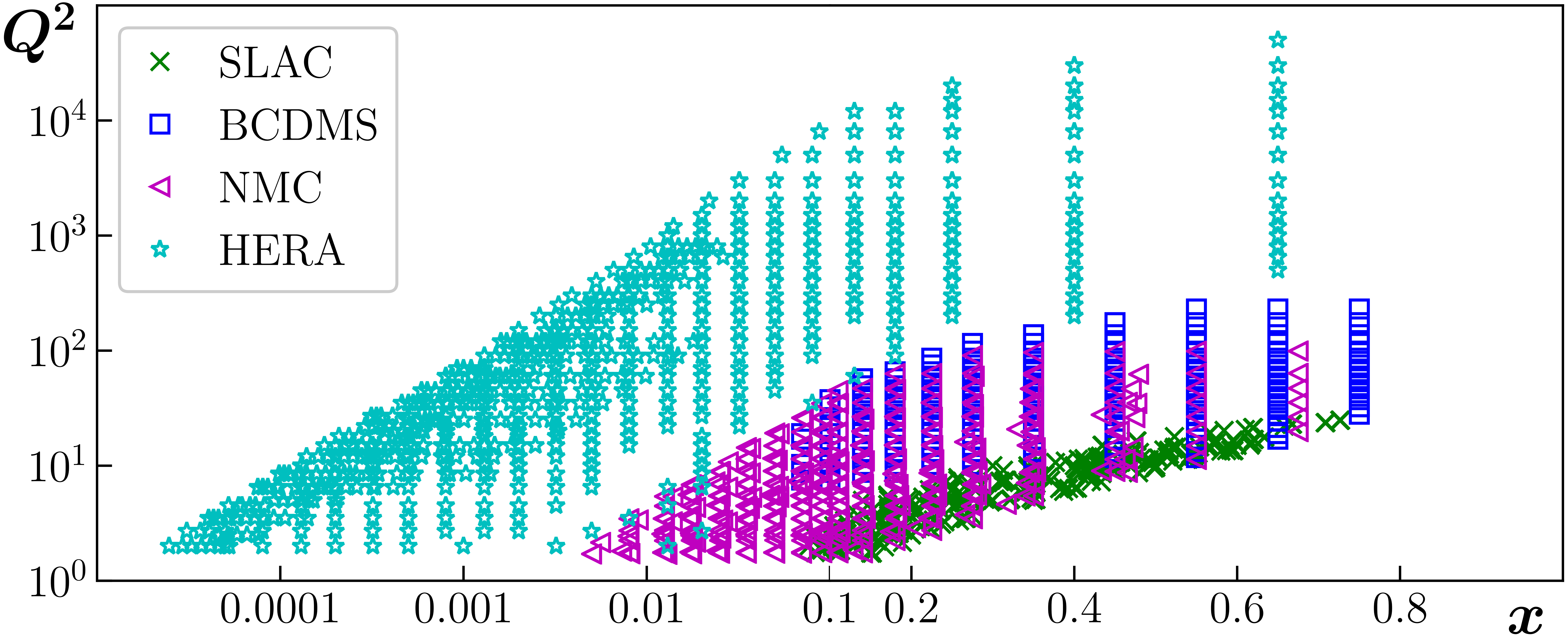
STRUCTURE OF THE NUCLEON

Quantum Chromodynamics

Quantum probability distributions (QPD) characterize the internal structure of a nucleon

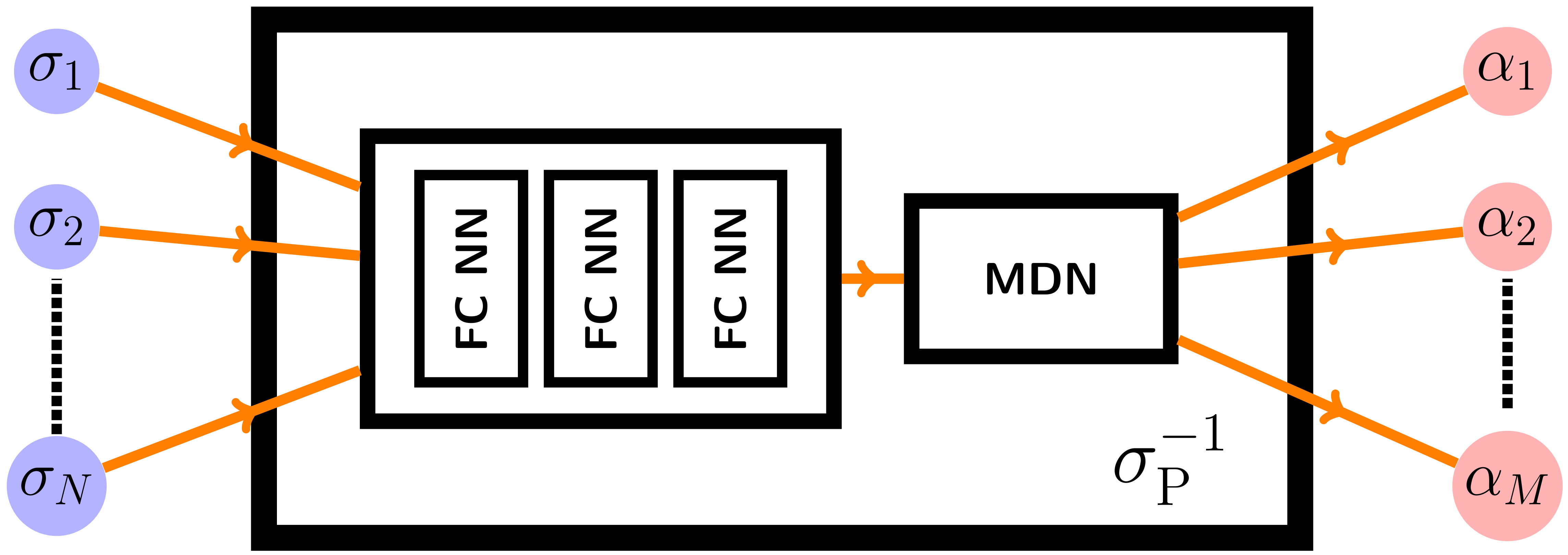
Can we prediction QPD parameters directly from experimental cross section data?

