

## 1 Question 1

We expect high cosine similarities for embeddings of nodes in the same connected components, and low cosine similarities for nodes in different connected components.

## 2 Question 2

$$X_1 = AX_2$$

With

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$X_1$  and  $X_2$  represent the same vector in different bases, and  $A$  is the transition matrix between these bases. The two representations carry the same information.

## 3 Question 3

At each passing layer, the information from a node is passed to its neighbors. The connectivity of the graph is encoded in the adjacency matrix, therefore multiplying  $A$  and  $X$  is equivalent to an homogeneous diffusion of  $X$  information among all its neighbors.

If only one message passing layer is used, the diffusion is limited, and we can't study properties of the graph properly. On the other hand, if the number of message passing layers is greater than the number of nodes in the graph, the information is completely diffused in the graph, and the information is lost.

## 4 Question 4

$$\begin{aligned} X_1 W_0 &= \begin{pmatrix} 0.5 & -0.2 \\ 0.5 & -0.2 \\ 0.5 & -0.2 \\ 0.5 & -0.2 \end{pmatrix} \\ AX_1 W_0 &= \begin{pmatrix} 0.5 & -0.2 \\ 1 & -0.4 \\ 1 & -0.4 \\ 0.5 & -0.2 \end{pmatrix} \\ Z_0 = f(AX_1 W_0) &= \begin{pmatrix} 0.5 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0.5 & 0 \end{pmatrix} \\ Z_0 W_1 &= \begin{pmatrix} 1.5 & 0 \\ -0.4 & 0 \\ 0.8 & 0 \\ 0.25 & 0 \end{pmatrix} \\ AZ_0 W_1 &= \begin{pmatrix} -0.4 & 0 \\ 2.3 & 0 \\ -0.15 & 0 \\ 0.8 & 0 \end{pmatrix} \\ Z_1 &= \begin{pmatrix} 0 & 0 \\ 2.3 & 0 \\ 0 & 0 \\ 0.8 & 0 \end{pmatrix} \end{aligned}$$