8 Supplementary Material

8.1 Additional visualisation results

(a) sigmoid(N) (b) tanh(E) (c) Matrix M
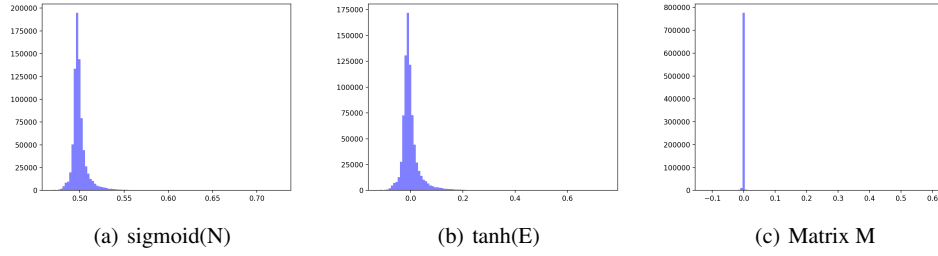
Figure 3: Visualization at $d = 200$ on SearchQA dataset.

(a) sigmoid(N) (b) tanh(E) (c) Matrix M
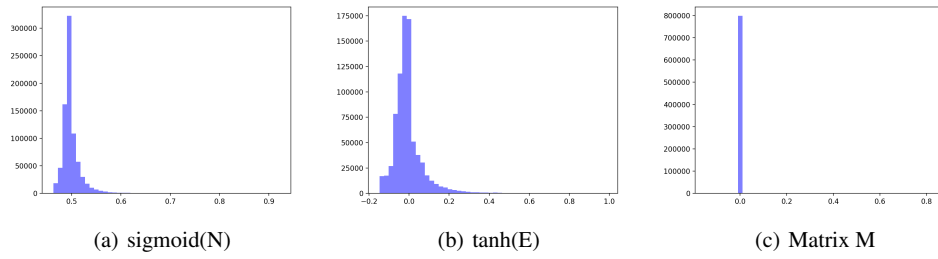
Figure 4: Visualization at $d = 200$ on Quasar-T dataset.

12