The Goldreich-Levin Theorem

Noah Stephens-Davidowitz

June 7, 2023

(In my undergraduate cryptography course, I usually skip this lecture, as it's a bit difficult. Nevertheless, it's a beautiful and important theorem with a beautiful proof. These notes follow closely notes of Yael Kalai from a similar lecture that she gave when we co-taught a cryptography class together at MIT.)

1 Recap of hardcore predicates

Remember that a *hardcore predicate* of a function $f : \{0,1\}^* \to \{0,1\}^*$ is an efficiently computable predicate $P : \{0,1\}^* \to \{0,1\}$ such that "it is difficult to compute $P(x)$ given $f(x)$." More formally, for any PPT $\mathcal{A}$,

$$\Pr_{x \sim \{0,1\}^n}[b \leftarrow \mathcal{A}(1^n, f(x)),\ b = P(x)] \le 1/2 + \varepsilon(n)$$

for negligible $\varepsilon(n)$.

We saw that the existence of a hardcore predicate for an injective function f implies the existence of commitment schemes and secure coin-flipping protocols. So, we already know that hardcore predicates pretty useful. And, better still, we will use hardcore predicates to build secret-key encryption (via pseudorandom generators and pseudorandom functions; it's a long journey. . .).

We mentioned (but did not prove) that the most significant bit of the group exponentiation function is hardcore, if the discrete logarithm is hard over the group in question. Indeed, there are by now many results proving that certain predicates are hardcore for specific one-way functions. In particular, these proofs exploit specific properties of the one-way function f in question in order to show that, if there exists an efficient algorithm that guesses $P(x)$ given only $f(x)$ with probability that is non-negligibly larger than $1/2$, then there also exists an efficient algorithm that inverts f with non-negligible probability.

These theorems are useful when you're interested in working with a specific one-way function. But, it would be much more convenient to show some *universal* way to get a hardcore predicate from any one-way function.

The Goldreich-Levin Theorem essentially does that, with a very simple predicate. More formally, it shows how to trivially modify any *injective* one-way function into one with a hardcore predicate. (You can also think of the theorem as giving a "distribution of hardcore predicates," i.e., $P_w : \{0,1\}^* \to \{0,1\}$ such that it is difficult to compute $P_w(x)$ given $f(x)$ for uniformly random $x, w \in \{0,1\}^n$.) The theorem is (of course) due to Goldreich and Levin, in 1989 [GL89].

2 Basic probability tools

Below, we will need three basic probabilistic tools below: the union bound, the Chernoff bound (also called the Chernoff-Hoeffding bound), and Chebyshev's inequality for pairwise independent random variables.

All of these bounds are quite important in their own right—in cryptography, theoretical computer science, and more broadly. So, they're quite good to know, and it's worth spending some time to introduce them.

2.1 Union bound

Union bound is really quite simple, but it comes up so often that it's worth giving it a name. In particular, union bound says that "the probability that either event A *or* event B happens cannot be larger than the sum of the probabilities of events A and B." It is often convenient to apply it to a long list of events $E_1, \ldots, E_\ell$ simultaneously.

Claim 2.1 (Union bound)**.** *For any events $E_1, \ldots, E_\ell$,*

$$\Pr[\exists i,\ E_i \text{ occurs }] \le \sum_i \Pr[E_i]\ .$$

We use this *a lot*, often without even bothering to name it. It is called the union bound presumably because it bounds the probability that the *union* of all of the events happens.

The gigantic benefit of the union bound is that it comes with no restrictions whatsoever on the events E_i. In particular, we don't need any notion of independence whatsoever.

2.2 Chernoff bound

The Chernoff bound (or, at least, the version that we will work with) says that if you have many independent[1] bits $z_1, \ldots, z_m \in \{0,1\}$ with $\Pr[z_i = 1] = p$, then their average value $\frac{1}{m}\sum z_i$ will be very close to p with high probability. This can be used, e.g., to approximate p, or to determine whether the z_i are biased towards 0 or biased towards 1. (The constant 10 in the theorem is not optimal.)

Notice that it is extremely important that the random variables z_i are *independent.* E.g., if we instead took $(z_1, \ldots, z_m) \sim \{0^m, 1^m\}$ (i.e., if with probability 1/2 all of the z_i are zero simultaneously and with probability 1/2 they are all one simultaneously), then be bound would clearly be false.

