The dream XOR lemma is false

Emanuele Viola*

June 8, 2026

Abstract

I refute the dream XOR lemma as stated in a 1995 paper by Goldreich, Nisan, and Wigderson. I also give a counterexample to the XOR lemma for low-degree polynomials.

Yao's fundamental XOR lemma (cf [GNW95, Vio26, Vio]) states that if a function $f : [2]^n \to [2]$ has *correlation* $\leq \epsilon$ with circuits of size s, meaning $\mathbb{E}_{x\in[2]^n}(-1)^{g(x)+f(x)} \leq \epsilon$ for every such circuit g, then the XOR of k independent copies of f, denoted

$$f^{\oplus k} : [2]^{kn} \ni (x_1, x_2, \ldots, x_k) \to f(x_1) \oplus f(x_2) \oplus \cdots \oplus f(x_k) \in [2],$$

has correlation about ϵ^k with circuits of size close to s. Here $[2] = \{0, 1\}$. The intuition behind the XOR lemma is that since the inputs x_i are independent, the maximum correlation with $f^{\oplus k}$ can be achieved by a circuit of the type $g(x_1) \oplus g(x_2) \oplus \cdots \oplus g(x_k)$. We shall see that this intuition is wrong. At least four different proofs of the XOR lemma are available, see [GNW95, Vio26, Vio]. Levin [Lev87, GNW95] proves it for $k = 2$, then iterates. He shows that $f^{\oplus 2}$ has correlation $\leq \epsilon^2 + \gamma$ with circuits of size $(\gamma/n)^c s - n^c$, for any γ. Every occurrence of "c" denotes a possibly different real > 0. All available proofs incur a similar loss of γ in circuit size. Goldreich, Nisan, and Wigderson [GNW95] note that Levin's proof

> *still contains some slackness; specifically, the closest one wants to get to the "obvious" bound [ϵ^2], the more one loses in terms of the complexity bounds (i.e., bounds on circuit size). [...] We do not know how to remove this slackness. We even do not know if it can be reduced "a little" as follows.*

They then state [GNW95] a *dream XOR lemma* where this loss is mitigated to a power independent of the target correlation. A specific instance is that $f^{\oplus 2}$ has correlation $\leq \epsilon^2 + 2^{-\log^2 n}$ with circuits of size s/n^b for a constant b. They credit Rudich for the observation that this dream version is false for *black-box proofs*. A similar question was raised by several other researchers, for example Impagliazzo concludes his landmark hard-core set paper [Imp95] thus:

*Supported by NSF grant CCF-2430026.

Why in all Yao-style arguments is there a trade-off between resources and probability, rather than a real increase in the hardness in the problem? If f is hard for resources R, the parity of many copies of f should still be hard for resources R, not just some slightly smaller bound.

Variants of the dream XOR lemma have been studied in many other works including [BGI08, BIK+22, BC25, BCS24, TYY11, DJMW12].

In this paper we show that the dream XOR lemma is false.

Theorem 1. *Let n and k be even and let S be a set of functions from $[2]^n$ to $[2]$. There is $f : [2]^n \to [2]$ that has correlation $\leq 2^{-n/2} \cdot c \log(2^{n/2} \cdot |S|)$ with any function in S, but $f^{\oplus k}$ has correlation $\geq \left(2^{-n/2}\right)^{k/2}$ with a quadratic polynomial modulo 2 in kn variables.*

To refute the dream XOR lemma let S be the functions computed by circuits of size n^a, for large enough constant a. Then, say, $c \log(2^{n/2} \cdot |S|) \leq n^{a+1}$ for large enough n. We can apply Theorem 1 to, say, the first $20a \log n$ bits and obtain f with correlation $\epsilon \leq n^{-10a} \cdot cn^{a+1} \leq cn^{-9a+1}$ with S; but the correlation of $f^{\oplus 2}$ with a circuit of size $ca \log n$ is $\geq n^{-10a}$.

