activation_l2_distance
Mean ± Std across pairs