Theorem 2.2 (Chernoff-Hoeffding bound)**.** *Let $z_1, \ldots, z_m \in \{0,1\}$ be independent random variables with $\Pr[z_j = 1] = p$, and let $Z := \sum z_j$. Then, for every $\eta > 0$,*

$$\Pr[|Z/m - p| \ge \eta] \le 2e^{-\eta^2 m/10}\ .$$

Here, we give a proof for completeness (and because I really like the proof).

[1] Remember that we say that random variables $x_1, \ldots, x_m$ are *independent* if "the distribution of x_i does not depend on the values taken by $\{x_i\ :\ i \ne j\}$." Specifically, for any $y_1, \ldots, y_m$,

$$\Pr[x_1 = y_1, x_2 = y_2, \ldots, x_m = y_m] = \Pr[x_1 = y_1]\Pr[x_2 = y_2]\cdots\Pr[x_m = y_m]\ .$$

Proof. We will prove that $\Pr[Z/m - p \geq \eta] \leq e^{-\eta^2 m/10}$. A nearly identical proof shows that $\Pr[Z/m - p \leq -\eta] \leq e^{-\eta^2 m/10}$. And, the result then follows by union bound. (In other words, the event $|Z/m - p| \geq \eta$ is equivalent to the event that $Z/m - p \geq \eta$ *or* $Z/m - p \leq -\eta$, so by union bound, the probability of $|Z/m - p| \geq \eta$ is at most the sum of these probabilities.) For simplicity, we assume $p = 1/2$. The proof for arbitrary p is similar but requires a bit more work.

The idea is to study the random variable $e^{tZ} = \prod_j e^{tz_j}$ for some $t \in \mathbb{R}$ that we will choose later, depending on η. Because the z_j are independent, we have that the expectation of the product is the product of the expectations, i.e.,

$$\mathbb{E}\Big[\prod_j e^{tz_j}\Big] = \prod_j \mathbb{E}[e^{tz_j}] \; .$$

But, these expectations are easy to compute. Specifically, $\mathbb{E}[e^{tz_j}] = e^t/2 + 1/2$. Therefore, $\mathbb{E}[e^{tZ}] = (e^t+1)^m/2^m$.

On the other hand, we also have

$$\mathbb{E}[e^{tZ}] \geq \Pr[Z \geq m/2 + \eta m] \cdot e^{tm/2 + t\eta m} \; .$$

(This is known as Markov's inequality, i.e., that for a non-negative random variable $X \geq 0$ and $t \geq 0$, $\mathbb{E}[X] \geq t \Pr[X \geq t]$. We typically use it in the form $\Pr[X \geq t] \leq \mathbb{E}[X]/t$.) Rearranging, we see that

$$\Pr[Z \geq m/2 + \eta m] \leq \Big(\frac{e^t + 1}{e^{t/2 + t\eta}}\Big)^m / 2^m \; .$$

Finally, we choose t to minimize the right-hand side. In particular, taking $e^t := (1/2+\eta)/(1/2-\eta)$ (noting that we may $\eta < 1/2$, since the theorem is trivial otherwise), we have

$$\Pr[Z \geq pm + \eta m] \leq \Big((1+2\eta)^{1/2+\eta}(1-2\eta)^{1/2-\eta}\Big)^{-m} \leq e^{-m\eta^2/3} \; ,$$

where the last inequality follows from the fact that $(1+2\eta)^{1/2+\eta}(1-2\eta)^{1/2-\eta}/e^{-\eta^2/10}$ is a decreasing function of η for $0 \leq \eta \leq 1/2$. □

Crucially, the Chernoff bound only works for *independent* z_i. A very very common mistake that people make in cryptography (and any field that uses a lot of probability) is to apply a result that requires independence when they do not have it.

2.3 Chebyshev's inequality

Chebyshev's inequality (or, again, the version that we will see) gives a weaker bound than the Chernoff bound, but it requires a weaker assumption. Specifically, Chebyshev's inequality only requires that the z_i are *pairwise independent*, that is, that $\Pr[z_i = 1 \text{ and } z_j = 1] = p^2$ for all $i \neq j$. However, it could still be the case that, say, $\Pr[z_i = 1 \text{ and } z_j = 1 \text{ and } z_k = 1] = 0$. For example, consider (z_1, z_2, z_3) that are sampled uniformly from $\{(z_1, z_2, z_3) \in \{0,1\}^3 \; : \; z_1 \oplus z_2 = z_3\}$. These z_j are pairwise independent but *not* independent.

