8 Appendix

8.1 Key relevant results from [4]

In order to make this manuscript self containing we include in this section key relevant lemmas, corollaries and definitions from [4].

Lemma 8.1 (Lemma 2.5, [4])**.** *Let S be an index set of cardinality s. For any level j of the dyadic splitting, $j = 0, \ldots, \lceil \log_2 s \rceil - 1$, the set S is decomposed into disjoint sets each having cardinality $Q_j = \lceil \frac{s}{2^j} \rceil$ or $R_j = Q_j - 1$. Let q_j sets have cardinality Q_j and r_j sets have cardinality R_j, then*

$$q_j = s - 2^j \cdot \left\lceil \frac{s}{2^j} \right\rceil + 2^j, \quad \text{and} \quad r_j = 2^j - q_j. \tag{132}$$

Lemma 8.2 (Lemma 2.3, [4])**.** *Let $B,\ B_1,\ B_2 \subset [n]$ where $|B_1| = b_1$, $|B_2| = b_2$, $B = B_1 \cup B_2$ and $|B| = b$. Also let B_1 and B_2 be drawn uniformly at random, independent of each other, and define $P_n(b, b_1, b_2) := \mathit{Prob}\left(|B_1 \cap B_2| = b_1 + b_2 - b\right)$, then*

$$P_n(b, b_1, b_2) = \binom{b_1}{b_1 + b_2 - b}\binom{n - b_1}{b - b_1}\binom{n}{b_2}^{-1}. \tag{133}$$

Definition 8.1. *$P_n(x, y, z)$ defined in (133) satisfies the upper bound*

$$P_n(x, y, z) \le \pi(x, y, z) \exp(\psi_n(x, y, z)) \tag{134}$$

with bounds of $\pi(x, y, z)$ given in Lemma 8.3.

Lemma 8.3. *For $\pi(x, y, z)$ and $P_n(x, y, z)$ given by (134) and (133) respectively, if $\{y, z\} < x < y + z$, $\pi(x, y, z)$ is given by*

$$\left(\frac{5}{4}\right)^4 \left[\frac{yz(n-y)(n-z)}{2\pi n(y+z-x)(x-y)(x-z)(n-x)}\right]^{\frac{1}{2}}, \tag{135}$$

otherwise $\pi(x, y, z)$ has the following cases.

$$\left(\frac{5}{4}\right)^3 \left[\frac{y(n-z)}{n(y-z)}\right]^{\frac{1}{2}} \quad \text{if} \quad x = y > z; \tag{136}$$

$$\left(\frac{5}{4}\right)^3 \left[\frac{(n-y)(n-z)}{n(n-y-z)}\right]^{\frac{1}{2}} \quad \text{if} \quad x = y + z; \tag{137}$$

$$\left(\frac{5}{4}\right)^2 \left[\frac{2\pi z(n-z)}{n}\right]^{\frac{1}{2}} \quad \text{if} \quad x = y = z. \tag{138}$$

Lemma 8.4. *Define*

$$\psi_n(x, y, z) := y \cdot H\left(\frac{x-z}{y}\right) + (n-y) \cdot H\left(\frac{x-y}{n-y}\right) - n \cdot H\left(\frac{z}{n}\right), \tag{139}$$

then for $n > x > y$ we have that

$$\text{for } y > z \quad \psi_n(x, y, y) \le \psi_n(x, y, z) \le \psi_n(x, z, z); \tag{140}$$

$$\text{for } x > z \quad \psi_n(x, y, y) > \psi_n(z, y, y); \tag{141}$$

$$\text{for } 1/2 < \alpha \le 1 \quad \psi_n(x, y, y) < \psi_n(\alpha x, \alpha y, \alpha y). \tag{142}$$

Corollary 8.1. *If $n > 2y$, then $\pi(y, y, y)$ is monotonically increasing in y.*

The following bound, used in [4], is deducible from an asymptotic series for the logarithms Stirling approximation of the factorial (!)

$$\frac{16e^{N\mathcal{H}(p)}}{25\sqrt{2\pi p(1-p)N}} \leq \binom{N}{pN} \leq \frac{5e^{N\mathcal{H}(p)}}{4\sqrt{2\pi p(1-p)N}} . \tag{143}$$

8.2 Derivation of Inequalities

8.2.1 Inequality 64

By Lemma 8.1, the left hand side (LHS) of (64) is equal to the following.