Theorem 1 is also relevant to XOR lemmas for low-degree polynomials modulo 2. This model is of central interest because obtaining strong correlation bounds for low-degree polynomials is *necessary* for long-sought progress on circuit complexity, rigidity, communication complexity, and various Fourier conjectures, see [Vio22]. I have arrived at the results in this paper by thinking about XOR lemmas for quadratic polynomials. Standard proofs of the XOR lemma cannot be applied to low-degree polynomials (see [GSV18] and the discussion there), but XOR lemmas have been obtained via *ad hoc* techniques. Viola [Vio06] (cf [VW08]) employs the *degree norm* to prove an XOR lemma for low-degree polynomials. However, the correlation decay is weaker than for circuits. Chattopadhyay, Hatami, Hosseini, Lovett, and Zuckerman [CHH+20] prove an XOR lemma for low-degree polynomials which has strong decay but only applies when f is a *resilient* function. This raises the question whether there is an XOR lemma for low-degree polynomials that has strong decay and holds for every function. Theorem 1 gives a negative answer whenever $n \geq c_d$. As hinted, the XOR lemma in [Vio06] is related to properties of the degree norm. The results in this paper, combined with [Vio06], imply a nearly quadratic gap between correlation with quadratic polynomials and the corresponding degree norm. By known connections [Lov12] this implies gaps for other quantities in additive combinatorics. I omit precise statements since these connections are tangential to this paper. The only other example I am aware of that counters the XOR lemma unconditionally is the observation that the XOR lemma is not true for alternating circuits augmented with few majority gates, because parity has correlation about $1/\sqrt{n}$ with those circuits (cf [Vio]).

In the remainder we prove Theorem 1. It suffices to prove it for $k = 2$. To set the stage let us consider a naive approach to proving a counterexample. For every function $f : \{0,1\}^n \to \{0,1\}$, we can achieve correlation $\geq 2^{-n}$ with $f(x) \oplus f(y)$ simply by outputting 0 if $x = y$ and a uniform bit otherwise. Thus, if f has correlation substantially smaller than $2^{-n/2}$ with a class of functions, this would provide a counterexample to the dream XOR lemma. However, a simple consequence of the analysis of boolean functions is that any f in fact has correlation $\geq 2^{-n/2}$ with some parity function.

To remedy this, it is in hindsight natural to consider *bent* functions that have the smallest possible correlation with any parity. We seek a large class of bent functions among which we can hope to find one that doesn't correlate with S. Let $n := 2m$ and consider the Maiorana-McFarland functions (cf [CM16])

$$f_\pi : [2]^{2m} \ni (x, y) \to x.\pi(y) \in [2]$$

where . is inner product modulo 2, and $\pi : [2]^m \to [2]^m$ is an involution ($\pi\pi x = x$) with no fixed points ($\pi x \neq x$). Equivalently, π is a matching of the complete graph on $[2]^m$ with no self loops. Let $M := 2^m$.

First we note that, for every π, the quadratic polynomial modulo 2

$$p : [2]^{4m} \ni (x, y, x', y') \to x.y' + x'.y \in [2]$$

has correlation $\geq 1/M$ with $f_\pi(x, y) + f_\pi(x', y')$. Indeed, the correlation is

$$\left|\mathbb{E}(-1)^{x.(\pi(y)\oplus y')+x'.\pi(y')+x'.y}\right|.$$

This is 0 if $\pi(y) \neq y'$. But when they are equal we also have $\pi\pi(y) = y = \pi(y')$ using that π is an involution, and so the correlation is 1. Hence the correlation equals $\mathbb{P}[\pi(y) = y']$ which is $1/M$.

Second we claim that for any fixed function $h : [2]^n \to [2]$ one has

$$\mathbb{P}_\pi[C_\pi(h) \geq t/M] \leq 2 \cdot 2^{-ct} \tag{1}$$

for $t \geq c \log M$, where $C_\pi(h)$ is the correlation between h and f_π, and the probability is over all matchings as above. The theorem then follows by taking a union bound over the functions in S, for which $t \geq c \log |S|$ suffices.