More generally, random variables $r_1, \ldots, r_m$ are pairwise independent if for any $i \neq j$ and any s_i, s_j, $\Pr[r_i = s_i \text{ and } r_j = s_j] = \Pr[r_i = s_i]\Pr[r_j = s_j]$.

Theorem 2.3 (Chebyshev's inequality)**.** *Let* $z_1, \ldots, z_m \in \{0,1\}$ *be* pairwise *independent random variables with* $\Pr[z_j = 1] = p$*, and let* $Z := \sum z_j$*. Then, for every* $\eta > 0$*,*

$$\Pr[|Z/m - p| \geq \eta] \leq p(1-p)/(\eta^2 m) \leq 1/(4\eta^2 m) \; .$$

Again, we provide a proof for completeness. It's quite a bit simpler than the proof of the Chernoff bound.

Proof. The trick here is to study $\mathbb{E}[(Z - mp)^2]$ (i.e., the variance of Z). Notice that

$$\mathbb{E}[(Z-mp)^2] = \mathbb{E}\Big[\Big(\sum_j z_j - mp\Big)^2\Big] = \mathbb{E}\Big[\sum_j z_j^2 + \sum_{i \neq j} z_i z_j - 2mp\sum_j z_j + m^2p^2\Big] = mp + \sum_{i\neq j} \mathbb{E}[z_i z_j] - m^2p^2 \; .$$

Since the z_i, z_j are pairwise independent, we have

$$\sum_{i \neq j} \mathbb{E}[z_i z_j] = \sum_{i \neq j} \mathbb{E}[z_i]\,\mathbb{E}[z_j] = m(m-1)p^2 \; .$$

Therefore,

$$\mathbb{E}[(Z - mp)^2] = mp(1-p) \; .$$

(You might recognize this as the variance of a Binomial random variable.) Finally, (again, using Markov's inequality) we have

$$mp(1-p) = \mathbb{E}[(Z-mp)^2] \geq \Pr[|Z/m - p| \geq \eta] \cdot (\eta m)^2 \; .$$

Rearranging gives

$$\Pr[|Z/m - p| \geq \eta] \leq p(1-p)/(\eta^2 m) \leq 1/(4\eta^2 m) \; ,$$

as needed. □

3 The Goldreich-Levin Theorem

We will need some notation. For two bit strings $x, w \in \{0,1\}^n$ with $x = (x_1, \ldots, x_n)$ and $w = (w_1, \ldots, w_n)$, we write $\langle x, w\rangle := x_1w_1 \oplus \cdots \oplus x_nw_n$. In other words, $\langle x, w \rangle$ is the parity of the number of indices i such that $x_i = w_i = 1$. (This is also simply the inner product function over $\mathbb{F}_2^n$.)

Fix some function $f : \{0,1\}^* \to \{0,1\}^*$. Define $g(x, w) := (f(x), w)$ where x and w have the same length. (As defined here, g can only take inputs with even length, so that we can divide its input into two equal-length strings. We would ideally like g to take arbitrary bit strings as input, which we could do by just ignoring the last bit of the input if the input has odd length. But, we do not worry about this here.) Notice that g is one-way if (and only if) f is one-way. Notice also that g is injective if and only if f is injective, and g is a permutation if and only if f is a permutation.

Here is the theorem that we are going to prove.

Theorem 3.1 (Goldreich-Levin)**.** *If* f *is a one-way function, then the predicate* $P(x, w) := \langle w, x\rangle$ *is a hardcore predicate for* g*.*

(Below is a long explanation of the proof and where it came from. Skip to the last section if you just want to see the formal proof.)

We will of course prove this theorem via a reduction. So, suppose that there exists a PPT adversary $\mathcal{A}$ with

$$\Pr_{x,w\sim\{0,1\}^n}[b \leftarrow \mathcal{A}(1^n, f(x), w),\ b = \langle w, x\rangle] = 1/2 + \varepsilon(n)$$

for non-negligible $\varepsilon(n)$. We want to construct an adversary that takes as input (1^n and) $y^* := f(x^*)$ for $x^* \sim \{0,1\}^n$ and outputs x' such that $f(x') = y^*$ with non-negligible probability.