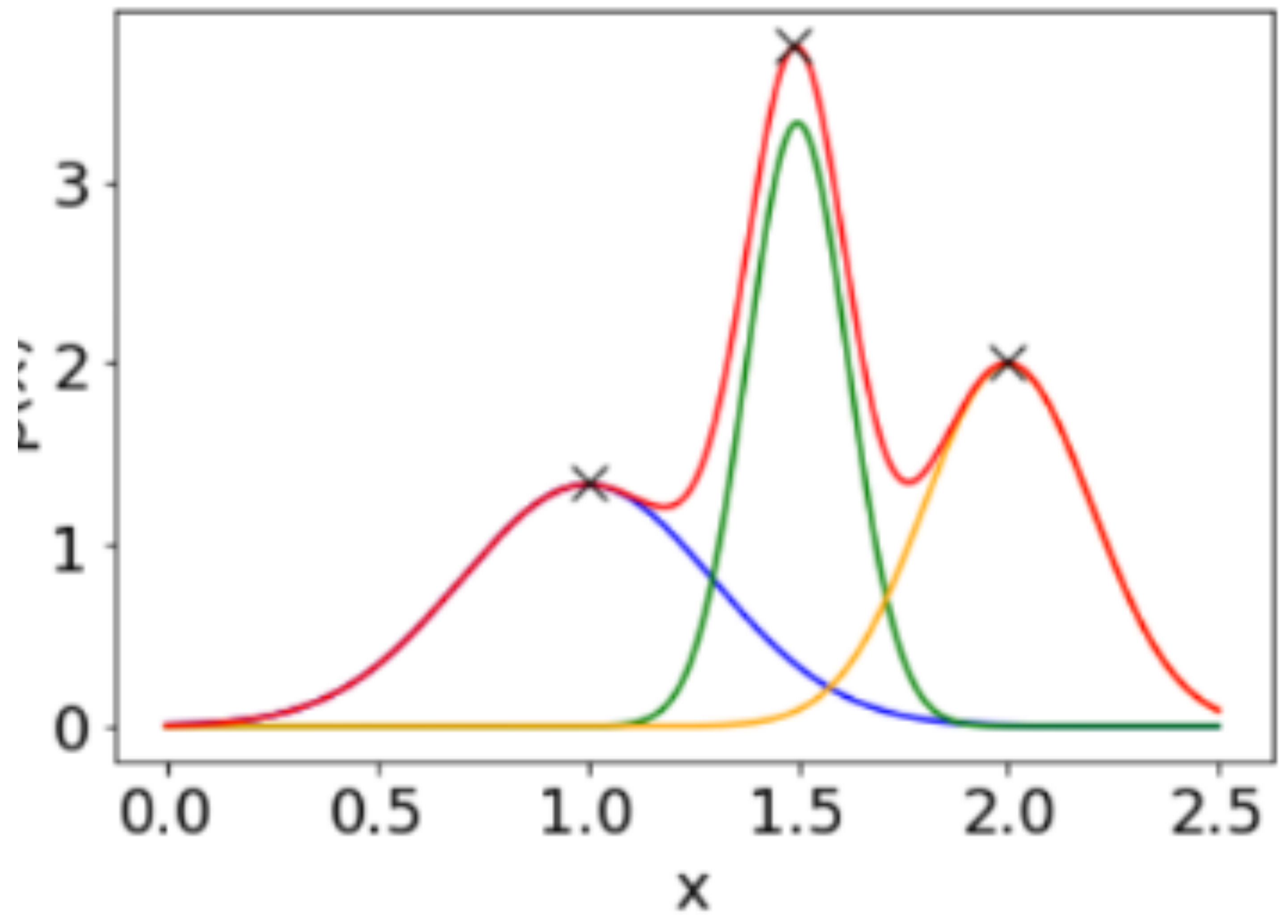




Q^2
SLAC
BCDMS
NMC
HERA
10^4
10^3
10^2
10^1
10^0
0.0001
0.001
0.01
0.1
0.2
0.4
0.6
0.8
x

MIXTURE DENSITY NETWORK

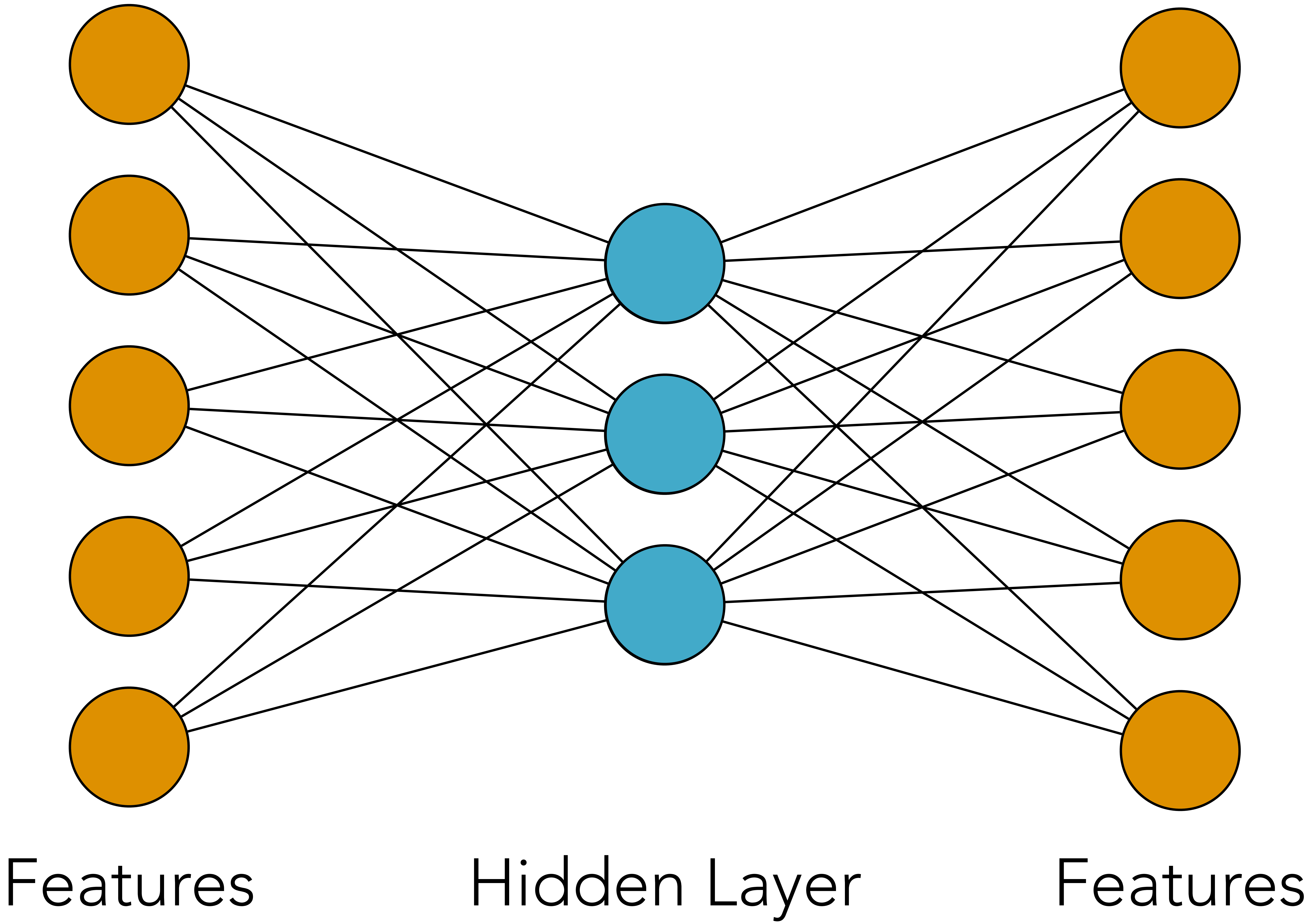




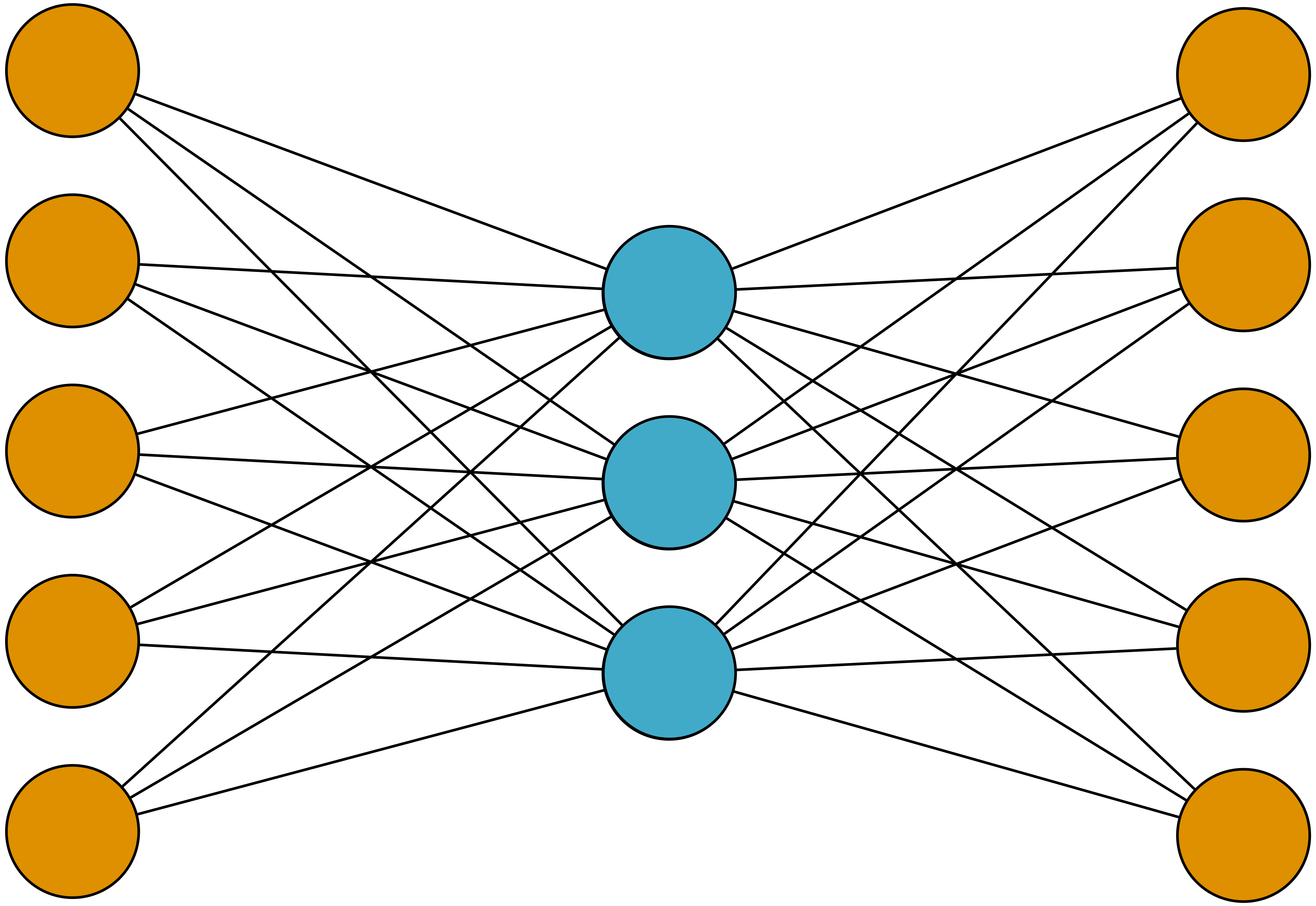
Output Layer Interpretation:

$$p(\mathbf{t}|\mathbf{x}) = \sum_{k=1}^{K} \pi_k(\mathbf{x}) \mathcal{N}\left(\mathbf{t}|\boldsymbol{\mu}_k(\mathbf{x}), \sigma_k^2(\mathbf{x})\right)$$

AUTOENCODER

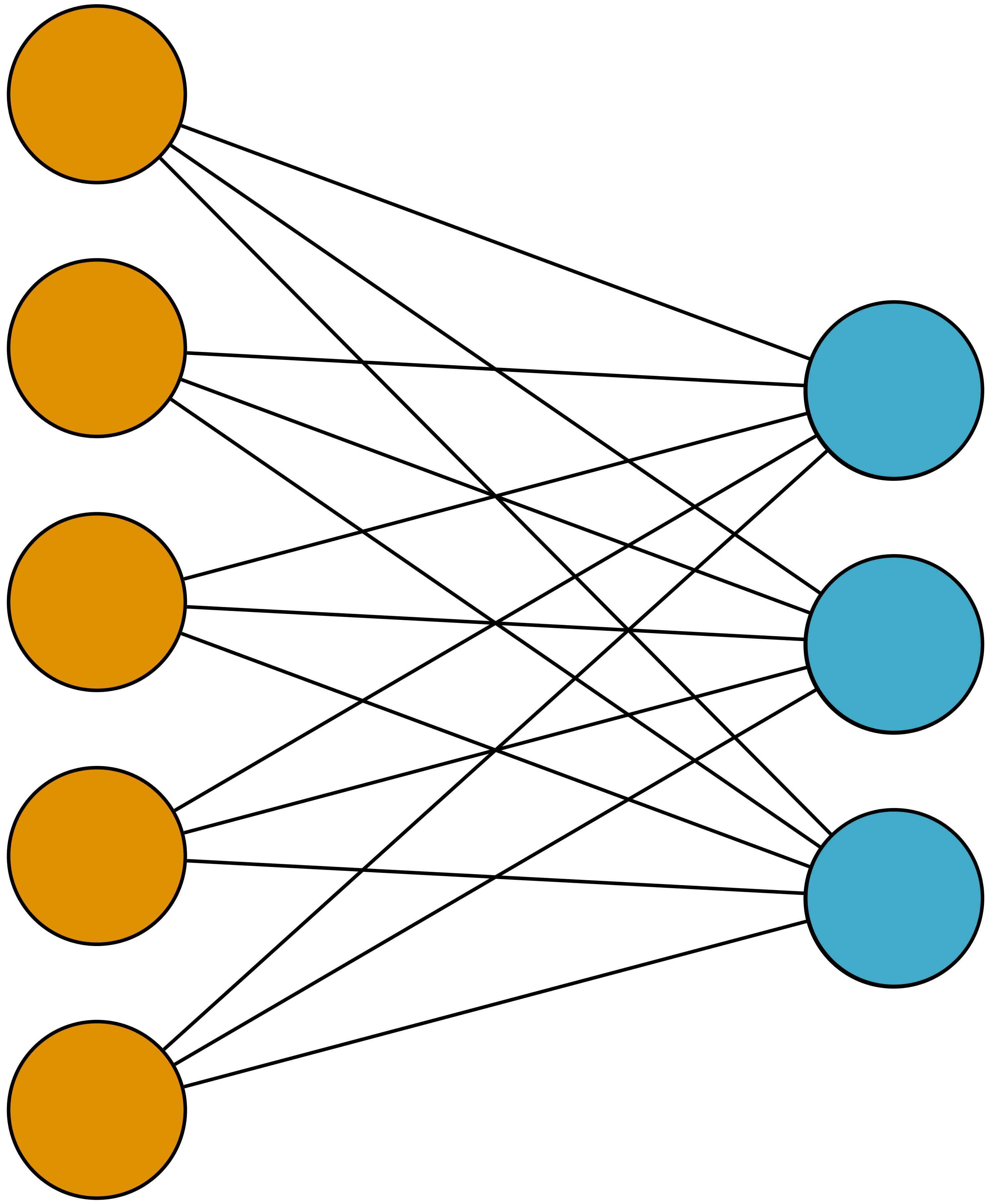


AUTOENCODER



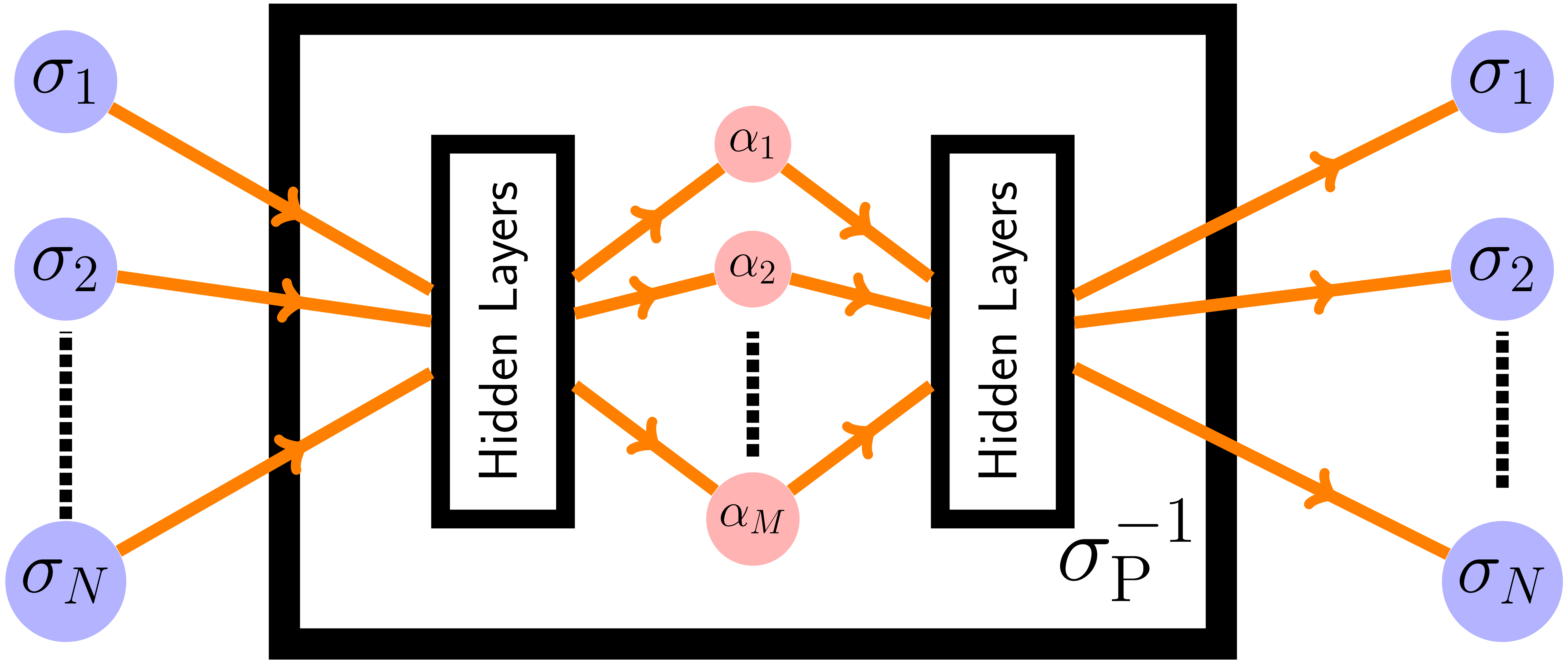
Features **LATENT SPACE** Features

AUTOENCODER: DIMENSIONALITY REDUCTION



Features **LATENT SPACE**

PARAMETER-SUPERVISED AUTOENCODER (PSA)