$$\begin{aligned} q_0\,(q_1 r_1)\cdot(q_2 r_2)\cdot(q_3 r_3)\cdots\left(q_{\lceil \log_2 s\rceil -2} r_{\lceil \log_2 s\rceil -2}\right) &= \left(s - \left\lceil \frac{s}{1} \right\rceil + 1\right)\cdot\left(s - 2\cdot\left\lceil \frac{s}{2}\right\rceil + 2\right) \\ \times\left(2 - \left(s - 2\cdot\left\lceil\frac{s}{2}\right\rceil + 2\right)\right) &\times\cdots\times\left(s - 2^{\lceil \log_2 s\rceil -2}\cdot\left\lceil \frac{s}{2^{\lceil \log_2 s\rceil -2}}\right\rceil + 2^{\lceil \log_2 s\rceil -2}\right) \\ &\times\left(2^{\lceil \log_2 s\rceil -2} - \left(s - 2^{\lceil \log_2 s\rceil -2}\cdot\left\lceil \frac{s}{2^{\lceil \log_2 s\rceil -2}}\right\rceil + 2^{\lceil \log_2 s\rceil -2}\right)\right) . \end{aligned} \tag{144}$$

We simplify (144) to get the following.

$$\begin{aligned} 1\cdot\left(s - 2\cdot\left\lceil\frac{s}{2}\right\rceil + 2\right)\cdot\left(2\cdot\left\lceil\frac{s}{2}\right\rceil - s\right)\cdot\left(s - 2^2\cdot\left\lceil\frac{s}{2^2}\right\rceil + 2^2\right)\cdot\left(2^2\cdot\left\lceil\frac{s}{2^2}\right\rceil - s\right)\times \\ \cdots\times\left(s - 2^{\lceil \log_2 s\rceil -2}\cdot\left\lceil \frac{s}{2^{\lceil \log_2 s\rceil -2}}\right\rceil + 2^{\lceil \log_2 s\rceil -2}\right)\cdot\left(2^{\lceil \log_2 s\rceil -2}\cdot\left\lceil \frac{s}{2^{\lceil \log_2 s\rceil -2}}\right\rceil - s\right) . \end{aligned} \tag{145}$$

We upper bound $-\lceil z\rceil$ by $-z$ and $\lceil z\rceil$ by $z+1$ to upper bound (145) as follows.

$$\begin{aligned} \left(s - 2\cdot\frac{s}{2} + 2\right)\cdot\left(2\left(\frac{s}{2}+1\right) - s\right)\cdot\left(s - 4\cdot\frac{s}{4} + 2\right)\cdot\left(4\left(\frac{s}{4}+1\right) - s\right)\times \\ \cdots\times\left(s - 2^{\lceil \log_2 s\rceil -2}\cdot\frac{s}{2^{\lceil \log_2 s\rceil -2}} + 2^{\lceil \log_2 s\rceil -2}\right)\cdot\left(2^{\lceil \log_2 s\rceil -2}\cdot\frac{s}{2^{\lceil \log_2 s\rceil -2}} - s\right) . \end{aligned} \tag{146}$$

The bound (146) is then simplified to the following.

$$(2\cdot 2)\cdot(4\cdot 4)\cdot(8\cdot 8)\times\cdots\times\left(2^{\lceil \log_2 s\rceil -2}\cdot 2^{\lceil \log_2 s\rceil -2}\right) = 2^2\cdot 4^2\cdot 8^2\times\cdots\times 2^{2\lceil \log_2 s\rceil -4} \tag{147}$$

$$= 4^1\cdot 4^2\cdot 4^3\cdots\times 4^{\lceil \log_2 s\rceil -2} \tag{148}$$

$$\leq 4^{\left(\sum_{i=1}^{\log_2 s - 1} i\right)} \tag{149}$$

$$= 4^{\frac{1}{2}(\log_2 s - 1)\cdot\log_2 s} = 2^{(\log_2 s - 1)\cdot\log_2 s} . \tag{150}$$

In (149) we upper bound $\lceil \log_2 s\rceil$ by $\log_2 s + 1$; while in the LHS of (150) we computed the summation of a finite arithmetic series. After some algebraic manipulations of logarithms we end up with the RHS of (150), which simplifies to (64).

