NN: Jet Momentum Magnitude

Regression Results
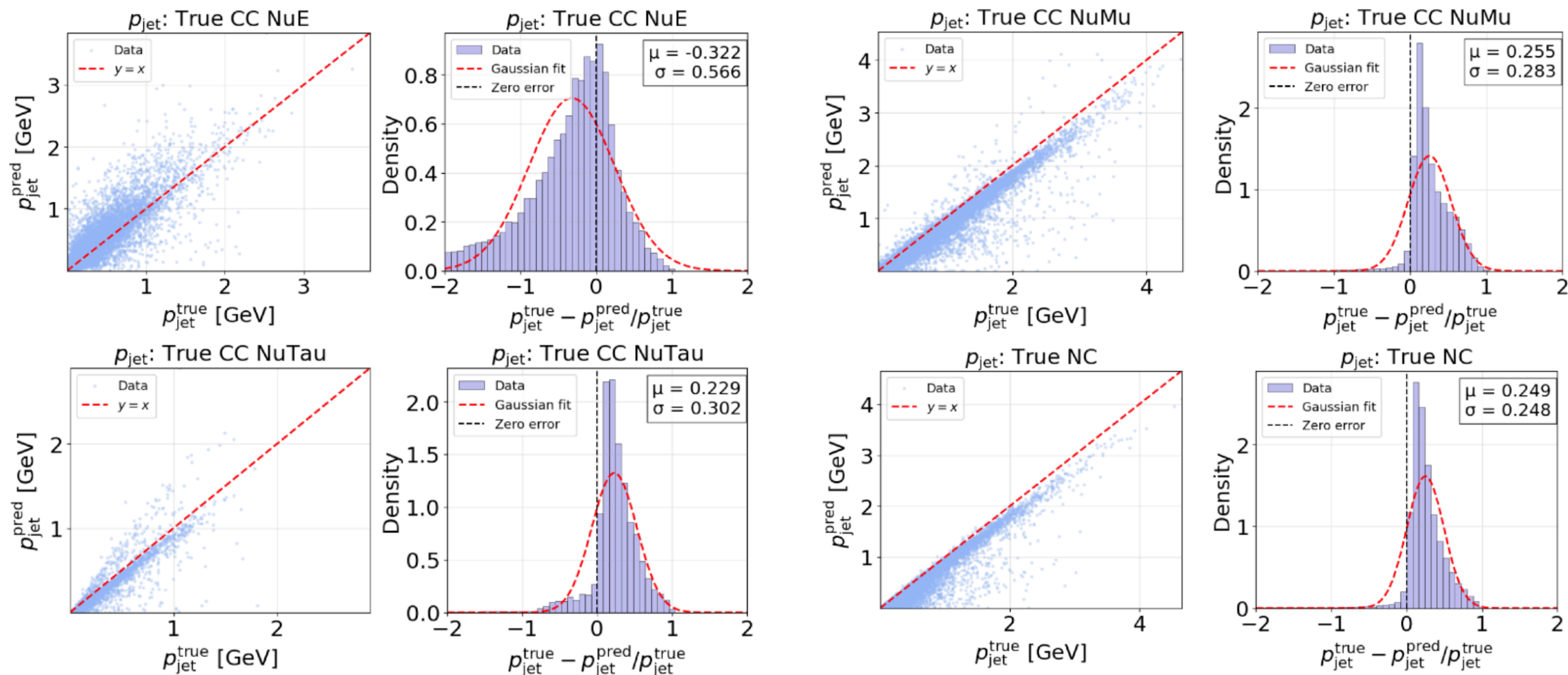
Pre Training

Transformer 2

Patched input

Masked input

Non-Masked embedding

