Divergence for different configurations attcnn_1 (prior) (n = 32, FA) --- attcnn_1 (post) (n = 32, FA) - cnn_1 (prior) (n = 32, FA) --- cnn_1 (post) (n = 32, FA) 10⁰ cnn_2 (prior) (n = 32, FA) --- cnn_2 (post) (n = 32, FA) cno_1 (prior) (n = 32, FA) --- cno_1 (post) (n = 32, FA) — INS_ref (prior) (n = 32, FA) Face-average --- INS_ref (post) (n = 32, FA) 10^{-4}