

## A PROOFS

### A.1 Proof of Lemma 1

PROOF. In Algorithm 1, the  $\lfloor \frac{\rho}{2} \rfloor$  nodes that hold the fewest replicas are included when selecting  $\mathbf{n}_\alpha$  to ensure storage balance. Thus, generating  $\lfloor \frac{\rho}{2} \rfloor$  good placement schemes ensures that each node is included in a good placement scheme at least once.  $\square$

### A.2 Proof of Lemma 2 and 3

#### A.2.1 Proof of Lemma 2.

PROOF. Let  $g_{\alpha^*} = \sum_{n_Y \in \mathbf{n}_\alpha^*} g(n_Y, \mathcal{N}_i \setminus \mathbf{n}_\alpha^*, \mathcal{A})$ , which is equal to the sum of complement degrees from  $\mathbf{n}_\alpha^*$  to the remainder of nodes in  $\mathcal{N}_i$ . The initial complement degree for any  $n_Y \in \mathbf{n}_\alpha^*$  under  $\mathcal{A}$  is  $|\mathcal{N}_i| - \lfloor \frac{\rho}{2} \rfloor$ . Additionally, each time a vertex is included in  $\mathbf{n}_\alpha$ , its complement degree under  $\mathcal{A}$  decreases by at most  $\lfloor \frac{\rho}{2} \rfloor$ , and it can be selected into  $\mathbf{n}_\alpha$  at most  $\omega$  times. Thus the lower bound is  $g_{\alpha^*} \geq \lfloor \frac{\rho}{2} \rfloor \left( |\mathcal{N}_i| - \lfloor \frac{\rho}{2} \rfloor (1 + \omega) \right)$ . According to the pigeonhole principle, when  $g_{\alpha^*} > \left( \lfloor \frac{\rho}{2} \rfloor - 1 \right) \left( |\mathcal{N}_i| - \lfloor \frac{\rho}{2} \rfloor \right)$ , there must exist a  $n_i \in \mathcal{N}_i$  s.t.  $g(n_i, \mathbf{n}_\alpha^*, \mathcal{A}) = \lfloor \frac{\rho}{2} \rfloor$ . Finally, this Lemma is proven by combining the two inequations above.  $\square$

#### A.2.2 Proof of Lemma 3.

PROOF. Let  $g_\alpha = \sum_{n_Y \in \mathbf{n}_\alpha} g(n_Y, \mathcal{N}_j, \mathcal{A})$ , which is equal to the sum of complement degrees from  $\mathbf{n}_\alpha$  to all nodes in  $\mathcal{N}_j$ . Following the same processes as Lemma 2, we obtain an inequation for the lower bound  $g_\alpha \geq \left( \lfloor \frac{\rho}{2} \rfloor + 1 \right) \left( |\mathcal{N}_j| - \left( \lfloor \frac{\rho}{2} \rfloor + 1 \right) \omega \right)$  and another inequation for the pigeonhole principle  $g_\alpha > \lfloor \frac{\rho}{2} \rfloor |\mathcal{N}_j|$ . This lemma is proven by combining the two inequations above.  $\square$

### A.3 Proof of Lemma 4

PROOF. In the worst case, generating  $\omega$  schemes fills  $\rho$  nodes, as though all  $\omega$  schemes select the same nodes. Since Algorithm 1 always avoids selecting duplicated node combinations to maximize the scatter width, the rate at which the number of filled nodes increases will not be faster than the aforementioned worst case.  $\square$

### A.4 Proof of Lemma 5

PROOF. According to Hall's marriage theorem [17], if for any subset of the shard set  $R$ , the sum of the number of leader replicas that can be held by its associated nodes exceeds its size, each shard can be matched with a node. Since the commonly configured replication factor  $\rho > 1$ , which means for any shard, the sum of the number of leader replicas that can be held by its associated nodes is  $\rho > 1$ , the above condition is satisfied.  $\square$