Extração de aleatoriedade a partir de fontes defeituosas

Domingos Dellamonica Junior

Dissertação apresentada
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para
obtenção do título de Mestre
em
Ciências

Área de Concentração: Ciência da Computação
Orientador: Yoshiharu Kohayakawa

Durante a elaboração deste trabalho o autor recebeu auxílio financeiro da FAPESP

São Paulo, março de 2007

Extração de aleatoriedade a partir de fontes defeituosas

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Domingos Dellamonica Junior e aprovada pela Comissão Julgadora.

São Paulo, 27 de março de 2007

Banca examinadora:

Membros titulares
Prof. Dr. Yoshiharu Kohayakawa (orientador) – IME/USP
Prof. Dr. Marcos Kiwi – Universidad de Chile
Prof. Dr. Arnaldo Mandel – IME/USP

Membros suplentes
Prof. Dr. Jair Donadelli Junior – UFPR
Profa. Dra. Cristina Gomes Fernandes – IME/USP

Agradecimentos

Agradeço a meus pais, irmã, amigos e namorada, pela paciência e apoio tão necessários em qualquer grande projeto. Agradeço igualmente a meu orientador, não só por apresentar-me esta área de pesquisa, mas também por muitas boas idéias compartilhadas e sua estimada amizade.

É a vocês que dedico este trabalho,

Domingos Dellamonica Jr.

Resumo

Recentemente, Barak et al. (2004) exibiram construções de *extratores* e *dispersores* determinísticos (funções computáveis em tempo polinomial) com parâmetros melhores do que era anteriormente possível. Introduzimos os conceitos envolvidos em tal trabalho e mencionamos suas aplicações; em particular, mostramos como é possível obter cotas muito melhores para o problema *Ramsey bipartido* (um problema bem difícil) utilizando as construções descritas no artigo.

Também apresentamos resultados originais para melhorar tais construções. Tais idéias são inspiradas no trabalho de Anup Rao (2005) e utilizam o recente êxito de Jean Bourgain (2005) em obter extratores que quebram a "barreira 1/2".

Abstract

Recently, Barak et al. (2004) constructed explicit deterministic *extractors* and *dispersers* (these are polynomial-time computable functions) with much better parameters than what was known before. We introduce the concepts involved in such a construction and mention some of its applications; in particular, we describe how it is possible to obtain much better bounds for the *bipartite Ramsey* problem (a very hard problem) using the machinery developed in that paper.

We also present some original results that improve on these constructions. They are inspired by the work of Anup Rao (2005) and uses the recent breakthrough of Jean Bourgain (2005) in obtaining 2-source extractors that break the "1/2-barrier".

Sumário

Capítulo 1

Introdução

> "Any one who considers arithmetical methods of producing random digits is, of course, in a state of sin."
>
> *John von Neumann*[1]

A aplicação de aleatoriedade em computação, incluindo o uso de algoritmos aleatórios, estruturas aleatórias, protocolos de comunicação, entre tantos, foi uma revolução na ciência da computação. Em muitos casos, algoritmos têm alta probabilidade de sucesso e/ou baixo tempo esperado de execução.

Em geral, a análise de tais algoritmos depende fortemente do uso de uma distribuição uniforme de bits. Na prática, no entanto, os meios de obtenção de bits aleatórios não possuem garantia alguma de tal uniformidade. Meios físicos de obtenção desses bits (como diodos que produzem ruído quântico) são lentos e também não possuem garantia de uniformidade (apesar de que isso é comumente omitido comercialmente). Note que é impossível avaliar, utilizando testes estatísticos, se uma dada fonte consegue gerar, digamos, 200 bits aleatórios. Isso porque o espaço de possíveis seqüências tem tamanho $2^{200} > 10^{60}$ e qualquer amostragem (prática) terá consideravelmente menos que 10^{60} elementos.

Talvez o primeiro a propor extração de aleatoriedade a partir de fontes viesadas foi von Neumann [vN51], que considerou extração a partir de bits vindos de lançamentos de moedas desbalanceadas porém independentes. A hipótese de independência, porém, é forte demais e modelos mais fracos foram sugeridos posteriormente por Blum [Blu86], que considerou fontes cujos bits eram gerados por uma cadeia de Markov finita. Santha e Vazirani [SV86], Chor e Goldreich [CG85] sucessivamente refinaram definições e propuseram soluções de extração de aleatoriamente. Finalmente, Zuckerman [Zuc90] propôs um modelo baseado em *min-entropia* que generaliza todos os anteriores e é, de certa forma, um requisito mínimo para que a extração seja possível.

A idéia central do que discutiremos neste texto é a de extração de aleatoriedade a partir de uma ou mais *fontes aleatórias defeituosas*. Queremos obter uma *saída* estatisticamente *próxima* a distribuição uniforme.

Para cada fonte, associamos uma distribuição de probabilidade. Diremos que uma distribuição $\mathcal{X}$ sobre palavras binárias de comprimento n (i.e., distribuição sobre $\{0,1\}^n$) tem *min-entropia*

[1] "Qualquer um que considere métodos aritméticos para a produção de dígitos aleatórios está claramente cometendo um pecado." (tradução livre).

pelo menos k se toda palavra tem probabilidade no máximo 2^{-k} em $\mathcal{X}$. Observe que, quanto maior a min-entropia de uma distribuição, mais próxima da distribuição uniforme ela se torna. Em particular, se a min-entropia de $\mathcal{X}$ for n, temos que $\mathcal{X}$ é uniforme em $\{0,1\}^n$.

1.1 Extratores e Dispersores

A condição de uma distribuição ter alta min-entropia, apesar de parecer forte, não determina com precisão o comportamento da mesma. Por exemplo, $\mathcal{X}$ poderia ser uma distribuição uniforme sobre um subconjunto de $\{0,1\}^n$ de tamanho 2^k (os elementos fora do subconjunto teriam probabilidade nula). Neste caso, como poderíamos saber qual subconjunto é o suporte de $\mathcal{X}$? Lembre-se que dispomos de um poder computacional limitado (com isso queremos dizer que apenas algoritmos polinomiais são de interesse).

Se dispomos de apenas uma fonte desse tipo e um algoritmo determinístico, não conseguimos extrair sequer um bit aleatório, como pode-se observar a partir do seguinte exemplo. Seja $f\colon \{0,1\}^n \to \{0,1\}$ uma função qualquer. Existe $b \in \{0,1\}$ tal que $f^{-1}(b)$ possui pelo menos 2^{n-1} elementos. Defina X como uma fonte que é uniforme em $f^{-1}(b)$. Por definição, $f(X) = b$ com probabilidade 1 e X tem min-entropia $\geq n-1$.

Precisamos então de algo a mais para podermos extrair aleatoriedade. Historicamente, a primeira abordagem na literatura consistia de funções que recebem uma *semente* consistindo de bits genuinamente aleatórios e uma amostra vinda de uma fonte defeituosa com alguma min-entropia (ou seja, um algoritmo probabilístico).

Recentemente, diversos avanços permitiram obter aleatoriedade a partir de fontes defeituosas independentes de baixa min-entropia sem a necessidade de sementes genuinamente aleatórias. Por ser essa linha de pesquisa mais recente e menos conhecida, o enfoque desta dissertação será sobre extratores e dispersores que não utilizam sementes. Extratores com semente são apresentados em mais detalhes no capítulo 9.

Uma função $f\colon \{0,1\}^{n\times l} \to \{0,1\}^m$ é um *extrator* de l fontes com requerimento de min-entropia k se, para quaisquer l fontes independentes $\mathcal{X}_1,\ldots,\mathcal{X}_l$, cada uma com min-entropia $\geq k$, a distribuição $f(\mathcal{X}_1,\ldots,\mathcal{X}_l)$ é próxima a distribuição uniforme em $\{0,1\}^m$. *Dispersores* são outros objetos de interesse particular, mais fracos que extratores, e no entanto úteis para desaleatorização de certos algoritmos probabilísticos. Se, para algum $0 \leq \varepsilon < 1$, tivermos $|\mathrm{supp}(f(\mathcal{X}_1,\ldots,\mathcal{X}_l))| \geq (1-\varepsilon)\,2^m$ então f é um dispersor de l fontes com requerimento de min-entropia k com parâmetro de erro ε.

Apesar do forte enfoque algorítmico dado nos parágrafos anteriores, é interessante frisar que a obtenção de extratores possui diversas aplicações além da desaleatorização de algoritmos, como por exemplo, a construção de poderosos *grafos expansores* [WZ99, HLW06], que por sua vez podem ser empregados na codificação e decodificação [SS96] de *códigos corretores de erro*[2]. Há também aplicações em criptografia [KZ03, GRS04, KRVZ06] , onde deseja-se, por exemplo, que mesmo que uma parte dos dados secretos usados na criptografia seja descoberta por um inimigo, ele ainda não consiga, eficientemente, descriptografar o conteúdo. Alguns desses extratores são específicos para fontes com estrutura mais bem definida do que as tratadas aqui (que só precisam possuir uma quantidade de min-entropia mínima).

A pesquisa em extratores e outras objetos relacionados tem sido tão intensa que foi necessário escolher trabalhos representativos e de impacto ([BIW04, BKS+05]) para discuti-los em profun-

[2] A relação entre extratores e códigos corretores de erros é uma de mão-dupla [TSZ04]. (Veja o extrator de Trevisan na seção 9.3.)

didade deixando em um nível superficial a apresentação de importantes trabalhos recentes de Raz [Raz04] e Anup Rao [Rao05], que melhoram diversos dos resultados aqui contidos.

1.2 Contribuições deste trabalho

Nesta dissertação pudemos reunir uma parte da extensa literatura recente sobre extração de aleatoriedade. A área é tão dinâmica que é difícil se manter atualizado, sendo que parte substancial do que se encontra aqui já foi melhorado (os dispersores do capítulo 8 são um exemplo [BRSW06]). No entanto, acreditamos que o foco não ficou restrito a métodos isolados: as técnicas utilizadas podem ter sido refinadas em artigos mais recentes, porém as idéias fundamentais—como o uso de teoria aditiva dos números na obtenção de condensadores, a relação entre códigos corretores de erros e extratores, o mecanismo *desafio-resposta*—permanecem no cerne da área e são tais idéias que esperamos ter organizado de forma sistemática aqui.

Além de uma ampla visão de tais idéias fundamentais e técnicas utilizadas, contribuímos com alguns resultados originais (veja o capítulo 10) que melhoram as construções discutidas [BKS+05]. Tais melhoras foram possíveis graças a avanços posteriores ao artigo [BKS+05], nos trabalhos de Raz [Raz04], Bourgain [Bou05] e Rao [Rao05]. Soubemos, em comunicação com Prof. Avi Wigderson, que algumas dessas melhoras (com relação ao tamanho da saída e redução de erro nos extratores e dispersores) aparecem em um artigo em preparação devido a Rao.

Capítulo 2

Definições e Resultados Preliminares

2.1 Notação

A seguinte tabela tem a finalidade de ser uma referência rápida para toda a notação usada nesta dissertação.

Tabela 2.1: Notação

Notação	Significado
$\mathbf{P}[\cdot]$	probabilidade de um evento.
$\mathbf{E}[\cdot]$	esperança de uma variável aleatória.
U_m	distribuição uniforme sobre o conjunto $\{0,1\}^m$.
$X \sim \mathcal{D}$	X é uma variável aleatória de distribuição $\mathcal{D}$; $X \sim Y$ indica que X e Y têm a mesma distribuição
$\text{supp}(X)$	suporte da distribuição X, ou seja $\{x \mid \mathbf{P}[X = x] > 0\}$
$a \in_R A$	o elemento a é escolhido de A aleatoriamente (se A é uma distribuição, então a escolha segue a distribuição; se A é um conjunto então a escolha é uniforme sobre A).
$a_1, a_2, \ldots, a_m \leftarrow_R \mathcal{D}$	os elementos $a_1, \ldots, a_m$ são escolhidos de forma independente de acordo com a distribuição $\mathcal{D}$.
$x_{i\ldots j}$	Dado $x = x_1 x_2 \ldots x_t$, e $1 \leq i \leq j \leq t$, $x_{i\ldots j} = x_i x_{i+1} \ldots x_j$.
$a \equiv b + c$	define o lado esquerdo como sendo igual ao lado direito.
$\mathbf{e}_j$	vetor cujas coordenadas são todas nulas exceto pela j-ésima coordenada, que vale 1.

Tabela 2.1: Notação

Notação	Significado
$\Gamma_G(\cdot)$	vizinhança de um vértice ou conjunto de vértices no grafo G (i.e., dado $v \in V(G)$, temos $\Gamma_G(v) = \{u \in V \mid \{u,v\} \in E(G)\}$ e, para todo $S \subseteq V(G)$, temos $\Gamma_G(S) = \bigcup_{v \in S} \Gamma_G(v)$).
$[n]$	conjunto $\{1, 2, \ldots, n\}$ (convencionamos $[0] = \emptyset$).
$A + B$	conjunto $\{a + b \mid (a,b) \in A \times B\}$.
$A \cdot B$	conjunto $\{a \cdot b \mid (a,b) \in A \times B\}$.
kA	conjunto $\underbrace{A + \cdots + A}_{k \text{ vezes}}$, ou seja $\{a_1 + \cdots + a_k \mid a_1, \ldots, a_k \in A\}$.
A^k	conjunto $\underbrace{A \cdots A}_{k \text{ vezes}}$.
$\mathrm{DTIME}\big(g(n)\big)$	classe das linguagens (problemas de decisão) que podem ser decididas deterministicamente em tempo $g(n)$.

2.2 Preliminares

Começaremos com algumas definições e resultados preliminares utilizados na versão preliminar de revista do artigo [BKS+05].

Definição 2.2.1 (Variáveis aleatórias, fontes, min-entropia e razão de entropia)**.** *Trataremos de* variáveis aleatórias (v.a.) *sobre* $\{0,1\}^n$*. Chamaremos tais variáveis de* fontes de n-bits*. A* min-entropia *de uma variável X, denotada por $H^\infty(X)$ é definida como*

$$\log_2 \frac{1}{\max_a\{\mathbf{P}[X = a]\}}.$$

A razão de entropia *de uma fonte de n-bits X, denotada por $r(X)$ é dada por $H^\infty(X)/n$. Em alguns casos, será conveniente tratar a distribuição induzida por X como um vetor $\mathbf{x} \in \mathbb{R}^{\{0,1\}^n}$ tal que $x_a = \mathbf{P}[X = a]$ para todo $a \in \{0,1\}^n$.*

*Chamaremos X de δ-*fonte *se $r(X) \geq \delta$.*

Definição 2.2.2 (Fontes planas, combinações convexas)**.** *O* suporte *de uma variável aleatória X, denotado por* $\mathrm{supp}(X)$*, é o conjunto $\{a \mid \mathbf{P}[X = a] > 0\}$. Se $\mathbf{P}[X = a] = \mathbf{P}[X = a']$ para todo $a, a' \in \mathrm{supp}(X)$ dizemos que X é uma* fonte plana*. Observe que se X é uma fonte plana de min-entropia k então X é uma variável aleatória uniforme sobre um conjunto de tamanho 2^k.*

Sejam $X_1, \ldots, X_j$ variáveis aleatórias e $\alpha_1, \ldots, \alpha_j$ números não-negativos somando 1. Dizemos que uma fonte de n-bits Y é combinação convexa de $\{X_i\}_{i=1}^j$ com coeficientes $\{\alpha_i\}_{i=1}^j$ *se $\mathbf{y} = \sum_{i=1}^j \alpha_i \mathbf{x}_i$.*

Proposição 2.2.3. *Toda fonte com min-entropia $\geq k$ (com 2^k inteiro) é combinação convexa de fontes planas de min-entropia $\geq k$.*

Observe que se 2^k não for inteiro, podemos reduzir levemente o valor de k de forma a torná-lo inteiro.

Demonstração. Considere os vetores correspondentes a todas as distribuições com min-entropia $\geq k$. O conjunto de todos esses vetores é um politopo $P \subseteq \mathbb{R}^{\{0,1\}^n}$. De fato, as restrições do politopo são todas lineares: coordenadas em $[0, 2^{-k}]$; soma das coordenadas igual a 1. Resta-nos mostrar que os vértices do politopo correspondem a fontes planas de min-entropia k.

Suponha que algum $\mathbf{x}$, correspondente a uma fonte plana de min-entropia k, seja combinação convexa não-trivial de dois pontos distintos $\mathbf{y}, \mathbf{z} \in P$. Suponha, sem perda de generalidade, que a é tal que $y_a < z_a$. Observe que, para algum $0 < \alpha < 1$, temos $x_a = \alpha y_a + (1-\alpha) z_a > 0$. Como $x_a > 0$, temos $a \in \text{supp}(X)$ e, como X é fonte plana, temos $x_a = 2^{-k}$. Por outro lado, $\alpha y_a + (1-\alpha) z_a < 2^{-k}$, um absurdo.