At a high level, there's only one thing to try. Presumably, $\mathcal{A}'$ is going to choose $w_1, \ldots, w_\ell \in \{0,1\}^n$ in some way, call $\mathcal{A}$ on input $(f(x^*), w_i)$, and somehow use the resulting output of $\mathcal{A}$ to try to compute x.

However, before we even try to fill in the details, we already have a problem. Our assumption only guarantees that $\mathcal{A}$ has some advantage $\varepsilon(n)$ on input $(f(x), w)$ for uniformly random $w \in \{0,1\}^n$ *and* $x \in \{0,1\}^n$. But, we want to call $\mathcal{A}$ on input $(f(x^*), w_i)$ for many different $w_i \in \{0,1\}^n$, but one fixed value of $x^* \in \{0,1\}^n$. What if $\mathcal{A}$ refuses to cooperate on our particular value of x^*?

To capture this, we define the following set:

$$\mathsf{GOOD} := \{x \in \{0,1\}^n \ :\ \Pr_{w\sim\{0,1\}^n}[b \leftarrow \mathcal{A}(f(x), w),\ b = \langle w, x\rangle] \geq 1/2 + \varepsilon(n)/2\}\ .$$

In other words, if $x^* \in \mathsf{GOOD}$, then $\mathcal{A}$ has non-negligible advantage in guessing $\langle w, x^*\rangle$ for random w (but fixed x^*).

Claim 3.2.

$$\Pr_{x\sim\{0,1\}^n}[x \in \mathsf{GOOD}] \geq \varepsilon(n)/2\ .$$

Proof. We have

$$\begin{aligned}1/2 + \varepsilon(n) &= \Pr_{x,w\in\{0,1\}^n}[b = \langle w, x\rangle \text{ and } x \in \mathsf{GOOD}] + \Pr[b = \langle w, x\rangle \text{ and } x \notin \mathsf{GOOD}]\\ &\leq \Pr[x \in \mathsf{GOOD}] + 1/2 + \varepsilon(n)/2\ ,\end{aligned}$$

where we have used the fact that for $x \notin \mathsf{GOOD}$, the probability of success is less than $1/2 + \varepsilon(n)/2$ by definition. The result follows by rearranging. □

In particular, the above claim tells us that, if $\varepsilon(n)$ is non-negligible, then $x^* \in \mathsf{GOOD}$ with non-negligible probability. Therefore, it suffices to invert f with non-negligible probability in the special case when $x^* \in \mathsf{GOOD}$.

In the following subsections, we will work our way up to the proof of the actual theorem by proving the theorem in certain special cases.

3.1 First special case: 100% success

First, suppose that for $x \in \mathsf{GOOD}$,

$$\Pr_{w\in\{0,1\}^n}[b \leftarrow \mathcal{A}(1^n, f(x), w),\ b = \langle w, x\rangle] = 1\ .$$

(This is a very strong assumption!) Then, we can have $\mathcal{A}'$ behave as follows. Let $e_i := (0, \ldots, 0, 1, 0, \ldots, 0) \in \{0,1\}^n$ be the bit string with a 1 in its ith coordinate and zeros elsewhere. So, $\mathcal{A}'$ simply runs the algorithm $\mathcal{A}$ n times, on input $(1^n, y^*, e_1)$, $(1^n, y^*, e_2)$, $\ldots$, $(1^n, y^*, e_n)$, receiving as output bits $x'_1, \ldots, x'_n$. The adversary $\mathcal{A}'$ then simply outputs $x' := (x'_1, \ldots, x'_n)$.

Clearly this version of $\mathcal{A}'$ is PPT if $\mathcal{A}$ is PPT. By our (ridiculously strong) assumption, if $x^* \in \mathsf{GOOD}$, then $x'_i = \langle x^*, e_i \rangle$, i.e., in this case $x' = x^*$. So,

$$\Pr[x' = x^*] \geq \Pr[x^* \in \mathsf{GOOD}] \geq \varepsilon(n)/2 \ ,$$

where we have used Claim 3.2. In particular, if $\mathcal{A}$ has non-negligible advantage, then so does $\mathcal{A}'$.

So, in this very very special case, the theorem is true.

Since this special case was so very special, and since the reduction in this case was so easy, it's worth playing with it a bit to get some idea for how we might modify it. Perhaps the most ridiculous aspect of this reduction is that we assumed that $\mathcal{A}$ is well-behaved on *specific* choices of w. E.g., the above reduction would fail even if $\mathcal{A}$ gave correct output for every choice of w *except* e_1. This is silly.

