The aim of this exercise session is to understand Ambainis’ quantum walk algorithm for element distinctness. Assume that we are given an array of integers $x_1, x_2, \ldots, x_N$. A collision is a pair of distinct $i, j$ such that $x_i = x_j$. How many elements do we have to query in order to find a collision (or decide that all elements are distinct)? Classically this takes $\Omega(N)$ queries. In contrast, there is a quantum walk algorithm that makes only $O(N^{2/3})$ queries to the input, which is optimal.

The quantum algorithm searches over elements $\mathcal{Y} = (Y, x_Y)$, with $Y \subseteq [N]$ a size-$k$ subset of indices and $x_Y$ the list of integers $x_j$ with index $j \in Y$. We call an element $Y$ marked if $x_Y$ contains a collision.

**Exercise 1.** Let $n = \binom{N}{k}$ denote the number of elements and $m$ the number of marked elements. Show that $m/n \in \Omega(k^2/N^2)$.

We consider a graph $G$ with vertex set $V$ indexed by the elements $\mathcal{Y}$. There is an edge between $\mathcal{Y} = (Y, x_Y)$ and $\mathcal{Y}' = (Y', x_{Y'})$ if the subsets $Y$ and $Y'$ differ in exactly one element (i.e., we can obtain $Y'$ from $Y$ by replacing one index). The resulting graph $G$ has $n = \binom{N}{k}$ vertices and is $k(n - k)$-regular. It corresponds to a so-called Johnson graph, and one can show that its spectral gap is $\delta \in \Omega(1/k)$ when $k \ll n$. Quantum walk search on this graph then has complexity

$$S + \sqrt{\frac{m}{n}} \left( \frac{1}{\sqrt{\delta}} \mathcal{U} + \mathcal{C} \right) = S + \frac{N}{k} \left( \sqrt{\delta} \mathcal{U} + \mathcal{C} \right).$$

It remains to bound the checking cost $\mathcal{C}$, the setup cost $S$, and the update cost $\mathcal{U}$ (specifically, the number of queries to the input for each operation).

**Exercise 2** (Checking cost). A basis state takes the form $|\mathcal{Y}\rangle = |Y, x_Y\rangle$. Given $|\mathcal{Y}\rangle$, how many queries does it take to check whether $\mathcal{Y}$ is marked?

**Exercise 3** (Setup cost). How many queries does the operation $|0\rangle \langle Y, 0| \mapsto |Y\rangle = |Y, x_Y\rangle$ require? Use this to bound the query cost of the setup $|0\rangle \mapsto |\pi\rangle = \frac{1}{\sqrt{n}} \sum_{\mathcal{Y}} |\mathcal{Y}\rangle$.

**Exercise 4** (Update cost). The row oracle $V$ maps $|0\rangle |\mathcal{Y}\rangle$ to

$$V |0\rangle |\mathcal{Y}\rangle = \frac{1}{\sqrt{k}} \sum_{Y' \sim Y} |\mathcal{Y}'\rangle |\mathcal{Y}\rangle,$$

where $Y' \sim Y$ indicates that there is an edge between $\mathcal{Y}$ and $\mathcal{Y}'$ in $G$. Show that $V$ can be implemented with $O(1)$ queries. Use $V$ to implement a quantum walk on $G$.

If we set $k = N^{2/3}$, then this yields a quantum algorithm with total query complexity

$$S + N^{1/3} \left( N^{1/3} \mathcal{U} + \mathcal{C} \right) \in O(N^{2/3}).$$