Agora mostraremos que se $\mathbf{x}$ é vértice então ele corresponde a uma fonte plana de min-entropia k. Suponha que este não é o caso. Logo, existem dois elementos a, b distintos tais que $0 < x_a, x_b < 2^{-k}$. Porém, isso implica que para algum $\varepsilon > 0$, vale que $\mathbf{y} \equiv \mathbf{x} + \varepsilon(\mathbf{e}_a - \mathbf{e}_b)$ e $\mathbf{z} \equiv \mathbf{x} - \varepsilon(\mathbf{e}_a - \mathbf{e}_b)$ estão no politopo, contradizendo o fato de $\mathbf{x} = (\mathbf{y} + \mathbf{z})/2$ ser vértice. □

Definição 2.2.4 (Distância Estatística)**.** *Sejam X e Y variáveis aleatórias com valores em um conjunto discreto Λ. A* distância estatística $\text{dist}(X, Y)$ *é definida como*

$$\frac{1}{2} \sum_{x \in \Lambda} \left| \mathbf{P}[X = x] - \mathbf{P}[Y = x] \right| = \max_{S \subseteq \Lambda} \left| \mathbf{P}[X \in S] - \mathbf{P}[Y \in S] \right|.^{*} \tag{2.1}$$

Observe que $\text{dist}(X, Y) \in [0, 1]$. *Diremos que X é ε-próximo de Y (em muitas vezes, "Y" é trocado por "uniforme", ou seja, a distribuição uniforme com mesmo domínio) se* $\text{dist}(X, Y) \leq \varepsilon$.

2.2.1 Por que usar min-entropia?

É interessante frisar que a definição de min-entropia não é arbitrária ou desenhada para que soluções sejam criadas. De certa forma, a exigência de uma cota inferior para a min-entropia de X é um requisito mínimo. De fato, suponha que tenhamos uma função $E\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$, com $m > d$, tal que para qualquer $A \subseteq \{0,1\}^m$ e alguma constante pequena $\varepsilon > 0$, vale que

$$\left| \mathbf{P}_{x \in_R X, y \in_R U_d}[E(x, y) \in A] - |A| \, 2^{-m} \right| \leq \varepsilon. \tag{2.2}$$

A equação acima é equivalente a dizer que $\text{dist}\Big(E(X, U_d), U_m\Big) \leq \varepsilon$. Pelas hipóteses, E é uma função que extrai aleatoriedade da fonte X usando bits genuinamente aleatórios. Nesta dissertação lidaremos principalmente com o caso (mais difícil) onde não temos uma fonte genuinamente aleatória mas apenas duas ou mais fontes fracas como X.

Note que é natural desejarmos $m > d$ bits de saída pois, do contrário, o "investimento" de aleatoriedade em E tem "retorno" negativo. Tome $A = \Big\{ z \mid \mathbf{P}[E(X, U_d) = z] \geq 2 \times 2^{-m} \Big\}$. Segue da definição de A que $\mathbf{P}[E(X, U_d) \in A] \geq |A| \, 2^{1-m}$. Por outro lado, pela equação (2.2)

*Para ver que as duas definições acima são equivalentes, observe que o máximo a direita é atingido pelo conjunto $S = \Big\{ x \in \Lambda \mid \mathbf{P}[X = x] > \mathbf{P}[Y = x] \Big\}$.

obtemos $|A|\,2^{1-m} - |A|\,2^{-m} \leq \varepsilon$, donde segue $|A| \leq \varepsilon\, 2^m$ e $\mathbf{P}[E(X, U_d) \in A] \leq 2\varepsilon$. Logo, existe $y \in \{0,1\}^d$ tal que $\mathbf{P}[E(X, y) \in A] \leq 2\varepsilon$. Para todo x tal que $z = E(x, y) \notin A$ vale que

$$\begin{aligned}
\mathbf{P}[X = x] &\leq \mathbf{P}[E(X, y) = z] \\
&= \mathbf{P}[E(X, Y) = z \mid Y = y] \\
&= \mathbf{P}[E(X, Y) = z, Y = y] \Big/ \mathbf{P}[Y = y] \\
&\leq \mathbf{P}[E(X, Y) = z] \Big/ \mathbf{P}[Y = y] \\
&\leq 2^{-m+1+d}.
\end{aligned}$$

A massa de probabilidade atribuída aos x tais que $E(x, y) \in A$ é limitada por 2ε, demonstrando que a distribuição X é $O(\varepsilon)$-próxima de uma distribuição com min-entropia $\geq m - d - 1$.

Proposição 2.2.5. *Sejam X, Y v.a. sobre um conjunto Λ e $f\colon \Lambda \to \Lambda'$. Então* $\operatorname{dist}(f(X), f(Y)) \leq \operatorname{dist}(X, Y)$.

Demonstração.

$$\begin{aligned}
\operatorname{dist}(f(X), f(Y)) &= \frac{1}{2} \sum_{x' \in \Lambda'} \Big| \mathbf{P}[f(X) = x'] - \mathbf{P}[f(Y) = x'] \Big| \\
&= \frac{1}{2} \sum_{x' \in \Lambda'} \Big| \mathbf{P}[X \in f^{-1}(x')] - \mathbf{P}[Y \in f^{-1}(x')] \Big| \\
&\leq \frac{1}{2} \sum_{x' \in \Lambda'} \sum_{x \in f^{-1}(x')} \Big| \mathbf{P}[X = x] - \mathbf{P}[Y = x] \Big| \\
&= \operatorname{dist}(X, Y).
\end{aligned}$$

□

Proposição 2.2.6. *Sejam X, Y variáveis aleatórias sobre um conjunto Λ. Seja Z uma v.a. sobre um conjunto discreto Λ' e $f\colon \Lambda \to \Lambda'$. Se* $\operatorname{dist}(X, Y) \leq \varepsilon_1$ *e* $\operatorname{dist}(f(Y), Z) \leq \varepsilon_2$ *então temos* $\operatorname{dist}(f(X), Z) \leq \varepsilon_1 + \varepsilon_2$.

Demonstração.

$$\begin{aligned}
&\frac{1}{2} \sum_{z \in \Lambda'} \Big| \mathbf{P}[f(X) = z] - \mathbf{P}[Z = z] \Big| \\
&= \frac{1}{2} \sum_{z \in \Lambda'} \Big| \mathbf{P}[f(X) = z] - \mathbf{P}[f(Y) = z] + \mathbf{P}[f(Y) = z] - \mathbf{P}[Z = z] \Big| \\
&\leq \frac{1}{2} \sum_{z \in \Lambda'} \Big| \mathbf{P}[f(X) = z] - \mathbf{P}[f(Y) = z] \Big| \\
&+ \frac{1}{2} \sum_{z \in \Lambda'} \Big| \mathbf{P}[f(X) = z] - \mathbf{P}[Z = z] \Big| \\
&\leq \varepsilon_1 + \varepsilon_2.
\end{aligned}$$

□

Proposição 2.2.7. *Seja X uma distribuição sobre um conjunto Λ e suponha que $\sum_i \alpha_i X_i$ seja uma combinação convexa igual a X. Seja $f\colon \Lambda \to \Lambda'$ e Y uma distribuição sobre Λ'. Então, se* $\mathrm{dist}(f(X_i), Y) \leq a$ *para todo i, temos* $\mathrm{dist}(f(X), Y) \leq a$.

Demonstração.

$$\begin{aligned}
\mathrm{dist}(f(X), Y) &= \frac{1}{2}\sum_{x\in\Lambda'}\left|\mathbf{P}[f(X) = x] - \mathbf{P}[Y = x]\right| \\
&= \frac{1}{2}\sum_{x\in\Lambda'}\left|\sum_i \alpha_i\Big(\mathbf{P}[f(X_i) = x] - \mathbf{P}[Y = x]\Big)\right| \\
&\leq \sum_i \alpha_i \frac{1}{2}\sum_{x\in\Lambda'}\left|\mathbf{P}[f(X_i) = x] - \mathbf{P}[Y = x]\right| \\
&\leq a.
\end{aligned}$$

□

A proposição acima ajuda a nos mostrar que em muitos casos, podemos provar que os extratores funcionam para fontes planas e estender o resultado para qualquer fonte com perda negligenciável nos parâmetros.

Corolário 2.2.8. *Seja $f\colon \{0,1\}^{n_1+\cdots+n_k} \to \{0,1\}^m$ e Y uma distribuição sobre $\{0,1\}^m$. Suponha que para toda coleção de fontes* planas *independentes $X_1, \ldots, X_k$ sobre, respectivamente, $\{0,1\}^{n_1}, \ldots, \{0,1\}^{n_k}$, com $H^\infty(X_j) \geq \delta n_j$, temos* $\mathrm{dist}(f(X_1, \ldots, X_k), Y) \leq a$. *Então para quaisquer $(\delta + o(1))$-fontes independentes $X'_1, \ldots, X'_k$ temos* $\mathrm{dist}(f(X'_1, \ldots, X'_k), Y) \leq a$. *(Onde o termo $o(1) \to 0$ quando $\min_j\{n_j\} \to \infty$.)*

Demonstração. Para todo $j = 1, \ldots, k$ defina $\delta_j \geq \delta$ como o menor valor tal que $2^{\delta_j n_j}$ é inteiro. Sejam $X'_1, \ldots, X'_k$ fontes com $H^\infty(X'_j) \geq \delta_j n_j$. (Note que $\delta_j - \delta \to 0$ quando $n_j \to \infty$) Pela proposição 2.2.3, podemos decompor cada fonte X'_j como combinação convexa de fontes planas $X'_j = \sum_{i=1}^{r_j} \alpha_{j,i} X'_{j,i}$ onde $H^\infty(X'_{j,i}) \geq \delta_j n_j$.

Pela hipótese, para todo $(u_1, \ldots, u_k) \in [r_1] \times \cdots \times [r_k]$, temos $\mathrm{dist}(f(X'_{1,u_1}, \ldots, X'_{k,u_k}), Y) \leq a$. Como $X = (X_1, \ldots, X_k)$ é combinação convexa das fontes $\Big\{(X'_{1,u_1}, \ldots, X'_{k,u_k}) \mid (u_1, \ldots, u_k) \in [r_1] \times \cdots \times [r_k]\Big\}$, o corolário segue a partir da proposição 2.2.7. □

Proposição 2.2.9. *Sejam X e Y variáveis aleatórias discretas. Tais v.a. satisfazem* $\mathrm{dist}(X, Y) \leq \varepsilon$ *se e somente se existem variáveis aleatórias X' e Y' cujas distribuições são iguais a de X e Y respectivamente e ambas são definidas sobre* o mesmo *espaço de probabilidade Ω de forma que $\mathbf{P}_\omega[X'(\omega) \neq Y'(\omega)] \leq \varepsilon$.*

Demonstração. Suponha que existam X', Y' satisfazendo o enunciado. Para $A \in \{X', Y'\}$, defina $\Omega_{A,x} \equiv A^{-1}(x) = \{\omega \in \Omega \mid A(\omega) = x\}$. Primeiramente vamos provar que

$$\mathbf{P}_\omega[X'(\omega) \neq Y'(\omega)] = \sum_{x\in\Lambda} \mathbf{P}[\omega \in \Omega_{X',x} \triangle \Omega_{Y',x}]/2. \tag{2.3}$$

Observe que

$$\begin{aligned}\mathbf{P}_\omega[X'(\omega) \neq Y'(\omega)] &= \mathbf{P}\Big[\{\omega \mid X'(\omega) \neq Y'(\omega)\}\Big] \\ &= \sum_{x\in\Lambda} \mathbf{P}\Big[\{\omega \mid X'(\omega) = x, Y'(\omega) \neq x\}\Big] \\ &= \sum_{x\in\Lambda} \mathbf{P}[\Omega_{X',x} \setminus \Omega_{Y',x}].\end{aligned}$$

Por simetria, $\mathbf{P}_\omega[X'(\omega) \neq Y'(\omega)] = \sum_{x\in\Lambda} \mathbf{P}[\Omega_{Y',x} \setminus \Omega_{X',x}]$, donde segue a equação (2.3). Como

$$\begin{aligned}|\mathbf{P}[X = x] - \mathbf{P}[Y = x]| &= |\mathbf{P}[\omega \in \Omega_{X',x}] - \mathbf{P}[\omega \in \Omega_{Y',x}]| \\ &\leq \mathbf{P}[\omega \in \Omega_{X',x} \triangle \Omega_{Y',x}],\end{aligned}$$

segue que $\operatorname{dist}(X, Y) \leq \mathbf{P}_\omega[X'(\omega) \neq Y'(\omega)]$.

Vamos mostrar a existência de um Ω e variáveis X', Y' satisfazendo o enunciado. Tome $\Omega = [0, 1]$ com a medida usual associada. Definiremos indutivamente conjuntos Ω_{X',x_i} e Ω_{Y',x_i}, com $1 \leq i \leq |\Lambda|$, da seguinte forma. Suponha, sem perda de generalidade, que

$$\sum_{j<i} \mathbf{P}[X = x_j] \geq \sum_{j<i} \mathbf{P}[Y = x_j]$$

(inicialmente, nenhum x_j está definido e a inequação acima é trivialmente verdadeira). Tome $x_i \in \Lambda \setminus \{x_1, \ldots, x_{i-1}\}$ tal que $\mathbf{P}[X = x_i] \leq \mathbf{P}[Y = x_i]$ e tome Ω_{X',x_i} um subconjunto de $\Omega \setminus \bigcup_{j<i} \Omega_{X',x_j}$ de medida $\mathbf{P}[X = x_i]$. Seja $q = \mathbf{P}[Y = x_i] - \mathbf{P}[X = x_i]$. Se q for menor ou igual a medida de $S = \bigcup_{j<i} \Omega_{X',x_j} \setminus \bigcup_{j<i} \Omega_{Y',x_j}$, escolha $R \subseteq S$ de medida q e tome $\Omega_{Y',x_i} = \Omega_{X',x_i} \cup R$. Caso contrário, tome $\Omega_{Y',x_i} = \Omega_{X',x_i} \cup S \cup R'$, onde $R' \subseteq \Omega \setminus \bigcup_{j<i} \Omega_{X',x_j}$ faz com que a medida de Ω_{Y',x_i} seja $\mathbf{P}[Y = x_i]$.

Por construção, teremos que $X \sim X'$ e $Y \sim Y'$. Ademais,

$$|\mathbf{P}[\omega \in \Omega_{X',x}] - \mathbf{P}[\omega \in \Omega_{Y',x}]| = \mathbf{P}[\omega \in \Omega_{X',x} \triangle \Omega_{Y',x}],$$

portanto, segue a igualdade $\mathbf{P}_\omega[X'(\omega) \neq Y'(\omega)] = \operatorname{dist}(X, Y)$ para nossas escolhas de Ω, X' e Y'. □

Definição 2.2.10 (Blocos)**.** *Seja X uma fonte de n-bits. Seja $(X_1, \ldots, X_n)$ a representação de X como um vetor de bits. Para qualquer conjunto $S \subseteq [n]$, definimos o* S-bloco *de X, como sendo $X_S = (X_i)_{i\in S}$. Dizemos que X_S é um bloco em X ou um sub-bloco de X.*

Observe que, para qualquer $S \subseteq [n]$, temos $H^\infty(X_S) \geq H^\infty(X) - (n - |S|)$. Se $\operatorname{dist}(X, Y) \leq \varepsilon$ então $\operatorname{dist}(X_S, Y_S) \leq \varepsilon$ para qualquer escolha de $S \subseteq [n]$.

