For convenience, we can introduce a binary random variable X_i which takes the value 1 if and only if $h_{\pi_i}(C_1) = h_{\pi_i}(C_2)$, and 0 otherwise.

It has been demonstrated in class that the similarity between sets C_1 and C_2, denoted as $J(C_1, C_2)$, is equivalent to the probability of $h_{\pi_i}(C_1) = h_{\pi_i}(C_2)$, or $X_i = 1$, represented by p. The estimated similarity $\hat{J}(C_1, C_2)$ is calculated by determining the proportion of matches under k different permutations of the hash function, expressed as

$$\hat{J}(C_1, C_2) = \frac{1}{k} \sum_{i=1}^{k} \mathbf{1}[h_{\pi_i}(C_1) = h_{\pi_i}(C_2)] = \frac{X_1 + X_2 + \cdots + X_k}{k}$$

Given that X_i follows a Bernoulli distribution, we can derive that the expectation of $\hat{J}(C_1, C_2)$, $\mathbb{E}(\hat{J}(C_1, C_2))$, equals $p = J(C_1, C_2)$, and its variance is $D = D(\hat{J}(C_1, C_2)) = \frac{p(1-p)}{k}$. Utilizing Chebyshev's Inequality, we obtain

$$\Pr(|\mathbb{E}(\hat{J}) - \hat{J}| \leq \lambda\sqrt{D}) \geq 1 - \frac{1}{\lambda^2}$$

Let $\delta = \frac{1}{\lambda^2}$, $\lambda\sqrt{D} = \epsilon$, then $\delta = \frac{p(1-p)}{k\epsilon^2}$, which meets the condition $k = O\left(\frac{1}{\epsilon^2\delta}\right)$, that is

$$\Pr(|J - \hat{J}| \leq \epsilon) \geq 1 - \delta$$

QED.