8.2.2 Inequality 65

Again by Lemma 8.1, the left hand side (LHS) of (65), i.e. $(a_{Q_0}a_{Q_1}a_{R_1}a_{Q_2}a_{R_2}a_{Q_3}a_{R_3}\cdots a_3 a_2)^{1/2}$ is equal to the following.

$$\left(a_{\lceil \frac{s}{2^0}\rceil}a_{\lceil \frac{s}{2^1}\rceil}a_{\lceil \frac{s}{2^1}\rceil -1}a_{\lceil \frac{s}{2^2}\rceil}a_{\lceil \frac{s}{2^2}\rceil -1}a_{\lceil \frac{s}{2^3}\rceil}a_{\lceil \frac{s}{2^3}\rceil -1}\times\cdots\times a_{\lceil \frac{s}{2^{\lceil \log_2 s\rceil -2}}\rceil}a_{\lceil \frac{s}{2^{\lceil \log_2 s\rceil -2}}\rceil -1}\right)^{1/2} . \tag{151}$$

Given the monotonicity of $a_{(\cdot)}$ in terms of its subscripts, which indicate cardinalities of sets. Due to the nestedness of the sets due to the dyadic splitting, we upper bound $a_{\lceil \frac{s}{2^j} \rceil - 1}$ by $a_{\frac{s}{2^j}}$, and $a_{\lceil \frac{s}{2^j} \rceil}$ by $a_{\frac{s}{2^j}+1}$, resulting in the following upper bound for (151).

$$\left[a_s a_{\left(\frac{s}{2}+1\right)} a_{\frac{s}{2}} a_{\left(\frac{s}{4}+1\right)} a_{\frac{s}{4}} a_{\left(\frac{s}{8}+1\right)} a_{\frac{s}{8}} \times \cdots \times a_{\left(\frac{s}{2^{\lceil \log_2 s \rceil - 2}}+1\right)} a_{\frac{s}{2^{\lceil \log_2 s \rceil - 2}}} \right]^{1/2} \tag{152}$$

$$\leq \left[a_s a_{\left(\frac{s}{2}+1\right)} a_{\frac{s}{2}} a_{\left(\frac{s}{4}+1\right)} a_{\frac{s}{4}} a_{\left(\frac{s}{8}+1\right)} a_{\frac{s}{8}} \times \cdots \times a_{\left(\frac{s}{2^{\log_2 s - 2}}+1\right)} a_{\frac{s}{2^{\log_2 s - 2}}} \right]^{1/2}. \tag{153}$$

In (153) we used the fact that $2^{\log_2 s - 2}$ is a lower bound to $2^{\lceil \log_2 s \rceil - 2}$. We fix $a_s = (1-\epsilon)ds =: cs$ and we require expansion to hold for all $|\mathcal{S}| \leq s$, i.e. $a_{s'} = cs'$ for all $s' \leq s$. Thus we can re-write (153) as follows.

$$\left[a_s \left(\frac{cs}{2} + c \right) \frac{cs}{2} \left(\frac{cs}{4} + c \right) \frac{cs}{4} \left(\frac{cs}{8} + c \right) \frac{cs}{8} \times \cdots \times \left(\frac{cs}{2^{\log_2 s - 2}} + c \right) \frac{cs}{2^{\log_2 s - 2}} \right]^{1/2} \tag{154}$$

$$= \left[a_s \left(\frac{a_s}{2} + c \right) \frac{a_s}{2} \left(\frac{a_s}{4} + c \right) \frac{a_s}{4} \left(\frac{a_s}{8} + c \right) \frac{a_s}{8} \times \cdots \times \left(\frac{a_s}{2^{\log_2 s - 2}} + c \right) \frac{a_s}{2^{\log_2 s - 2}} \right]^{1/2}. \tag{155}$$

In (155) we substitute a_s for cs. Next we factor a_s out in all the brackets to have the following.

$$\left[a_s a_s \left(\frac{1}{2} + \frac{c}{a_s} \right) a_s \left(\frac{1}{2} \right) a_s \left(\frac{1}{4} + \frac{c}{a_s} \right) a_s \left(\frac{1}{4} \right) a_s \left(\frac{1}{8} + \frac{c}{a_s} \right) a_s \left(\frac{1}{8} \right) \times \right.$$
$$\left. \cdots \times a_s \left(\frac{1}{2^{\log_2 s - 2}} + \frac{c}{a_s} \right) a_s \left(\frac{1}{2^{\log_2 s - 2}} \right) \right]^{1/2}. \tag{156}$$

In total we have twice $(\log_2 s - 2)$ plus 1 factors of a_s. We use this and the fact that $c/a_s = 1/s$ to simplify (156) to (157), which further simplifies to (158) by rearranging the terms in (157).