Definição 2.2.11 (Sub-fontes)**.** *Sejam X e X' fontes de n-bits. Dado $0 \leq \alpha \leq 1$, dizemos que X' é uma* sub-fonte de X de densidade α *se $\mathbf{x} = \alpha\mathbf{x}' + (1 - \alpha)\mathbf{z}$ para alguma fonte de n-bits Z. Utilizaremos a notação $X' \subseteq X$ para dizer que X' é sub-fonte de X (para alguma densidade α).*

Proposição 2.2.12. *Se $X_1, X_2, \ldots, X_m \subseteq X$ têm densidade pelo menos $\alpha > 0$ então qualquer combinação convexa $Y = \sum_{i=1}^m \beta_i X_i$ é sub-fonte de X com densidade α.*

Demonstração. É simples observar que se $X_i \subseteq X$ tem densidade pelo menos α, então existe uma fonte Z_i tal que $\mathbf{x} = \alpha\mathbf{x}_i + (1 - \alpha)\mathbf{z}_i$. Segue que

$$\alpha\Big(\sum_{i=1}^m \beta_i \mathbf{x}_i\Big) = \sum_{i=1}^m \beta_i\Big\{\mathbf{x} - (1 - \alpha)\mathbf{z}_i\Big\} = \mathbf{x} - (1 - \alpha)\sum_{i=1}^m \beta_i \mathbf{z}_i,$$

obtendo $\mathbf{x} = \alpha\mathbf{y} + (1 - \alpha)\mathbf{z}$, para $\mathbf{z} = \sum_{i=1}^m \beta_i \mathbf{z}_i$. □

Definição 2.2.13 (Propriedades de variáveis aleatórias)**.** *Seja $\mathcal{P}$ uma coleção de fontes. Dizemos que uma variável X* está a distância ε de $\mathcal{P}$ *ou que X* é ε-próximo de ter a propriedade $\mathcal{P}$ *se existe uma variável $Y \in \mathcal{P}$ tal que* $\mathrm{dist}(X, Y) \leq \varepsilon$*. Denotamos por $\mathcal{P}_\varepsilon$ o conjunto de variáveis ε-próximo de $\mathcal{P}$. Por exemplo, dizemos que X é ε-próximo de ter min-entropia $\geq k$ se existe uma variável Y de min-entropia $\geq k$ tal que* $\mathrm{dist}(X, Y) \leq \varepsilon$.

Definição 2.2.14 (Blocos e fontes $\mathcal{P}$-em-algum-lugar)**.** *Intuitivamente, dizemos que uma fonte X é $\mathcal{P}$-*em-algum-lugar *se algum bloco de X está em $\mathcal{P}$. Mais formalmente, se X é uma fonte de $(n \times l)$-bits, podemos considerá-la como uma concatenação de l blocos consecutivos de comprimento n, digamos $X = (X_1, \ldots, X_l)$. Um* seletor $I = I(X)$ *para X é uma variável aleatória tomando valores em $[l]$.*

Se existir um seletor I para X tal que a fonte X_I tenha a propriedade $\mathcal{P}$, dizemos que X é $\mathcal{P}$-em-algum-lugar.

Proposição 2.2.15. *Seja $\mathcal{P}$ uma propriedade de fontes e $X = (X_1, \ldots, X_l)$ como na definição 2.2.14. Se X é $\mathcal{P}_\varepsilon$-em-algum-lugar então X é ε-próximo de ser $\mathcal{P}$-em-algum-lugar.*

Demonstração. Seja I um seletor para X tal que $X_I \in \mathcal{P}_\varepsilon$, ou seja, existe $Y \in \mathcal{P}$ que é ε-próximo de X_I. Note que, pela proposição 2.2.9, podemos supor que X_I e Y são variáveis definidas no mesmo espaço de probabilidade e que $\mathbf{P}[X_I \neq Y] \leq \varepsilon$.

Defina X' como sendo a seguinte fonte de $(n \times l)$-bits,

$$X' = (X'_1, \ldots, X'_l),$$

onde $X'_i = Y$ se $I = i$ e $X'_i = X_i$, caso contrário. Claramente, $\mathbf{P}[X \neq X'] = \mathbf{P}[X_I \neq Y] \leq \varepsilon$ e $X'_I = Y \in \mathcal{P}$. Portanto, X' é $\mathcal{P}$-em-algum-lugar e, pela proposição 2.2.9, segue que $\mathrm{dist}(X, X') \leq \varepsilon$, concluindo a demonstração. □

Definição 2.2.16 (Probabilidade de Colisão)**.** *Seja $\mathcal{D}$ uma distribuição de probabilidade. Definimos* $\mathrm{cp}(\mathcal{D}) = \mathbf{P}_{x,y \leftarrow_R \mathcal{D}}[x = y]$ *onde x e y são escolhidos de forma independente. Note que se $\mathcal{D}(x)$ é a probabilidade de x em $\mathcal{D}$ então*

$$\mathrm{cp}(\mathcal{D}) = \sum_x \mathcal{D}(x)^2 = \|\mathcal{D}\|_2^2.$$

Proposição 2.2.17. *Seja X uma distribuição de probabilidade que é dada por uma combinação convexa de distribuições $X_1, \ldots, X_m$. Então*

$$\mathrm{cp}(X) \leq \max\Big\{\mathrm{cp}(X_1), \ldots, \mathrm{cp}(X_m)\Big\}.$$

Demonstração. É suficiente provar o caso $m = 2$ e o caso geral segue por indução usando-se o fato que $\alpha_1 X_1 + \cdots + \alpha_m X_m = (1 - \alpha_m)Y + \alpha_m X_m$ onde $Y = \beta_1 X_1 + \cdots + \beta_{m-1} X_{m-1}$ com $\beta_i = \alpha_i/(1 - \alpha_m)$ para todo i. Claramente Y é combinação convexa de $X_1, \ldots, X_{m-1}$.

Seja então $Z = \alpha X + (1-\alpha)Y$ uma combinação convexa de distribuições. Seja $\mathbf{x}$ e $\mathbf{y}$ a representação vetorial das distribuições X e Y. Temos

$$\begin{aligned} \mathrm{cp}(Z) &= \|\alpha\mathbf{x} + (1-\alpha)\mathbf{y}\|^2 \\ &= \alpha^2\|\mathbf{x}\|^2 + (1-\alpha)^2\|\mathbf{y}\|^2 + 2\alpha(1-\alpha)\langle \mathbf{x}, \mathbf{y}\rangle \\ &\leq \alpha^2 \,\mathrm{cp}(X) + (1-\alpha)^2 \,\mathrm{cp}(Y) + 2\alpha(1-\alpha)\sqrt{\mathrm{cp}(X)\,\mathrm{cp}(Y)} \\ &\leq \max\{\mathrm{cp}(X), \mathrm{cp}(Y)\}. \end{aligned} \tag{2.4}$$

A primeira desigualdade segue por Cauchy-Schwarz. A partir de $\mathrm{cp}(X), \mathrm{cp}(Y)$, $\sqrt{\mathrm{cp}(X)\,\mathrm{cp}(Y)} \leq \max\{\mathrm{cp}(X), \mathrm{cp}(Y)\}$ obtemos a segunda desigualdade. □

Proposição 2.2.18. *Seja X uma distribuição satisfazendo $\mathrm{cp}(X) \leq (KL)^{-1}$, com $K, L \geq 1$. Então X está a distância estatística $L^{-1/2}$ de ter min-entropia pelo menos $\log K$.*

Demonstração. Seja $\mathbf{x}$ a representação vetorial de X. Defina $S = \{a \mid x_a \geq 1/K\}$ e $S^c = \mathrm{supp}(X) \setminus S$. Observe que se $S = \emptyset$ então $H^\infty(X) \geq \log K$ e nada temos a provar. Se $S^c = \emptyset$ então $\mathrm{cp}(X) \geq K^{-1}$, o que implica $L = 1$ e, novamente, nada temos a provar. Podemos definir distribuições X' e X'' tais que $\mathrm{supp}(X') = S$ e $\mathrm{supp}(X'') = S^c$ da seguinte forma. Tome $x'_a = x_a \Big/ \sum_{b \in S} x_b$ pra todo $a \in S$ e $x_a = 0$ para $a \notin S$. Analogamente tome $x''_a = x_a \Big/ \sum_{b \in S^c} x_b$ se $a \in S^c$ e $x''_a = 0$ e $a \notin S^c$.

Observe que X é combinação convexa de X' e X'', digamos $X = \alpha X' + (1-\alpha)X''$. De (2.4) seque que $1/(KL) \geq \mathrm{cp}(X) \geq \alpha^2 \,\mathrm{cp}(X') \geq \alpha^2/K$. Portanto, $\alpha \leq 1/\sqrt{L}$. Note que $H^\infty(X'') \geq \log K$. Como $\mathrm{dist}(X, X'') \leq \alpha \leq 1/\sqrt{L}$, concluímos a prova do proposição. □

Lema 2.2.19. *Seja X uma distribuição sobre um conjunto de tamanho m. Se $\mathrm{cp}(X) \leq (1 + 4\varepsilon^2)/m$ então X é ε-próximo de uniforme.*

Demonstração. Seja $\mathbf{x}$ a representação vetorial de X e $\mathbf{u}$ o vetor com m coordenadas iguais a $1/m$. Sem perda de generalidade, suponha que $S = \mathrm{supp}(X) = [m]$. Por Cauchy-Schwarz segue que

$$\|\mathbf{x} - \mathbf{u}\|_1^2 \leq m \sum_{i=1}^{m} (x_i - u_i)^2 = m\Big(\sum_{i=1}^{m} x_i^2 - \frac{2}{m}\sum_{i=1}^{m} x_i + \sum_{i=1}^{m} \frac{1}{m^2}\Big) = 4\varepsilon^2.$$

Segue que $\mathrm{dist}(X, U) = \|\mathbf{x} - \mathbf{u}\|_1/2 \leq \varepsilon$. □

O lema a seguir é conhecido como lema XOR e é geralmente atribuído a Vazirani. Há um *survey* de Goldreich [Gol95] a respeito de tal lema.

Lema 2.2.20. *Seja X uma distribuição sobre $\mathbb{Z}_2^n$. Defina*

$$\mathrm{maxbias}(X) = \max_{\mathbf{0} \neq v \in \mathbb{Z}_2^n} \Big\{ \Big| \mathbf{P}[\langle X, v\rangle = 0] - \mathbf{P}[\langle X, v\rangle = 1] \Big| \Big\}.$$

Para todo k, seja U_k a distribuição uniforme sobre $\mathbb{Z}_2^k$. Temos

$$\mathrm{dist}(X, U_n) \leq \sqrt{\sum_{\mathbf{0} \neq v \in \mathbb{Z}_2^n} \mathrm{dist}(\langle X, v\rangle, U_1)^2} \leq 2^{n/2} \cdot \mathrm{maxbias}(X). \tag{2.5}$$

Começamos provando um simples lema auxiliar.

Lema 2.2.21. *Seja $\mathbf{0} \neq v \in \{0,1\}^n$. Então*

$$\sum_{w \in \{0,1\}^n} (-1)^{\langle v, w\rangle} = 0.$$

Demonstração. A identidade a seguir é imediata:

$$\sum_{w\in\{0,1\}^n}(-1)^{\langle v,w\rangle} = \#\Big\{w\in\{0,1\}^n \;\Big|\; \langle v,w\rangle = 0 \pmod 2\Big\} - \#\Big\{w\in\{0,1\}^n \;\Big|\; \langle v,w\rangle = 1 \pmod 2\Big\}.$$

Seja $m = |\mathrm{supp}(v)| \geq 1$. O número de vetores w no primeiro conjunto da equação acima é

$$\sum_{j=0}^{\infty}\binom{m}{2j}2^{n-m} = 2^{m-1}2^{n-m} = 2^{n-1},$$

provando assim o lema. □

Esboço da demonstração do lema 2.2.20. Seguiremos a exposição de [Gol95]. Para toda base ortonormal B de $\mathbb{R}^{2^n}$ e todo $r \geq 1$, definimos a (r,B)-*norma* de $v \in \{0,1\}^n$ como

$$N_r^B(v) = \Big(\sum_{w\in B}\big|\langle v,w\rangle\big|^r\Big)^{1/r}.$$

Seja K a base canônica de $\mathbb{R}^{2^n}$. É simples verificar que, sendo $\mathbf{x}$ o vetor de probabilidades da distribuição X e $\mathbf{u}$ o vetor correspondente a U_n, temos

$$\mathrm{dist}(X,U_n) = \frac{1}{2}N_1^K(\mathbf{x}-\mathbf{u}). \tag{2.6}$$

Seja $F' = \Big\{\mathbf{b}_v = \big((-1)^{\langle v,w\rangle}\big)_{w\in\{0,1\}^n} \mid v\in\{0,1\}^n\Big\}$. O lema 2.2.21 claramente implica que os vetores de F' são dois-a-dois ortogonais (pois se $v \neq v'$ então $\langle \mathbf{b}_v, \mathbf{b}_{v'}\rangle = \sum_{w\in\{0,1\}^n}(-1)^{\langle v+v',w\rangle} = 0$). Podemos normalizar os vetores de F' simplesmente dividindo-os por $2^{n/2}$. Seja F a base ortonormal obtida.

As seguintes desigualdades valem para qualquer $v \in \{0,1\}^n$. Para toda base ortonormal A,

$$N_1^A(v) \leq 2^{n/2}N_2^A(v). \tag{2.7}$$

Para todo par de bases ortonormais A e B,

$$N_2^A(v) = N_2^B(v). \tag{2.8}$$

Temos

$$\begin{aligned} N_2^F(\mathbf{x}-\mathbf{u}) &= 2^{-n/2}\Big\{\sum_{v\in\{0,1\}^n}\langle \mathbf{b}_v, \mathbf{x}-\mathbf{u}\rangle^2\Big\}^{1/2} \\ &= 2^{-n/2}\Big\{\sum_{v\in\{0,1\}^n}\Big(\sum_{w\in\{0,1\}^n}(-1)^{\langle v,w\rangle}(x_w-u_w)\Big)^2\Big\}^{1/2} \end{aligned}$$

Como $\sum_w(x_w-u_w) = 0$, podemos eliminar o índice $v = \mathbf{0}$ do somatório. Além disso, quando $v \neq \mathbf{0}$ temos $\sum_w(-1)^{\langle v,w\rangle}u_w = 2^{-n}\sum_w(-1)^{\langle v,w\rangle} = 0$ (lema 2.2.21). Portanto,

$$\begin{aligned} N_2^F(\mathbf{x}-\mathbf{u}) &= 2^{-n/2}\Big\{\sum_{\mathbf{0}\neq v\in\{0,1\}^n}\Big(\sum_{w\in\{0,1\}^n}(-1)^{\langle v,w\rangle}x_w\Big)^2\Big\}^{1/2} \\ &= 2^{-n/2}\Big\{\sum_{\mathbf{0}\neq v\in\{0,1\}^n}\Big(\mathbf{P}[\langle X,v\rangle = 0] - \mathbf{P}[\langle X,v\rangle = 1]\Big)^2\Big\}^{1/2} \\ &= 2^{-n/2}\Big\{\sum_{\mathbf{0}\neq v\in\{0,1\}^n}4\,\mathrm{dist}\Big(\langle X,v\rangle, U_1\Big)^2\Big\}^{1/2}. \end{aligned}$$

Das equações (2.6), (2.7) e (2.8), obtemos

$$\operatorname{dist}(X, U_n) \leq \Big\{ \sum_{\mathbf{0} \neq v \in \{0,1\}^n} \operatorname{dist}\Big(\langle X, v\rangle, U_1\Big)^2 \Big\}^{1/2},$$

a desigualdade que queríamos provar. □

Agora definiremos objetos centrais deste estudo.

Definição 2.2.22 (Extratores)**.** *Uma função* $\texttt{ext}\colon \{0,1\}^{n\times l} \to \{0,1\}^m$ *é chamada de* extrator de l fontes com requerimento de min-entropia k, entrada de n-bits, saída de m-bits e distância estatística ε *se, para quaisquer fontes independentes com n-bits cada, digamos* $X_1, \ldots, X_l$*, todas com min-entropia* $\geq k$*, vale que*

$$\operatorname{dist}\Big(\texttt{ext}(X_1, \ldots, X_l), U_m\Big) \leq \varepsilon.$$

Definição 2.2.23 (Dispersores)**.** *Uma função* $\texttt{disp}\colon \{0,1\}^{n\times l} \to \{0,1\}^m$ *é chamada de* dispersor de l fontes com requerimento de min-entropia k, entrada de n-bits, saída de m-bits e parâmetro de erro ε *se, para quaisquer fontes independentes com n-bits cada, digamos* $X_1, \ldots, X_l$*, todas com min-entropia* $\geq k$*, vale que*

$$\Big|\operatorname{supp}\Big(\texttt{disp}(X_1, \ldots, X_l)\Big)\Big| \geq (1-\varepsilon)2^m.$$

Quando $\varepsilon = 0$ *temos um* dispersor sem-erro.

Capítulo 3

Resultados de Barak, Impagliazzo e Wigderson

Em [BIW04], é mostrado que, para todo $\delta > 0$ existe uma construção explícita de extratores que utilizam $l = \Theta\Big((1/\delta)^{\Theta(1)}\Big)$ (observe que l é constante) δ-fontes independentes de n-bits e cuja saída consiste de n-bits que estão a distância estatística 2^{-n} da distribuição uniforme U_n. O resultado é o seguinte teorema.

