Randomness Extractors Seminar **December 29, 2016**

Explicit Construction of 2-source Extractors - Mathematics Background

Leader: KM Chung **Notes: Chi-Ning Chou**

We are going to see some useful mathematics lemmas and inequalities for constructing 2-source extractors.

1 Overview

This note summarize the mathematical techniques in the proof of [CZ15].

2 Some Inequalities

2.1 Useful inequalities for probabilistic argument

When dealing with and-or tree, we often need to deal with probability in the form $(1-\frac{x}{n})^n$. The following inequality provides a good approximation when x is small and n is large.

Claim (**2.6**)**.** *For any $n>1$ and $0\leq x\leq n$, we have*

$$e^{-x}(1-\frac{x^2}{n})\leq(1-\frac{x}{n})^n\leq e^{-x}. \tag{1}$$

Proof. Observe that

$$\ln(1-\frac{x}{n})=-\frac{x}{n}-\frac{(x/n)^2}{2!}-\frac{(x/n)^3}{3!}-\cdots. \tag{2}$$

Thus, $\ln(1-\frac{x}{n})\leq -x/n$ and we have the upper bound $(1-\frac{x}{n})^n\leq e^{-x}$. As to the lower bound, apply Taylor's expansion on $\ln(1-\frac{x^2}{n})$.

$$\ln(1-\frac{x^2}{n})=-\frac{x^2}{n}-\frac{(x^2/n)^2}{2!}-\frac{(x^2/n)^3}{3!}-\cdots. \tag{3}$$

As we have

$$\begin{cases} -\frac{x}{n}\leq-\frac{x}{n} & \text{, the first term} \\ -\frac{x^{2(k-1)}}{(k-1)\cdot n^k}\leq-\frac{x^k}{k\cdot n^k} & \text{, the } k\text{th term} \end{cases}$$

The lower bound is proved. □

Claim (A)**.** *For $0<\delta<\ln 2$ and $0\leq x\leq 1$, we have $e^{\delta x}\leq 1+x$.*

Proof. When $x=0$, $e^{\delta x}=1+x=1$, and when $x=1$, $e^{\delta x}\leq 1+x=2$. By the convexity of $e^{\delta x}$ and the linearity of $1+x$, the inequality holds. □

Claim (B)**.** *For any $x\in\mathbb{R}$, $e^{-x}\leq 1-x$.*

2.2 Useful inequalities for combinatoric argument

Claim (Weierstrass product inequality). *Let $0 \leq a_1, \dots, a_n \leq 1$ be n arbitrary numbers in $[0,1]$. we have*

$$\prod_{i\in[n]}(1-a_i) \geq 1 - \sum_{i\in[n]} a_i. \tag{4}$$

Proof. This can be simply proved by induction. □

Claim (inclusion-exclusion principle, union bound/Bonferroni inequality). *Let $A_1, \dots, A_n$ be n events in universe Ω. We have*

$$\mathbb{P}[\cup_{i\in[n]}A_i] = \sum_{c\in[n]} \sum_{1\leq i_1<i_2<\cdots<i_c\leq n} \mathbb{P}[\cap_{g\in[c]}A_{i_g}]. \tag{5}$$

Specifically, for any $a < n/2$.

$$\sum_{c\in[2a]} \sum_{1\leq i_1<i_2<\cdots<i_c\leq n} \mathbb{P}[\cap_{g\in[c]}A_{i_g}] \leq \mathbb{P}[\cup_{i\in[n]}A_i] \leq \sum_{c\in[2a+1]} \sum_{1\leq i_1<i_2<\cdots<i_c\leq n} \mathbb{P}[\cap_{g\in[c]}A_{i_g}]. \tag{6}$$

Proof. This can be simply proved by induction. □

2.3 Janson's inequality

Consider the situation where there are several positively correlated[1] error events. The goal is to bound the probability of none of the error events happening. In such situation, Janson's inequality provides a good approximation when the correlation among error events are small.