LATENT SPACE

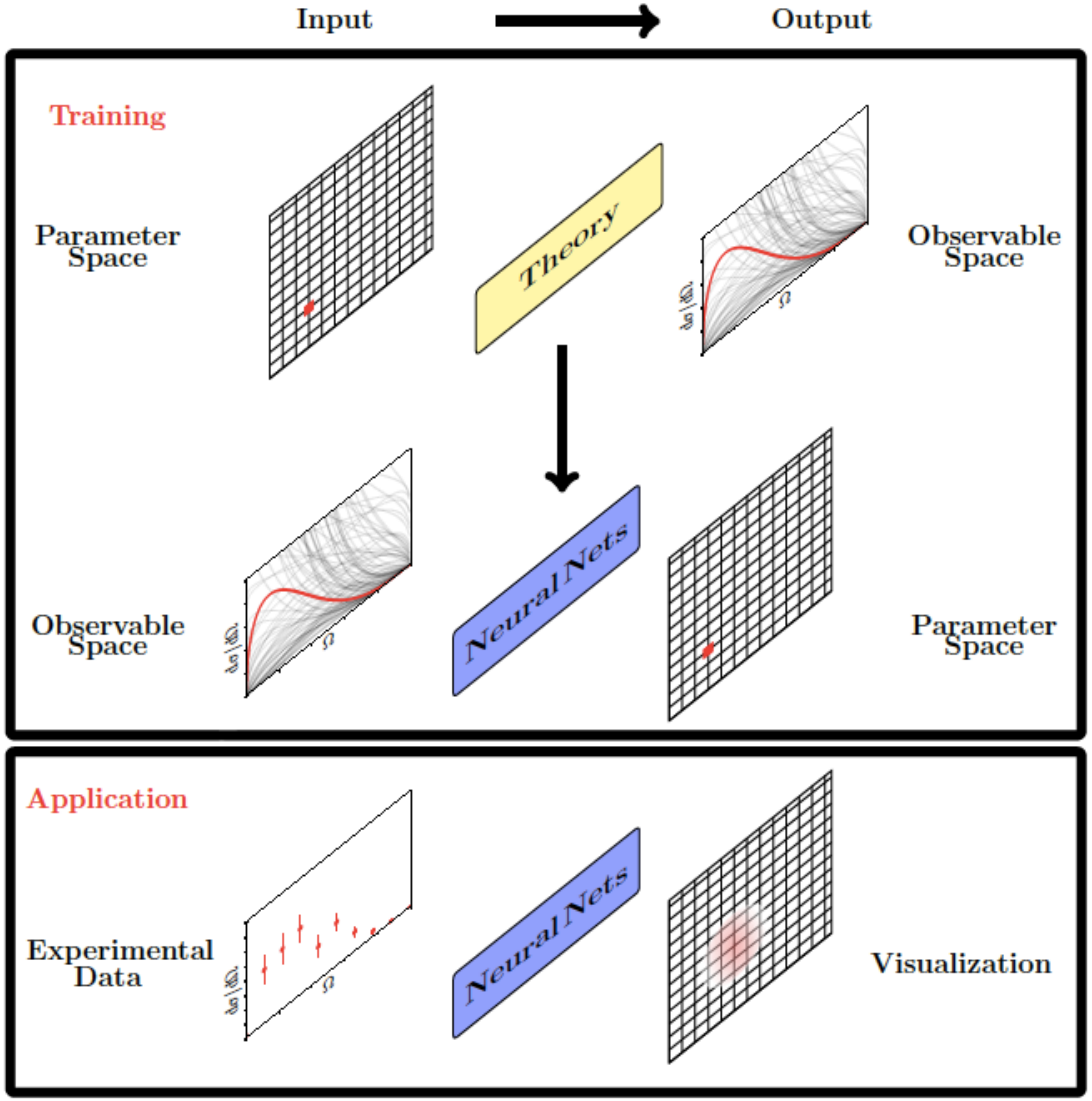
parameters

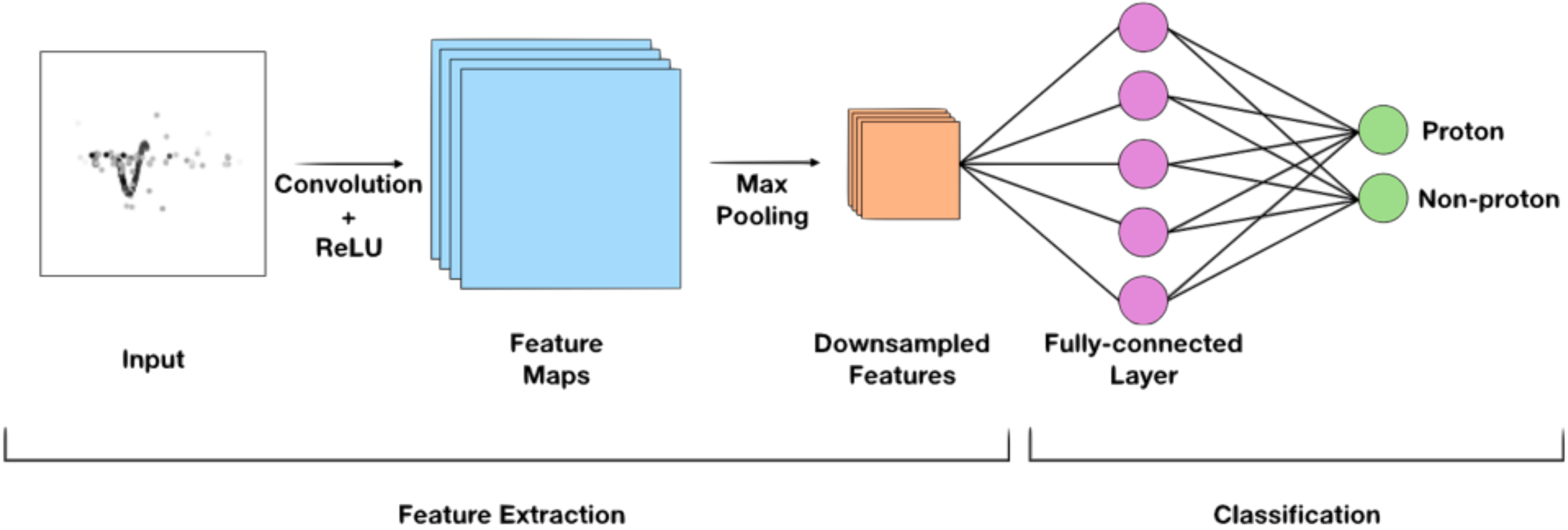
STRUCTURE OF THE NUCLEON

Quantum Chromodynamics

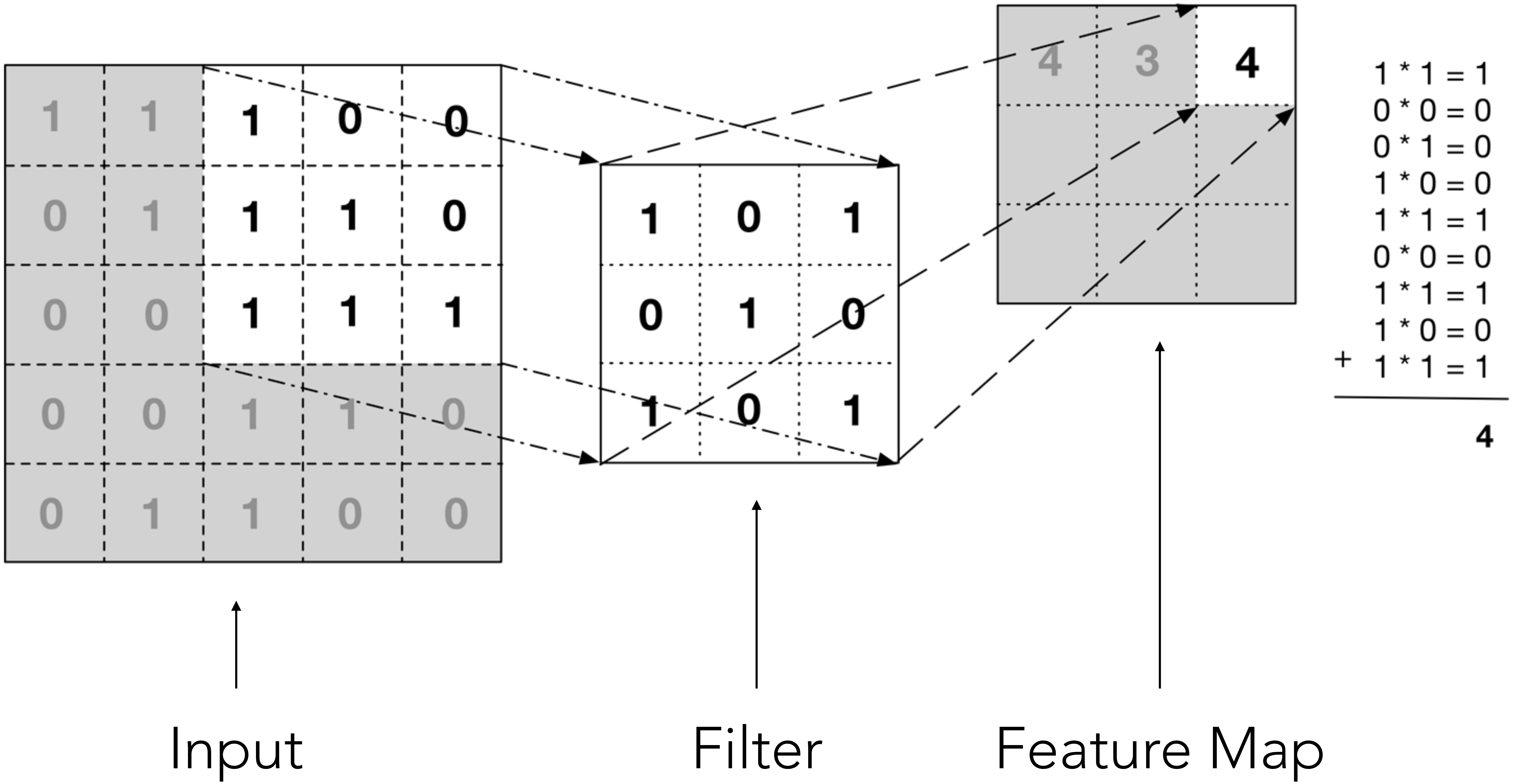
Quantum probability distributions (QPD) characterize the internal structure of a nucleon

Can we prediction QPD parameters directly from experimental cross section data?

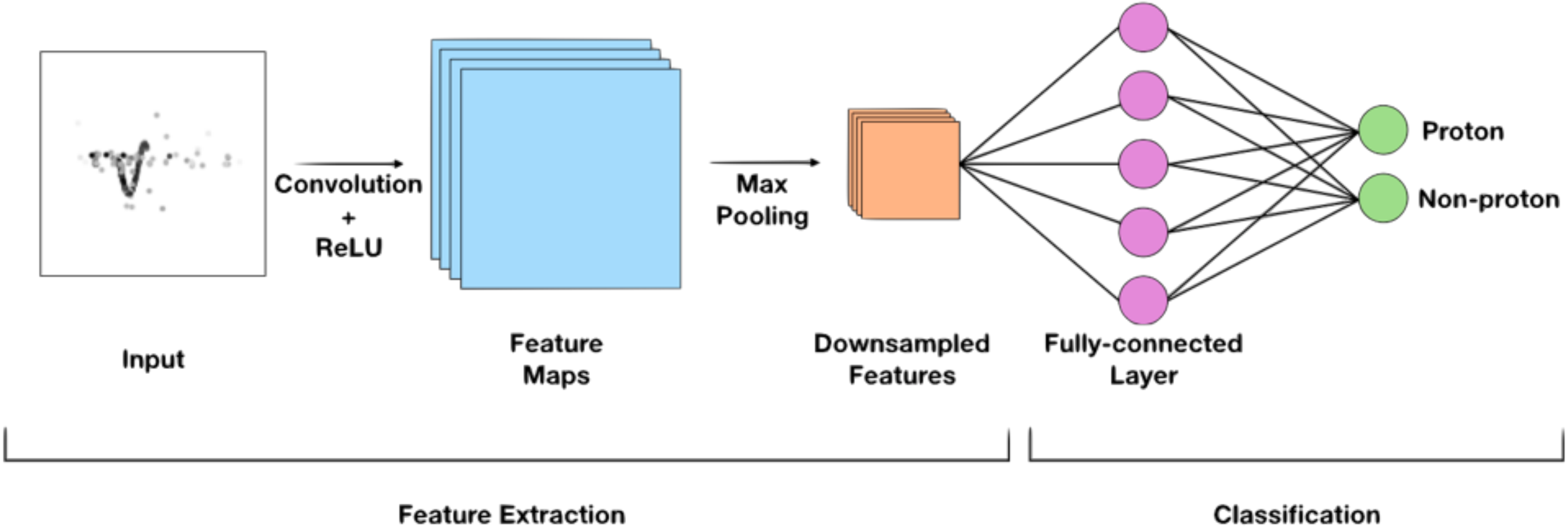



Convolution
+
ReLU
Max
Pooling
Proton
Non-proton
Input
Feature
Maps
Downsampled
Features
Fully-connected
Layer
Feature Extraction
Classification