Teorema 3.0.24 (Extrator de BIW)**.** *Existem constantes $c_1, c_2 \geq 1$ tais que para qualquer constante $\delta > 0$, e todo parâmetro $l \geq L(\delta) = c_1(1/\delta)^{c_2}$, existe uma função* $\texttt{ext}\colon \{0,1\}^{n\times l} \to \{0,1\}^n$ *com a seguinte propriedade: para toda coleção de δ-fontes independentes $X_1, \ldots, X_l$, todas com n-bits, vale que*

$$\mathrm{dist}\Big(\texttt{ext}(X_1, \ldots, X_l), U_n\Big) \leq 2^{-\Omega(n)}.^{\dagger}$$

Ademais, `ext` *é uma função polinomialmente computável.*

Recentemente, Anup Rao [Rao05] obteve melhorias significativas, basicamente reduzindo o requerimento de min-entropia linear para polinomial.

Teorema 3.0.25 (Extrator de Rao)**.** *Para toda constante $c > 0$ existe uma constante c' tal que para todo n, k com $k = k(n) = \Omega(\log^4 n)$ existe um algoritmo polinomial* $\texttt{ext}\colon \{0,1\}^{n\times u} \to \{0,1\}^k$ *com $u \leq c' \log n / \log k$ tal que se $X_1, \ldots, X_u$ são fontes independentes com min-entropia pelo menos k, então*

$$\mathrm{dist}\Big(\texttt{ext}(X_1, \ldots, X_u), U_k\Big) \leq n^{-c}.$$

Observe que se $k = n^\alpha$ com α constante, o número de fontes exigidas é constante, uma melhora substancial em relação ao Teorema 3.0.24.

O lema a seguir é o bloco fundamental para a construção dos extratores e dispersores em [BIW04]. Observamos que, em [BIW04], é provado um resultado com cotas melhores para o caso de corpos de ordem prima enquanto que o lema abaixo trata de corpos finitos GF(2^p) com p primo.

[†]Observamos que é possível tornar o erro menor que 2^{-cn} para qualquer constante c usando o XOR de saídas correspondentes a coleções disjuntas de l fontes (isso implica em multiplicar o número de δ-fontes independentes usadas por uma constante). Tal redução de erro é usada por exemplo, por Raz [Raz04], para obtenção de condensadores de semente constante.

Lema 3.0.26. *Existe uma constante absoluta $\alpha > 0$ tal que, para todo $0 < \delta < 1$ fixado, para quaisquer variáveis aleatórias independentes A, B e C sobre o corpo* $\mathrm{GF}(2^p)$ *(para um primo $p \geq p_0(\delta)$), cada uma delas com razão de entropia δ, a distribuição $A \cdot B + C$ é $2^{-\alpha\delta^2 p}$-próxima de ter razão de entropia pelo menos* $\min\{\delta + \alpha\delta^2, 0.99\}$.

A demonstração do lema acima utiliza resultados recentes de teoria aditiva dos números incluídos no capítulo 4.

A partir do lema 3.0.26, podemos obter extratores. Para isso, vamos definir uma família de funções $\mathtt{ext}^i \colon \mathbb{F}^{3^i} \to \mathbb{F}$ onde $\mathbb{F}$ é um corpo finito. Definimos $\mathtt{ext}^0 \equiv \mathrm{id}$ e, para $i \geq 0$,

$$\mathtt{ext}^{i+1}(x_1, x_2, x_3) = \mathtt{ext}^i(x_1)\,\mathtt{ext}^i(x_2) + \mathtt{ext}^i(x_3),$$

onde x_1, x_2, x_3 são vetores com 3^i coordenadas (em $\mathbb{F}$) cada.

Esboço da prova do teorema 3.0.24. Seja $\delta > 0$ uma constante fixada. Seja $k = k(\delta) = 3^i$ o número de δ-fontes independentes necessárias para que o lema 3.0.26 nos garanta que $\mathtt{ext}^i \colon \mathbb{F}^k \to \mathbb{F}$ seja tal que para quaisquer δ-fontes independentes $X_1, \ldots, X_k$ sobre $\mathbb{F}$, vale que

$$H^\infty\Big(\mathtt{ext}^i(X_1, \ldots, X_k)\Big) \geq 0.99.$$

Sejam $X_1, \ldots, X_{2k}$ uma coleção de δ-fontes independentes. Aplicando o extrator de Hadamard (veja o teorema 6.0.14) às fontes $Y = \mathtt{ext}^i(X_1, \ldots, X_k)$ e $Z = \mathtt{ext}^i(X_{k+1}, \ldots, X_{2k})$ obtemos uma saída $2^{-\Omega(n)}$-próxima de uniforme com $\Omega(n)$ bits. (Para que a saída tenha n bits, basta repetir a idéia com mais fontes independentes e concatenar a saída.) □

Capítulo 4

Teoria Aditiva dos Números

Na construção dos extratores de Barak, Impagliazzo e Wigderson [BIW04], diversos resultados de teoria aditiva dos números foram empregados. Muitos desses resultados são de grande interesse por si só e também têm aplicações na construção dos extratores de 3 fontes [BKS+05] que descreveremos no capítulo 7.

Nesta seção abordaremos os resultados necessários para o desenvolvimento dos extratores acima mencionados. Tentaremos simplificar ao máximo as provas. Observamos que, o livro recém publicado de Tao e Vu [TV06, Capítulo 2, 6] contém diversos resultados desta seção, incluindo algumas generalizações.

4.1 Cotas soma-produto e um lema de Gowers

Um dos resultados que abordaremos é a seguinte estimativa *soma-produto* descrita em [BKT04].

Teorema 4.1.1. *Seja $0 < \delta < 1$ e $\mathbb{F}$ um corpo finito para o qual $|\mathbb{F}|^{1/k}$ não é inteiro para todo $k \in \{2, 3, \ldots, 2/\delta\}$. Existe uma constante absoluta $c_0 > 0$ tal que, para todo $A \subseteq \mathbb{F}$ com $|\mathbb{F}|^{\delta} < |A| < |\mathbb{F}|^{1-\delta}$, temos*

$$\max\Big\{|A + A|, |A \cdot A|\Big\} \geq |A|^{1+c_0\delta}.$$

Observe que com p primo e $\delta > 2/p$, o corpo $\mathbb{F} = \mathrm{GF}(2^p)$ satisfaz o teorema acima. Basta observar que $(2^p)^{1/k} = 2^{p/k} = m$ com m inteiro se e somente se $m = 2^c$ para algum natural c. Disso segue que $p = kc$ e, como $k \in \{2, 3, \ldots, 2/\delta\}$, devemos ter $k = p$ e $c = 1$. Para provar o lema 3.0.26 é suficiente verificar a conclusão do teorema 4.1.1 para corpos $\mathrm{GF}(2^p)$ com p primo.

O seguinte lema trata de uma estimativa para a cardinalidade de conjuntos-soma (veja [Nat96, p. 219]).

Lema 4.1.2. *Sejam A e B subconjuntos não-vazios de um grupo aditivo tal que $|A + B| \leq K \min\{|A|, |B|\}$ para algum K. Então, para todo $c, d \geq 1$, temos*

$$|cA - dA| \leq K^{c+d}|A|,$$

onde $cA = \underbrace{A + \cdots + A}_{c \text{ vezes}}$.

Há uma versão multiplicativa do lema acima, a saber, vale o seguinte lema.

Lema 4.1.3. *Sejam A e B subconjuntos não-vazios de um grupo aditivo tal que $|A \cdot B| \leq K \min\{|A|, |B|\}$ para algum K. Então, para todo $c, d \geq 1$, temos*

$$|A^c/(A^d)| \leq K^{c+d}|A|,$$

onde $A^c = \underbrace{A \cdot A \cdots A}_{c \text{ vezes}}$.

O lema a seguir foi provado no caso particular em que $A = B$ por Gowers [Gow98, Proposition 12]. Uma versão com conjuntos diferentes é devida a Bourgain [Bou99]. Há uma generalização desses resultados para um número arbitrário (porém constante) de subconjuntos em [SSV05]. Todos os trabalhos acima citados seguem uma linha combinatória onde o problema é modelado como um problema em grafos. Seguiremos de perto a exposição de [SSV05].

Teorema 4.1.4. *Sejam A e B subconjuntos de um grupo com $|A| \leq |B| = M$. Suponha que exista um grafo bipartido $\mathcal{G}$ com classes A e B tal que $|E(\mathcal{G})| \geq M^2/K$ para algum K e $S = S(\mathcal{G}) = \{a + b \mid (a, b) \in E(\mathcal{G})\}$ satisfaz $|S| \leq CM$ para algum C.*[1] *Então podemos obter conjuntos $A' \subseteq A$ e $B' \subseteq B$ tais que $|A'| \geq M/(16K^2)$, $|B'| \geq M/(4K)$ e $|A' + B'| \leq 2^{12}C^3K^5M$. Ademais, para cada elemento $a + b \in A' + B'$, há pelo menos $2^{-12}K^{-5}M^2$ triplas ordenadas $(s_1, s_2, s_3) \in S^3$ tais que $a + b = s_1 - s_2 + s_3$.*

Um corolário simples do teorema 4.1.4 segue.

Corolário 4.1.5. *Nas mesmas condições do teorema 4.1.4 (supondo a existência de um grafo $\mathcal{G}$ como acima), podemos obter $A' \subseteq A$ e $B' \subseteq B$ tais que para cada $a + b \in A' + B'$ há pelo menos $2^{-20}K^{-8}C^{-3}M^5$ sêxtuplas ordenadas $(a_1, a_2, a_3, b_1, b_2, b_3) \in A^3 \times B^3$ satisfazendo*

$$a + b = (a_1 + b_1) - (a_2 + b_2) + (a_3 - b_3).$$

Defina $r_X(n) = \#\{(a, b) \in X \mid a + b = n\}$. Para não carregar a notação, tomaremos $r(\cdot) = r_{A \times B}(\cdot)$. Por conveniência, abusaremos da notação e identificaremos grafos com seus respectivos conjuntos de arestas.

Demonstração. Tome $S' = \{n \in S \mid r_{\mathcal{G}}(n) \geq M/(2KC)\}$ e forme um grafo $\mathcal{G}' \subseteq \mathcal{G}$ com $\mathcal{G}' = \{(a, b) \in \mathcal{G} \mid a + b \in S'\}$. Claramente $S(\mathcal{G}') \subseteq S' \subseteq S$. Como $\mathcal{G} \setminus \mathcal{G}' = \bigcup_{n \in S \setminus S'} \{(a, b) \in \mathcal{G} \mid a + b = n\}$ possui no máximo $|S \setminus S'|M/(2KC) < M^2/(2K)$ arestas, segue que $|\mathcal{G}'| \geq M^2/(2K)$.

Então, aplicando o teorema 4.1.4 obtemos A' e B' tais que todo $a + b \in A' + B'$ pode ser representado por $2^{-12}(2K)^{-5}M^2$ triplas $(s_1, s_2, s_3) \in S'^3$. Por definição, cada s_i pode ser representado como $s_i = a_i + b_i$ por pelo menos $M/(2KC)$ pares $(a_i, b_i) \in A \times B$. Logo, há pelo menos $2^{-20}K^{-8}C^{-3}M^5$ maneiras de representar $a + b$ como enunciado no corolário. □

A prova do teorema 4.1.4 se baseia no seguinte lema em grafos.

Lema 4.1.6. *Seja $\mathcal{G} \subseteq A \times B$ um grafo no qual $|B| \leq |A| = M$ e $|\mathcal{G}| \geq M^2/K$. Então podemos obter conjuntos $A' \subseteq A$ e $B' \subseteq B$ com $|A'| \geq M/(16K^2)$ e $|B'| \geq M/(4K)$ tais que para todo $(a, b) \in A' \times B'$ existem pelo menos $2^{-12}K^{-5}M^2$ passeios de comprimento 3 de a até b em $\mathcal{G}$.*

[1] No teorema, K e C não precisam ser constantes.

Demonstração. Primeiramente, apague todos os vértices de B com grau $\leq M/(2K)$ (observe que o grau médio em B é pelo menos M/K) e, por conveniência, mantenha a mesma notação (portanto, agora $|\mathcal{G}| \geq M^2/(2K)$). Claramente $\mathbf{E}_{a\in_R A}\Big[d(a)\Big] = |\mathcal{G}|/|A| \geq M/(2K)$. O conjunto B' será subconjunto de $\Gamma(a)$ para algum $a \in A$.

Dados $u, w \in B$ chamamos o par (u, w) de *ruim* se $N(u, w) = |\Gamma(u) \cap \Gamma(w)| \leq M/(128K^3)$. Seja Y_a o número de pares ruins em $\Gamma(a) \times \Gamma(a)$ para $a \in A$. Seja $\mathcal{B}$ a coleção de pares ruins. Então

$$\begin{aligned}\mathbf{E}_{a\in_R A}[Y_a] &= \sum_{a\in A} \frac{1}{|A|} \sum_{(u,w)\in\mathcal{B}} \chi_{\Gamma(u)\cap\Gamma(w)}(a) \\ &= \frac{1}{M} \sum_{(u,w)\in\mathcal{B}} N(u, w) \leq \frac{|\mathcal{B}|}{M} \frac{M}{128K^3} \leq \frac{M^2}{128K^3}.\end{aligned} \tag{4.1}$$

Seja S_a o conjunto de elementos de $\Gamma(a)$ que formam par ruim com pelo menos $M/(32K^2)$ outros elementos de $\Gamma(a)$. Seja $Z_a = |S_a|$. Temos que

$$Z_a \frac{M}{32K^2} \leq Y_a.$$

Segue que $\mathbf{E}_{a\in_R A}[Z_a] \leq M/(4K)$. Por linearidade da esperança, segue que $\mathbf{E}_{a\in_R A}[d(a) - Z_a] \geq M/(4K)$ e, portanto, existe um $a \in A$ tal que $|\Gamma(a) \setminus S_a| \geq M/(4K)$. Tome $B' = \Gamma(a) \setminus S_a$ para tal escolha de a.

Defina A' como sendo o conjunto dos vértices em A com pelo menos $M/(16K^2)$ vizinhos em B'. O número de arestas incidentes a B' é pelo menos $|B'|\,M/(2K) \geq M^2/(8K^2)$ (pois no início apagamos todo vértice de B com grau $< M/(2K)$). Como cada vértice de A' tem no máximo $|B'| \leq M$ vizinhos em B' e cada vértice de $A \setminus A'$ tem no máximo $M/(16K^2)$ vizinhos em B', segue que,

$$|A'|\,M + M\frac{M}{16K^2} \geq M^2/(8K^2),$$

e, portanto, $|A'| \geq M/(16K^2)$, como queríamos.

Resta mostrarmos quantos passeios existem entre quaisquer vértices $a \in A'$ e $b \in B'$. Por construção, existem pelo menos $M/(32K^2)$ vértices $b' \in \Gamma(a)$ tais que $N(b, b') \geq M/(128K^3)$. Logo há pelo menos $M/(128K^3)$ escolhas para $a' \in \Gamma(b) \cap \Gamma(b')$. Claramente (a, b', a', b) é um passeio de a até b e, portanto, há pelo menos $M^2/2^{12}K^5$ passeios de comprimento 3 entre a e b. □

Prova do teorema 4.1.4. Vamos demonstrar que o lema 4.1.6 implica o teorema 4.1.4. Considere os conjuntos A' e B' obtidos pelo lema. Seja (a, b', a', b) um passeio de comprimento 3 em $\mathcal{G}$ com $a \in A'$ e $b \in B'$. Observe que

$$a + b = (a + b') - (a' + b') + (a' + b)$$

e $a + b', a' + b', a' + b \in S$. Portanto cada passeio de comprimento 3 de $a \in A'$ até $b \in B'$ pode ser mapeado univocamente a uma tripla ordenada $(s_1, s_2, s_3) \in S^3$ tal que $a+b = s_1-s_2+s_3$. Como há no máximo C^3M^3 triplas em S^3 e pelo menos $2^{-12}K^{-5}M^2$ passeios distintos para cada par $(a, b) \in A'\times B'$, o teorema segue. □

Precisaremos de alguns resultados de [BKT04] para provarmos o teorema 4.1.1. Dado um valor K, usaremos a notação $X \lessapprox Y$ ($Y \gtrapprox X$) para dizer que existe uma constante C (que não

depende de K) tal que $X \leq CK^C Y$. Se $X \lessapprox Y$ e $Y \lessapprox X$ então diremos que $X \approx Y$ (neste caso, para alguma constante C e $\alpha = CK^C$, temos $\alpha^{-1}Y \leq X \leq \alpha Y$ e $\alpha^{-1}X \leq Y \leq \alpha X$). Observamos que K pode não ser uma constante absoluta e depender do tamanho dos conjuntos em questão. Outro ponto importante é que para qualquer constante absoluta l, se $X_1 \lessapprox X_2 \lessapprox \cdots \lessapprox X_l$ então $X_1 \lessapprox X_l$.