Next we prove (1). We reduce the question to a concentration inequality for the weight of a matching selected uniformly from all possible matchings in a weighted complete graph. To prove concentration we use the method of bounded variances, Equation (8.5) and Problem 8.14 from [DP09]. We sample the permutation π by selecting the edges $X_1, X_2, \ldots, X_{M/2}$ in the corresponding matching. Let

$$w(y \to z) := \mathbb{E}_x(-1)^{x.z+h(x,y)}$$

be the contribution to the correlation when y is fixed and $\pi(y) = z$. And let

$$w(\{y, z\}) := w(y \to z) + w(z \to y)$$

be the undirected contribution along edge $\{y, z\}$. We let $w(\{y\}) := 0$ though note that $w(y \to y)$ may not be zero. For edges $x_1, x_2, \ldots, x_i$ we also write $w(x_1, x_2, \ldots, x_i)$ for $w(x_1) + w(x_2) + \cdots + w(x_i)$. The correlation between h and the function f corresponding to π is just the sum of the weights, up to scaling:

$$M \cdot C_\Pi(h) = w\left(X_1, X_2, \ldots, X_{M/2}\right) =: W \tag{2}$$

where the sum is over all edges in (the graph corresponding to) π.

To apply the method we need the following bounds on W:

Lemma 2. *We have* $\mathbb{E}[W] \leq 3$. *Also, conditioned on any fixed outcome* $x_{<i}$ *of* $X_{<i}$, *the variable* $W_i := \mathbb{E}_{X_{>i}} w(x_1, x_2, \ldots, x_{i-1}, X_i, X_{i+1}, \ldots, X_{M/2})$ *enjoys:*

1. *Boundedness:* $|W_i - \mathbb{E}[W_i]| := b \leq c$,
2. *Variance bound:*$varW_i := v_i \leq c/(M - 2(i-1))$.

The method of bounded variances then says the following.

$$\mathbb{P}_{\Pi}[W \geq \mathbb{E}[W] + t] \leq 2 \cdot \exp\left(-\frac{t^2}{c\sum_{i \leq M/2} v_i + cbt}\right).$$

By Lemma 2 and the harmonic sum we have

$$\sum_{i \leq M/2} v_i \leq c \sum_{i \leq M} 1/i \leq c \log M.$$

Hence for $t \geq c \log M$ the exponent is $\geq ct$ and it suffices that $t \geq c \log(M|S|)$ to afford the union bound over $2|S|$ functions. Hence there is π s.t., for every h, $W < \mathbb{E}[W] + c \log(M \cdot |S|)$. Using the bound on $\mathbb{E}[W]$ from Lemma 2, the rhs is $\leq c \log(M \cdot |S|)$. By (2) the correlation bound follows by dividing by M, concluding the proof of Theorem 1 assuming Lemma 2.

To bound W_i as in Lemma 2 we first need to bound w.

Claim 3. We have:

1. $|w(x \to y)| \leq 1$ for every x, y.
2. For every y, $\sum_{z \in [2]^m} w(y \to z) = (-1)^{h(0,y)} \in \{-1, 1\}$.
3. For every $R \subseteq [2]^m$, $\sum_{y \in R, z \in R} w^2(\{y, z\}) \leq 4|R|$.

Proof. We have

1. By definition.
2. By linearity of expectation the lhs equals $\mathbb{E}_x \sum_{z \in [2]^m} (-1)^{x.z + h(x,y)}$. When $x \neq 0$ the sum over z is 0. Otherwise it is M. The result follows.
3. The lhs is at most

$$\sum_{y,z \in R} (w(y \to z) + w(z \to y))^2 \leq 2 \sum_{y,z \in R} \left(w^2(y \to z) + w^2(z \to y)\right) \leq 4 \sum_{y \in R, z \in R} w^2(y \to z).$$

We get an upper bound if we sum over all $z \in [2]^m$. Expanding the square and using the definition of $w(y \to z)$ we obtain the upper bound

$$\sum_{y \in R, z \in [2]^m} \mathbb{E}_{x,x'} (-1)^{x.z + h(x,y) + x'.z + h(x',y)} \leq \sum_{y \in R} \mathbb{E}_{x,x'} \sum_{z \in [2]^m} (-1)^{x.z + h(x,y) + x'.z + h(x',y)}.$$

Similarly to 2., when $x \neq x'$ the sum over z is 0 because of the term $(x \oplus x').z$. Otherwise the sum over z is M. Hence we get an upper bound of $4|R| \cdot M \cdot 1/M = 4|R|$.