DISCRETE CONVOLUTION

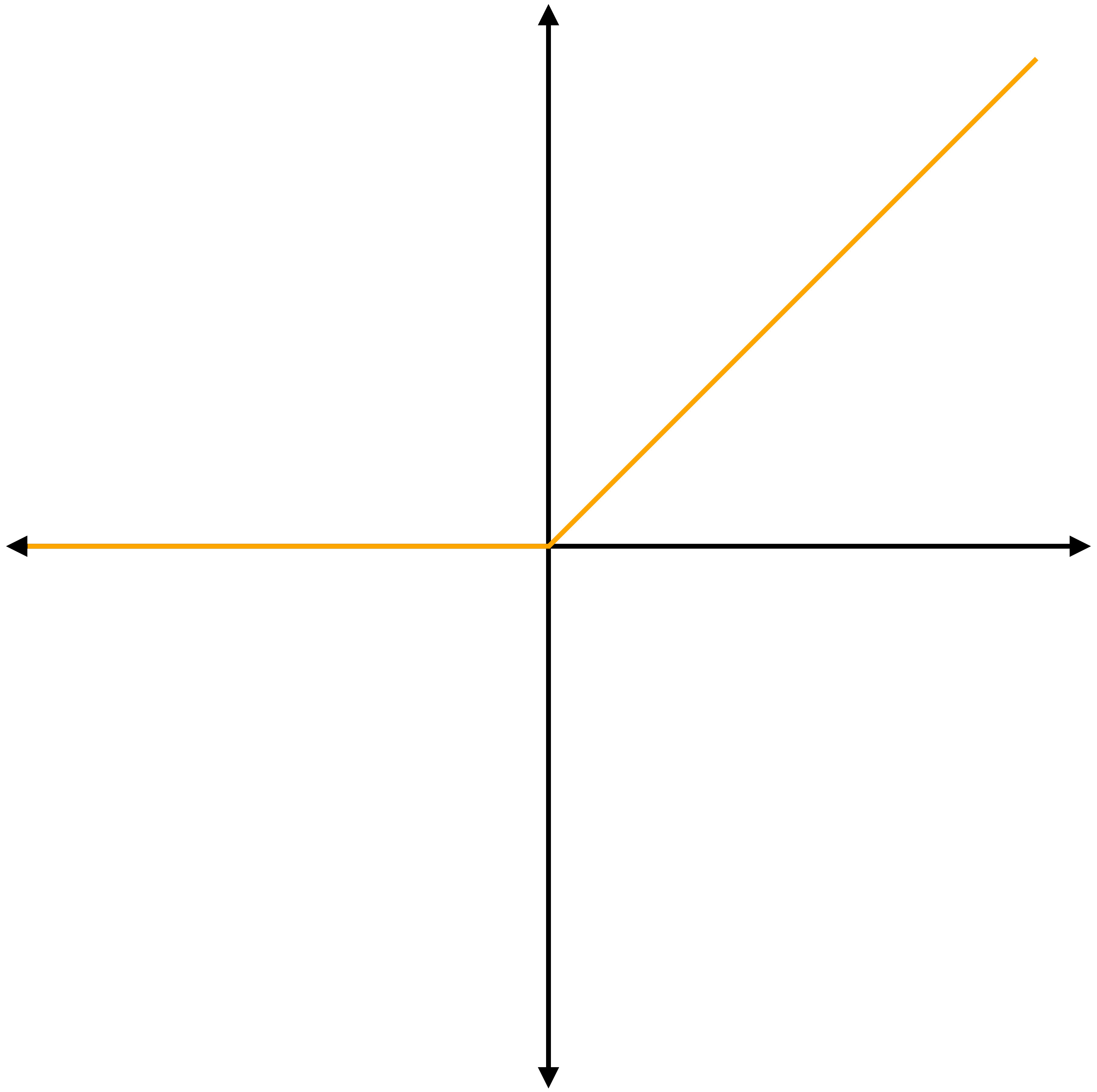


ADAPTED FROM *DEEP LEARNING*, ADAM GIBSON & JOSH PATTERSON



Input
Convolution
+
ReLU
Feature
Maps
Max
Pooling
Downsampled
Features
Fully-connected
Layer
Proton
Non-proton
Feature Extraction
Classification

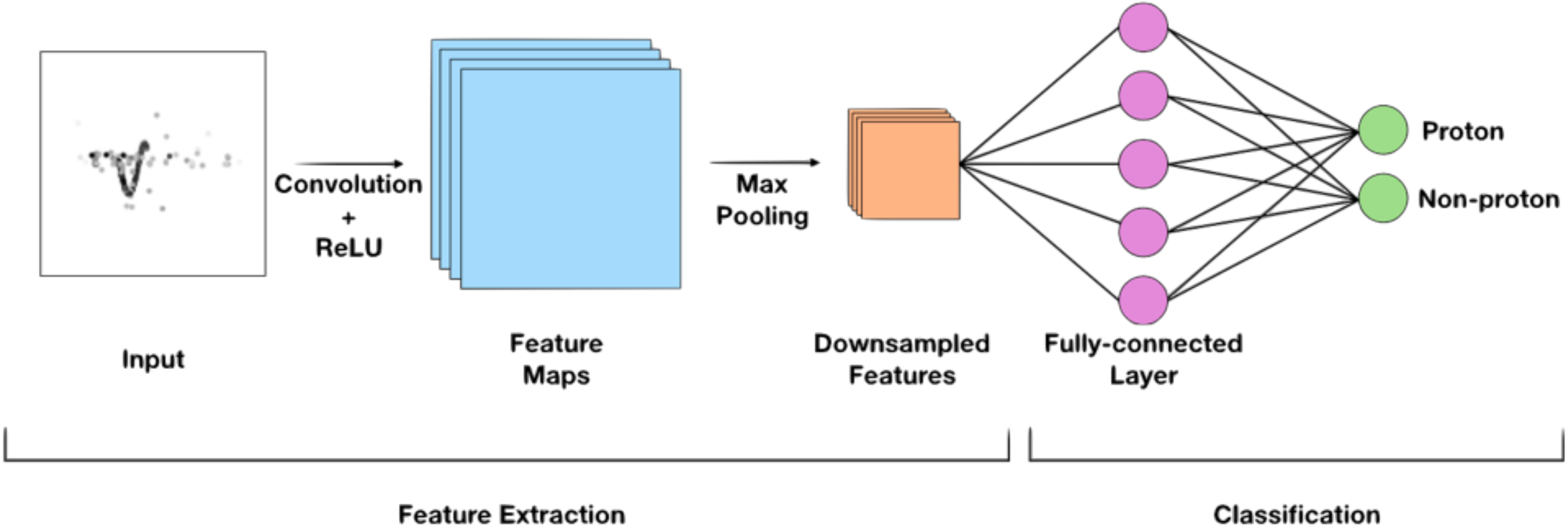
RECTIFIED LINEAR UNIT (ReLU)



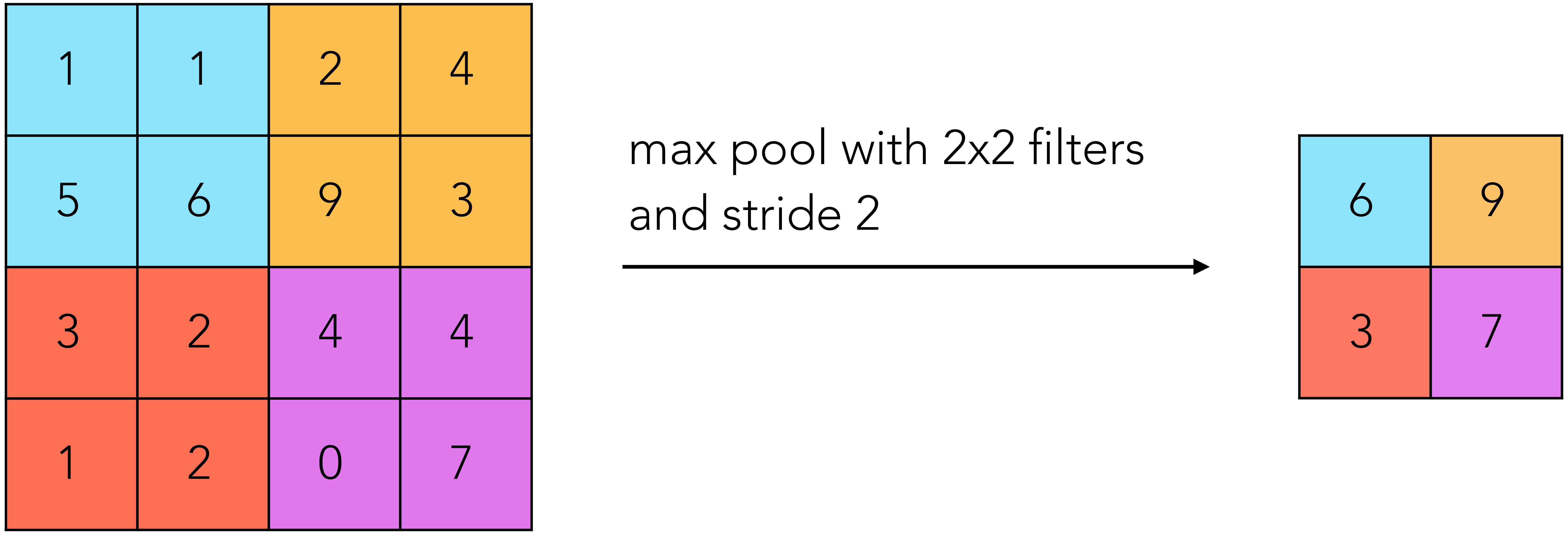

1156

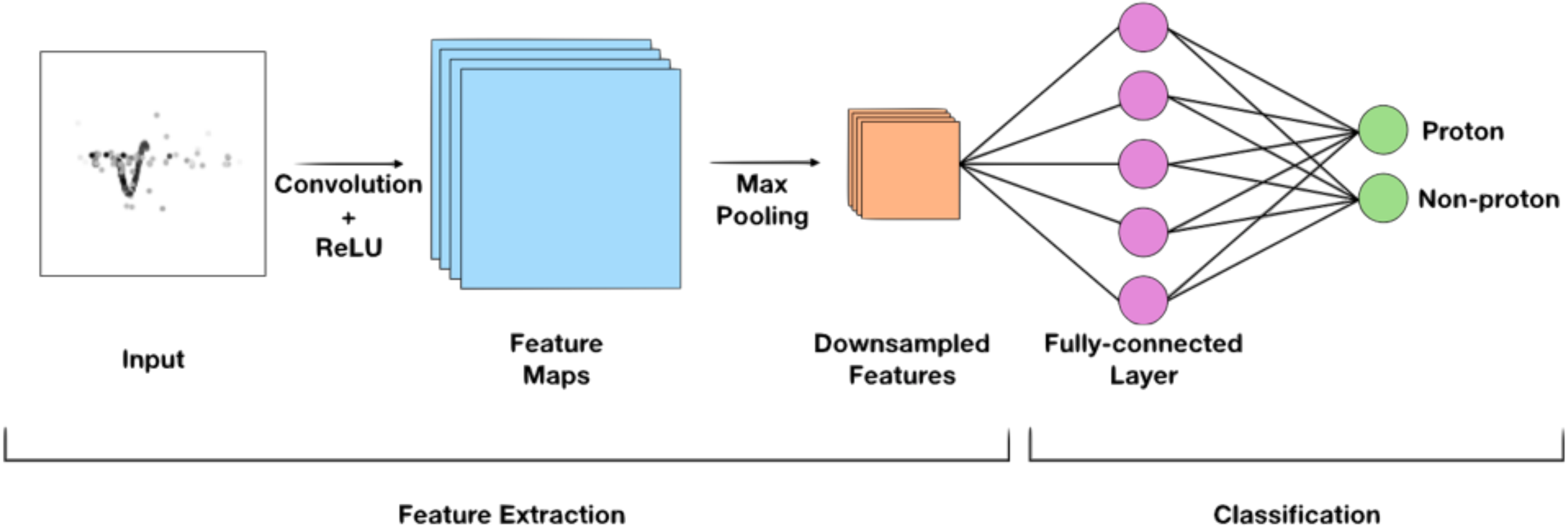
-1	-1	-1	-1	-1
-1	-1	-1	-1	-1
5	5	5	5	5
-1	-1	-1	-1	-1
-1	-1	-1	-1	-1