Lema 4.1.7. *Seja A um subconjunto não-vazio de um corpo finito $\mathbb{F}$ tal que*

$$\max\{|A+A|, |A \cdot A|\} \leq K|A|.$$

Então existe $A' \subseteq A$ com $|A'| \approx |A|$ tal que

$$|A' \cdot A' - A' \cdot A'| \lessapprox |A'|.$$

Demonstração. Vamos supor, sem perda de generalidade que $|A| \gg 1$ e que $0 \notin A$. (Observe que remover 0 não afeta significativamente as cotas do lema.)

Note que o lema 4.1.2 implica $|A - A| \leq K^2|A|$ e então podemos aplicar o corolário 4.1.5 tomando $\mathcal{G} = A \times (-A)$ (grafo bipartido completo). Dessa forma, obtemos conjuntos $D, E \subseteq A$ com $|D|, |E| \approx |A|$ tal que todo elemento de $d - e \in D - E$ pode ser representado de $\gtrapprox |A|^5$ maneiras na forma

$$d - e = (a_1 - a_2) - (a_3 - a_4) + (a_5 - a_6),$$

com $a_i \in A$ para $i = 1, \ldots, 6$.

Defina $F = A \cdot A \cdot A/(A \cdot A)$. Seja $f \in F$ um elemento arbitrário. É simples observar que todo elemento $(d-e)f \in (D-E) \cdot F$ pode ser representado de $\gtrapprox |A|^5$ maneiras na forma

$$(d-e)f = (b_1 - b_2) - (b_3 - b_4) + (b_5 - b_6),$$

com $b_i \in A \cdot F$ para $i = 1, \ldots, 6$. Portanto, como $|F| \lessapprox |A|$, pela versão multiplicativa do lema 4.1.2, segue que

$$|(D-E) \cdot F| \lessapprox |A|. \tag{4.2}$$

Agora refinaremos D e E. Como $|D|, |E| \approx |A| \approx |A \cdot A|$ segue que $|D \cdot E| \approx |D|, |E|$ e, portanto, pela versão multiplicativa do corolário 4.1.5 existem $D' \subseteq D$ e $E' \subseteq E$ tais que $|D'|, |E'| \approx |A|$ e todo elemento em $D' \cdot E'$ tem $\gtrapprox |A|^5$ representações do tipo

$$\frac{d_1 e_1 d_3 e_3}{d_2 e_2}, \text{ com } d_1, d_2, d_3 \in D \text{ e } e_1, e_2, e_3 \in E.$$

Sejam $d, d' \in D$ e $e, e' \in E$. Pelo princípio da casa dos pombos, existem $d_2 \in D$ e $e_2 \in E$ tais que há $\gtrapprox |A|^3$ soluções para

$$de = \frac{d_1 e_1 d_3 e_3}{d_2 e_2}, \text{ com } d_1, d_3 \in D \text{ e } e_1, e_3 \in E.$$

Podemos então expressar $de - d'e'$ como $x_1 - x_2 + x_3 - x_4$, onde

$$\begin{aligned}
x_1 &= (d_1 - e')e_1 d_3 e_3 (d_2 e_2)^{-1} \\
x_2 &= e'(d' - e_1) d_3 e_3 (d_2 e_2)^{-1} \\
x_3 &= d'e'(d_3 - e_2) e_3 (d_2 e_2)^{-1} \\
x_4 &= e_2 d'e'(d_2 - e_3)(d_2 e_2)^{-1}.
\end{aligned}$$

Para d, d', e, e', d_2, e_2 fixos, o mapeamento $(d_1, d_3, e_1, e_3) \mapsto (x_1, x_2, x_3, x_4)$ é injetor. Note que o mapeamento induzido $e_3 \mapsto x_4$ é injetor e, conseqüentemente, o mapeamento induzido $(d_3, e_3) \mapsto (x_3, x_4)$ também é injetor. Basta repetir o argumento mais duas vezes. Como todo $x_i \in (D-E)\cdot F$, verificamos que cada $de - d'e' \in D'\cdot E' - D'\cdot E'$ admite $\gtrsim |A|^3$ representações de $de - d'e'$ na forma $x_1 - x_2 + x_3 - x_4$ com $x_i \in (D-E)\cdot F$. Logo, de (4.2) segue que

$$|D'\cdot E' - D'\cdot E'| \lesssim |A|.$$

Como temos $|D'\cdot E'| \approx |D|$, segue pelo lema 4.1.3 que $|D'/E'| \approx |D|$. Pelo princípio da casa dos pombos, há um elemento $z \in D'/E'$ para o qual existem $\approx |D|$ pares $(x, y) \in D' \times E'$ tais que $x/y = z$. Observe que se $x/y = x'/y'$ então $(x, y) = (x', y')$ ou $x \neq x', y \neq y'$. Isso significa que para tal z, temos $|D' \cap zE'| \approx |D| \approx |A|$. Tomando $A' = D' \cap zE'$ provamos o lema, pois $|A'\cdot A' - A'\cdot A'| \leq |D'\cdot E' - D'\cdot E'| \lesssim |A|$. □

4.2 Estimativas para Somas Iteradas e Conjuntos-Produto

Nesta subseção, complementaremos o resultado do lema 4.1.7.

Lema 4.2.1. *Seja A um subconjunto não-vazio de um corpo finito $\mathbb{F}$ com $1 \in A$. Suponha que, para algum $K \geq 1$, temos*

$$|A\cdot A - A\cdot A| \leq K|A|.$$

Então, para qualquer polinômio P sobre diversas variáveis e com coeficientes inteiros, temos

$$P(A, A, \ldots, A) \leq CK^C|A|,$$

onde a constante C só depende de P.

Vamos introduzir uma notação. Diremos que A está *essencialmente contido* em B, e escreveremos $A \Subset B$, se existe X com $|X| \lesssim 1$ tal que $A \subseteq X + B$. Diremos que $x \in \mathbb{F}$ é *bom* se $x\cdot A \Subset A - A$. Temos três propriedades importantes da relação "essencialmente contido".

Transitividade. Suponha que $A \Subset B \Subset C$. Então existem X_1 e X_2, tais que $|X_1| \leq C_1K^{C_1}, |X_2| \leq C_2K^{C_2}$, $A \subseteq X_1 + B$ e $B \subseteq X_2 + C$. Portanto, $A \subseteq (X_1 + X_2) + C$, onde $|X_1 + X_2| \leq C_1C_2K^{C_1+C_2} \leq C_3K^{C_3}$. Segue que

$$A \Subset B \Subset C \text{ implica } A \Subset C \tag{4.3}$$

Aditividade. Se $A \Subset B$ e $C \Subset D$ então existem X_1 e X_2 tais que $|X_1| \leq C_1K^{C_1}, |X_2| \leq C_2K^{C_2}$, $A \subseteq X_1 + B$ e $C \subseteq X_2 + D$. Logo $A + C \subseteq (X_1 + X_2) + (B + D)$ e, pelo mesmo argumento acima, concluímos

$$A \Subset B \text{ e } C \Subset D \text{ implica } A + C \Subset B + D. \tag{4.4}$$

Elementos bons. Seja $X = \{x_1, \ldots, x_s\}$ um conjunto com $s = |X| \lesssim 1$ elementos bons. Temos $X\cdot A = x_1\cdot A \cup x_2\cdot A \cup \cdots \cup x_s\cdot A$. Por hipótese, para todo $i = 1, \ldots, s$ existe Y_i com $|Y_i| \lesssim 1$ tal que $x_i\cdot A \subseteq Y_i + (A - A)$. Segue que $X\cdot A \subseteq Y + A - A$, onde $Y = \bigcup_{i=1}^s Y_i$ claramente satisfaz $|Y| \leq \sum_{i=1}^s |Y_i| \lesssim s \lesssim 1$. Portanto,

$$\text{se } X \subseteq \mathbb{F} \text{ é tal que } |X| \lesssim 1 \text{ e todo } x \in X \text{ é bom então } X\cdot A \Subset A - A \tag{4.5}$$

O seguinte lema é devido a Plünnecke-Ruzsa [Ruz99].

Lema 4.2.2. *Sejam A e B subconjuntos de $\mathbb{F}$ tais que $|A+B| \lessapprox |A|$ ou $|A-B| \lessapprox |A|$. Então $B \Subset A-A$.*

Demonstração. Como $A-A$ é simétrico (ie, $x \in A-A$ se e somente se $-x \in A-A$), basta mostrar o caso em que $|A+B| \lessapprox |A|$. Seja $X \subseteq B$ um conjunto maximal com a propriedade que os membros da família $\{A+x \mid x \in X\}$ são dois-a-dois disjuntos. Como todo conjunto da família tem tamanho $|A|$ e são contidos em $A+B$, pela hipótese de disjunção, segue $|X||A| \leq |A+B|$ e, portanto, $|X| \lessapprox 1$. Como X é maximal, para todo $b \in B$, o conjunto $A+b$ intersecta pelo menos um $A+x$ com $x \in X$. Logo, $b \in x+A-A$ e, portanto, $B \Subset A-A$, como queríamos. □

Proposição 4.2.3. *Temos os seguintes fatos.*

- *Todo elemento do conjunto A do lema 4.2.1 é bom.*
- *Se $x, y \in \mathbb{F}$ são bons então $x+y$, $x-y$ e xy são bons.*

(As constantes implícitas envolvidas variam para cada ocorrência.)

Demonstração. Primeiramente vamos mostrar que todo elemento de A é bom. Como $1 \in A$, temos

$$|A \cdot A - A| \leq |A \cdot A - A \cdot A| \leq K|A|.$$

Portanto, segue do lema 4.2.2, que

$$A \cdot A \Subset A-A, \tag{4.6}$$

o que implica que todo elemento de A é bom.

Suponha que x e y sejam bons. Então existem X e Y com $|X|, |Y| \lessapprox 1$ tais que $x \cdot A \subseteq X+A-A$ e $y \cdot A \subseteq Y+A-A$. Logo

$$(x+y) \cdot A \subseteq x \cdot A + y \cdot A \subseteq (X+Y) + (A-A+A-A).$$

Donde segue que $(x+y) \cdot A \Subset 2A-2A$ (pois $|X+Y| \lessapprox 1$). Por outro lado, como $|A-A| \leq |A \cdot A - A \cdot A| \leq K|A|$, pelo lema 4.1.2, segue que $|3A-2A| \lessapprox |A|$. Tomando $B = 2A-2A$ no lema 4.2.2 (com $|A+B| \lessapprox |A|$) obtemos

$$B = 2A-2A \Subset A-A. \tag{4.7}$$

Como $(x+y) \cdot A \Subset B$, segue, por transitividade, que $(x+y) \cdot A \Subset A-A$. Um argumento análogo mostra que $x-y$ é bom.

Resta mostrar que xy é bom. Como $x \cdot A \Subset A-A$, temos

$$xy \cdot A \Subset y \cdot A - y \cdot A \Subset 2A-2A,$$

donde, por (4.7), concluímos que xy é bom. □

Agora mostraremos como iterar a proposição acima. Primeiramente vamos provar que $A^k \Subset A-A$ para todo $k \geq 1$. Com isso, pela transitividade (cf. (4.3)) e a aditividade (cf. (4.4)) da relação $\Subset$, mais a equação 4.7, segue que $P(A, A, \ldots, A) \Subset A-A$. Isso claramente prova o lema 4.2.1 pois a existência de um conjunto X com $|X| \lessapprox 1$ tal que $P(A, \ldots, A) \subseteq X+A-A$ implica $|P(A, \ldots, A)| \lessapprox |A-A| \leq K|A|$.

Prova do Lema 4.2.1. Vamos provar por indução que $A^k \Subset A - A$ para todo $k \geq 1$. O caso $k = 1$ é trivial (pois, dado $a \in A$, temos $A \subseteq \{a\} + A - A$), o caso $k = 2$ é a equação (4.6). Suponha que $k > 2$ e que $A^{k-1} \Subset A - A$. Logo, existe X com $|X| \lessapprox 1$ tal que

$$A^{k-1} \subseteq X + A - A.$$

Claramente podemos supor que $X \subseteq A^{k-1} - A + A$. Pela proposição 4.2.3, temos que todo elemento de X é bom. Multiplicando por A e usando (4.5) obtemos

$$A^k \subseteq X \cdot A + A^2 - A^2 \Subset 3A - 3A.$$

Usando a equação (4.7) sucessivamente, concluímos que $3A - 3A \Subset A - A$ e a prova do lema está completa. □

4.3 Prova do Teorema 4.1.1

Os lemas abaixo serão usados na prova do teorema 4.1.1. Chamaremos de *expressão racional* uma expressão envolvendo apenas as operações $+, -, \cdot, /$ e variáveis sem coeficientes, por exemplo $(x_1 - x_2)/(x_2 \cdot x_3 + x_4)$. O tamanho de uma expressão é o número de operações na mesma.

Lema 4.3.1. *Existe uma expressão racional $r(\cdot)$ de tamanho fixo tal que, para todo δ, todo corpo finito $\mathbb{F}$ tal que $|\mathbb{F}|^{1/k}$ não é inteiro para $k \in [2, 2/\delta]$ e todo $B \subseteq \mathbb{F}$ com $|B| \geq |\mathbb{F}|^{\delta}$, temos*

$$|r(B, \dots, B)| \geq \min\Big\{|B|^{1+\delta}, |\mathbb{F}|\Big\}.$$

Começamos provando uma asserção mais simples.

Lema 4.3.2. *Seja $\mathbb{F}$ um corpo finito e $B \subseteq \mathbb{F}$ e $k \in \mathbb{N}$ ($k \geq 2$) tal que $|\mathbb{F}|^{1/k} < |B| \leq |\mathbb{F}|^{1/(k-1)}$. Defina[2] $D = (B - B)/(B - B)$. Então $|D| \geq |\mathbb{F}^{1/(k-1)}|$.*

Demonstração. Suponha que $|D| < |\mathbb{F}^{1/(k-1)}|$. Então podemos encontrar $s_1 \in \mathbb{F} \setminus D$. De forma análoga, para $j = 2, \dots, k-1$ podemos obter s_j tal que

$$s_j \in \mathbb{F} \setminus \Big(D + s_1 \cdot D + \cdots + s_{j-1} \cdot D\Big).$$

Considere o mapa $f : B \times B \times \dots B \to \mathbb{F}$ dado por $(x_0, \dots, x_{k-1}) \mapsto x_0 + s_1 x_1 + \cdots + s_{k-1} x_{k-1}$. Como o domínio tem tamanho $|B|^k > |\mathbb{F}|$, o mapa não pode ser injetivo e, portanto, existem $\mathbf{x}, \mathbf{x}'$ diferentes cuja imagem é o mesmo elemento de $\mathbb{F}$. Seja i o maior índice tal que $x_i \neq x_i'$. Claramente

$$(x_0 - x_0') + s_1(x_1 - x_1') + \cdots + s_{i-1}(x_{i-1} - x_{i-1}') = s_i(x_i - x_i').$$

Dividindo por $x_i - x_i'$ concluímos que $s_i \in D + s_1 \cdot D + \cdots + s_{i-1} \cdot D$, uma contradição. □

Prova do lema 4.3.1. Sejam $\mathbb{F}, B$ e δ como no enunciado do lema. Tome k tal que $|\mathbb{F}|^{1/k} < |B| \leq |\mathbb{F}|^{1/(k-1)}$. Observe que devemos ter $\delta \leq 1/(k-1)$. Seja $D = (B - B)/(B - B)$. Se $D = \mathbb{F}$ então, tomando $r(a, b, c, d) = (a - b)/(c - d)$, concluímos o lema. Caso contrário, tome $E = (D - D)/(D - D)$.

[2] Com isso queremos dizer que os elementos de D são da forma $(d_1 - d_2)/(d_3 - d_4)$ e $d_3 \neq d_4$.

Como $|D| \geq |\mathbb{F}|^{1/(k-1)}$ e o lado direito desta desigualdade, por hipótese, não é inteiro, temos na verdade uma desigualdade *estrita*. Logo,

$$|E| \geq |\mathbb{F}|^{1/(k-2)} = |\mathbb{F}|^{\frac{1}{k-1}\left(1+\frac{1}{k-2}\right)} \geq |B|^{1+\frac{1}{k-2}} \geq |B|^{1+\delta}.$$

Neste caso, tome $r(x_1,\ldots,x_{16}) = \frac{r'(x_1,\ldots,x_4)-r'(x_5,\ldots,x_8)}{r'(x_9,\ldots,x_{12})-r'(x_{13},\ldots,x_{16})}$ com $r'(a,b,c,d) = (a-b)/(c-d)$ para concluir o lema. □

Lema 4.3.3. *Para toda expressão racional $r(\cdot)$, existe uma constante C (que depende de r) tal que se $A \subseteq \mathbb{F}$ satisfaz $|A+A|, |A\cdot A| \leq |A|^{1+\rho}$, para $\rho > 0$ suficientemente pequeno, então existe $B \subseteq \mathbb{F}$ com $|B| \geq |A|^{1-C\rho}$ mas $|r(B,\ldots,B)| \leq |B|^{1+C\rho}$.*

A prova do lema seguirá da seguinte asserção.

