

Homework is due by **7am of Sep 28**. Send by email to both “regev” (under the cs.nyu.edu domain) and “mgeorgiou@nyu.edu” with subject line “CSCI-GA 3210 Homework 2” and name the attachment “YOUR NAME HW2.tex/pdf”. No need for a printed copy. Start early!

Instructions. Solutions must be typeset in \LaTeX (a template for this homework is available on the course web page). Your work will be graded on *correctness*, *clarity*, and *conciseness*. You should only submit work that you believe to be correct; if you cannot solve a problem completely, you will get significantly more partial credit if you clearly identify the gap(s) in your solution. It is good practice to start any long solution with an informal (but accurate) “proof summary” that describes the main idea.

You are expected to read all the hints either before or after submission, but before the next class.

You may collaborate with others on this problem set and consult external sources. However, you must **write your own solutions**. You must also **list your collaborators/sources** for each problem.

1. (*Statistical distance*.) Recall that given two distributions over a (finite) set Ω , their statistical distance (also known as variational or L_1 distance) is defined as

$$\Delta(X, Y) := \frac{1}{2} \sum_{\omega \in \Omega} |X(\omega) - Y(\omega)| .$$

- (3 points) Show that Δ defines a metric (see here for the definition).
- (3 points) Show that the following is an equivalent definition:

$$\Delta(X, Y) := \sup_{A \subseteq \Omega} |X(A) - Y(A)| ,$$

where $X(A)$ denotes the probability of X to be in A , and similarly for $Y(A)$. Give an “operational” interpretation to this definition (i.e., in terms of an algorithm trying to distinguish X and Y).

- (c) (3 points) Let D_0 and D_1 be two distributions over the same support Ω . Suppose that we play the following game with an algorithm \mathcal{A} . First, we pick at random a bit $b \leftarrow \{0, 1\}$ and then we pick $x \leftarrow D_b$ and we give x to \mathcal{A} . Finally, \mathcal{A} returns a bit $\mathcal{A}(x)$. It wins if the bit returned is equal to b . Show that the highest success probability in this game is exactly $\frac{1}{2} + \frac{1}{2}\Delta(D_0, D_1)$.
- (d) (0 points) (Not for submission) Revisit Question (2d) from Homework 1 in light of this question

2. (*Weak vs strong one-way functions*.[♦]) Recall that we say that $f : \{0, 1\}^n \rightarrow \{0, 1\}$ is a one-way function if there is an efficient algorithm for computing it, and moreover, for any PPT algorithm I ,

$$\Pr_{x \in \{0,1\}^n} [I(1^n, f(x)) \in f^{-1}(f(x))] \in \text{negl}(n) , \quad (0.1)$$

where the 1^n is simply a convenient hack for allowing I to run in time $\text{poly}(n)$ (which would not be the case otherwise if the output of f happens to be short). One can also consider a variant of this definition, known as a *weak* one-way function, saying that there exists a constant $c > 0$ such that for any PPT I , Equation (0.1) holds with $< 1 - n^{-c}$ instead of $\in \text{negl}(n)$. As their names suggest, any (strong) one-way function is also a weak one-way function (make sure you see why). Can you construct an example of a weak one-way function that is not a strong one-way function? (You can assume that strong one-way functions exist) Can you think of a way to create a strong one-way function from a weak one-way function?

[♦]Again, this is a question meant to encourage you to think; you are not required to solve it fully, but you are required to demonstrate that you thought about it seriously.

3. (Fun with one-way functions.)

- (a) (2 points) Assume we modify the definition of a one-way function by allowing the adversary to output a *list* of supposed preimages, and he wins if at least one of them is a valid preimage (and as before the winning probability of any efficient adversary should be negligible). How does this modified definition compare with the original one? Formally prove your answer.
- (b) (2 points)² For a security parameter n , define $f : \{2^{n-1}, \dots, 2^n\} \rightarrow \{1, \dots, 2^{2n}\}$ by $f(x) = x^2$ (over the integers). Is it a one-way function? (Rabin's function is similar, except it's done in \mathbb{Z}_N)
- (c) (4 points)³ Suppose that $f : \{0, 1\}^* \rightarrow \{0, 1\}^*$ is such that $|f(x)| \leq c \log|x|$ for every $x \in \{0, 1\}^*$, where $c > 0$ is some fixed constant. (Here $|\cdot|$ denotes the length of a string.) Prove that f is *not* a one-way function. (You may use a *non-uniform* inverter in your solution; for one bonus point, use a uniform one. Click to read about uniform vs non-uniform (ID 84841))
- (d) (5 points)² Assume $g : \{0, 1\}^n \rightarrow \{0, 1\}^n$ is a one-way function. Is the function $f : \{0, 1\}^{2n} \rightarrow \{0, 1\}^{2n}$ defined by $f(x_1, x_2) = (g(x_1), g(x_1 \oplus x_2))$ necessarily also a one-way function?
- (e) (3 points) (bonus⁴) Show that there exists a one-way function $f : \{0, 1\}^n \rightarrow \{0, 1\}^n$ for which the function $f'(x) := f(x) \oplus x$ is *not* one-way. You can assume the existence of a one-way function $g : \{0, 1\}^n \rightarrow \{0, 1\}^n$ for all n . I need a hint for 1/2 points! (ID 82778)

4. (6 points) (Worst-case to average-case reduction.³) Let N be the product of two distinct n -bit primes, and suppose there is an efficient algorithm \mathcal{A} that computes square roots on a noticeable fraction of quadratic residues mod N :

$$\Pr_{y \leftarrow \mathbb{QR}_N^*} [\mathcal{A}(N, y) \in \sqrt{y} \bmod N] = \delta \geq 1/\text{poly}(n).$$

Construct an efficient algorithm \mathcal{B} that, using \mathcal{A} as an oracle, computes the square root of *any* $y \in \mathbb{QR}_N^*$ with *overwhelming* probability (solely over the random coins of \mathcal{A} and \mathcal{B}). That is, for every $y \in \mathbb{QR}_N^*$, it should be the case that

$$\Pr[\mathcal{B}^{\mathcal{A}}(N, y) \in \sqrt{y} \bmod N] = 1 - \text{negl}(n).$$

Explain in your own words why such reductions are known as worst-case to average-case reductions.

5. (PRG) Try to think how to precisely define the property that a function $f : \{0, 1\}^n \rightarrow \{0, 1\}^{n+1}$ satisfies that $f(U)$ “looks” like a uniform string in $\{0, 1\}^{n+1}$ where U is sampled uniformly from $\{0, 1\}^n$. There is no need to write down your solution: just think about it in preparation for Monday's class. Such efficiently computable functions are known as *pseudorandom generators*.

²A question from Dodis's class

³A question from Peikert's class

⁴By Bao Feng, as appears in Goldreich's book