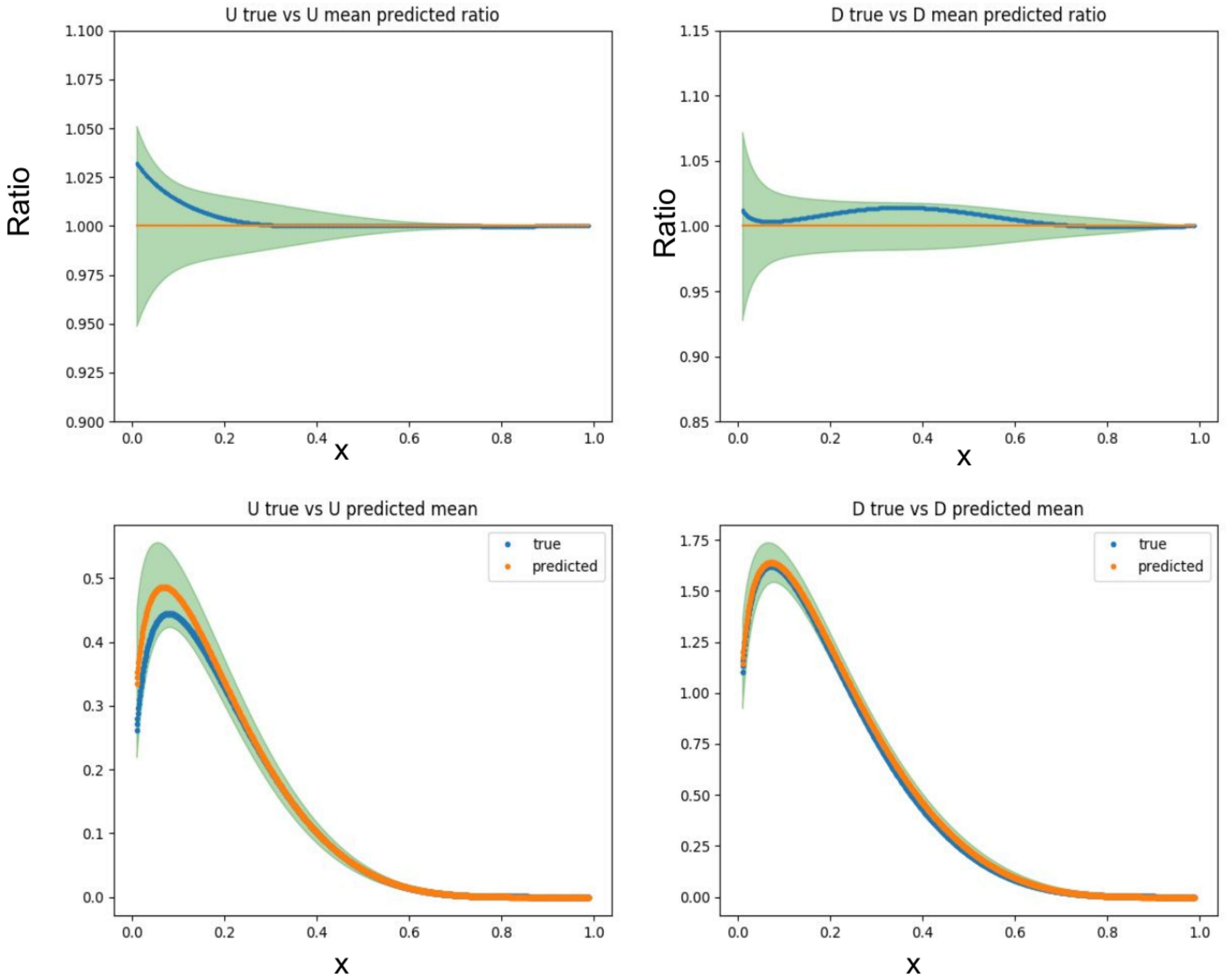
-1	-1	5	-1	-1
-1	-1	5	-1	-1
-1	-1	5	-1	-1
-1	-1	5	-1	-1
-1	-1	5	-1	-1


Convolution
+
ReLU
Max
Pooling
Proton
Non-proton
Input
Feature
Maps
Downsampled
Features
Fully-connected
Layer
Feature Extraction
Classification

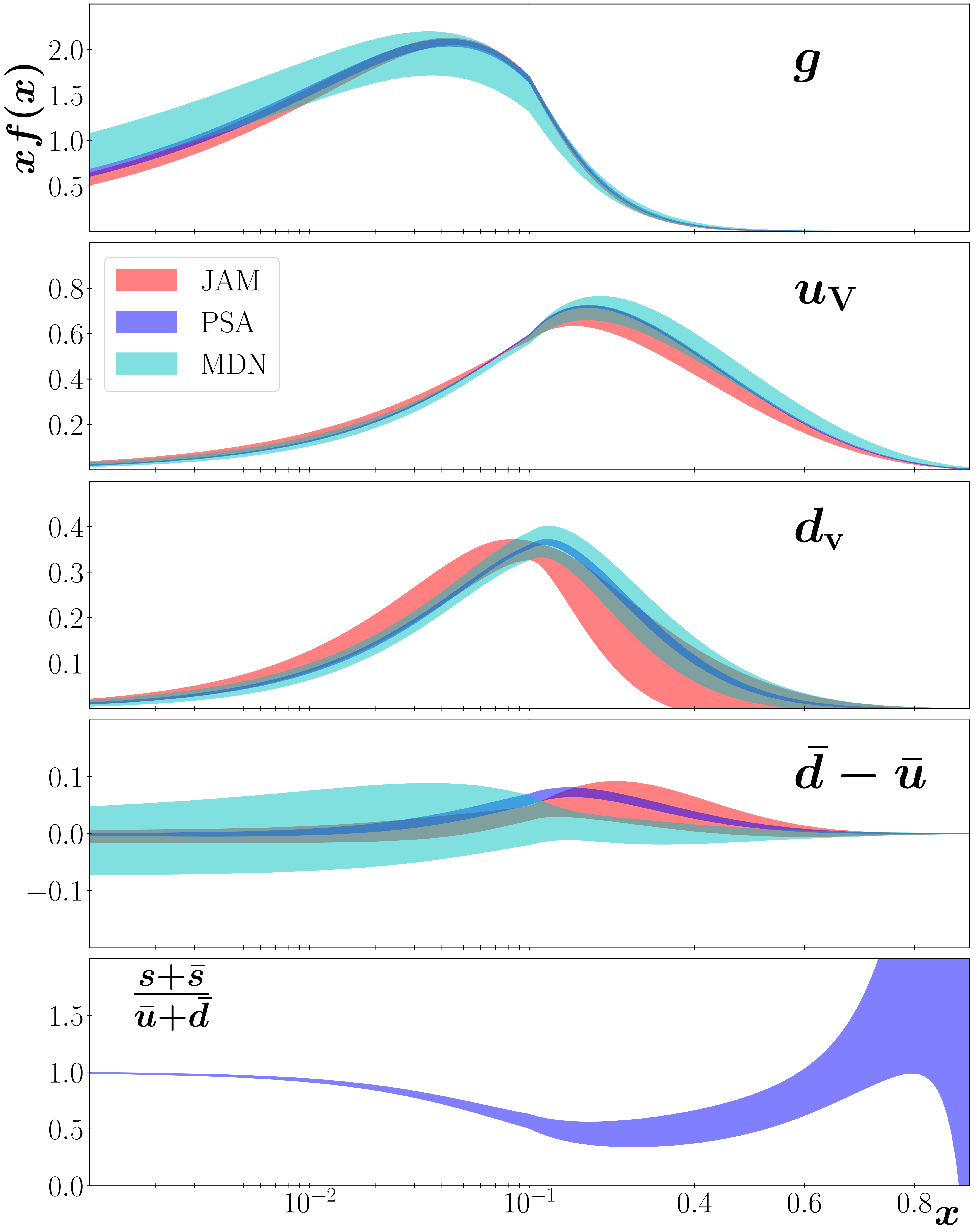
MAX POOLING


1 1 2 4
5 6 9 3
3 2 4 4
1 2 0 7
max pool with 2x2 filters
and stride 2
6 9
3 7

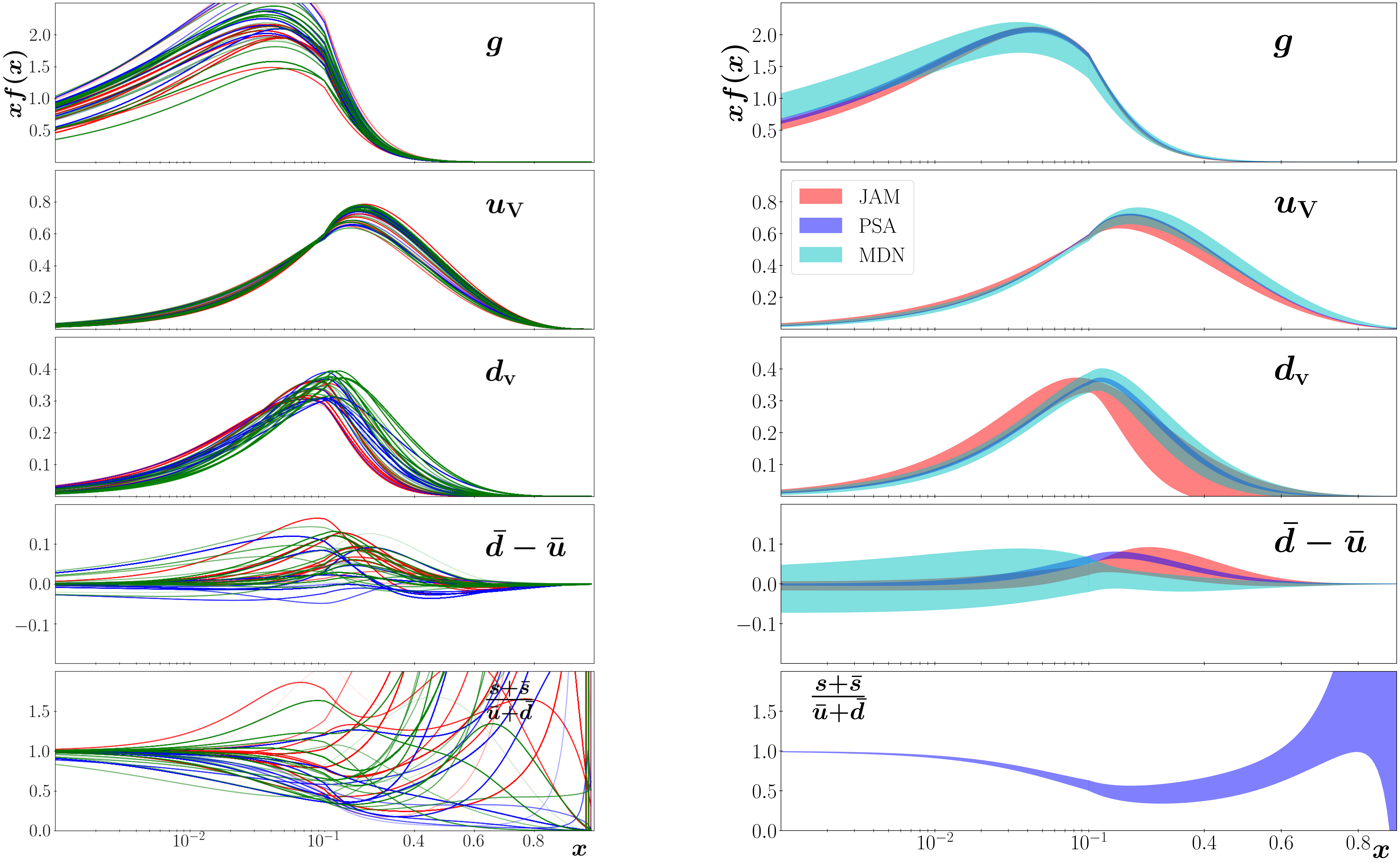

Convolution
+
ReLU
Max
Pooling
Proton
Non-proton
Input
Feature
Maps
Downsampled
Features
Fully-connected
Layer
Feature Extraction
Classification



