Simple XOR lemma

Emanuele Viola*

March 2, 2026

The *xor lemma* says that if a function has small correlation with circuits of a certain size, then the xor of many independent copies of the function has much smaller correlation (with circuits of related size). The *correlation* between two functions $f, g : [2]^n \to \{-1, 1\}$ is defined as $|\mathbb{E}_x[f(x)g(x)]|$, where $[2] = \{0, 1\}$. The xor lemma has had significant impact in theoretical computer science, with applications in complexity theory, pseudorandomness, and cryptography. For background we refer to [GNW95] or [Vio]. Several proofs of the xor lemma with various parameters are known. Three are given in [GNW95], two of which also appear in [Vio].

I give an alternative proof of the xor lemma which may provide a simple explanation of why xor-ing decreases correlation. Also, the new proof appears to improve parameters in some settings. For example, we can decrease the correlation by a constant factor with only a constant-factor loss in size, whereas other proofs appear to lose a factor which depends on the original correlation. To highlight the simplicity of the argument we begin with a version whose proof requires no background.

Theorem 1. *Let $f : [2]^n \to \{-1, 1\}$ be a function. Suppose there is a circuit $C : [2]^n \times [2]^n \to \{-1, 1\}$ of size s s.t. $\mathbb{E}_{x,y\in[2]^n} C(x, y)f(x)f(y) \geq \epsilon$. Then there is a circuit C' of size $\leq 3s + 5$ s.t. $\mathbb{E}_{x\in[2]^n} C'(x)f(x) \geq \alpha$ where α is the solution of $3\epsilon - \alpha^3 = 2\alpha$.*

One can verify that $\alpha \geq 1 - 0.8(1 - \epsilon)$ for $\epsilon \geq 1/2$, and $\alpha \geq 1.2\epsilon$ for $\epsilon \leq 1/2$.

Proof. Let $C'(x) := \mathrm{Maj}\left(C(x, y_1)f(y_1), C(x, y_2)f(y_2), C(x, y_3)f(y_3)\right)$ for uniform $y_1, y_2, y_3 \in [2]^n$. Because $\mathrm{Maj}_3(x_1, x_2, x_3) = \frac{1}{2}(\sum_i x_i - \prod_i x_i)$ as a real polynomial, we have

$$\begin{aligned}
\mathbb{E}_{x,y_1,y_2,y_3} C'(x)f(x) &= \mathbb{E}_{x,y_1,y_2,y_3} \frac{1}{2}\Big(\sum_i C(x, y_i)f(y_i)f(x) - f(x)\prod_i C(x, y_i)f(y_i)\Big) \\
&= \frac{3}{2}\mathbb{E}_{x,y\in[2]^n} C(x, y)f(x)f(y) - \frac{1}{2}\mathbb{E}_x\left[f(x)\mathbb{E}_y^3 C(x, y)f(y)\right]. \\
&\geq \frac{3\epsilon}{2} - \frac{1}{2}\max_x \left|\mathbb{E}_y^3 C(x, y)f(y)\right|.
\end{aligned}$$

If there is x s.t. $|\mathbb{E}_y C(x, y)f(y)| \geq \alpha$ the conclusion holds with $C'(y) := C(x, y)$. Otherwise, the last displayed expression is $\geq 3\epsilon/2 - \alpha^3/2 = \alpha$. We can fix the y_i to maintain the inequality. □

*Supported by NSF grant CCF-2430026.

Next we consider the natural extension.

Theorem 2. *Let $f : [2]^n \to \{-1,1\}$ be a function. Suppose there is a circuit $C : [2]^n \times [2]^n \to \{-1,1\}$ of size s s.t. $\mathbb{E}_{x,y\in[2]^n} C(x,y)f(x)f(y) \geq \epsilon$. Then there is a circuit C' of size $\leq c\sqrt{1/\epsilon}s$ s.t. $\mathbb{E}_{x\in[2]^n} C'(x)f(x) \geq \epsilon^{1-c}$ for all $\epsilon \leq c$.*

Every occurrence of 'c' denotes a possibly different real > 0.

Proof. Let $C'(x) := \text{Maj}\left(C(x,y_1)f(y_1), C(x,y_2)f(y_2), \ldots, C(x,y_k)f(y_k)\right)$ where the y_i are uniform in $[2]^n$. We again write Maj as a polynomial $m(z_1,z_2,\ldots,z_k) = \sum_S m_S \prod_{i\in S} z_i$. The polynomial is symmetric, and the coefficients of monomials of degree 2 are zero. The coefficient of z_i is $\geq c/\sqrt{k}$ for every i, and the sum of the absolute values of the coefficients of degree d is $\leq k^{d/2}$ in absolute value (see e.g. [Vio]). Hence

$$\begin{aligned}
\mathbb{E}_{x,y_1,y_2,\ldots,y_k} C'(x)f(x) &= \mathbb{E}\left[\left(\sum_{d=0}^{k}\sum_{S:|S|=d} m_S \prod_{i\in S} C(x,y_i)f(y_i)\right) f(x)\right] \\
&\geq c\sqrt{k}\mathbb{E}_{x,y} C(x,y)f(x)f(y) + \mathbb{E}_x f(x) \sum_{d\geq 3}^{k}\sum_{S:|S|=d} m_S \mathbb{E}^d[C(x,y)f(y)] \\
&\geq c\sqrt{k}\epsilon - \max_x \sum_{d\geq 3}^{k} k^{d/2} \left|\mathbb{E}^d[C(x,y)f(y)]\right|.
\end{aligned}$$

We set, say, $k := \sqrt{1/\epsilon}$. If there is x s.t. $|\mathbb{E}[C(x,y)f(y)]| \geq \sqrt{k}\epsilon$ we are done. Otherwise each summand is $\leq (k\epsilon)^d$ and using $k\epsilon \leq 1/2$ the sum is $\leq 2(k\epsilon)^3$. So the whole correlation is $\geq c\sqrt{k}\epsilon - 2(k\epsilon)^3 \geq c\epsilon^{4/5} - 2\epsilon^{3/2} \geq \epsilon^{1+c}$. We can fix the y_i to maintain the inequality. □

Starting from a function on n bits that has correlation $\leq 1-1/n^a$ with circuits of a certain size, these results can be used to construct a function on n^{c_a} bits that has exponentially small correlation with circuits of related size. We can first apply theorem 1 $c_a \log n$ times to construct a function with correlation $\leq 1-c$, and then again apply theorem 2 $c_a \log n$ times.

References

[GNW95] Oded Goldreich, Noam Nisan, and Avi Wigderson. On Yao's XOR lemma. Technical Report TR95–050, *Electronic Colloquium on Computational Complexity*, March 1995. www.eccc.uni-trier.de/.

[Vio] Emanuele Viola. *Mathematics of the impossible: The complexity of computation.* Cambridge University Press. 2023-2026, draft on the author's webpage.