So, here's a slightly more clever version of this reduction that works even if $\mathcal{A}$ fails on some negligible fraction of w. Instead of calling $\mathcal{A}$ on input (y^*, e_i), we can call it on input (y^*, w) and $(y^*, w \oplus e_i)$, receiving as output bits $b_{i,0}$ and $b_{i,1}$ If $\mathcal{A}$ gives us the correct output, $b_{i,0} = \langle x^*, w \rangle$ and $b_{i,1} = \langle x^*, w \oplus e_i \rangle$, then notice that the ith bit x^*_i of x^* is exactly $b_{i,0} \oplus b_{i,1}$. (Here, we are using the linearity of the inner product: $\langle x, w \rangle \oplus \langle x, w' \rangle = \langle x, w \oplus w' \rangle$).

Below, we will use this idea quite a bit.

3.2 Second special case: 76% success

Now, suppose that for $x \in \mathsf{GOOD}$,

$$\Pr_{w \in \{0,1\}^n}[b \leftarrow \mathcal{A}(1^n, f(x), w),\ b = \langle w, x \rangle] = 3/4 + \delta(n) \ , \tag{1}$$

where $\delta(n) > 0$ is non-negligible. (This is still quite a strong assumption, but not as ridiculous as the 100% success assumption that we made earlier.)

Claim 3.3. *If Eq. (1) holds and $x^* \in \mathsf{GOOD}$, then for all $i \in [n]$,*

$$\Pr_{w \sim \{0,1\}^n}[b_1 \leftarrow \mathcal{A}(1^n, f(x^*), w),\ b_2 \leftarrow \mathcal{A}(1^n, f(x^*), w \oplus e_i),\ b_1 \oplus b_2 = x^*_i] \geq 1/2 + 2\delta(n) \ .$$

Proof. We have

$$\begin{aligned}\Pr[b_1 \oplus b_2 \neq x^*_i] &\leq \Pr[b_1 \neq \langle w, x^* \rangle \text{ or } b_2 \neq \langle w \oplus e_i, x^* \rangle] \\ &\leq \Pr[b_1 \neq \langle w, x^* \rangle] + \Pr[b_2 \neq \langle w \oplus e_i, x^* \rangle] \ .\end{aligned}$$

(This last inequality is the union bound, Claim 2.1. It says that the probability of at least one event E_j happening is bounded by the sum of the probabilities of the E_j.) Finally, by Eq. (1), the above probability is $1/2 - 2\delta(n)$, as needed. (Here, we have used the fact that $w \oplus e_i$ is uniformly random if w is uniformly random.) □

So, Claim 3.3 gives us a way to compute a bit (namely, the bit $b_1 \oplus b_2$) that equals the ith bit of our input x_i^* with probability $1/2 + 2\delta(n)$. But, if we just use this bit as our guess for x_i^* for each i, then it is very unlikely that we will guess *all* bits of x_i^* correctly. We need some way to amplify our probability of success, so that we can find a procedure that guesses the ith bit correctly with probability close to one.

Of course, the obvious thing to do is to run the procedure suggested by Claim 3.3 many times, giving us many guesses $x'_{i,1}, \ldots, x'_{i,m} \in \{0,1\}$ for x_i^* and then to take the majority of the $x'_{i,j}$ as our true guess x'_i for x_i^*. Intuitively, if we use sufficiently many guesses $x'_{i,j}$, this procedure should succeed with high probability.

The Chernoff bound, Theorem 2.2, makes this precise. In particular, let $z_j \in \{0,1\}$ be the random variable that equals 1 if $x'_{i,j} = x_i^*$ and 0 otherwise. By Claim 3.3, we see that $\Pr[z_j = 1] \geq 1/2 + 2\delta(n)$. Therefore, if we take the number m of guesses $x'_{i,j}$ to be $m := \lceil 100n/\delta(n)^2 \rceil$, then by Chernoff bound (assuming that we sample the random choice w independently each time that we do this), we see that

$$\Pr\Big[\frac{1}{m} \cdot \sum z_j \geq 1/2\Big] \leq 2e^{-(2\delta(n))^2 m/10} \leq e^{-n} \; .$$

Therefore, by repeating the procedure from Claim 3.3 m times, we can guess x_i^* correctly with probability $1 - e^{-n}$.