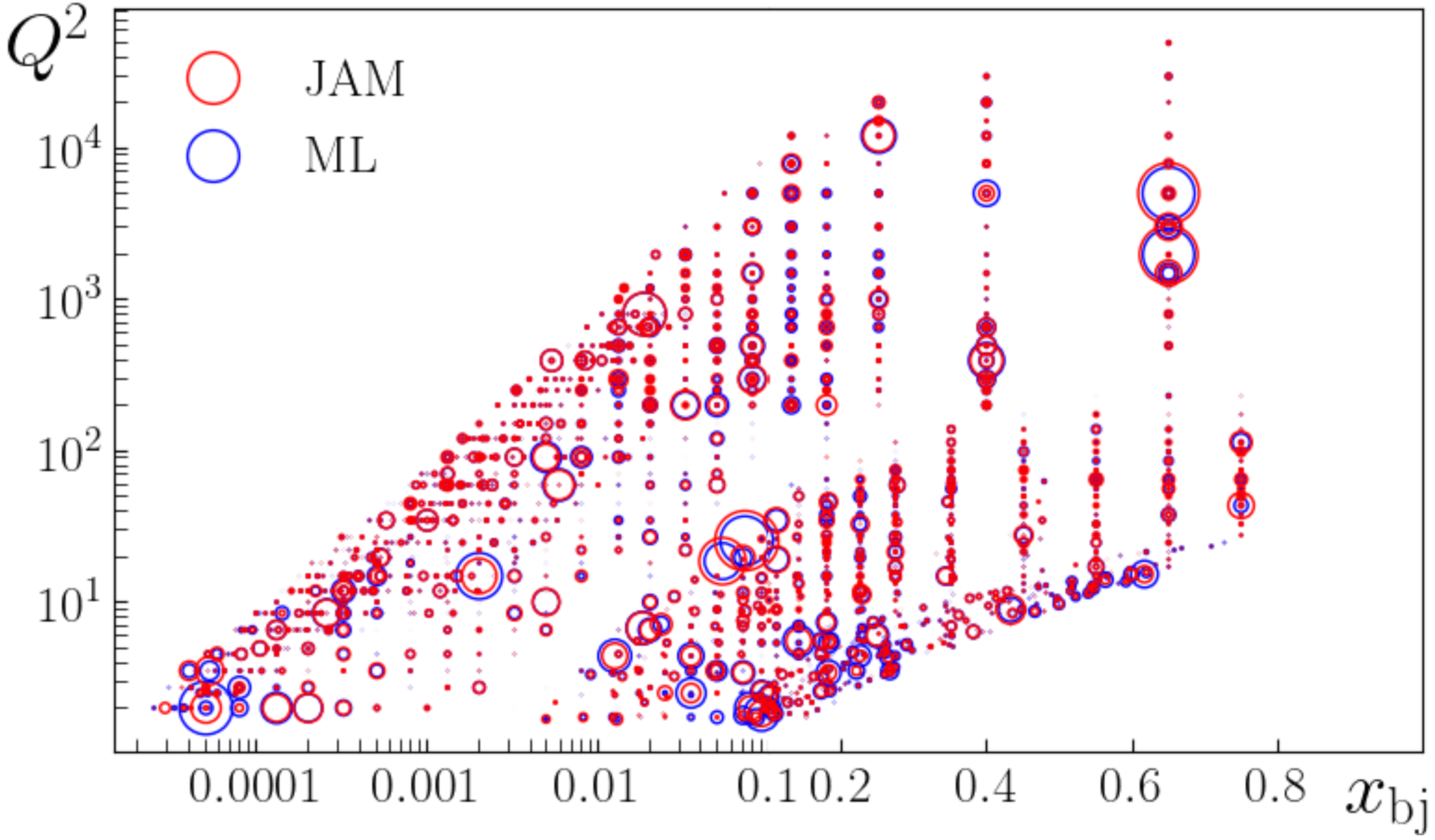
U true vs U mean predicted ratio
D true vs D mean predicted ratio
U true vs U predicted mean
D true vs D predicted mean
Ratio
x
true
predicted



$xf(x)$
g
u_V
d_v
$\bar{d}-\bar{u}$
$\frac{s+\bar{s}}{\bar{u}+\bar{d}}$
JAM
PSA
MDN
x



$xf(x)$
g
u_V
d_v
$\bar{d}-\bar{u}$
$\frac{s+\bar{s}}{\bar{u}+\bar{d}}$
JAM
PSA
MDN
10^{-2}
10^{-1}
0.4
0.6
0.8
x



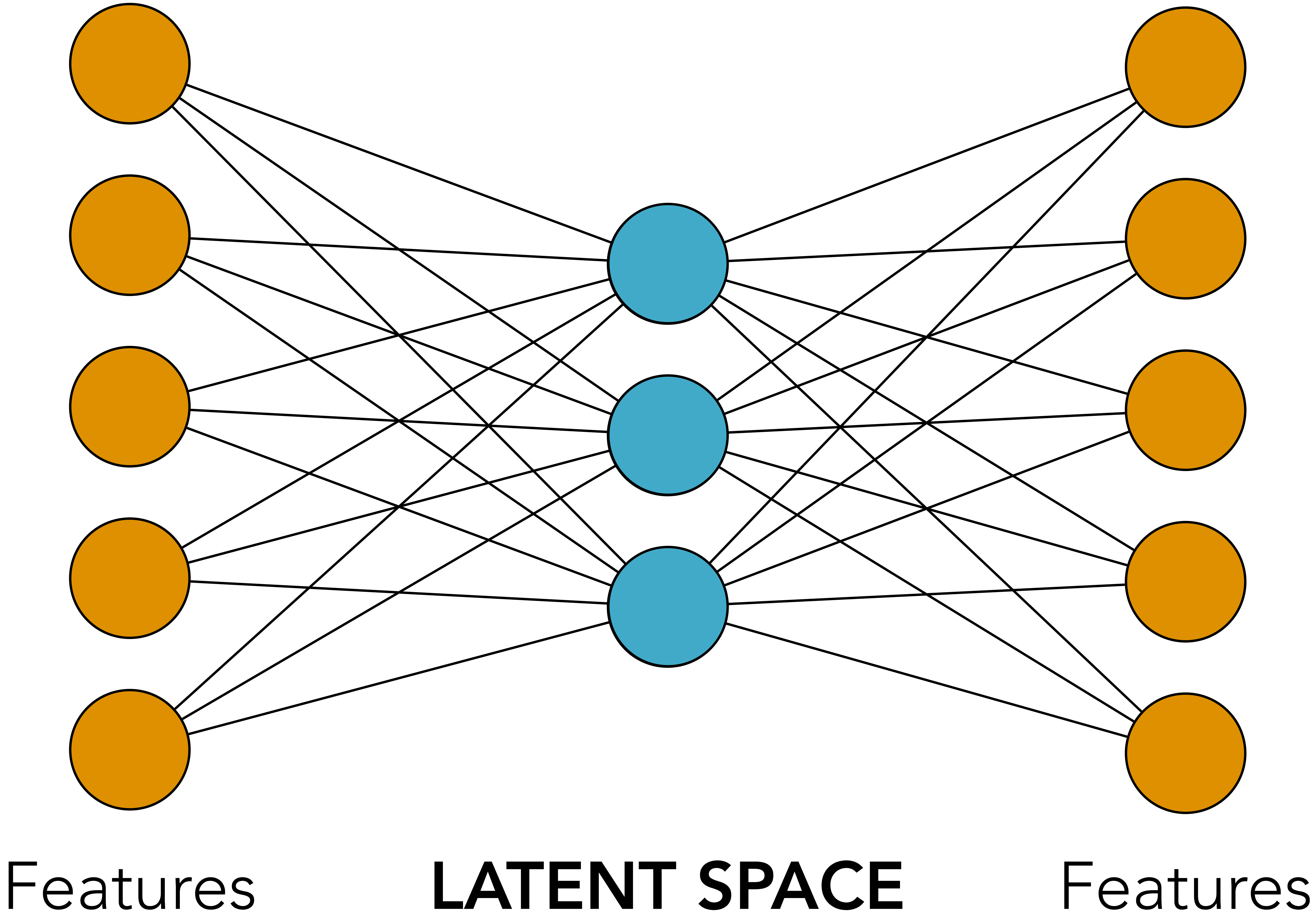
Q^2
JAM
ML
10^4
10^3
10^2
10^1
0.0001
0.001
0.01
0.1
0.2
0.4
0.6
0.8
x_{bj}

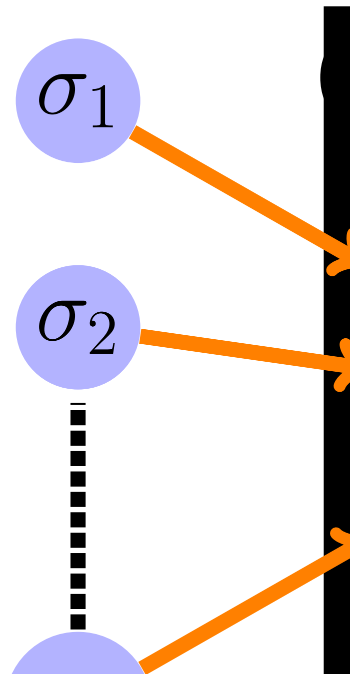
Application 2: Can we use machine learning to **simulate data**?

