# Reformulation of Lemma 4.2

## Old Lemma 4.2

Let $S$ be a finite set with power set $\mathcal{P}(S)$. Let $F$ be a distribution on $\mathcal{P}(S)$ and let $X = \{X_i\}_{i\geq 1}$, $Y = \{Y_i\}_{i\geq 1}$ both be i.i.d $F$. Then if $M, N$ are non-negative random variables such that $M \geq_{ST} N$, then:

$$\left|\bigcup_{i=0}^{M} X_i\right| \geq ST \left|\bigcup_{i=0}^{N} Y_i\right|.$$

## New Lemma 4.2

Let $S$ be a finite set with power set $\mathcal{P}(S)$. Let $F$ be a distribution on $\mathcal{P}(S)$ and let $X = \{X_i\}_{i\geq 1}$, $Y = \{Y_i\}_{i\geq 1}$ both be i.i.d $F$. Then if $M, N$ are random subsets of an index set $I$ such that $|M| \geq_{ST} |N|$, then:

$$\left|\bigcup_{i\in M} X_i\right| \geq ST \left|\bigcup_{i\in N} Y_i\right|.$$

## Proof

We want to show that

$$\mathbb{P}\left\{\left|\bigcup_{i\in M} X_i\right| \geq z\right\} \geq \mathbb{P}\left\{\left|\bigcup_{i\in N} Y_i\right| \geq z\right\}, \quad \forall z.$$

We first observe that, since the sequence $X$ is i.i.d. and that therefore the size of the union depends only on the number of sets in the union, by conditioning on the size of the random set $M$, $\mathbb{P}\left\{\left|\bigcup_{i\in M} X_i\right| \geq z\right\} = \sum_{j=0}^{\infty} \mathbb{P}\left\{\left|\bigcup_{i=0}^{j} X_i\right| \geq z\right\} \cdot \mathbb{P}\{|M| = j\}$. This implies that

$$\mathbb{P}\left\{\left|\bigcup_{i\in M} X_i\right| \geq z\right\} = \mathbb{E}_{|M|}\left[\mathbb{P}\left\{\left|\bigcup_{i=0}^{|M|} X_i\right| \geq z\right\}\right].$$

We therefore need to show that

$$\mathbb{E}_{|M|}\left[\mathbb{P}\left\{\left|\bigcup_{i=0}^{|M|} X_i\right| \geq z\right\}\right] \geq \mathbb{E}_{|N|}\left[\mathbb{P}\left\{\left|\bigcup_{i=0}^{|N|} Y_i\right| \geq z\right\}\right].$$

### CLAIM 1:

Let $f_X(K, z) = \mathbb{P}\left\{\left|\bigcup_{i=0}^{|M|} X_i\right| \geq z\right\}$. For any given fix $z$, $f_X(K, z)$ is a non-decreasing function of $K$.

### CLAIM 2:

Let $X \geq_{ST} Y$, and let $g$: $\mathbb{R} \mapsto \mathbb{R}$ be non-decreasing, then $\mathbb{E}[g(X)] \geq \mathbb{E}[g(Y)]$.

Fix an arbitrary $z$, using **claims 1** and **2**, we can write

$$\mathbb{E}_{|M|}\left[\mathbb{P}\left\{\left|\bigcup_{i=0}^{|M|} X_i\right| \geq z\right\}\right] = \mathbb{E}_{|M|}\left[f_X(|M|, z)\right] \geq$$
$$\mathbb{E}_{|N|}\left[f_X(|N|, z)\right] = \mathbb{E}_{|N|}\left[\mathbb{P}\left\{\left|\bigcup_{i=0}^{|N|} Y_i\right| \geq z\right\}\right].$$

To conclude the proof, we need to establish both **claims 1** and **2**.

## Proof of CLAIM 1:

We note that

$$\bigcup_{i=0}^{k} X_i \subset \left(X_{k+1} \cup \bigcup_{i=0}^{k} X_i\right), \quad \text{almost surely.}$$

Therefore by induction, for any $k_1 \leq k_2$, we have that $f_X(k_2, z) \geq f_X(k_1, z)$ for any $z \in \mathbb{N}$.

## Proof of CLAIM 2:

Recall Strassen's Theorem:

> The random variable $X$ is stochastically larger than the random variable $Y$ if and only if there exists a coupling $(\hat{X}, \hat{Y})$ of $X, Y$ such that
>
> $$\mathbb{P}\left\{\hat{X} \geq \hat{Y}\right\} = 1.$$

Let $(\hat{X}, \hat{Y})$ be that coupling of $X, Y$, then almost surely $\hat{X} \geq \hat{Y}$ which implies that $g(\hat{X}) \geq g(\hat{Y})$ almost surely since $g$ is non-decreasing. But since $g(X)$ has the distribution as $g(\hat{X})$, and $g(Y)$ has the distribution as $g(\hat{Y})$, we have that $(g(\hat{X}), g(\hat{Y}))$ is a coupling of $(g(X), g(Y))$ such that $\mathbb{P}\left\{g(\hat{X}) \geq g(\hat{Y})\right\} = 1$, implying that $g(X) \geq_{ST} g(Y)$. Once more we use Strassen's Theorem to say that, since $g(X)$ has the distribution as $g(\hat{X})$, and $g(Y)$ has the distribution as $g(\hat{Y})$:

$$\mathbb{E}[g(X)] = \mathbb{E}[g(\hat{X})] \geq \mathbb{E}[g(\hat{Y})] = \mathbb{E}[g(Y)], \quad \text{which proves the claim.}$$