Lema 4.3.4. *Para todo $k > 0$ existe uma constante $C = C(k)$ tal que para todo $\rho > 0$, se $A \subseteq \mathbb{F}$ satisfaz $|A+A|, |A\cdot A| \leq |A|^{1+\rho}$, então existe $B \subseteq \mathbb{F}$ tal que $|A|^{1-C\rho} \leq |B| \leq |A|$ mas $|B^k - B^k| \leq |A|^{1+C\rho}$.*

Demonstração. Aplicando o lema 4.1.7, obtemos $B' \subseteq A$ tal que $|B'| \approx |A|$ e $|B'\cdot B' - B'\cdot B'| \lessapprox |B'|$. Tome $\zeta \in \mathbb{F}$ tal que $1 \in \zeta \cdot B'$ e $B = \zeta \cdot B'$. Observe que $|B| = |B'|$ e $B\cdot B - B\cdot B = \zeta^2(B'\cdot B' - B'\cdot B')$. Usando o lema 4.2.1 concluímos que $|B^k - B^k| \lessapprox |B|$ (pois a expressão $B^k - B^k$ está fixada e tem tamanho limitado por uma constante). Tomando C suficientemente grande concluímos a prova do lema (note que só há dependência em k e nenhuma dependência em $K = |A|^\rho$). □

Prova do lema 4.3.3. Seja $r(\cdot)$ uma expressão racional. É sempre possível expressar r como quociente de dois polinômios. Seja então $r = p/q$ com p e q polinômios. Suponha que há no máximo k' monômios cada um com grau no máximo k' em p e o mesmo valha para q. Sejam B o conjunto e C' a constante obtidos pelo lema 4.3.4 para $k = 2k'^2$. Note que podemos supor que $0 \in B$ (e $1 \in B$ por construção) mudando apenas a constante C'. Observe que

$$r(B,\ldots,B) \subseteq \{0\} \cup \frac{k'B^{k'} \setminus \{0\}}{k'B^{k'} \setminus \{0\}}.$$

Como $|B| \geq |A|^{1-C'\rho}$, segue que $|B^k - B^k| \leq |A|^{1+C'\rho} \leq |B|^{1+2C'\rho} \leq |B^k|^{1+2C'\rho}$. Podemos então tomar $K = |B^k - B^k|/|B^k| \leq |B^k|^{2C'\rho}$ no lema 4.1.2 para obtermos $|kB^k| \leq K^k|B^k| \leq |B^k|^{1+2kC'\rho} \leq |B|^{1+C''\rho}$ para uma constante C'' diferente. Dada a escolha de k e o fato que $0,1 \in B$, segue que $kB^k \supseteq (k'B^{k'})\cdot(k'B^{k'})$. Aplicando a versão multiplicativa do lema 4.1.2 (para $k'B^{k'} \setminus \{0\}$) concluímos que $r(B,\ldots,B) \leq |B|^{1+C\rho}$ para uma constante C que só depende de r. □

Prova do teorema 4.1.1. Seja r a expressão racional obtida no lema 4.3.1. Seja C a constante obtida para r no lema 4.3.3. Se para δ, $\mathbb{F}$ e A nas condições do teorema, ambos $A+A$ e $A\cdot A$ tem tamanho $\leq |A|^{1+\delta/(10C)}$ então, pelo lema 4.3.3 existe um conjunto $B \subseteq \mathbb{F}$ tal que $|\mathbb{F}|^{1-\delta/2} > |A| \geq |B| \geq |A|^{1-\delta/10} > |\mathbb{F}|^{\delta/2}$ tal que

$$|r(B,\ldots,B)| \leq |B|^{1+\delta/10} < \min\left\{|B|^{1+\delta/2}, |\mathbb{F}|\right\},$$

o que contradiz o lema 4.3.1. □

4.4 Prova do Lema 3.0.26

Primeiramente vamos estabelecer uma relação entre probabilidade de colisão e o teorema 4.1.4 devido a Gowers.

Dados conjuntos A e B abusaremos da notação e diremos que $\mathrm{cp}(A+B)$ é a probabilidade de colisão da variável aleatória $x+y$, onde $x \in_R A$ e $y \in_R B$ são escolhidos de maneira uniforme e independente.

Lema 4.4.1. *Sejam A e B subconjuntos de um grupo tal que $|A| = |B| = M$ e seja K tal que $\mathrm{cp}(A+B) \geq (KM)^{-1}$. Então existe $\mathcal{G} \subseteq A \times B$ tal que $|\mathcal{G}| \geq M^2/(K+1)$ e $S = S(\mathcal{G}) = \{a+b \mid (a,b) \in \mathcal{G}\}$ satisfaz $|S| \leq KM$.*

Demonstração. Seja $S^* = A + B$. Suponha que ordenamos os elementos de S^* como $\{s_1, \ldots, s_t\}$, com $t = |S^*|$ de forma que $r(s_1) \geq r(s_2) \geq \cdots \geq r(s_t)$. (Lembrando que $r_X(n) = \#\{(a,b) \in X \mid a+b = n\}$ e $r(\cdot) = r_{A\times B}(\cdot)$.)

Se $t \leq KM$, o lema é imediato. Caso contrário, defina $S' = \{s_1, \ldots, s_{KM}\}$ e tome $\mathcal{G} = \{(a,b) \in A \times B \mid a+b \in S'\}$. Claramente $S(\mathcal{G}) \subseteq S'$. Resta mostrarmos que $|\mathcal{G}|$ é grande.

Como A e B são subconjuntos de um grupo e tem tamanho limitado por M, segue que $r(s_1) \leq M$. Por definição, temos $r(s_1) + \cdots + r(s_t) = M^2$ e $|\mathcal{G}| = r(s_1) + \cdots + r(s_{KM})$. Pela condição sobre a probabilidade de colisão, temos

$$\frac{1}{KM} \leq \mathrm{cp}(A+B) = M^{-4} \sum_{s_i} r(s_i)^2.$$

Podemos quebrar o somatório acima em dois, obtendo

$$\begin{aligned}
\sum_{s_i} r(s_i)^2 &= \sum_{i=1}^{KM} r(s_i)^2 + \sum_{i=KM+1}^{t} r(s_i)^2 \\
&\leq M \sum_{i=1}^{KM} r(s_i) + r(s_{KM}) \sum_{i=KM+1}^{t} r(s_i) \\
&\leq M\,|\mathcal{G}| + \frac{|\mathcal{G}|}{KM} M^2 \\
&= M\,|\mathcal{G}| \frac{K+1}{K}.
\end{aligned}$$

Donde segue que $|\mathcal{G}| \geq M^2/(K+1)$ como queríamos. □

Juntando o lema acima ao teorema 4.1.4 obtemos o seguinte teorema.

Teorema 4.4.2. *Sejam A e B subconjuntos de um grupo aditivo satisfazendo $\mathrm{cp}(A+B) \geq (KM)^{-1}$ e $|A|, |B| \geq M/L$. Então podemos obter $A' \subseteq A$ e $B' \subseteq B$ tais que $|A'|, |B'| \geq M/(64K^2L^3)$ e $|A' + B'| \leq 2^{17}K^8L^4M$.*

Demonstração. Sem perda de generalidade, suponha que $|A| \leq |B|$. Primeiramente, observamos que $\mathrm{cp}(A+B) \leq |B|^{-1} \leq |A|^{-1}$. Isso segue de

$$\begin{aligned}
(|A|\,|B|)^2\,\mathrm{cp}(A+B) &= \sum_{n \in A+B} r(n)^2 \leq |A| \sum_{n \in A+B} r(n) \\
&= |A|^2\,|B|.
\end{aligned}$$

Provaremos que existe $b' \in B$ tal que $\mathrm{cp}(A + B \setminus \{b'\}) \geq \mathrm{cp}(A + B)$. Isso nos permite supor que $|A| = |B| = M/L$ e, ao aplicarmos o lema 4.4.1, obtemos $\mathcal{G} \subseteq A \times B$ tal que $|\mathcal{G}| \geq |A|^2/(KL+1)\}$ e $|S(\mathcal{G})| \leq K|A|$. Aplicando o teorema 4.1.4 concluímos que existem $A' \subseteq A$ e $B' \subseteq B$, com $|A'|, |B'| \geq |A|/\{16(KL+1)^2\}M/(64K^2L^3)$ e $|A' + B'| \leq 2^{12}K^3(KL+1)^5|A| \leq 2^{17}K^8L^4M$, concluindo o teorema.

Defina $X_b = \#\{(a_1, b, a_2, b_2) \in (A\times B)^2 \mid a_1 + b = a_2 + b_2\}$. Observe que $(|A|\,|B|)^2 \,\mathrm{cp}(A+B) = \sum_{b\in B} X_b$. Tome $b' \in B$ com X_b mínimo e defina $B' = B \setminus \{b\}$. É evidente que $X_{b'} \leq |A|^2|B| \,\mathrm{cp}(A + B)$. Observe que

$$(|A|\,|B'|)^2 \,\mathrm{cp}(A + B') = \sum_{b\in B} X_b - 2X_{b'} + |A|. \tag{4.8}$$

De fato, em (4.8) estamos contando toda quádrupla $(a_1, b_1, a_2, b_2) \in (A \times B)^2$ com $a_1 + b_1 = a_2 + b_2$; depois (usando simetria) eliminamos todas aquelas que possuem $b_1 = b'$ ou $b_2 = b'$; finalmente, compensamos as quádruplas que foram eliminadas duas vezes, a saber (a_1, b', a_1, b'), para todo $a_1 \in A$. Temos

$$\begin{aligned}
\mathrm{cp}(A + B') &= \frac{\sum_{b\in B} X_b - 2X_{b'} + |A|}{(|A|\,|B| - |A|)^2} \\
&\geq \frac{(|A|\,|B|)^2 \,\mathrm{cp}(A + B) - 2|A|^2|B| \,\mathrm{cp}(A + B) + |A|^2 \,\mathrm{cp}(A + B)}{(|A|\,|B|)^2 - 2|A|^2\,|B| + |A|^2} \\
&= \mathrm{cp}(A + B).
\end{aligned}$$

□

Fixando ε. De agora em diante, tomamos[3]

$$\varepsilon = \varepsilon(\delta) = \begin{cases} c_0 10^{-5}\delta > 0 & \text{se } \delta \in (0, 1/2], \\ c_0 10^{-5}(1 - \delta) & \text{se } \delta \in (1/2, 1). \end{cases} \tag{4.9}$$

Com tal escolha, de acordo com o teorema 4.1.1, para todo conjunto $X \subseteq \mathrm{GF}(2^p)$ de tamanho $|X| \geq M^{1-10^4\varepsilon}$, com $M = 2^{\delta p} = |\mathrm{GF}(2^p)|^\delta$, vale $\max\{|X + X|, |X \cdot X|\} \geq M^{1+10^4\varepsilon}$. Em todas as demonstrações estaremos supondo que M é suficiente grande e $\mathbb{F} = \mathrm{GF}(2^p)$.

Lema 4.4.3. *Seja $A \subseteq \mathbb{F}$ tal que $|A| \geq M^{1-10\varepsilon}$ e $|A \cdot B| \leq M^{1+999\varepsilon}$ para algum B com $|B| \geq M^{1-300\varepsilon}$. Então, para todo conjunto $C \subseteq \mathbb{F}$, com $|C| = M$ temos $\mathrm{cp}(A + C) < M^{-1-10\varepsilon}$.*

Demonstração. Suponha que $\mathrm{cp}(A + C) \geq M^{-1-10\varepsilon}$. Pelo teorema 4.4.2, podemos obter $A' \subseteq A$ e $C' \subseteq C$ tais que $|A'|, |C'| \geq M^{1-300\varepsilon}$ e $|A' + C'| \leq M^{1+300\varepsilon} \leq M^{600\varepsilon} \min\{|A'|, |C'|\}$. Pelo lema 4.1.2, $|A' + A'| \leq M^{1+1200\varepsilon}$.

Então, para nossa escolha de ε, o teorema 4.1.1 nos garante que $|A' \cdot A'| \geq M^{1+10^4\varepsilon}$. Com isso, obtemos $|A' \cdot B| \geq M^{1+1000\varepsilon}$ (caso contrário, teríamos $|A' \cdot B| \leq M^{1300\varepsilon} \min\{|A'|, |B|\}$, o que, pelo lema 4.1.2, contradiz $|A' \cdot A'| \geq M^{1+10^4\varepsilon}$). Mas $A' \subseteq A$ e segue que $|A \cdot B| \geq M^{1+1000\varepsilon}$, contradizendo a hipótese inicial. □

[3] Observe que quanto mais próximo δ estiver de 1, menor é o valor de ε, evidenciando a dificuldade de trocarmos a constante 0.99 do lema 3.0.26 por $1 - o(1)$ enquanto mantemos α como uma constante absoluta.

Diremos que um conjunto A é *aditivamente amigável* (+-amigável) se a conclusão do lema acima (ie, para todo C de tamanho M temos $\mathrm{cp}(A + C) \leq M^{-1-10\varepsilon}$). Observe que o lema 4.4.3 admite uma versão multiplicativa (a demonstração apenas troca o papel da multiplicação e da adição na prova acima). De forma análoga, diremos que A é *multiplicativamente amigável* ($\times$-amigável) se para todo C de tamanho M temos $\mathrm{cp}(A \cdot C) \leq M^{-1-10\varepsilon}$.

Lema 4.4.4. *Seja $A \subseteq \mathbb{F}$ com $|A| = M$. Então existem dois conjuntos disjuntos A_+ e $A_\times$ tais que*

- *ou A_+ é vazio ou é +-amigável;*
- *ou $A_\times$ é vazio ou é $\times$-amigável;*
- $|A \setminus (A_+ \cup A_\times)| \leq M^{1-\varepsilon}$.

Demonstração. A prova seguirá a partir de repetidas aplicações do teorema 4.1.4. Começaremos com $A_+ = \emptyset$ e $A_\times = A$ e manteremos os invariantes (i) A_+ e $A_\times$ particionam A, (ii) A_+ é +-amigável ou vazio. Se num dado passo $|A_\times| < M^{1-\varepsilon}$ então faça $A_\times \leftarrow \emptyset$ e o lema segue.

Suponha que $A_\times$ não é $\times$-amigável (caso contrário, o lema segue imediatamente). Então existe B tal que $\mathrm{cp}(A_\times \cdot B) \geq M^{-1-10\varepsilon}$. Usando a versão multiplicativa do teorema 4.4.2 obtemos $A' \subseteq A_\times$ e $B' \subseteq B$ tal que $|A'|, |B'| \geq M^{1-30\varepsilon}$ e $|A' \cdot B'| \leq M^{1+90\varepsilon}$. Do lema 4.4.3, concluímos que A' é +-amigável. Faça $A_+ \leftarrow A_+ \cup A'$ e $A_\times \leftarrow A_\times \setminus A'$. É evidente que o invariante (i) é mantido. Para o invariante (ii), verifica-se que a união de dois conjuntos +-amigáveis disjuntos é também +-amigável. Isso segue da proposição 2.2.17 observando que a distribuição $(A_+ \cup A') + C$ é combinação convexa das distribuições $A_+ + C$ e $A' + C$. □

Lema 4.4.5. *Sejam $A, B, C \subseteq \mathrm{GF}(2^p)$ com $|A| = |B| = |C| = M$. Então $A \cdot B + C$ é $M^{-\varepsilon}$-próximo de ter probabilidade de colisão no máximo $M^{-1-\varepsilon}$.*

Demonstração. Primeiro quebramos o conjunto A em $(A_+, A_\times)$ de acordo com o lema 4.4.4. Suponha que A_+ e $A_\times$ são não-vazios. Seja $X_+ = A_+ \cdot B + C$ e $X_\times = A_\times \cdot B + C$. Observe que $X = A \cdot B + C$ é uma combinação convexa de X_+ e $X_\times$. No caso em que um dos conjuntos A_+ e $A_\times$ é vazio, então X é $M^{-\varepsilon}$-próximo de X_+ ou $X_\times$ (já que o conjunto não vazio obtido da quebra de A possui todos os elementos de A com exceção de no máximo $M^{1-\varepsilon} = M^{-\varepsilon}|A|$ elementos). Para finalizar, mostraremos que se A_* é não vazio então X_* tem probabilidade de colisão no máximo $M^{-1-\varepsilon}$ (onde $* \in \{+, \times\}$).

