# Lecture 37: Noise Operator

# Overview

- Today we shall introduce the basics of the "noise operator"
- This operator is crucial to one of the most powerful technical tools in Fourier Analysis, namely, the Hypercontractivity

# Noise Operator

- Let $\mathbb{N}_\varepsilon$ be a probability distribution over the sample space $\{0,1\}^n$ such that

$$\mathbb{P}\left[\mathbb{N}_\varepsilon = x\right] = (1-\varepsilon)^{n-|x|}\varepsilon^{|x|}$$

Here $|x|$ represents the number of 1s in $x$ (or, equivalently, the Hamming weight of $x$)

- Intuitively, imagine a channel through which $0^n$ is being fed as input. The channel converts each bit independently as follows. It converts $0 \mapsto 1$ with probability $\varepsilon$; and $1 \mapsto 0$ with probability $(1-\varepsilon)$. Note that the probability of the output being $x$ is $(1-\varepsilon)^{n-|x|}\varepsilon^{|x|}$
- Our objective is to prove that

$$\text{bias}_{\mathbb{N}_\varepsilon}(S) = (1-2\varepsilon)^{|S|}$$

We shall prove this result using a highly modular and elegant approach

- For $1 \leqslant i \leqslant n$, let $\mathbb{N}_{\varepsilon,i}$ be the probability distribution defined below

$$\mathbb{P}\left[\mathbb{N}_{\varepsilon,i} = x\right] = \begin{cases} (1-\varepsilon), & \text{if } x = 0^n \\ \varepsilon, & \text{if } x = \delta_i \\ 0, & \text{otherwise} \end{cases}$$

- Intuitively, $0^n$ is fed through a channel. All bits except the $i$-th bit is left unchanged. The $i$-th bit is converted as follows. It maps $0 \mapsto 1$ with probability $\varepsilon$; and $1 \mapsto 0$ with probability $(1-\varepsilon)$.

- Let us compute the bias of this distribution. For any $S \in \{0,1\}^n$, note that, if $S_i = 0$, we have

$$\mathrm{bias}_{\mathbb{N}_{\varepsilon,i}}(S) = 1$$

For any $S \in \{0,1\}$, if $S_i = 1$, we have

$$\mathrm{bias}_{\mathbb{N}_{\varepsilon,i}}(S) = (1-\varepsilon) - \varepsilon = (1-2\varepsilon)$$

- Succinctly, we can express this as

$$\mathrm{bias}_{\mathbb{N}_{\varepsilon,i}}(S) = (1-2\varepsilon)^{S_i}$$

- So, we can conclude that

$$\mathrm{bias}_{\bigoplus_{i=1}^n \mathbb{N}_{\varepsilon,i}}(S) = (1-2\varepsilon)^{\sum_{i=1}^n S_i} = (1-2\varepsilon)^{|S|}$$

- It is left as an exercise to prove that the distribution $\mathbb{N}_\varepsilon$ is identical to the distribution $\bigoplus_{i=1}^n \mathbb{N}_{\varepsilon,i}$

# Noisy Version of a Function

- Let $f\colon \{0,1\}^n \to \mathbb{R}$ be any function
- Define the noisy version of $f$ as follows

$$\widetilde{f}(x) = T_\rho(x) := \mathbb{E}\left[f(x+e)\colon e \sim \mathbb{N}_\varepsilon\right],$$

where $\rho = 1 - 2\varepsilon$

- So, we have

$$\widetilde{f}(x) = \sum_{e \in \{0,1\}^n} \mathbb{N}_\varepsilon(e) f(x+e) = N(\mathbb{N}_\varepsilon * f)$$

Equivalently, we have $\widetilde{f} = \mathbb{N}_\varepsilon \oplus f$ (we emphasize that $f$ need not be a probability distribution to use the notation of $\oplus$ of two functions)

- Therefore, we get

$$\text{bias}_{\widetilde{f}}(S) = \text{bias}_{\mathbb{N}_\varepsilon}(S) \cdot \text{bias}_f(S) = \rho^{|S|}\text{bias}_f(S)$$

- That is, we conclude that

$$\widehat{T_\rho(f)}(S) = \rho^{|S|}\widehat{f}(S)$$