So, the following reduction will succeed in this special case. The adversary $\mathcal{A}'$ will sample $w_{i,j} \in \{0,1\}^n$ uniformly at random for $i = 1, \ldots, n$ and $j = 1, \ldots, m = \lceil 100n/\delta(n)^2 \rceil$. For each i, j, it calls the adversary $\mathcal{A}$ on input $(1^n, y^*, w_{i,j})$ and then again on input $(1^n, y^*, w_{i,j} \oplus e_i)$, receiving as output $b_{i,j,0}$ and $b_{i,j,1}$. It sets $x'_{i,j} := b_{i,j,0} \oplus b_{i,j,1}$, and for each i, it sets x'_i to be the majority of the $x'_{i,j}$. Finally, it outputs $x' := (x'_1, \ldots, x'_n)$.

Since $\delta(n)$ is non-negligible and $\mathcal{A}$ runs in polynomial time, $\mathcal{A}'$ also runs in polynomial time. Furthermore, by the argument above, if $x^* \in \mathsf{GOOD}$, then $x'_i = x_i^*$ with probability at least $1 - e^{-n}$, so that in this case $\Pr[x' = x^*] \geq 1 - ne^{-n}$.[2] Finally, since $x^* \in \mathsf{GOOD}$ with probability at least $\varepsilon(n)/2$ (Claim 3.2), we see that the advantage of $\mathcal{A}'$ is at least $\varepsilon(n)(1 - ne^{-n})/2$, which is non-negligible if $\varepsilon(n)$ is non-negligible, as needed.

3.3 The actual proof (informally)

In this section, we describe the actual proof, but we are more interested in explaining where it comes from then in writing a formal proof. In the next section, I actually wrote the formal proof as a nice succinct, formal proof—without all the commentary.

The reason that we needed a success probability of greater than 75% above is because, in order to get a decent guess for a *single* bit $x_i^* = \langle e_i, x^* \rangle$, we needed to combine guesses for *two* bits $\langle w, x^* \rangle$ *and* $\langle w \oplus e_i, x^* \rangle$. An adversary that has success probability only slightly larger than 1/2 could choose to make her guess wrong for one of $\langle w, x^* \rangle$ or $\langle w \oplus e_i, x^* \rangle$ for almost all choices of w. (E.g., the adversary could be always be correct when the ith bit is 0, but be wrong with probability $1 - 1/n$ when the ith bit is 1. This adversary still has non-negligible advantage, but our guess for the bit $\langle e_i, x^* \rangle$ would be wrong with probability $1 - 1/n$ in this case!)

The Goldreich-Levin reduction starts with the following (ridiculous!) idea. Suppose that for each pair $(w, w \oplus e_i)$, we simply guess the bit $\langle w, x^* \rangle$ ourselves, and we only use the adversary $\mathcal{A}$ to

[2] This is one of the (many) time in which we have applied union bound without explicitly mentioning it. In particular, we skipped a step here in which we write $\Pr[x' \neq x^*] = \Pr[x'_1 \neq x_1^* \text{ or } x'_2 \neq x_2^* \text{ or } \ldots \text{ or } x'_n \neq x_n^*] \leq \sum_i \Pr[x'_i \neq x_i^*] \leq ne^{-n}$.

compute $\langle w \oplus e_i, x^* \rangle$. If we could somehow manage to always guess the bit $\langle w, x^* \rangle$ correctly, then the adversary would guess $\langle w \oplus e_i, x^* \rangle$ correctly with probability $1/2 + \varepsilon(n)/2$ (for $x^* \in \mathsf{GOOD}$). This would be enough to make the above argument go through.

Of course, we obviously cannot hope to guess $\langle w, x^* \rangle$ ourselves with probability better than random—this is exactly the task that we needed $\mathcal{A}$ for in the first place! What we can, of course, trivially do is guess each of these bits with success probability $1/2$. But, in the above procedure, for each i we needed to choose a total of m different values of $w \in \{0,1\}^n$. And, if we just guess $\langle w, x^* \rangle$ randomly, then the probability that we will get them *all* right is just 2^{-m}.