□

Proof of Lemma 2. We have

$$\mathbb{E}[W] = \sum_{i=1}^{M/2} \mathbb{E}[w(X_i)] = \frac{M}{2} \cdot \frac{1}{\binom{M}{2}} \sum_{\{u,v\} \subseteq [2]^m} w(\{u,v\}) = \frac{1}{M-1} \sum_{u \neq v} w(u \to v).$$

For fixed u the sum of $w(u \to v)$ over all v is ≤ 1 by Claim 3.(2). Also $|w(u \to u)| \leq 1$ by Claim 3.(1). Hence the sum is at most 2 and the expectation is $\leq 2M/(M-1) \leq 3$ for $M \geq c$.

For the claims on W_i we first give an expression for W_i. Let R be the remaining set of $r = M - 2(i-1)$ nodes not touching $x_{<i}$. If $r = 2$ then W_i is constant and the desired bounds are immediate; so let $r \geq 4$. Write $X_i = \{U, V\}$. Let $w(x) := \sum_{y \in R} w(\{x,y\})$ be the sum of the weights of the edges touching x. We can write

$$W_i = w(x_1, \ldots, x_{i-1}) + w(X_i) + \frac{1}{r-3} \left(\sum_{\{x,y\} \subseteq R} w(x,y) - w(U) - w(V) + w(X_i) \right), \quad (3)$$

because after X_i is selected each remaining edge is chosen with prob. $1/(r-3)$. In turn we can write

$$W_i = c_{x_{<i}} + Y_i$$

where $c_{x_{<i}}$ is a function of $x_{<i}$ only and

$$Y_i := \frac{1}{r-3} \left((r-2) w(X_i) - w(U) - w(V) \right).$$

To verify boundedness note

$$|W_i - \mathbb{E}[W_i]| = |Y_i - \mathbb{E}[Y_i]| \leq |Y_i| + \mathbb{E}[|Y_i|].$$

So it suffices to show that $|Y_i| \leq c$. Indeed, $|w(X_i)| \leq 2$ by Claim 3.(1). Also $w(U) + w(V)$ is a sum of $\leq cr$ terms, each of which has absolute value ≤ 2 again by Claim 3.(1).

Finally, we verify the variance bound. We have

$$\mathrm{var} W_i = \mathrm{var} Y_i = \frac{1}{(r-3)^2} \mathrm{var} \left((r-2) w(X_i) - w(U) - w(V) \right).$$

Using $\mathrm{var}(Z) \leq \mathbb{E} Z^2$ and $(x+y+z)^2 \leq 3(x^2+y^2+z^2)$ we bound the variance in the rhs by 3 times

$$(r-2)^2 \mathbb{E} w^2(X_i) + \mathbb{E} \left(w^2(U) + w^2(V) \right).$$

The first expectation is

$$\leq \binom{r}{2}^{-1} \sum_{u \in R, v \in R} \left(w(\{u,v\}) \right)^2 \leq \binom{r}{2}^{-1} cr \leq \frac{c}{r}$$

by Claim 3.(3). □

The second expectation is at most

$$\frac{1}{\binom{r}{2}} \sum_{u \in R, v \in R} \left(w^2(u) + w^2(v)\right) = \frac{2r}{\binom{r}{2}} \sum_{u \in R} w^2(u) = \frac{4}{r-1} \sum_{u \in R} \left(\sum_{x \in R} w(\{u,x\}) \right)^2$$
$$\leq c \sum_{u \in R, x \in R} w^2(\{u,x\}) \leq cr$$

by Claim 3.(3). Altogether, for $r \geq c$ the variance is

$$\leq \frac{c}{r^2} \left(r^2/r + cr\right) \leq c/r.$$

References

[BC25] Chris Brzuska and Geoffroy Couteau. On building fine-grained one-way functions from strong average-case hardness. *Journal of Cryptology*, 38(1), 2025. Article 8.