GENERATIVE MODELS

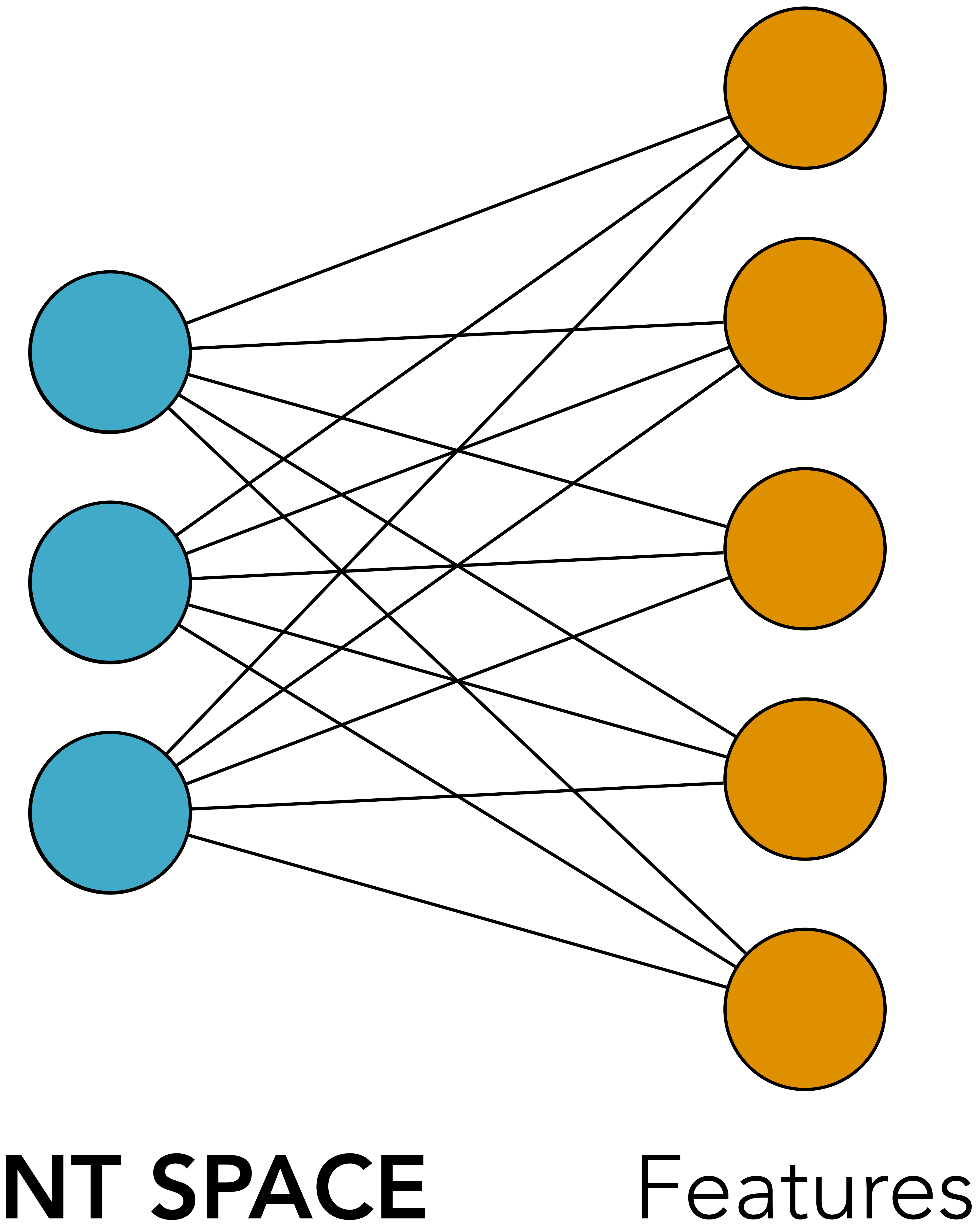
SIMULATION, EVENT GENERATION

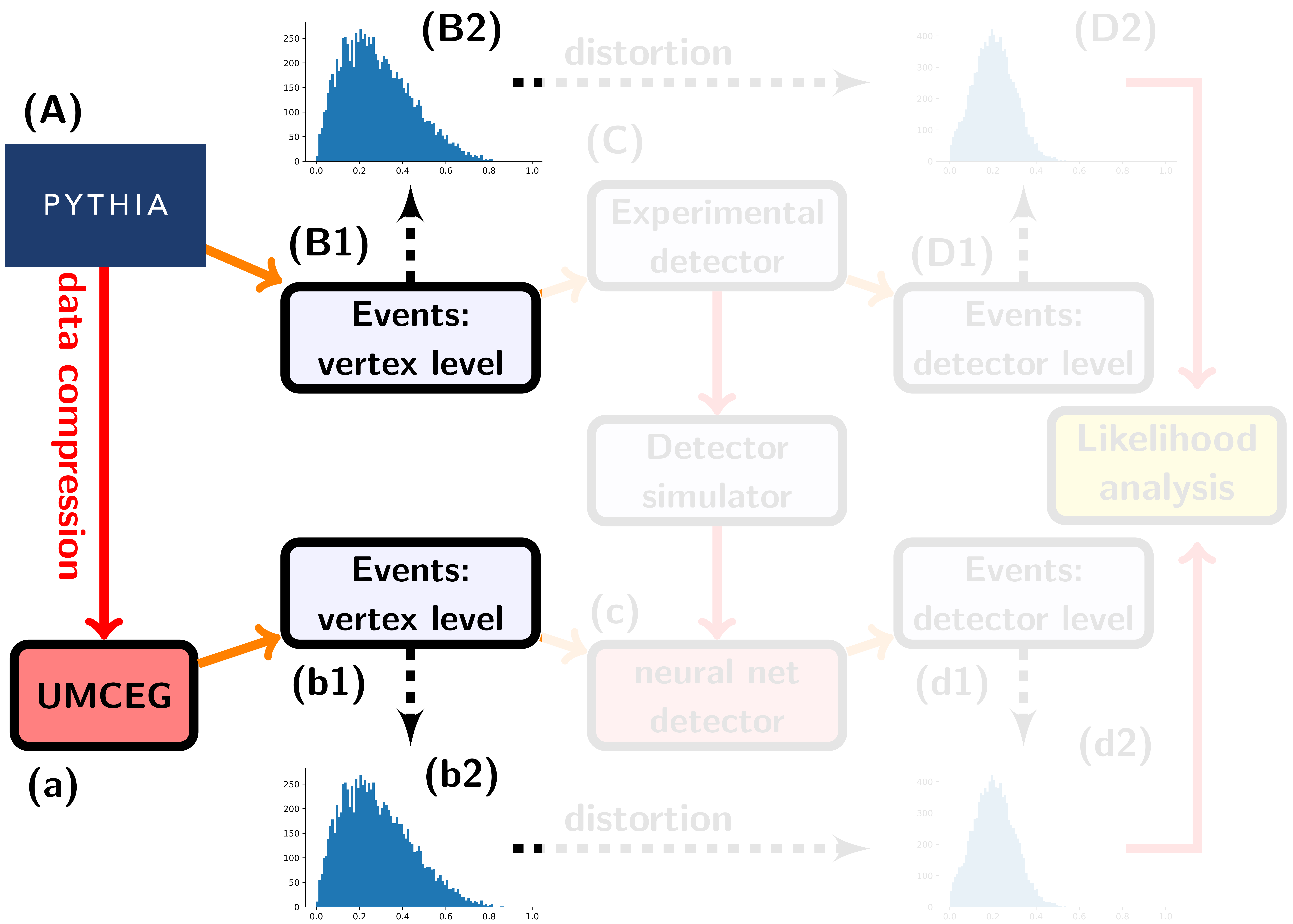
AUTOENCODER





AUTOENCODER: GENERATOR

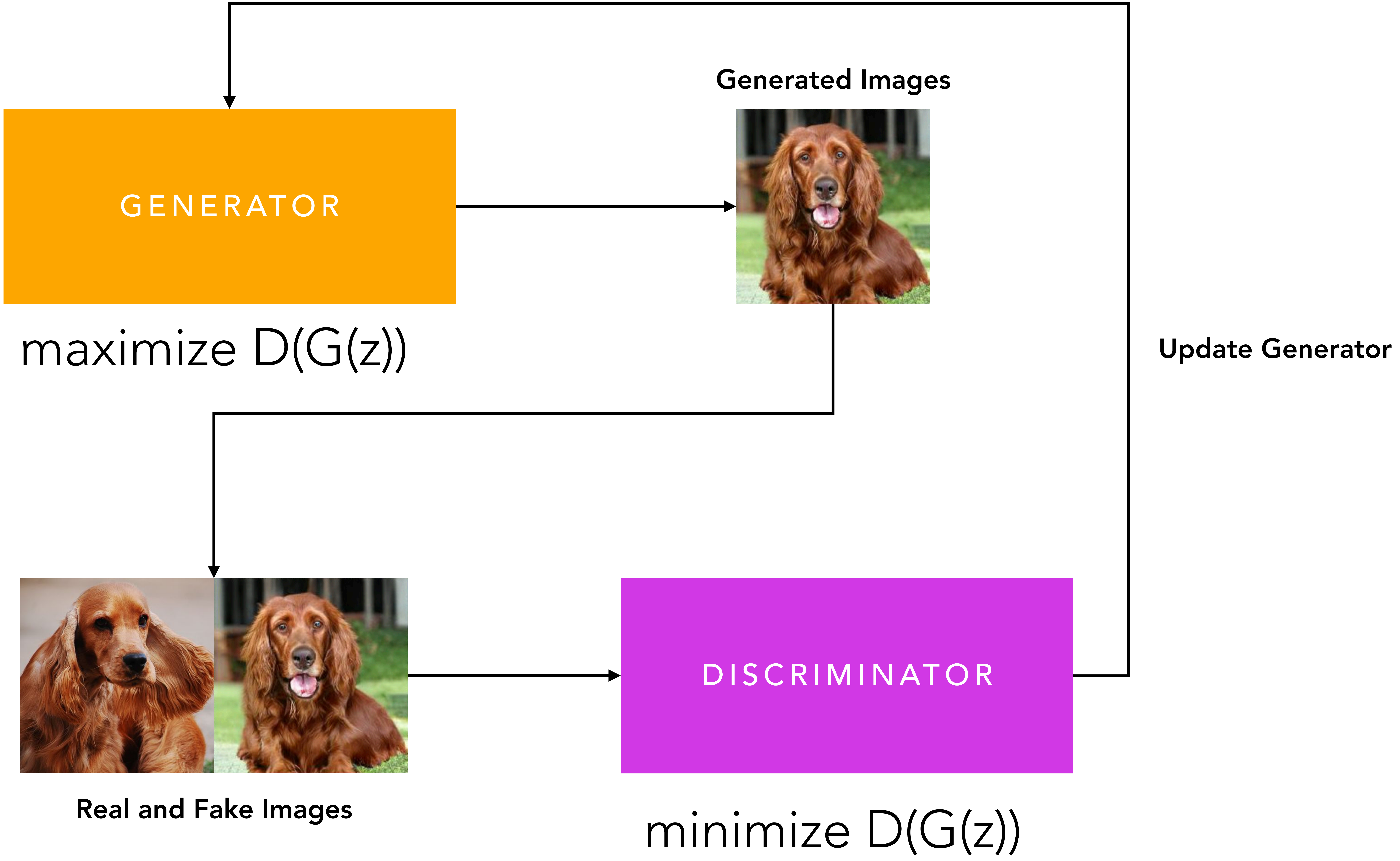




NOBUO SATO, JLAB

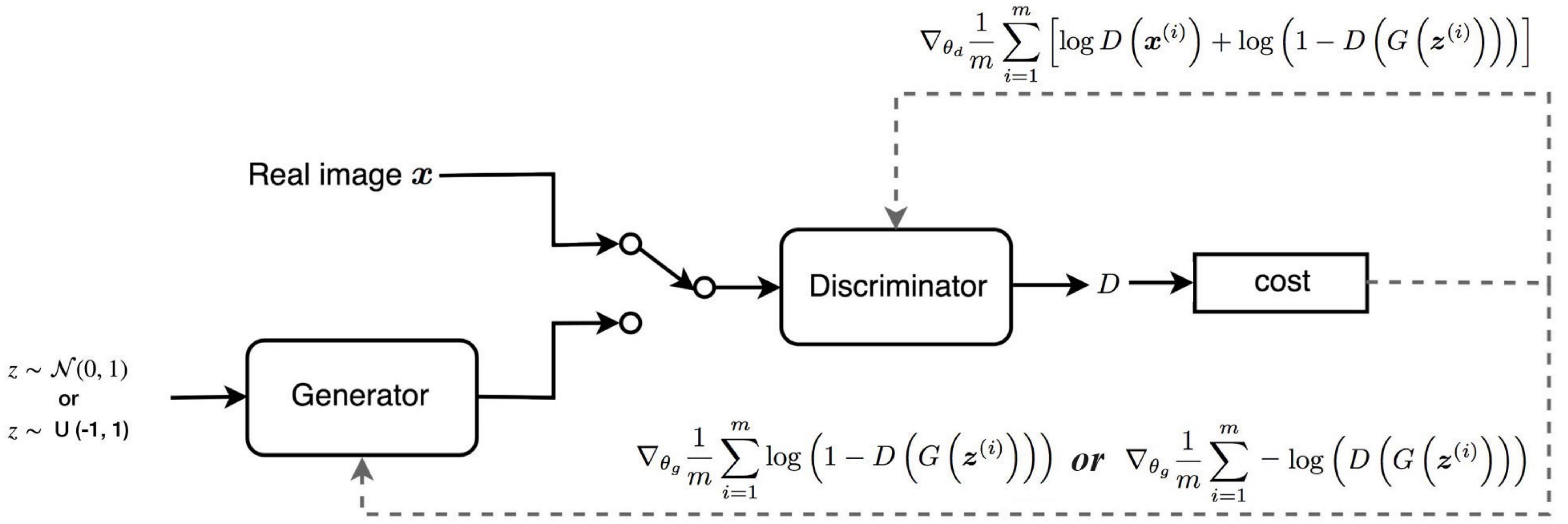
GENERATIVE ADVERSARIAL NETWORKS (GANS)

SIMULATION, EVENT GENERATION

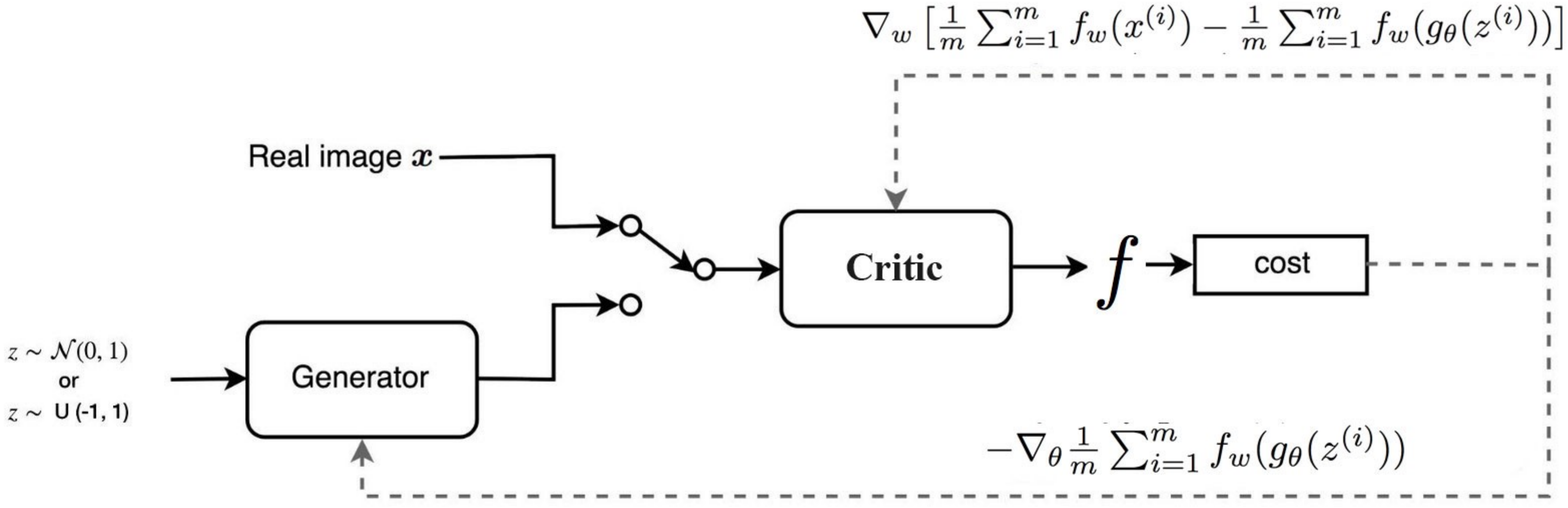


PHOTOS FROM KATRINA S; ANDREW B ET. AL., *BIGGAN*.

