Distribution of token ranks in QK circuit

Percentage of Model Vocabulary vs. Token rank

Query and Key Inputs:
$Q = MLP_0 (W_E)$
$K = MLP_0 (W_E)$

$Q = W_U$
$K = MLP_0 (W_E)$