$$\left[(a_s)^{2\log_2 s - 3} \left(\frac{1}{2} + \frac{1}{s} \right) \left(\frac{1}{2} \right) \left(\frac{1}{4} + \frac{1}{s} \right) \left(\frac{1}{4} \right) \left(\frac{1}{8} + \frac{1}{s} \right) \left(\frac{1}{8} \right) \times \right.$$
$$\left. \cdots \times \left(\frac{1}{2^{\log_2 s - 2}} + \frac{1}{s} \right) \left(\frac{1}{2^{\log_2 s - 2}} \right) \right]^{1/2}. \tag{157}$$

$$\left[(a_s)^{2\log_2 s - 3} \left(\frac{1}{2} \cdot \frac{1}{2^2} \cdot \frac{1}{2^3} \cdots \frac{1}{2^{\log_2 s - 2}} \right) \times \right.$$
$$\left. \cdots \times \left(\frac{1}{2} + \frac{1}{s} \right) \left(\frac{1}{2^2} + \frac{1}{s} \right) \left(\frac{1}{2^3} + \frac{1}{s} \right) \cdots \left(\frac{1}{2^{\log_2 s - 2}} + \frac{1}{s} \right) \right]^{1/2}. \tag{158}$$

We focus on bounding the second line of (158), ignoring the square-root for the moment, that is $\left(\frac{1}{2} + \frac{1}{s}\right) \left(\frac{1}{2^2} + \frac{1}{s}\right) \left(\frac{1}{2^3} + \frac{1}{s}\right) \times \cdots \times \left(\frac{1}{2^{\log_2 s - 2}} + \frac{1}{s}\right)$. This equals

$$\exp \left(\log \left[\left(\frac{1}{2} + \frac{1}{s} \right) \left(\frac{1}{2^2} + \frac{1}{s} \right) \left(\frac{1}{2^3} + \frac{1}{s} \right) \cdots \left(\frac{1}{2^{\log_2 s - 2}} + \frac{1}{s} \right) \right] \right) \tag{159}$$

$$= \exp \left(\log \left[\frac{1}{2} \left(1 + \frac{2}{s} \right) \right] + \log \left[\frac{1}{2^2} \left(1 + \frac{2^2}{s} \right) \right] + \cdots + \log \left[\frac{1}{2^{\log_2 s - 2}} \left(1 + \frac{2^{\log_2 s - 2}}{s} \right) \right] \right) \tag{160}$$

$$= \exp \left(\log \left[\frac{1}{2} \cdot \frac{1}{2^2} \cdot \frac{1}{2^3} \cdots \frac{1}{2^{\log_2 s - 2}} \right] + \log \left[\left(1 + \frac{2}{s} \right) \left(1 + \frac{2^2}{s} \right) \cdots \left(1 + \frac{2^{\log_2 s - 2}}{s} \right) \right] \right). \tag{161}$$

From (159) to (161), we used simple algebra involving logarithms. Upper bounding $\log(1+x)$ by x, since $\log(1+x) \le x$ for $|x| < 1$, we upper bounded the exponent involving the second log term to upper bound (161) by the following.

$$\left(\frac{1}{2} \cdot \frac{1}{2^2} \cdot \frac{1}{2^3} \cdots \frac{1}{2^{\log_2 s - 2}}\right) \times \exp\left(\frac{2}{s} + \frac{2^2}{s} + \frac{2^3}{s} + \cdots + \frac{2^{\log_2 s - 2}}{s}\right) \tag{162}$$

$$= \left(\frac{1}{2} \cdot \frac{1}{2^2} \cdot \frac{1}{2^3} \cdots \frac{1}{2^{\log_2 s - 2}}\right) \times \exp\left[\frac{1}{s}\left(\frac{s}{2} - 2\right)\right] \le \left(\frac{1}{2} \cdot \frac{1}{2^2} \cdot \frac{1}{2^3} \cdots \frac{1}{2^{\log_2 s - 2}}\right) e^{\frac{1}{2}}. \tag{163}$$

The exponent of the exponential on the right of (162) is a geometric series and this simplifies to the LHS bound of (163). The RHS bound of (163) is due to upper bounding $e^{1/2-2/s}$ by $e^{1/2}$. Using the bound in (163), we upper bound (158) by the following.

$$\left[e^{\frac{1}{2}} \cdot (a_s)^{2\log_2 s - 3} \left(\frac{1}{2} \cdot \frac{1}{2^2} \cdot \frac{1}{2^3} \cdots \frac{1}{2^{\log_2 s - 2}}\right)^2\right]^{1/2} = e^{\frac{1}{4}} \cdot (a_s)^{\log_2 s - \frac{3}{2}} \cdot \left[2^{-(1+2+\cdots+(\log_2 s - 2))}\right] \tag{164}$$

$$= \frac{1}{2} e^{\frac{1}{4}} \cdot (a_s)^{\log_2 s - \frac{3}{2}} \cdot s^{\frac{1}{2}\log_2 s + \frac{3}{2}}, \tag{165}$$

which is the bound in (65), hence concluding the derivation as required.