Caso $A_+ \neq \emptyset$ usando a proposição 2.2.17, concluímos que $\mathrm{cp}(X_+) \leq \max_{b \in B} \mathrm{cp}(A_+ b + C)$. No entanto,

$$\mathrm{cp}(A_+ b + C) = \mathrm{cp}(A_+ + Cb^{-1}) \leq M^{-1-\varepsilon}.$$

A igualdade segue pois $A_+ + Cb^{-1}$ é apenas uma permutação da distribuição $A_+ b + C$. A desigualdade segue pois A_+ é +-amigável.

Caso $A_\times \neq \emptyset$ usando argumentos similares ao caso anterior, temos

$$\mathrm{cp}(X_\times) \leq \max_{c \in C} \mathrm{cp}(A_\times \cdot B + c) = \mathrm{cp}(A_\times \cdot B) \leq M^{-1-\varepsilon}.$$

Donde segue que $\mathrm{cp}(X) \leq M^{-1-\varepsilon}$. Como X está a distância no máximo $M^{-\varepsilon}$ de $A \cdot B + C$, provamos o lema. □

Com isso é bem simples obter o lema 3.0.26.

Prova do lema 3.0.26. Sejam A, B e C variáveis sobre GF(2^p) com razão de entropia δ. Tome $\delta' \leq \delta_{\max} \equiv \min\{\delta, 0.999\}$ como o maior valor para o qual $M = 2^{\delta' p}$ é inteiro. Observe que quando $p \to \infty$, temos $\delta_{\max} - \delta' \to 0$.

Podemos expressar $A \cdot B + C$ como combinação convexa de distribuições $A' \cdot B' + C'$ onde $|A'| = |B'| = |C'| = M = 2^{\delta' p}$. O lema 4.4.5 nos garante a existência de X' tal que $\mathrm{dist}(X', A' \cdot B' + C') \leq M^{-\varepsilon}$ e $\mathrm{cp}(X') \leq M^{-1-\varepsilon}$. (Lembrando que $\varepsilon = \varepsilon(\delta')$ de acordo com a equação (4.9).) Pela proposição 2.2.18 (tomando $K = M^{1+\varepsilon/2}$ e $L = M^{\varepsilon/2}$), segue que X' está a distância $M^{-\varepsilon/4}$ de ter min-entropia pelo menos $\log M^{1+\varepsilon/2} = (1 + \varepsilon/2)\delta' p$.

Aplicando novamente a proposição 2.2.17, demonstramos que $A \cdot B + C$ está a distância $M^{-\varepsilon} + M^{-\varepsilon/4} \leq M^{-\varepsilon/8} = 2^{-\varepsilon\delta' p/8} \leq 2^{-\alpha\delta^2 p}$ (para uma constante absoluta $\alpha > 0$) de ter razão de entropia $(1 + \varepsilon/2)\delta'$. Observe que não teríamos α independente de δ se não limitássemos δ' a no máximo uma constante (no caso, 0.999).

Para p suficientemente grande, podemos garantir que, quando $\delta \geq 0.999$, temos $\delta' \geq 0.99$ e $A \cdot B + C$ está próximo de ter razão entropia pelo menos 0.99. Se $\delta < 0.999$ então podemos garantir que $\delta' \geq (1 - 1/p)\delta$. Tomando p suficientemente grande e uma constante $\alpha > 0$ apropriada, garantimos que $A \cdot B + C$ está próximo de ter razão entropia pelo menos $(1 + \alpha\delta)\delta = (1 + \varepsilon/2)\delta'$. □

Capítulo 5

Condensador de Semente Constante

Nesta seção mostraremos como são construídos *condensadores de semente constante*. Estes objetos são interessantes por si só, porém eles serão usados na construção dos extratores do capítulo 7.

Informalmente, um *condensador* é uma função que leva palavras de n-bits vindas de uma δ-fonte em várias palavras menores de m-bits tal que a distribuição de saída de uma das palavras menores tem razão de entropia maior que δ. Um condensador de *semente constante* é um condensador no qual o número de palavras menores é constante. Optaremos pela visão de que essas palavras menores são blocos de uma única palavra e que a saída do condensador é *condensada-em-algum-lugar*. A definição formal é como segue.

Definição 5.0.6. *Uma função* $\mathtt{con}\colon \{0,1\}^n \to \{0,1\}^{m\times l}$ *é chamada de* condensador-em-algum-lugar *com* fator de expansão $\rho = \rho(\delta)$ *e distância* $\varepsilon = \varepsilon(\delta)$ *se, para toda* δ*-fonte* X *de* n*-bits,* $\mathtt{con}(X)$ *é* ε*-próximo de ter em-algum-lugar razão de entropia pelo menos* $\min\{\rho\delta, 0.99\}$.

Vamos exibir a construção de um condensador que toma qualquer δ-fonte e cuja saída seja em-algum-lugar muito próxima de ter razão de entropia pelo menos 0.99. O centro dessa construção é o seguinte *condensador básico* que condensa δ-fontes a razão de entropia $\delta(1+\Omega(\delta))$ usando apenas quatro blocos.

Teorema 5.0.7 (Condensador Básico[1].)**.** *Seja* $\delta \gg n^{-1/4}$ *um valor fixado. Existe uma constante absoluta* $\beta > 0$ *e um condensador-em-algum-lugar, computável em tempo* $\mathrm{poly}(n)$, $\mathtt{bcon}\colon \{0,1\}^n \to \{0,1\}^{m\times 4}$, *com* $m = \big(1/3 + o(1)\big)n/3$, *que expande qualquer* δ*-fonte* X *de* $\rho = 1+\beta\delta$ *com distância* $\varepsilon = 2^{-\beta\delta^2 n}$.

Demonstração. Inicialmente suponha que $n = 3p$ para algum primo p. Depois trataremos do caso em que n não é dessa forma. Para $x = (x_1, x_2, x_3) \in \{0,1\}^{3p}$ defina

$$\mathtt{bcon}(x_1, x_2, x_3) = (x_1, x_2, x_3, x_1 \cdot x_2 + x_3),$$

Onde $x_1 \cdot x_2$ é a multiplicação de elementos do corpo $\mathrm{GF}(2^p)$.

Seja X uma δ-fonte sobre $\{0,1\}^{3p}$ e $\theta = \alpha\delta/1000$, onde α é a constante do lema 3.0.26. Vamos mostrar a existência de um seletor $I = I(X)$ tal que $\mathtt{bcon}(X)_I$ é $2^{-\theta\delta p}$-próximo de ter razão de entropia pelo menos $(1+\theta)\delta$.

[1] Ran Raz [Raz04] exibiu um condensador de semente constante utilizando de maneira engenhosa o teorema 3.0.24.

Para $i \in [4]$ defina $H_i = \left\{y \in \{0,1\}^p \mid \mathbf{P}[\mathtt{bcon}(X)_i = y] \geq 2^{-(1+\theta)\delta p}\right\}$ como o conjunto dos elementos "pesados" da distribuição $\mathtt{bcon}(X)_i$. Observe que $|H_i| \leq 2^{(1+\theta)\delta p}$. Suponha que

$$\mathbf{P}\Big[\bigvee_i \mathtt{bcon}(X)_i \notin H_i\Big] \geq 1 - 2^{-\theta\delta p}.$$

Neste caso, basta tomar $I(x)$ como sendo o menor índice para o qual $\mathtt{bcon}(x)_{I(x)} \notin H_{I(x)}$ se este existir, caso contrário, defina $I(x)$ arbitrariamente.

Para $i \in [4]$ defina $S_i = \left\{x \in \{0,1\}^n \mid I(x) = i \text{ e } \mathtt{bcon}(x)_i \notin H_i\right\}$. Obseve que $B \equiv \{0,1\}^n \setminus \bigcup_{i=1}^4 S_i$ é tal que $\mathbf{P}[X \in B] \leq 2^{-\theta\delta p}$. Para todo $y \in \{0,1\}^p$, temos

$$\begin{aligned}\mathbf{P}[\mathtt{bcon}(X)_{I(X)} = y] &= \mathbf{P}[\mathtt{bcon}(X)_{I(X)} = y \wedge X \in B] \\ &\quad + \sum_{i\,:\, y \notin H_i} \sum_{x \in S_i} \mathbf{P}[\mathtt{bcon}(x)_i = y]\mathbf{P}[X = x].\end{aligned} \tag{5.1}$$

Como $\sum_{x\in S_i} \mathbf{P}[\mathtt{bcon}(x)_i = y]\mathbf{P}[X = x] \leq \mathbf{P}[\mathtt{bcon}(X)_i = y] < 2^{-(1+\theta)\delta p}$ quando $y \notin H_i$, segue que

$$\mathbf{P}[\mathtt{bcon}(X)_{I(X)} = y] \leq \mathbf{P}[\mathtt{bcon}(X)_{I(X)} = y \wedge X \in B] + 2^{2-(1+\theta)\delta p}.$$

Seja Z uma $(2/p - (1+\theta)\delta)$-fonte sobre $\{0,1\}^p$ tal que $\mathbf{P}[Z = z] = 2^{2-(1+\theta)\delta p}$ para todo z satisfazendo $\mathbf{P}[\mathtt{bcon}(X)_{I(X)} = z] \geq 2^{2-(1+\theta)\delta p}$. A existência de tal fonte é simples de ser verificada. Note que a distância de $\mathtt{bcon}(X)_{I(X)}$ a Z é no máximo $\sum_{y\in\{0,1\}^p} \mathbf{P}[\mathtt{bcon}(X)_{I(X)} = y \wedge X \in B] = \mathbf{P}[X \in B] \leq 2^{-\theta\delta p}$. Sendo assim, provamos o teorema para alguma constante $\beta > 0$.

Suponha então que $\mathbf{P}\Big[\bigvee_i \mathtt{bcon}(X)_i \notin H_i\Big] < 1 - 2^{-\theta\delta p}$ e, sem perda de generalidade, tome $X = \sum_i \alpha_i X_i$ como uma combinação convexa de fontes planas de min-entropia pelo menos δn. Seja

$$B = \{x \in X \mid \mathtt{bcon}(x)_j \in H_j,\, j = 1, \ldots, 4\}.$$

Por hipótese,

$$2^{-\theta\delta p} \leq \mathbf{P}[X \in B] = \sum_i \alpha_i \mathbf{P}[X_i \in B] = \sum_i \alpha_i \frac{|\operatorname{supp}(X_i) \cap B|}{|\operatorname{supp}(X_i)|},$$

logo, podemos supor que $|\operatorname{supp}(X_1) \cap B| \geq 2^{-\theta\delta p}|\operatorname{supp}(X_1)| \geq 2^{-\theta\delta p}2^{\delta n} = 2^{(3-\theta)\delta p}$. Portanto,

$$\begin{aligned}&\#\{x = (x_1, x_2, x_3) \in \operatorname{supp}(X_1) \mid \\ &\qquad\qquad x \in H_1 \times H_2 \times H_3, x_1 \cdot x_2 + x_3 \in H_4\} \geq 2^{(3-\theta)\delta p}.\end{aligned} \tag{5.2}$$

Seja $X' = B \cap \operatorname{supp}(X_1)$ o conjunto acima. Observe que a equação (5.2) implica $|H_1| \geq 2^{(1-3\theta)\delta p}$ pois, do contrário, relembrando que $|H_i| \leq 2^{(1+\theta)\delta p}$, teríamos $|X'| \leq |H_1 \times H_2 \times H_3| < 2^{(3-\theta)\delta p}$, uma contradição. Analogamente, $|H_2|, |H_3| \geq 2^{(1-3\theta)\delta p}$. Defina A_1, A_2, A_3 como distribuições independentes e uniformes sobre H_1, H_2, H_3, respectivamente. Por construção, cada A_i é uma $(1-3\theta)\delta$-fonte sobre $\{0,1\}^p$. Pela equação (5.2), temos

$$\mathbf{P}[A_1 \cdot A_2 + A_3 \in H_4] \geq \frac{|X'|}{|H_1|\,|H_2|\,|H_3|} \geq 2^{-4\theta\delta p}. \tag{5.3}$$

Por outro lado, para toda $(1+10\theta)\delta$ fonte Z sobre $\{0,1\}^p$, temos

$$\mathbf{P}[Z \in H_4] \leq |H_4| 2^{-(1+10\theta)\delta p} \leq 2^{-9\theta\delta p}. \tag{5.4}$$

Das equações (5.3) e (5.4) concluímos que

$$\sum_{x\in\{0,1\}^p} \left|\mathbf{P}[A_1 \cdot A_2 + A_3 = x] - \mathbf{P}[Z = x]\right| \geq \mathbf{P}[A_1 \cdot A_2 + A_3 \in H_4] - \mathbf{P}[Z \in H_4] \geq 2^{-4\theta\delta p - 1}, \tag{5.5}$$

mas isso implica uma contradição com o lema 3.0.26.

Para o caso em que n não é da forma $n = 3p$ para um primo p, colocamos zeros no final das palavras até que essas tenham o menor tamanho $n' = 3p$. É fato que para n suficientemente grande, existe tal n' com $n' < n + 3n^{3/4}$. Como a entropia da fonte não muda com esse processo, a nova razão de entropia passa a ser $\delta n/n'$. No final obtemos uma saída com entropia $(1+\theta)\delta n/n' > 1 - \theta/2$, pois $\theta \gg n^{-1/4}$, o que conclui a demonstração do teorema, já que conseguimos um fator de expansão $1 + \Omega(\delta)$, e erro exponencialmente baixo. □

No que segue, mostraremos como compor um condensador de l blocos e expansão ρ e um condensador de l' blocos e expansão ρ' assim obtendo um condensador de $l'l$ blocos e expansão $\rho'\rho$. Combinado ao teorema 5.0.7, podemos obter um condensador com fator de expansão arbitrário (desde que a saída tenha razão de entropia menor que 0.99) usando apenas um número constante de blocos.

Lema 5.0.8 (Composição de Condensadores)**.** *Seja*

1. $\mathsf{con}\colon \{0,1\}^n \to \{0,1\}^{m\times l}$ *um condensador de expansão* ρ *e distância* ε*;*
2. $\mathsf{con}'\colon \{0,1\}^m \to \{0,1\}^{m'\times l'}$ *um condensador de expansão* ρ' *e distância* ε'*, onde* $\rho'\rho < 0.99$*.*

Defina o condensador de $l'l$*-blocos* $\mathsf{con}' \circ \mathsf{con}$ *como segue. Para todo* $(i,i') \in [l]\times[l']$ *o* (i,i')*-ésimo bloco de* $\mathsf{con}' \circ \mathsf{con}(x)$ *é* $\mathsf{con}'_{i'}(\mathsf{con}_i(x))$*. Então* $\mathsf{con}' \circ \mathsf{con}\colon \{0,1\}^m \to \{0,1\}^{m'\times l'l}$ *é condensador–em-algum-lugar para* δ*-fontes com expansão* $\rho'\rho$ *(onde* $\rho = \rho(\delta)$ *e* $\rho' = \rho'(\rho\delta)$*) e distância* $\varepsilon + \varepsilon'$*.*

Demonstração. Seja X uma δ-fonte (onde $\rho'\rho\delta < 0.99$, com $\rho = \rho(\delta)$ e $\rho' = \rho'(\rho\delta)$) e considere a variável aleatória $I = I(X)$ tal que $Y = \mathsf{con}(X)_I$ é ε-próximo de uma $\rho\delta$-fonte Z. Seja $I' = I(Z)$ tal que $\mathsf{con}'(Z)_{I'}$ é ε'-próximo de ter razão de entropia pelo menos $\rho'\rho\delta$. Pela proposição 2.2.6, $\mathsf{con}'(Y)_{I'} = \mathsf{con}' \circ \mathsf{con}(X)_{(I,I')}$ é $(\varepsilon_1 + \varepsilon_2)$-próxima de ter razão de entropia $\rho'\rho\delta$. □

Aplicando a composição de condensadores ao condensador básico do teorema 5.0.7 obtemos o seguinte corolário.