[BCS24] Balthazar Bauer, Geoffroy Couteau, and Elahe Sadeghi. Fine-grained non-interactive key exchange, revisited. In Leonid Reyzin and Douglas Stebila, editors, *Advances in Cryptology – CRYPTO 2024: 44th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18–22, 2024, Proceedings, Part II*, volume 14921 of *Lecture Notes in Computer Science*, pages 286–312, Cham, 2024. Springer.

[BGI08] Eli Biham, Yaron J. Goren, and Yuval Ishai. Basing weak public-key cryptography on strong one-way functions. In Ran Canetti, editor, *Theory of Cryptography*, volume 4948 of *Lecture Notes in Computer Science*, pages 55–72, Berlin, Heidelberg, 2008. Springer.

[BIK+22] Saikrishna Badrinarayanan, Yuval Ishai, Dakshita Khurana, Amit Sahai, and Daniel Wichs. Refuting the dream XOR lemma via ideal obfuscation and resettable MPC. In Dana Dachman-Soled, editor, *3rd Conference on Information-Theoretic Cryptography (ITC 2022)*, volume 230 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 10:1–10:21, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[CHH+20] Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, Shachar Lovett, and David Zuckerman. XOR lemmas for resilient functions against polynomials. In Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy, editors, *ACM Symp. on the Theory of Computing (STOC)*, pages 234–246. ACM, 2020.

[CM16] Claude Carlet and Sihem Mesnager. Four decades of research on bent functions. *Designs, Codes and Cryptography*, 78(1):5–50, 2016.

[DJMW12] Yevgeniy Dodis, Abhishek Jain, Tal Moran, and Daniel Wichs. Counterexamples to hardness amplification beyond negligible. In Ronald Cramer, editor, *Theory of Cryptography*, volume 7194 of *Lecture Notes in Computer Science*, pages 476–493, Berlin, Heidelberg, 2012. Springer.

[DP09] Devdatt Dubhashi and Alessandro Panconesi. *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, 2009.

[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/.

[GSV18] Aryeh Grinberg, Ronen Shaltiel, and Emanuele Viola. Indistinguishability by adaptive procedures with advice, and lower bounds on hardness amplification proofs. In *IEEE Symp. on Foundations of Computer Science (FOCS)*, 2018.

[Imp95] Russell Impagliazzo. Hard-core distributions for somewhat hard problems. In *IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 538–545, 1995.

[Lev87] Leonid A. Levin. One way functions and pseudorandom generators. *Combinatorica. A Journal of the János Bolyai Mathematical Society*, 7(4):357–363, 1987.

[Lov12] Shachar Lovett. Equivalence of polynomial conjectures in additive combinatorics. *Combinatorica*, 32(5):607–618, 2012.

[TYY11] Keisuke Tanaka, Akihiro Yamada, and Kenji Yasunaga. Weak oblivious transfer from strong one-way functions. In Xavier Boyen and Xiaofeng Chen, editors, *Provable Security*, volume 6980 of *Lecture Notes in Computer Science*, pages 34–51, Berlin, Heidelberg, 2011. Springer.

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

[Vio06] Emanuele Viola. New correlation bounds for GF(2) polynomials using Gowers uniformity. *Electronic Colloquium on Computational Complexity*, Technical Report TR06-097, 2006. www.eccc.uni-trier.de/.

[Vio22] Emanuele Viola. Correlation bounds against polynomials, a survey. 2022.

[Vio26] Emanuele Viola. Simple xor lemma. 2026.

[VW08] Emanuele Viola and Avi Wigderson. Norms, XOR lemmas, and lower bounds for polynomials and protocols. *Theory of Computing*, 4:137–168, 2008.