Theorem 1 (Jansons inequality). *Let Ω be a finite universal set and let $\mathcal{O}$ be a random subset of Ω constructed by picking each $h \in \Omega$ independently with probability p_h. Let $Q_1, \cdots, Q_\ell$ be arbitrary subsets of Ω, and let $\mathcal{E}_i$ be the event $Q_i \subseteq \mathcal{O}$. Define*

$$\Delta = \sum_{i<j} \mathbb{P}[\mathcal{E}_i \wedge \mathcal{E}_j],\ \ D = \prod_{i=1}^{\ell} \mathbb{P}[\bar{\mathcal{E}}_i]. \tag{7}$$

Assume that $\mathbb{P}[\mathcal{E}_i] \leq \tau$ for all $i \in [\ell]$. Then

$$D \leq \mathbb{P}[\wedge_{i=1}^{\ell}\bar{\mathcal{E}}_i] \leq De^{\frac{\Delta}{1-\tau}}. \tag{8}$$

Proof. In the very beginning of the proof, observe that

$$\mathcal{E}_i = \{\forall h \in Q_i,\ h \in \mathcal{O}\}, \tag{9}$$
$$\bar{\mathcal{E}}_i = \{\exists h \in Q_i,\ h \notin \mathcal{O}\}. \tag{10}$$

Now, let's use chain rule to expand $\mathbb{P}[\wedge_{i=1}^{\ell}\bar{\mathcal{E}}_i]$.

$$\mathbb{P}[\wedge_{i=1}^{\ell}\bar{\mathcal{E}}_i] = \prod_{i=1}^{\ell} \mathbb{P}[\bar{\mathcal{E}}_i | \wedge_{j=1}^{i-1} \bar{\mathcal{E}}_j]. \tag{11}$$

[1] If one event happens, the probability that the other will happen do no decrease.

First, notice that the event $\wedge_{j=1}^{i-1}\bar{\mathcal{E}}_j$ has a positive correlation on $\bar{\mathcal{E}}_i$ since the h missing in Q_j might also lie in Q_i which will increase the probability of $\bar{\mathcal{E}}_i$ to happen. Concretely,

$$\mathbb{P}[\wedge_{i=1}^{\ell}\bar{\mathcal{E}}_i] = \prod_{i=1}^{\ell}\mathbb{P}[\bar{\mathcal{E}}_i| \wedge_{j=1}^{i-1}\bar{\mathcal{E}}_j] \geq \prod_{i=1}^{\ell}\mathbb{P}[\bar{\mathcal{E}}_i]. \tag{12}$$

Thus, we provide a simple lower bound. Next, as to the upper bound, for any $i \in [\ell]$, divide $[i-1]$ into two parts according to if $\mathcal{E}_i$ is correlated to $\mathcal{E}_j$.

$$B_i := \{j \in [i-1] : \ Q_i \cap Q_j \neq \emptyset\}, \tag{13}$$
$$C_i := \{k \in [i-1] : \ Q_i \cap Q_k = \emptyset\}. \tag{14}$$

Consider lower bounding $\mathbb{P}[\mathcal{E}_i| \ \wedge_{j\in B_i}\bar{\mathcal{E}}_j \wedge_{k\in C_i}\bar{\mathcal{E}}_k]$.