Here is the really clever idea: suppose that we happen to correctly guess $\langle s_1, x^* \rangle$ *and* $\langle s_2, x^* \rangle$ for $s_1, s_2 \in \{0,1\}^n$. Notice that in this case we also know $\langle s_1 \oplus s_2, x^* \rangle = \langle s_1, x^* \rangle \oplus \langle s_2, x^* \rangle$. More generally, suppose that we happen to correctly guess $\langle s_1, x^* \rangle, \ldots, \langle s_\ell, x^* \rangle$. Then, for *all* subsets $S \subseteq \{1, \ldots, \ell\}$, we also know $\langle w_S, x^* \rangle = \bigoplus_{j \in S} \langle s_j, x^* \rangle$, where $w_S := \bigoplus_{j \in S} s_j$. There are 2^ℓ such subsets, so we learn 2^ℓ inner products $\langle r_S, x^* \rangle$.

This means that, in order to learn m different bits of the form $\langle w_S, x^* \rangle$, we only need to correctly guess $\ell \approx \log m$ bits of the form $\langle s_j, x^* \rangle$.

So, here is the actual reduction. $\mathcal{A}'$ will first choose random strings $s_i \in \{0,1\}^n$ and random "guess" bits c_j for $j = 1, \ldots, \ell$. Our hope is that $c_j = \langle s_j, x^* \rangle$. Then, for every (non-empty) subset $S \subseteq [\ell]$, we define $b_S := \bigoplus_{j \in S} c_j$ and $w_S := \bigoplus_{j \in S} s_j$. Then, for every $i = 1, \ldots, n$ and every subset $S \subseteq [\ell]$, $\mathcal{A}'$ runs the adversary $\mathcal{A}$ on input $(1^n, y^*, w_S \oplus e_i)$, receiving as output $b'_{i,S}$. $\mathcal{A}'$ sets x'_i to be the majority over all S of the $b'_{i,S} \oplus b_S$, and outputs $x' := (x'_1, \ldots, x'_n)$.

We now wish to prove that the above algorithm succeeds with non-negligible probability, assuming that $\varepsilon(n)$ is non-negligible, for some appropriate choice of $m = \mathrm{poly}(n, 1/\varepsilon(n))$. As before, we first notice that we may assume that $x^* \in \mathsf{GOOD}$. Next, we notice that with probability $2^{-\ell} \approx 1/m$, we have $c_j = \langle s_j, x^* \rangle$ for all $j = 1, \ldots, \ell$. This is non-negligible probability, so we may assume that each of the c_j's is "correct," i.e., $c_j = \langle s_j, x^* \rangle$ for all j. Of course, this also implies that $b_S = \langle w_S, x^* \rangle$ for all $S \subseteq [n]$!

Now, let $z_{i,S} \in \{0,1\}$ be the random variable that is one if $b'_{i,S} = \langle r_S, x^* \rangle$ and zero otherwise. Notice that $r_S \in \{0,1\}$ is in fact a uniformly random bit string (for non-empty S) and $x^* \in \mathsf{GOOD}$, so we have $\Pr_{r_S}[z_{i,S}] \geq 1/2 + \varepsilon(n)/2$. Therefore, if we could apply the Chernoff bound, we would immediately see that $x_i = x'_i$ with high probability.

However, we cannot apply the Chernoff bound because we did not choose $r_{\{1\}}, r_{\{2\}}, \ldots, r_{\{1,\ldots,m\}}$ independently. E.g., we have the relationship that $r_{\{1,2\}} = r_{\{1\}} \oplus r_{\{2\}}$, so clearly these are not independent random variables. They are, however, *pairwise independent.*

Claim 3.4. *For fixed i, the bits $z_{i,S} \in \{0,1\}$ are pairwise independent. Therefore,*

$$\Pr[x'_i \neq x^*_i \mid x^* \in \mathsf{GOOD} \text{ and } \forall j,\ c_j = \langle s_j, x^* \rangle] \leq 1/(4\varepsilon(n)^2 m)\ .$$

In particular, taking $m \geq n/\varepsilon(n)^2$, this probability is at most $1/(4n)$, so that with probability at least $1 - n/(4n) = 3/4$, all *bits will be correct simultaneously.*

Proof. It suffices to show that the bit strings r_S are pairwise independent. In other words, let $S \neq S'$ be distinct sets. Let $T := S \oplus S'$ be the set of elements that are in *one* of the sets S, S' but *not* both. Clearly, $r_S = r_{S'} \oplus r_T$. Since T is non-empty (because $S \neq S'$), r_T is a uniformly random bit string, and it follows immediately that r_S and $r_{S'}$ are independent, as needed.

