Lemma 1. *Having an arbitrary lists of binary values I and J, e.g. values in the ith and jth bit of bit strings $o \in X$, and list $\neg I$ of the negated values from I, the following holds for the Pearson correlation coefficient corr:*

$$corr(I, J) = -corr(\neg I, J).$$

Proof. Let us denote $|X|$ the length of arrays $I, J, \neg I$, and

- a the number of positions in lists I, J with bits 1 and 1,
- b the number of positions with bits 0 and 1 in these lists,
- c the number of positions with bits 1 and 0 in these lists,
- d the number of positions with bits 0 and 0 in these lists.

Please see, that $a+b+c+d = |X|$. The Pearson correlation coefficient $corr(I, J)$ is:

$$\begin{aligned} corr(I, J) &= \frac{|X| \cdot a - (a+b)(a+c)}{\sqrt{|X| \cdot (a+b) - (a+b)^2} \cdot \sqrt{|X| \cdot (a+c) - (a+c)^2}} \\ &= \frac{ad - bc}{\sqrt{(a+b)(c+d)} \cdot \sqrt{(a+c)(b+d)}} \end{aligned} \tag{1}$$

The numbers a, b, c, d stands for the following values after the negation the list I to list $\neg I$:

- b the number of positions in lists $\neg I, J$ with bits 1 and 1,
- a the number of positions with bits 0 and 1 in these lists,
- d the number of positions with bits 1 and 0 in these lists,
- c the number of positions with bits 0 and 0 in these lists.

Therefore the Pearson correlation coefficient $corr(\neg I, J)$ is equal to:

$$\begin{aligned} corr(\neg I, J) &= \frac{|X| \cdot b - (b+a)(b+d)}{\sqrt{|X| \cdot (b+a) - (b+a)^2} \cdot \sqrt{|X| \cdot (b+d) - (b+d)^2}} \\ &= \frac{bc - ad}{\sqrt{(a+b)(c+d)} \cdot \sqrt{(a+c)(b+d)}} \end{aligned} \tag{2}$$

Obviously, the Equations 1 and 2 describe opposite values, and thus $corr(I, J) = -corr(\neg I, J)$.