GAN (DCGAN)



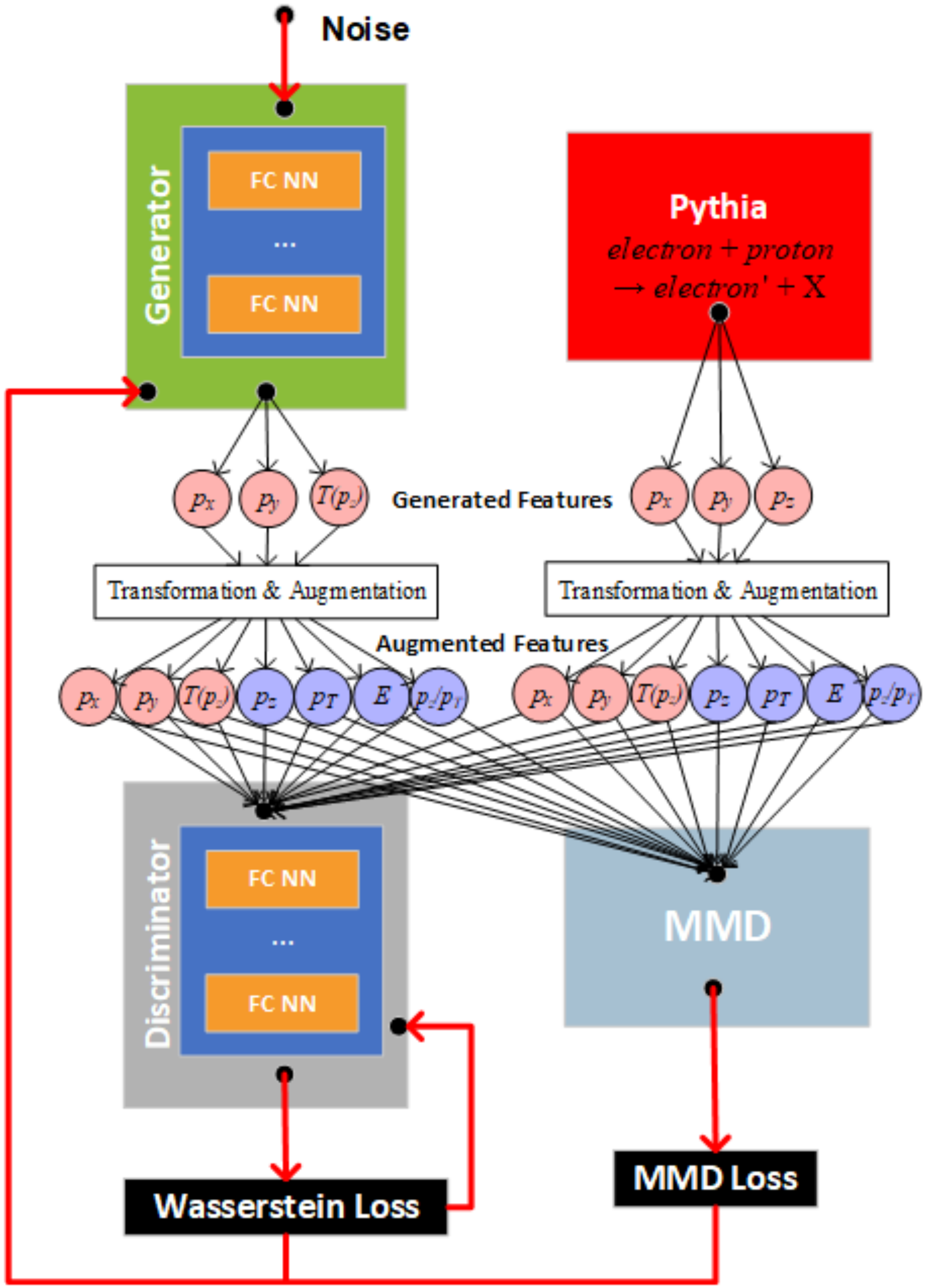
WGAN

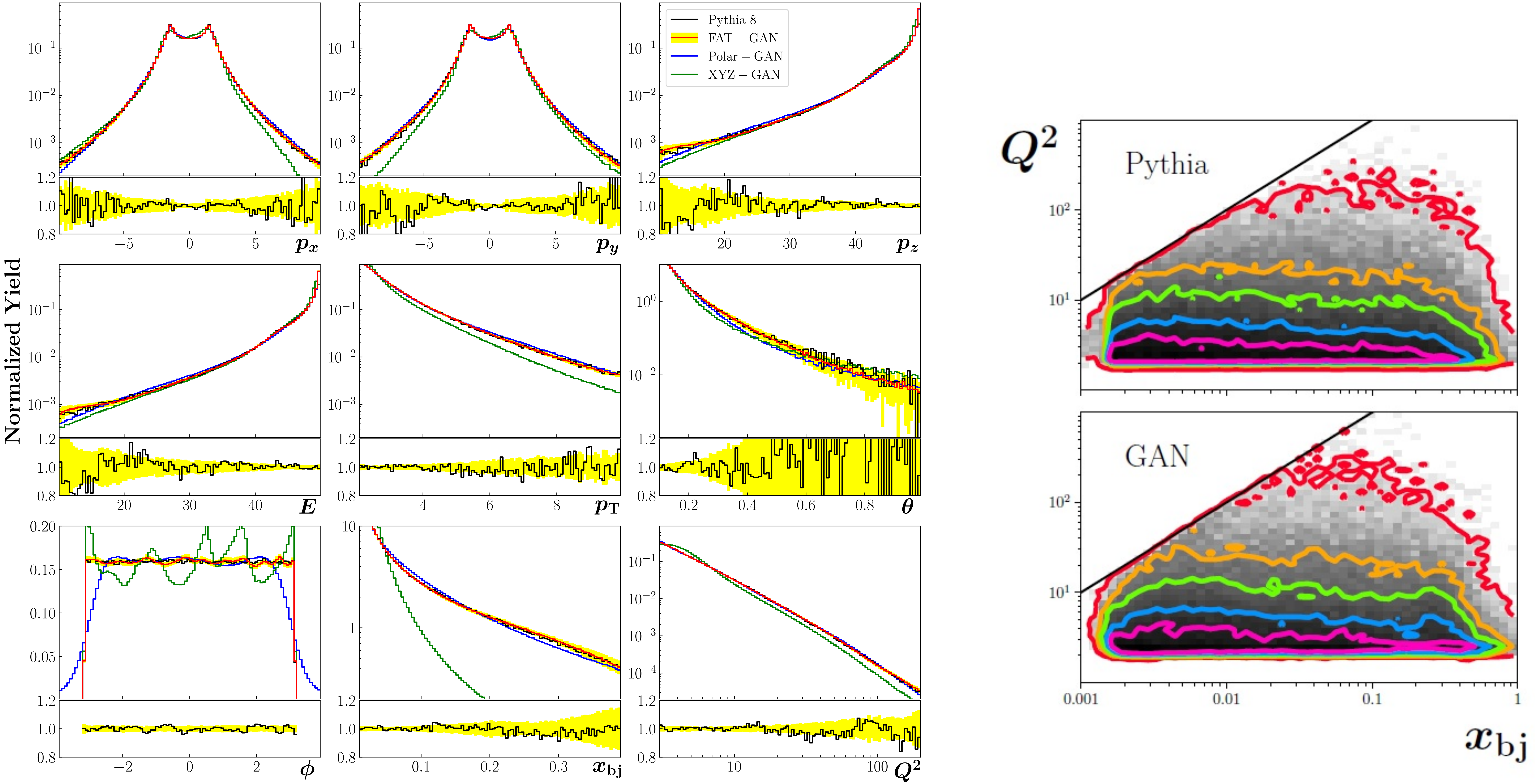


MAXIMUM MEAN DISCREPANCY (MMD) GAN

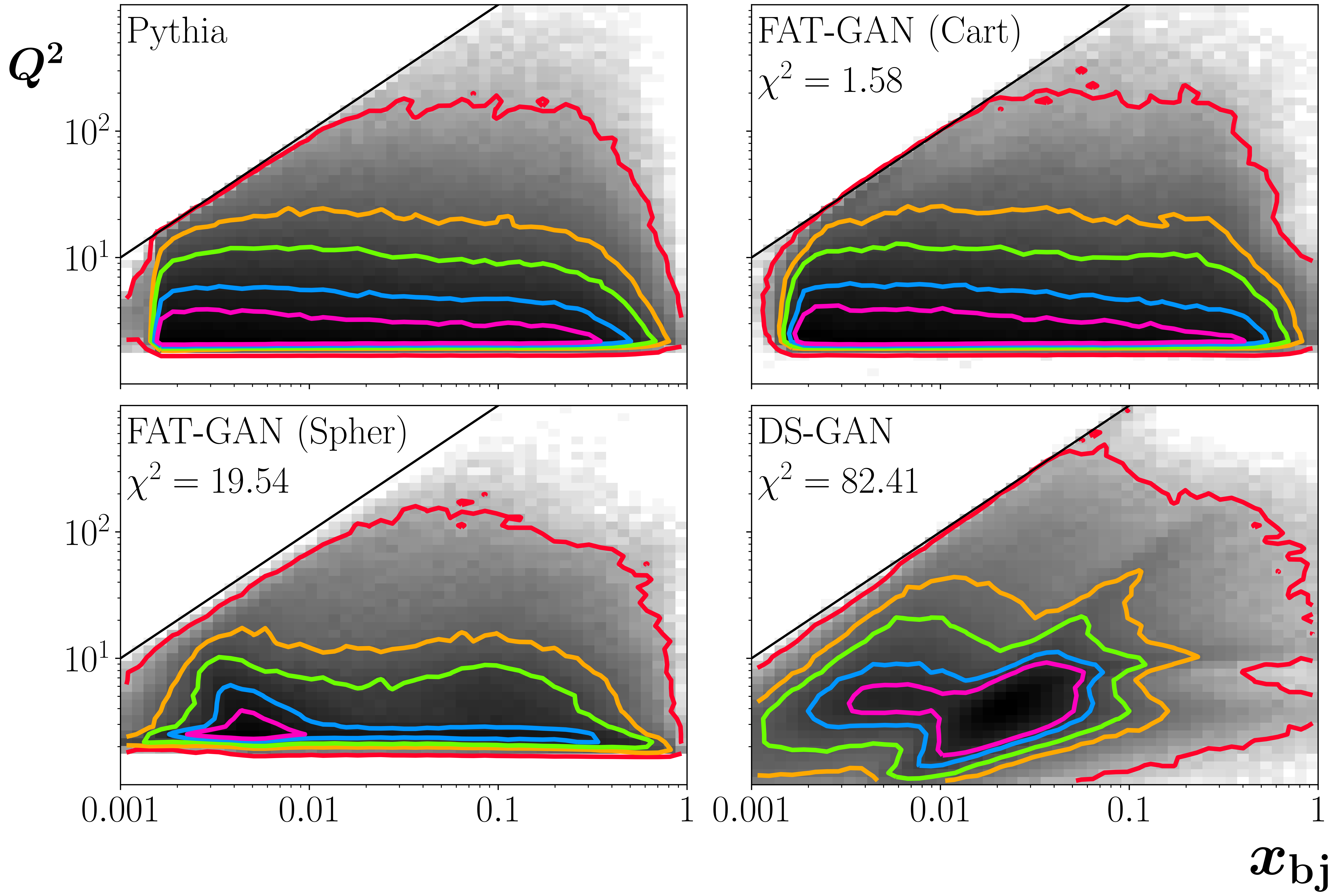
FAT-GAN

MMD: Critic loss: batch distribution matching





Pythia 8
FAT – GAN
Polar – GAN
XYZ – GAN
Normalized Yield
p_x
p_y
p_z
E
p_T
θ
ϕ
x_{bj}
Q^2
Pythia
GAN
x_{bj}



Pythia
FAT-GAN (Cart)
$\chi^2 = 1.58$
FAT-GAN (Spher)
$\chi^2 = 19.54$
DS-GAN
$\chi^2 = 82.41$
Q^2
10^2
10^1
0.001
0.01
0.1
1
x_{bj}

ACKNOWLEDGMENTS

- Raghu Ramanujan, Meg Houck, Eleni Tsitinidi, Jose Cruz, Andrew Hoyle, Michael Robertson, Evan Pritchard, Robert Solli, John Blue, Zach Nussbaum, Ryan Strauss, Jack Taylor

- JLab/ODU - Theory Center, Hall B: Nobuo Sato, Wally Melnitchouk, Yaohang Li, Yasir Alanazi, Manal Almaeen

ALPhA
DAVIDSON COLLEGE

FRIB

Jefferson Lab

DAVIDSON

NSCL