The displayed equation then follows from Chebyshev's inequality (Theorem 2.3). The "in particular" follows from union bound. □

The above then shows that if $\varepsilon(n) \geq 1/\mathrm{poly}(n)$ (for infinitely many n), then $m \leq \mathrm{poly}(n)$ (for infinitely many n), so that the above reduction runs in polynomial time

3.4 The formal proof

For completeness, here we give a formal, succinct proof of the theorem. (This is slightly more succinct than it should be because I refer back to Claims 3.2 and 3.4. But, those claims are themselves quite short.)

Proof of the Goldreich-Levin Theorem. Suppose that there exists a PPT adversary $\mathcal{A}$ with

$$\Pr_{x,w\in\{0,1\}^n}[b \leftarrow \mathcal{A}(f(x),w),\ b = \langle w,x\rangle] \geq 1/2 + 1/n^C \tag{2}$$

for some constant $C > 0$ for infinitely many values of n. Then, we construct an adversary $\mathcal{A}'$ that inverts f with non-negligible probability as follows. $\mathcal{A}'$ takes as input 1^n and $y^* := f(x^*)$ for uniformly random $x^* \in \{0,1\}^n$. Let $m := \lceil n^{C+1} \rceil$ and $\ell := \lceil \log m \rceil$. For $j = 1,\ldots,\ell$, $\mathcal{A}'$ samples $s_j \sim \{0,1\}^n$ and $c_j \sim \{0,1\}$ uniformly at random. For each non-empty subset $S \subseteq [\ell]$, let $b_S := \bigoplus_{j\in S} c_j$ and $w_S := \bigoplus_{j\in S} s_j$.

$\mathcal{A}'$ then calls the adversary $\mathcal{A}$ on input $(y^*, w_S \oplus e_i)$ for all non-empty subsets $S \subseteq [\ell]$ and all i, receiving as output $b'_{i,S} \in \{0,1\}$. It then sets x'_i to be the majority over all non-empty S of $b'_{i,S} \oplus b_S$. Finally, it outputs $x' := (x'_1,\ldots,x'_n)$.

Clearly, $\mathcal{A}'$ runs in polynomial time if $\mathcal{A}$ does. Let n be such that Eq. (2) holds. Then, it suffices to prove that $\Pr[x' = x^*] \geq 1/(16n^{2C+1})$. Indeed, let

$$\mathsf{GOOD} := \{x \in \{0,1\}^n \ :\ \Pr_{w\sim\{0,1\}^n}[b \leftarrow \mathcal{A}(f(x),w),\ b = \langle w,x\rangle] \geq 1/2 + 1/(2n^C)\}\ .$$

By Claim 3.2, we have

$$\Pr[x^* \in \mathsf{GOOD}] \geq 1/(2n^C)\ .$$

So,

$$\begin{aligned}\Pr[x' = x^*] &\geq \Pr[x^* \in \mathsf{GOOD} \text{ and } x' = x^*] \\ &\geq \frac{1}{2n^C} \cdot \Pr[x' = x^* \mid x^* \in \mathsf{GOOD}]\ .\end{aligned}$$

Furthermore,

$$\begin{aligned}\Pr[x' = x^* \mid x^* \in \mathsf{GOOD}] &\geq \Pr[x' = x^* \text{ and } \forall j, c_j = \langle s_j, x^*\rangle \mid x^* \in \mathsf{GOOD}] \\ &= 2^{-j} \Pr[x' = x^* \mid x^* \in \mathsf{GOOD} \text{ and } \forall j, c_j = \langle s_j, x^*\rangle] \\ &\geq \frac{1}{4n^{C+1}} \cdot \Pr[x' = x^* \mid x^* \in \mathsf{GOOD} \text{ and } \forall j, c_j = \langle s_j, x^*\rangle]\ .\end{aligned}$$

So, it suffices to show that, conditioned on $x^* \in \mathsf{GOOD}$ *and* $c_j = \langle s_j, x^*\rangle$ for all j, $\Pr[x = x] \geq 1/2$. But, notice that this is exactly Claim 3.4. □

References

[GL89] Oded Goldreich and Leonid A. Levin. A Hard-Core Predicate for all One-Way Functions. In *STOC*, 1989. 1