$$\begin{aligned}
\mathbb{P}[\mathcal{E}_i| \ \wedge_{j\in B_i}\bar{\mathcal{E}}_j \wedge_{k\in C_i}\bar{\mathcal{E}}_k] &= \frac{\mathbb{P}[\mathcal{E}_i \wedge_{j\in B_i}\bar{\mathcal{E}}_j \wedge_{k\in C_i}\bar{\mathcal{E}}_k]}{\mathbb{P}[\wedge_{j\in B_i}\bar{\mathcal{E}}_j \wedge_{k\in C_i}\bar{\mathcal{E}}_k]} && (15)\\
&= \frac{\mathbb{P}[\mathcal{E}_i]}{\mathbb{P}[\mathcal{E}_i]}\cdot\frac{\mathbb{P}[\mathcal{E}_i \wedge_{j\in B_i}\bar{\mathcal{E}}_j \wedge_{k\in C_i}\bar{\mathcal{E}}_k]}{\mathbb{P}[\wedge_{j\in B_i}\bar{\mathcal{E}}_j \wedge_{k\in C_i}\bar{\mathcal{E}}_k]} && (16)\\
(\because \mathbb{P}[\mathcal{E}_i] = \mathbb{P}[\mathcal{E}_i| \wedge_{k\in C_i}\bar{\mathcal{E}}_k]) &= \mathbb{P}[\mathcal{E}_i]\cdot\frac{\mathbb{P}[\wedge_{k\in C_i}\bar{\mathcal{E}}_k]}{\mathbb{P}[\mathcal{E}_i \wedge_{k\in C_i}\bar{\mathcal{E}}_k]}\cdot\frac{\mathbb{P}[\mathcal{E}_i \wedge_{j\in B_i}\bar{\mathcal{E}}_j \wedge_{k\in C_i}\bar{\mathcal{E}}_k]}{\mathbb{P}[\wedge_{j\in B_i}\bar{\mathcal{E}}_j \wedge_{k\in C_i}\bar{\mathcal{E}}_k]} && (17)\\
&= \mathbb{P}[\mathcal{E}_i]\cdot\mathbb{P}[\wedge_{j\in B_i}\bar{\mathcal{E}}_j|\mathcal{E}_i \wedge_{k\in C_i}\bar{\mathcal{E}}_k]\cdot\frac{\mathbb{P}[\wedge_{k\in C_i}\bar{\mathcal{E}}_k]}{\mathbb{P}[\wedge_{j\in B_i}\bar{\mathcal{E}}_j \wedge_{k\in C_i}\bar{\mathcal{E}}_k]} && (18)\\
&\geq \mathbb{P}[\mathcal{E}_i]\cdot\mathbb{P}[\wedge_{j\in B_i}\bar{\mathcal{E}}_j|\mathcal{E}_i \wedge_{k\in C_i}\bar{\mathcal{E}}_k] && (19)\\
&= \mathbb{P}[\mathcal{E}_i]\cdot\big(1-\mathbb{P}[\vee_{j\in B_i}\mathcal{E}_j|\mathcal{E}_i \wedge_{k\in C_i}\bar{\mathcal{E}}_k]\big) && (20)\\
(\because \text{union bound}) &\geq \mathbb{P}[\mathcal{E}_i]\cdot\big(1-\sum_{j\in B_i}\mathbb{P}[\mathcal{E}_j|\mathcal{E}_i \wedge_{k\in C_i}\bar{\mathcal{E}}_k]\big) && (21)\\
(\because \mathbb{P}[\mathcal{E}_i] = \mathbb{P}[\mathcal{E}_i| \wedge_{k\in C_i}\bar{\mathcal{E}}_k]) &= \mathbb{P}[\mathcal{E}_i] - \sum_{j\in B_i}\frac{\mathbb{P}[\mathcal{E}_i \wedge_{k\in C_i}\bar{\mathcal{E}}_k]}{\mathbb{P}[\wedge_{k\in C_i}\bar{\mathcal{E}}_k]}\cdot\frac{\mathbb{P}[\mathcal{E}_j \wedge \mathcal{E}_i \wedge_{k\in C_i}\bar{\mathcal{E}}_k]}{\mathbb{P}[\mathcal{E}_i \wedge_{k\in C_i}\bar{\mathcal{E}}_k]} && (22)\\
&= \mathbb{P}[\mathcal{E}_i] - \sum_{j\in B_i}\mathbb{P}[\mathcal{E}_i \wedge \mathcal{E}_j| \wedge_{k\in C_i}\bar{\mathcal{E}}_k] && (23)\\
(\because \wedge_{k\in C_i}\bar{\mathcal{E}}_k \text{ has negative correlation}) &\geq \mathbb{P}[\mathcal{E}_i] - \sum_{j\in B_i}\mathbb{P}[\mathcal{E}_i \wedge \mathcal{E}_j]. && (24)
\end{aligned}$$

Namely,

$$\begin{aligned}
\mathbb{P}[\bar{\mathcal{E}}_i| \ \wedge_{j\in B_i}\bar{\mathcal{E}}_j \wedge_{k\in C_i}\bar{\mathcal{E}}_k] &\leq 1-\mathbb{P}[\mathcal{E}_i] + \sum_{j\in B_i}\mathbb{P}[\mathcal{E}_i \wedge \mathcal{E}_j] && (25)\\
&= \mathbb{P}[\bar{\mathcal{E}}_i] + \sum_{j\in B_i}\mathbb{P}[\mathcal{E}_i \wedge \mathcal{E}_j] && (26)\\
&= \mathbb{P}[\bar{\mathcal{E}}_i]\big(1+\frac{\sum_{j\in B_i}\mathbb{P}[\mathcal{E}_i \wedge \mathcal{E}_j]}{\mathbb{P}[\bar{\mathcal{E}}_i]}\big) && (27)\\
&\leq \mathbb{P}[\bar{\mathcal{E}}_i]\cdot e^{\frac{\sum_{j\in B_i}\mathbb{P}[\mathcal{E}_i \wedge \mathcal{E}_j]}{1-\tau}}. && (28)
\end{aligned}$$

As $\Delta = \sum_{i\in[\ell]}\sum_{j\in B_i}\mathbb{P}[\mathcal{E}_i \wedge \mathcal{E}_j]$, we have

$$\mathbb{P}[\wedge_{i=1}^{\ell}\bar{\mathcal{E}}_i] \leq \Big(\prod_{i=1}^{\ell}\mathbb{P}[\bar{\mathcal{E}}_i]\Big)\cdot e^{\frac{\Delta}{1-\tau}} \tag{29}$$

□

3 Some technical details

In this section, I encapsulate several technical details in [CZ15] to independent problems.