Masked embedding

~50x50

~20

M0
M1
M2
M3
M4
M5
M6
M7

Encoder   Hierarchical
Transformer
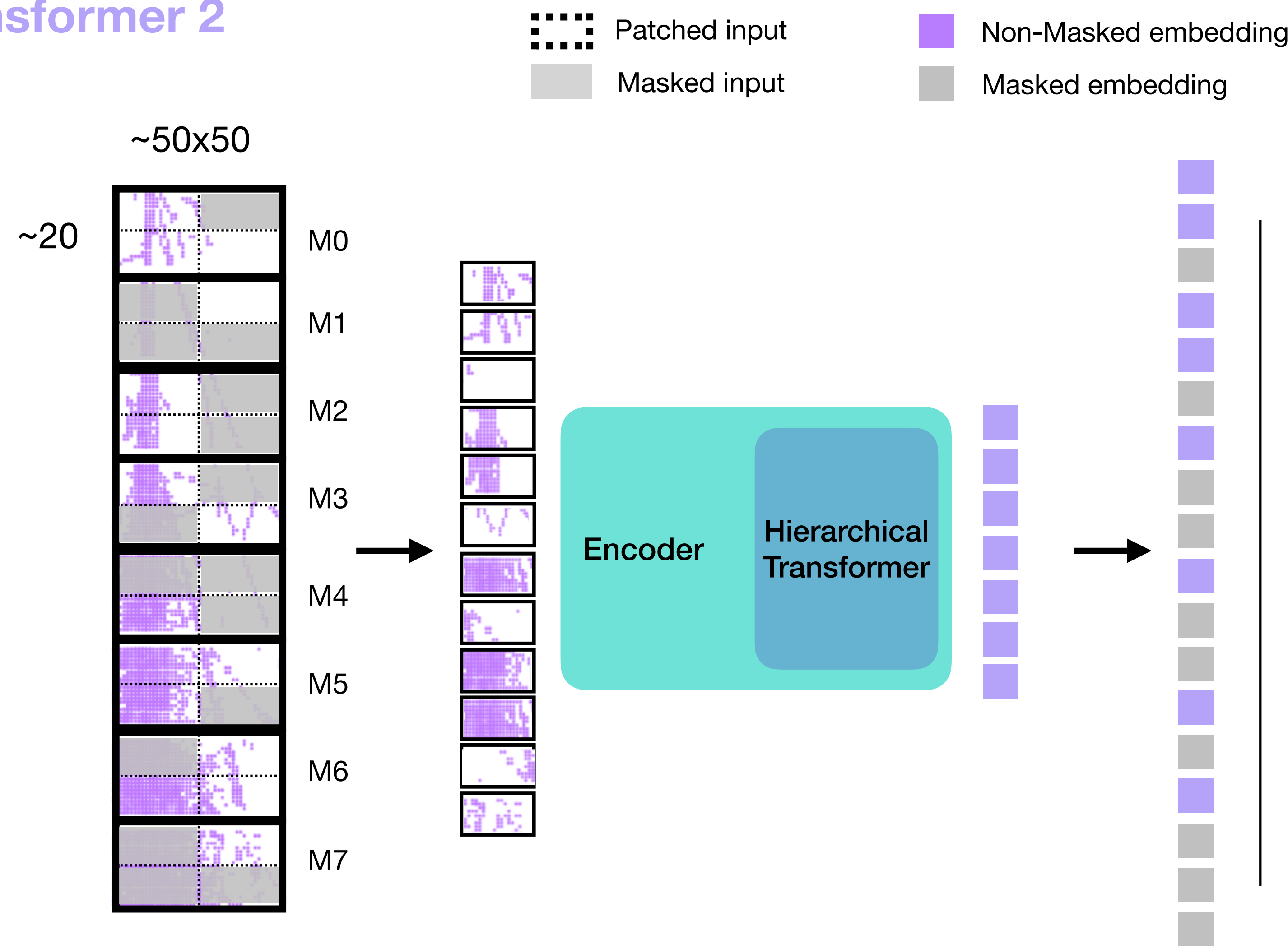
Contrastive Learning??? Bkup