Monosemanticity spectrum (encoder.layer2)