3.1 From layer to layer

In an and-or tree, we often encounter probability in the form $q = (1-p)^N$ where p is small and N is large. The goal is to estimate q. Basically, this is what Claim 5.10 and Claim 5.15 in [CZ15] are doing.

Claim. *Let $q = (1-p)^N$ where $|p - \frac{a}{N}| \leq \frac{a}{N}\cdot N^{-\epsilon}$ for some constants $a, \epsilon > 0$. Then, $|p - e^{-a}| \leq e^{-a}\cdot N^{-\epsilon'}$ for some $\epsilon' > 0$.*

Proof. First, consider the upper bound.

$$\begin{aligned} p &= (1-q)^N && (30)\\ (\because \text{statement in the claim}) &\leq \Big(1 - \frac{a}{N}(1-N^{-\epsilon})\Big)^N && (31)\\ (\because \text{Claim 2.6}) &\leq e^{-a(1-N^{\epsilon})} = e^{-a}\cdot e^{aN^{\epsilon}} && (32)\\ (\because \text{Claim A and take } aN^{-\epsilon} \leq \ln 2N^{-\epsilon_1}) &\leq e^{-a}(1+N^{-\epsilon_1}). && (33) \end{aligned}$$

Next, consider the lower bound.

$$\begin{aligned} p &= (1-q)^N && (34)\\ (\because \text{statement in the claim}) &\geq \Big(1 - \frac{a}{N}(1+N^{-\epsilon})\Big)^N && (35)\\ (\because \text{Claim 2.6}) &\geq e^{-a(1+N^{\epsilon})}(1 - \frac{a^2(1+N^{-\epsilon})^2}{N}) && (36)\\ (\because \text{Claim A}) &\geq e^{-a}(1-aN^{-\epsilon})(1 - \frac{a^2(1+N^{-\epsilon})^2}{N}) && (37)\\ (\because \text{take } \epsilon_2 > 0 \text{ properly}) &\geq e^{-a}(1-N^{-\epsilon_2}). && (38) \end{aligned}$$

Finally, take $\epsilon = \max\{\epsilon_1, \epsilon_2\}$ then the inequality holds. □

3.2 Estimating slightly positively correlated events

Now, it's time to prove Claim 5.20 by Janson's inequality. Recall that E_i denotes the event when $f^i_{TExt}(y) = 0$ for all $i \in [R]$ and p_3 is the probability of E_i to happen[2].

[2] By symmetry, the probability for every E_i is the same.

Claim (5.20). *There exists constant $\beta_1, \beta_2 > 0$ such that for any $c \leq s^{\beta_1}$ and arbitrary $1 \leq i_1 < \cdots < i_c \leq R$, the following holds:*

$$p_3^c \leq \mathbb{P}[\wedge_{g\in[c]} E_{i_g}] \leq p_3^c(1 + \frac{1}{M^{\beta_2}}). \tag{39}$$

Furthermore,

$$\binom{R}{c} p_3^c \leq S_c \leq \binom{R}{c} p_3^c(1 + \frac{1}{M^{\beta_2}}). \tag{40}$$

Proof. Before we formally manipulate with the inequality, let's first map the elements in Claim 5.20 to Janson's inequality in Table 1.

	Janson's inequality	Claim 5.20	
Universe	Ω	$[s]$	Bottom layer
Picked	$\mathcal{O}$	$\{z : \ y_z = 1, \ z \in [s]\}$	Bit set to 1
Probability	p_h	$1 - p_1$	$Ber(1 - p_1)$
Subset	Q_i	P_j^i	Block
Event	$\mathcal{E}_i$	$\mathcal{E}_{i,j}$	$f_{TExt}^{i,j}(y) = 1$
No error	$\wedge_i \bar{\mathcal{E}_i}$	$\wedge_{i\in[c],j\in[M]} \bar{\mathcal{E}_{i,j}}$	$f_{TExt}^i(y) = 0, \ \forall i \in [c]$

Table 1: Mapping between Janson's inequality and Claim 5.20.

