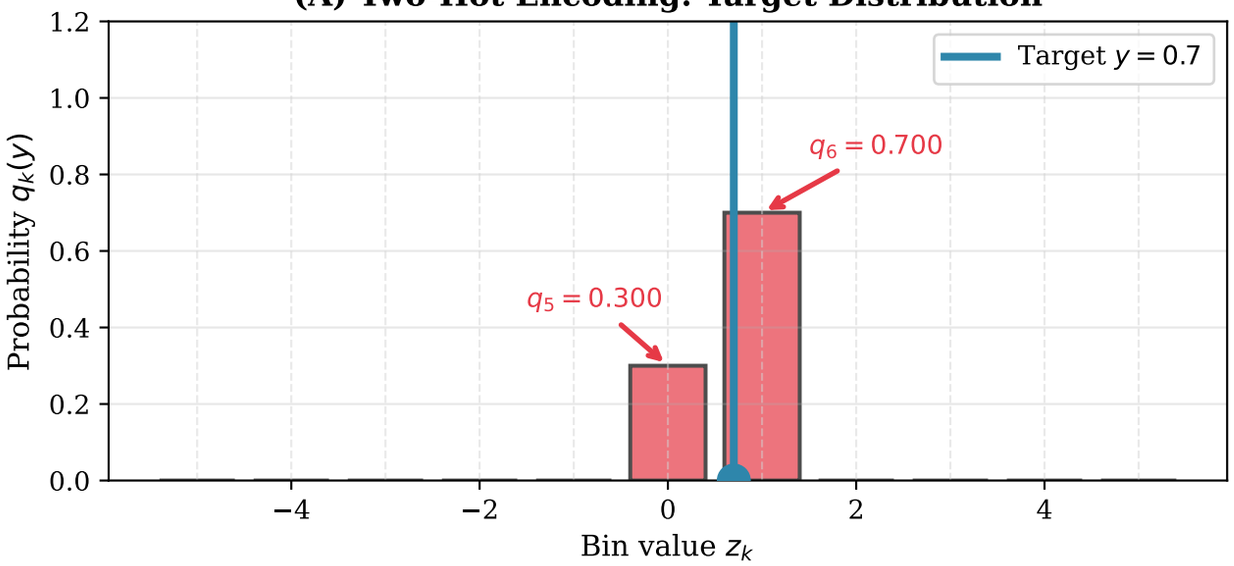
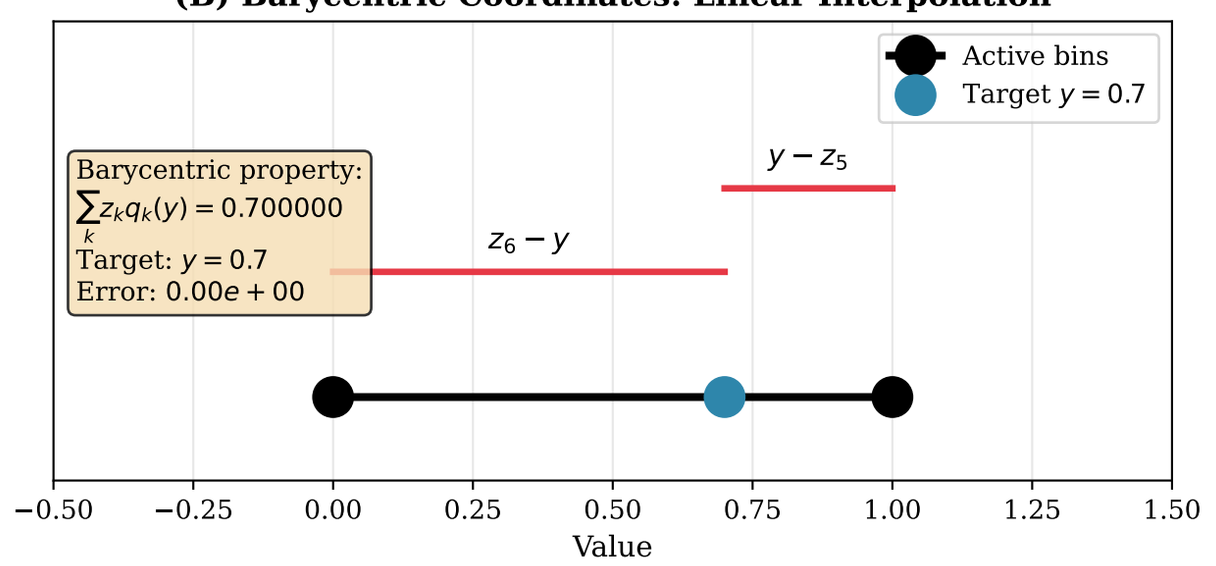
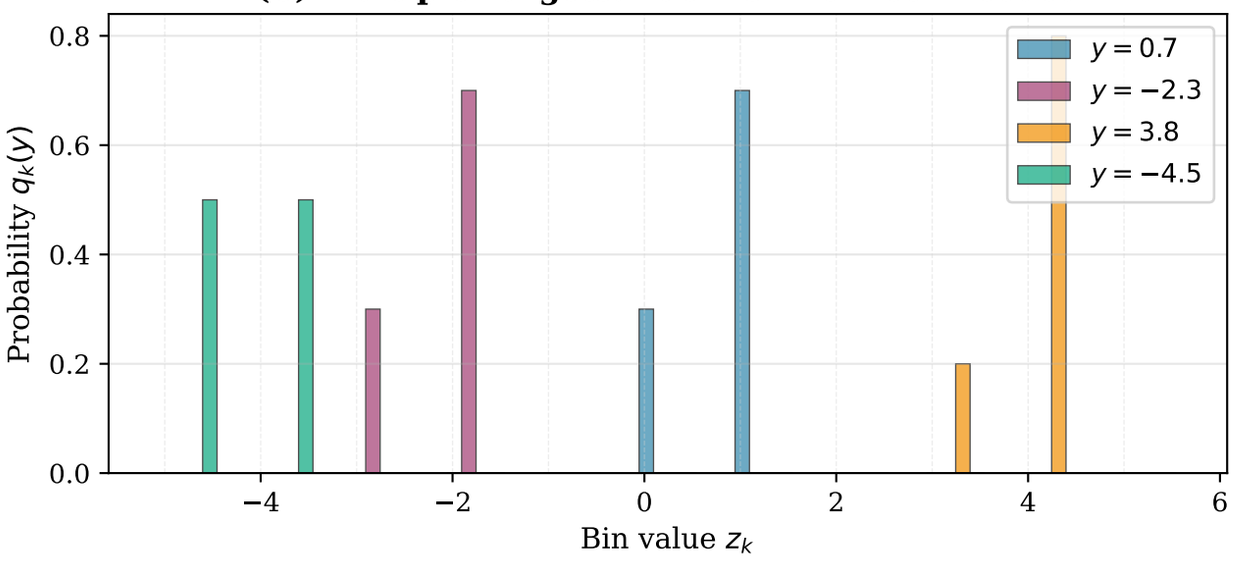
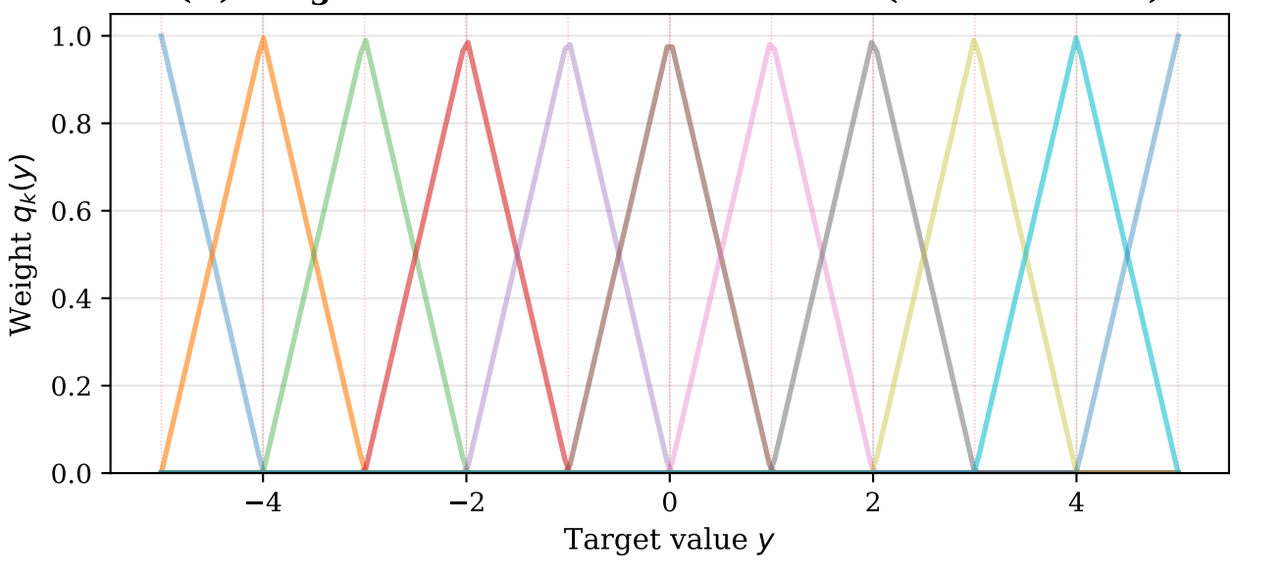
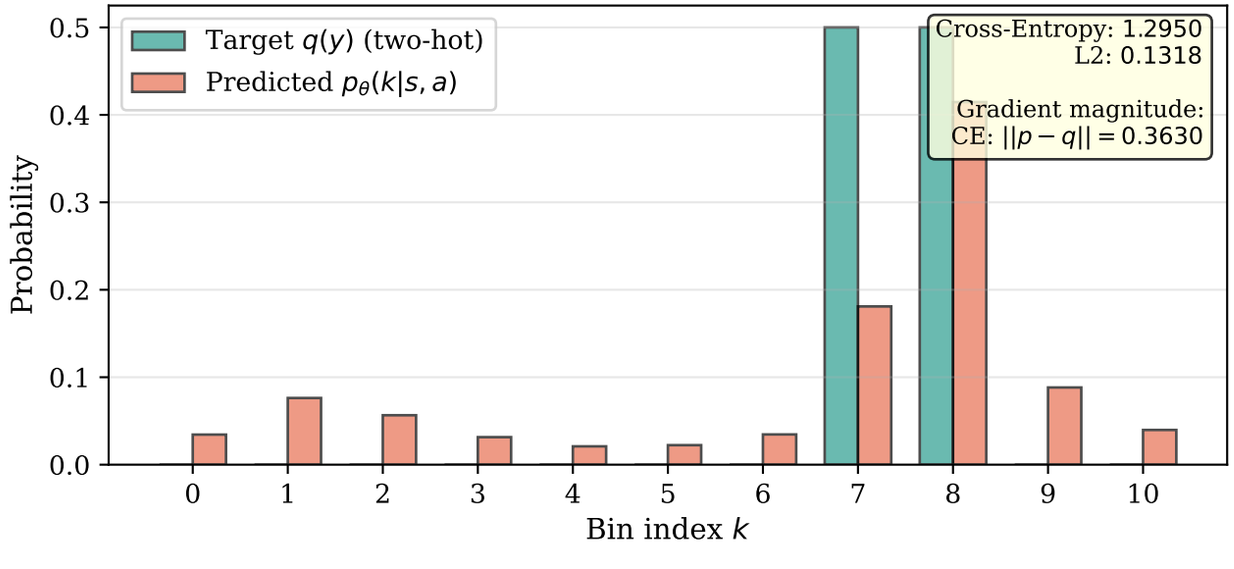
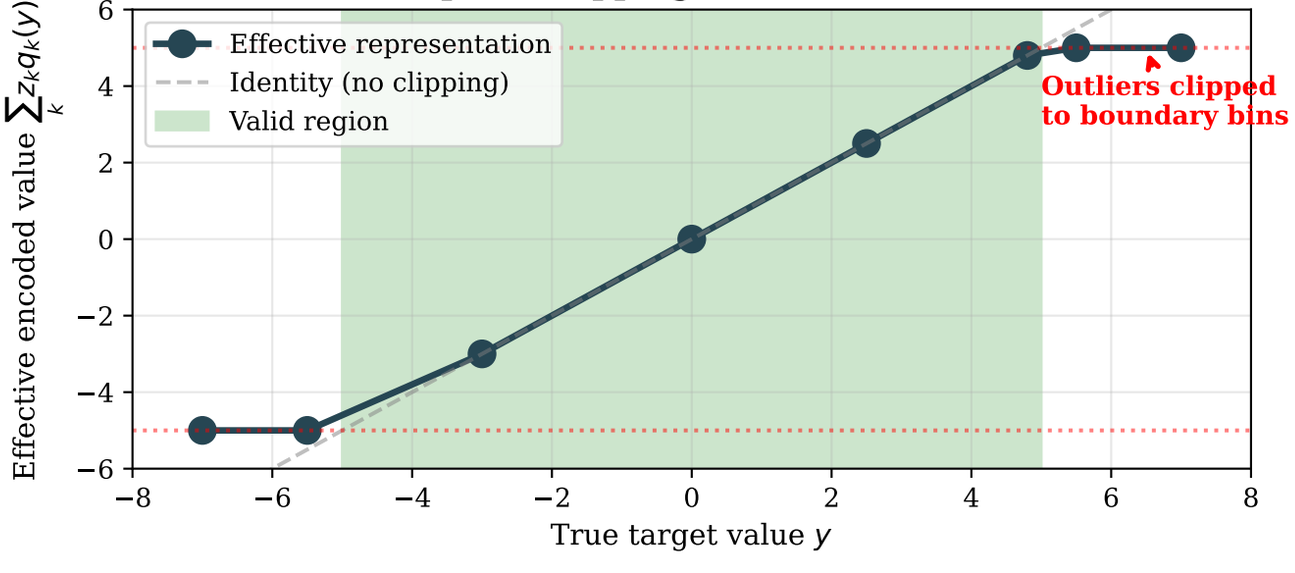

(A) Two-Hot Encoding: Target Distribution
Target $y = 0.7$
$q_6 = 0.700$
$q_5 = 0.300$
Probability $q_k(y)$
Bin value z_k


(B) Barycentric Coordinates: Linear Interpolation
Active bins
Target $y = 0.7$
Barycentric property:
$\sum_k z_k q_k(y) = 0.700000$
Target: $y = 0.7$
Error: $0.00e+00$
$y - z_5$
$z_6 - y$
Value


(C) Multiple Targets: Each is Two-Hot Encoded
$y = 0.7$
$y = -2.3$
$y = 3.8$
$y = -4.5$
Probability $q_k(y)$
Bin value z_k


(D) Weight Functions: Piecewise Linear (Hat Functions)
Weight $q_k(y)$
Target value y


(E) Cross-Entropy Loss: Target vs Predicted Distribution
Target $q(y)$ (two-hot)
Predicted $p_\theta(k|s,a)$
Cross-Entropy: 1.2950
L2: 0.1318
Gradient magnitude:
CE: $||p - q|| = 0.3630$
Probability
Bin index k


(F) Implicit Clipping: Outlier Robustness
Effective representation
Identity (no clipping)
Valid region
Outliers clipped to boundary bins
Effective encoded value $\sum_k z_k q_k(y)$
True target value y