Note that $\wedge_{i\in[c]} \mathcal{E}_i = \wedge_{i\in[c],j\in[M]} \bar{\mathcal{E}_{i,j}}$. Thus, what we need to do now is simply estimating D and Δ. First, D is trivial.

$$D = \prod_{i\in[c],j\in[M]} \mathbb{P}[\bar{\mathcal{E}_{i,j}}] = \Big((1 - p_2)^M\Big)^c = p_3^c. \tag{41}$$

To bound Δ, consider $i, i' \in [R]$ where $i \neq i'$ and arbitrary $j, j' \in [M]$. By Lemma 5.19, since the Trevisan extractor guarantees that $|P_j^i \cap P_{j'}^{i'}| \leq 0.9B$, we have $|P_j^i \cup P_{j'}^{i'}| \geq 1.1B$. Thus,

$$\mathbb{P}[\mathcal{E}_{i,j} \wedge \mathcal{E}_{i',j'}] = (1 - p_1)^{|P_j^i \cup P_{j'}^{i'}|} \geq (1 - p_1)^{1.1B} = p_2^{1.1}. \tag{42}$$

The last equality is because we let $p_2 = (1 - p_1)^B$. Furthermore, by the choice of parameter in the Section 3.1 of the note for resilient function, we can prove that $\Delta \leq M^{-\beta'}$ for some $\beta' > 0$. As τ can be simply picked as 0.5, we have $e^{\frac{D}{1-\tau}} \leq (1 + \frac{1}{M^{\beta_2}})$ for some $\beta_2 > 0$. Finally, plug D and δ into Janson's inequality, we have the desired inequality. $\square$

3.3 Wrap-up the proof of Lemma 5.5

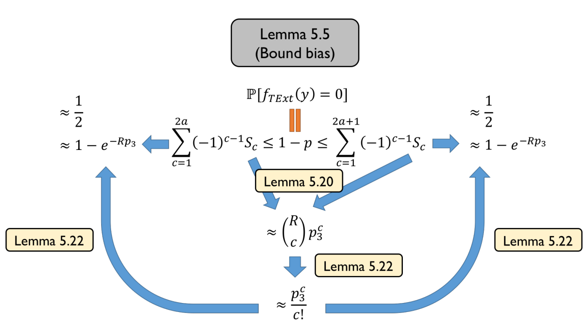

Figure 1: Proof structure of Lemma 5.5.

We sketch the proof structure of Lemma 5.5 in Figure 1. The only ingredient left is the following lemma.

Lemma (5.22)**.** *Take* $a = \lfloor s^{\beta_3} \rfloor$*, we have*

1. *For all* $c \in [a]$, $|S_c - \frac{(Rp_3)^c}{c!}| \leq \frac{1}{M^{\beta_2/2}}$*, and*
2. $|e^{-Rp_3} - \sum_{c\in[a]}(-1)^{c-1}S_c| \leq \frac{1}{M^{\beta_2}}$.

Proof. 1. The upper bound is easy.

$$S_c \leq \binom{R}{c} p_3^c(1+\frac{1}{M^{\beta_2}}) \leq \frac{R^c}{c!}p_3^c(1+\frac{1}{M^{\beta_2}}) \tag{43}$$

An the lower bound,

$$\begin{aligned} S_c \geq \binom{R}{c} p_3^c(1-\frac{1}{M^{\beta_2}}) &= \frac{R\cdot(R-1)\cdots(R-c+1)}{c!}p_3^c(1+\frac{1}{M^{\beta_2}}) && (44)\\ &= \frac{R\cdots(R-c+1)}{R\cdots R}\frac{(Rp_3)^c}{c!}(1+\frac{1}{M^{\beta_2}}) && (45)\\ (\because \text{Werierstrass product inequality}) &\geq (1-\frac{c^2}{R})\frac{(Rp_3)^c}{c!}(1+\frac{1}{M^{\beta_2}}) && (46)\\ (\because c \leq s^{\beta_3}) &\geq (1-\frac{1}{M^{\beta_2}})\frac{(Rp_3)^c}{c!}. && (47)\end{aligned}$$

2. Recall that the Taylor's expansion of $1-e^{-Rp_3}$ is

$$1-e^{-Rp_3} = Rp_3 - \frac{(Rp_3)^2}{2!} + \frac{(Rp_3)^3}{3!} - \cdots = \sum_{c=1}^{\infty}(-1)^{c-1}\frac{(Rp_3)^c}{c!}. \tag{48}$$

Combine the know results and some tedious calculation, we can prove the statement.

□

References

[CZ15] Eshan Chattopadhyay and David Zuckerman. Explicit two-source extractors and resilient functions. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 22, page 119, 2015.