Corolário 5.0.9. *Para todo* $0 < \delta < 1$ *existe* $l = 2^{\Theta(\delta^{-1})}$*,* $m = 2^{-\Theta(\delta^{-1})}n$ *e um condensador-em-algum-lugar computável em tempo* $2^{-\Theta(\delta^{-1})}\operatorname{poly}(n)$*, digamos,* $\mathsf{con}\colon \{0,1\}^n \to \{0,1\}^{m\times l}$*, que funciona com qualquer* δ*-fonte e tem saída* $2^{-\Theta(m)}$*-próxima de ter razão de entropia* 0.99 *em-algum-lugar.*

Demonstração. A prova segue a partir de repetidas composições de condensadores básicos. O teorema 5.0.7 nos garante a existência de uma família de condensadores básicos para δ fixado e todo r tal que $\delta \gg r^{-1/4}$. Denotaremos por bcon_r o condensador básico de tal família cuja entrada consiste de palavras de r-bits. Defina $\mathsf{con}^{(0)} \equiv \mathsf{bcon}_n$ e, para $i \geq 0$, defina indutivamente $\mathsf{con}^{(i+1)} = \mathsf{bcon}_{m_i} \circ \mathsf{con}^{(i)}$, onde m_i é o tamanho dos blocos da saída de $\mathsf{con}^{(i)}$. Observe que o número de blocos na saída de $\mathsf{con}^{(i)}$ é 4^i. Também segue do teorema 5.0.7 que $m_{i+1} = \big(1/3 + o(1)\big)m_i$ para todo $i \geq 0$ e $m_0 \geq \big(1/3 + o(1)\big)n$. Claramente, vale $m_i = 3^{-\Theta(i)}n$.

Primeiramente, vamos estimar o efeito de amplificação da composição de condensadores básicos. Suponha que $\mathsf{con}^{(i)}$ seja ε_i-próximo de ter razão de entropia pelo menos δ_i em-algum-lugar. Então, pelo lema 5.0.8, a saída de $\mathsf{con}^{(i+1)}$ é ε_{i+1}-próximo de ter razão de entropia pelo menos $\delta_{i+1} = \min\{(1+\beta\delta_i)\delta_i, 0.99\}$ em-algum-lugar, onde $\varepsilon_{i+1} = \varepsilon_i + 2^{-\beta\delta_i^2 m_i}$.

Considere intervalos $I_0, I_1, \ldots,$ nos quais $I_k = [2^k\delta, 2^{k+1}\delta)$. Suponha que $t = O(-\log\delta)$ é o primeiro intervalo a conter 0.99. Considere a seqüência $\{\delta_i\}_i$, onde δ_i é a razão de min-entropia garantida por $\mathsf{con}^{(i)}$. Se $\delta_j \in I_k$, então

$$\delta_{j+1} \geq \min\{0.99, \delta_j + \beta 2^{2k}\delta^2\}.$$

Segue que o número de índices j tais que $\delta_j \in I_k$ é no máximo $1 + 2^k\delta/(\beta 2^{2k}\delta^2) = 1 + 2^{-k}(\beta\delta)^{-1}$. O número total de índices da seqüência (até atingir o valor 0.99) pode ser limitado por

$$t + (\beta\delta)^{-1}\sum_{k=0}^{\infty} 2^{-k} = t + 2(\beta\delta)^{-1} = \Theta(\delta^{-1}).$$

Isso mostra que podemos tomar $\mathsf{con}^{(i)}$ como o condensador desejado para $i = \Theta(\delta^{-1})$. □

A seguinte proposição mostra que podemos obter, a partir de um condensador $\mathsf{con}\colon \{0,1\}^n \to \{0,1\}^{m\times l}$ um condensador cuja saída consiste de blocos com $m' < m$ bits (onde supomos sem perda de generalidade que m' divide m). Isso será feito de uma maneira bem trivial: cada bloco é quebrado em m/m' blocos de mesmo tamanho.

Proposição 5.0.10. *Seja X uma δ-fonte sobre $\{0,1\}^m$ e suponha que X seja a concatenação de k variáveis $X_1, \ldots, X_k$, cada uma de comprimento m'. Seja $\varepsilon > 0$. Então existe uma variável aleatória I sobre $[k]$ tal que X_I é $2^{-\varepsilon m}$-próximo de ter razão de entropia pelo menos $\delta - \varepsilon$.*

Demonstração. Sejam $Y_1, \ldots, Y_r$ fontes planas de min-entropia $\geq \delta$ tais que X é combinação convexa destas com coeficientes $\alpha_1, \ldots, \alpha_r$. Para todo $i \in [k]$ defina H_i como o conjunto dos elementos "pesados" de X_i, ou seja $H_i = \left\{y \in \{0,1\}^{m'} \;\middle|\; \mathbf{P}[X_i = y] \geq 2^{-(\delta-\varepsilon)m'}\right\}$. Observe que, para todo $j \in [r]$ vale

$$\mathbf{P}[Y_j \in H_1 \times \cdots \times H_k] \leq |H_1|\cdots|H_k|\,2^{-\delta m'} \leq \left(2^{(\delta-\varepsilon)m'}\right)^k 2^{-\delta m'} = 2^{-\varepsilon m}.$$

Donde segue que $\mathbf{P}[X \in H_1 \times \cdots \times H_k] \leq 2^{-\varepsilon m}$. Note que é simples obter uma distribuição Z que redistribui a massa de probabilidade dos conjuntos pesados de forma a ter min-entropia $\geq \delta$ e distância estatística de X limitada por $2^{-\varepsilon m}$. □

Usando a proposição 5.0.10 e o corolário 5.0.9 podemos obter um condensador final descrito no seguinte teorema.

Corolário 5.0.11 (Condensador final)**.** *Para todo $\delta > 0$ existe uma constante $\nu(\delta) = 2^{-\Theta(\log(1/\delta)/\delta)}$ tal que para toda constante $\beta \leq \nu$ existe uma constante $l = l(\delta, \beta)$ e um condensador-em-algum-lugar polinomialmente computável $\mathsf{con}\colon \{0,1\}^n \to \{0,1\}^{\beta n \times l}$ que recebe fontes cuja razão de entropia é δ e tem saída $2^{-\Theta(\beta n)}$-próxima de ter razão de entropia 0.97.*

Capítulo 6

Extrator-em-algum-lugar de 2 Fontes

Nesta seção mostraremos a construção de um extrator-em-algum-lugar de 2 fontes, ou seja, uma função polinomialmente computável que recebe como entrada duas δ-fontes de n-bits independentes (onde δ é uma constante arbitrariamente pequena, ou até mesmo levemente sub-constante) e tem saída uniforme-em-algum-lugar. Dizemos que uma distribuição X sobre $\{0,1\}^{m\times l}$ é uniforme-em-algum-lugar se existe uma variável aleatória $I = I(X)$ sobre $[l]$ tal que $X_I = U_m$. A seguir, alguns fatos simples sobre fontes uniformes-em-algum-lugar.

1. Se X é uma distribuição uniforme-em-algum-lugar sobre $\{0,1\}^{m\times l}$ então X tem min-entropia alta, i.e., $H^\infty(X) \geq m - \log l$. De fato, observe que, por construção, $H^\infty(X, I) \geq H^\infty(X_I)$. Como $H^\infty(X) \geq H^\infty(X, I) - \log |I|$, segue a afirmação. Note que ao considerarmos distribuições onde l é pequeno comparado a m, teremos $r(X) \geq 0.99$;

2. dado $m' < m$, a distribuição X' obtida a partir de X truncando cada bloco de X ao tamanho m' é uniforme-em-algum-lugar.

Definição 6.0.12 (Extrator-em-algum-lugar de duas-fontes)**.** *Dizemos que*

$$\texttt{s\_ext}_m \colon \{0,1\}^{n_1} \times \{0,1\}^{n_2} \to \{0,1\}^{m\times l}$$

é um extrator-em-algum-lugar *com requerimento de min-entropia δ e distância ε se $\texttt{s\_ext}_m(X, Y)$ é ε-próximo de uma distribuição uniforme-em-algum-lugar (pela proposição 2.2.15, isso é equivalente a dizer que $\texttt{s\_ext}_m(X, Y)$ é em-algum-lugar ε-próximo da distribuição uniforme). Omitiremos o índice m quando o tamanho da saída estiver dentro do contexto.*

Um elemento $y \in \{0,1\}^{n_2}$ fixado é bom *para X se $\texttt{s\_ext}(X, y)$ é ε-próximo de uma distribuição uniforme-em-algum-lugar. De forma análoga definimos quando um elemento $x \in \{0,1\}^{n_1}$ é bom para Y.*

Dizemos que $\texttt{s\_ext}$ é um extrator-em-algum-lugar forte *com os mesmos parâmetros acima se para todo X, Y como acima,*

$$\mathbf{P}_{x\in_R X}[x \text{ é bom para } Y] > 1 - \varepsilon, \text{ e } \mathbf{P}_{y\in_R Y}[y \text{ é bom para } X] > 1 - \varepsilon,$$

onde $x \in_R X$ significa que escolhemos x de acordo com a distribuição aleatória X.[1]

[1]Outra maneira (comum na literatura) de descrever um extrator forte, equivalente a definição acima é exigir que, mesmo que um adversário veja um dos parâmetros de entrada, ele não consegue distinguir (exceto com uma probabilidade muito pequena) a saída do extrator em relação a distribuição uniforme.

Teorema 6.0.13 (Extrator-em-algum-lugar)**.** *Para todo $\delta > 0$ fixado e $0 < \gamma \leq 1$, existem constantes $l, \beta, \eta > 0$ e um extrator-em-algum-lugar forte de l blocos $\texttt{s\_ext}\colon \{0,1\}^n \times \{0,1\}^{\gamma n} \to \{0,1\}^{m \times l}$, polinomialmente-computável, que funciona para todo $m \leq \beta n$ e fontes cuja razão de entropia é pelo menos δ e cuja distância é $\varepsilon \leq 2^{-\eta n}$*

A construção do extrator do teorema acima envolve o uso do condensador obtido no capítulo 5. Também será usado o seguinte extrator de duas fontes independentes com razão de entropia maior que $1/2$.

Teorema 6.0.14. *Para todo inteiro n existe $k = \Omega(n)$ e um extrator forte, polinomialmente computável, $\texttt{Had}\colon \{0,1\}^{n \times 2} \to \{0,1\}^k$ cuja saída é $2^{-\Omega(n)}$-próxima de uniforme para fontes de razão de entropia pelo menos 0.6.*

Mostramos uma construção explícita de extratores de Hadamard (e provamos o teorema 6.0.14) no apêndice A.

6.1 Prova do Teorema 6.0.13

Prova do Teorema 6.0.13. Dados parâmetros δ e γ, seja $\nu = \nu(\delta)$ a constante do corolário 5.0.11. Seja $n_1 = n, n_2 = \gamma n, \beta_1 = \gamma\nu$ e $\beta_2 = \nu$. Obtenha dois condensadores $\texttt{con}_j$ $(j = 1, 2)$ a partir do corolário 5.0.11 usando os parâmetros δ, n_j, β_j. Note que $\texttt{con}_j$ tem distância $\varepsilon_j = 2^{-\Theta(\beta_j n_j)}$ e sua saída consiste de l_j blocos de tamanho $n' = \beta_j n_j = \gamma\nu n$ (ou seja, os tamanhos são iguais para ambos os condensadores). Obtenha um extrator de duas fontes $\texttt{Had}_{n'}$ a partir do teorema 6.0.14. O extrator-em-algum-lugar construído terá $l = l_1 l_2$ blocos de saída, indexados por $(a, b) \in [l_1] \times [l_2]$, tais que

$$\texttt{s\_ext}(x_1, x_2)_{(a,b)} = \texttt{Had}_{n'}(\texttt{con}_1(x_1)_a, \texttt{con}_2(x_2)_b).$$

Por definição, cada bloco de saída tem saída de tamanho $\Omega(n') = \beta n$ para alguma constante $\beta > 0$. Sejam X_1 e X_2 duas δ-fontes sobre $\{0,1\}^{n_1}$ e $\{0,1\}^{n_2}$ respectivamente. Para $j = 1, 2$ seja I_j um seletor sobre $[l_j]$ tal que $\texttt{con}_j(X_j)_{I_j}$ é ε_j-próximo de ter razão de entropia 0.97 (observe que tais seletores existem pela definição de condensador-em-algum-lugar). Tomando $I = (I_1, I_2)$ e usando a proposição 2.2.6 concluímos que $\texttt{s\_ext}(X_1, X_2)_I = \texttt{Had}_{n'}(\texttt{con}_1(X_1)_{I_1}, \texttt{con}_2(X_2)_{I_2})$ é ε-próximo de ser uniforme-em-algum-lugar com $\varepsilon = \varepsilon_1 + \varepsilon_2 + 2^{-\Omega(n')} = 2^{-\Omega(n)}$. Ademais, como $\texttt{Had}_{n'}$ é um extrator forte, $\texttt{s\_ext}$ também é forte. Finalmente, pela proposição 2.2.5 e o fato 2, basta truncar os blocos para estes terem todos tamanho m e assim obter $\texttt{s\_ext}_m$. □

Capítulo 7

Extrator de 3 Fontes

A construção de um extrator que usa três δ-fontes será feita nesta seção.

Teorema 7.0.1. *Para quaisquer constantes $\delta, \varepsilon > 0$ e $m \in \mathbb{N}$. Existe n suficientemente grande e um extrator de 3-fontes* `3ext`$: \{0,1\}^{n\times 3} \to \{0,1\}^m$*, computável em tempo polinomial com relação a n, com requerimento de entropia δ e distância ε.*

As ferramentas usadas nessa construção consistem do extrator-em-algum-lugar de duas fontes construídos no capítulo 6 e de um extrator *ótimo* que recebe entradas bem pequenas. Tal extrator será construído a seguir utilizando-se, essencialmente, busca exaustiva.

7.1 Extrator Ótimo de 2 Fontes

Lema 7.1.1. *Para todos inteiros $d \geq d_0 = \Theta(1)$ e $m = \lfloor \log d \rfloor$ podemos obter em tempo $O\left(2^{5d^{14}}\right)$ um extrator* `opt`$: \{0,1\}^{d\times 2} \to \{0,1\}^m$ *cuja razão de entropia necessária é $6\log d$ e cuja distância é $1/d$.*

Demonstração. Seja $k = 6\log d, M = 2^m, D = 2^d$ e $K = 2^k = d^6$. Vamos mostrar, usando o Método Probabilístico , que existe uma matriz A, $D \times D$, com entradas em $\{0,1\}^m$ tal que $\texttt{opt}(x,y) = A_{xy}$ é o extrator procurado. Depois usaremos busca exaustiva para encontrar A.

Escolha A aleatoriamente de forma que os $D^2 m$ bits usados são K^2M-a-K^2M independentes. Chamaremos R de *retângulo* se $R = S \times T$ com $S, T \subseteq [D]$ e $|S| = |T| = K$. Seja R um retângulo e $\mathbf{0} \neq v \in \{0,1\}^m$, vamos provar que

$$\mathbf{P}\Big[\Big|\sum_{(i,j)\in R} (-1)^{\langle A_{ij}, v\rangle}\Big| > \frac{K^2}{dM}\Big] < \exp\Big\{-\Omega\Big(\frac{K}{dM}\Big)^2\Big\}. \tag{7.1}$$

Para cada $(i,j) \in R$ defina a variável aleatória $B_{ij} = (-1)^{\langle A_{ij}, v\rangle}$ e $B = \sum_{(i,j)\in R} B_{ij}$. Note que $\mathbf{E}[B_{ij}] = 0$ e, portanto, $\mathbf{E}[B] = 0$. Ademais, como os bits de A_R (sub-matriz induzida pelas coordenadas de R) são independentes, B é soma de variáveis independentes. A equação (7.1) segue utilizando-se a cota de Chernoff.

Usando a cota da união vemos que a probabilidade de existir um retângulo e um vetor v para o qual o evento de (7.1) falha é no máximo

$$\binom{D}{K}^2 M \cdot \exp\Big\{-\Omega\Big(\frac{K}{dM}\Big)^2\Big\} < \exp\Big\{3Kd - \Omega\Big(\frac{K}{dM}\Big)^2\Big\} \leq 1, \tag{7.2}$$

quando d é maior que uma constante d_0.

Conforme o corolário 2.2.8, podemos supor que as fontes são planas e isso acarreta em uma perda mínima nos parâmetros (de fato, aumentando levemente o valor de d_0, garantimos que (7.2) é satisfeita com k menor). Sejam X, Y fontes planas independentes de min-entropia k. Observe que $R = \mathrm{supp}(X) \times \mathrm{supp}(Y)$ é um retângulo e cada par de entrada é distribuído uniformemente sobre R. Isso significa que $\mathtt{opt}(X, Y)$ é um elemento de A_R escolhido uniformemente (evidentemente, uma mesma palavra de m bits pode aparecer várias vezes em A_R, ou seja, a saída de $\mathtt{opt}(X, Y)$ não é necessariamente uniforme). Então

$$\mathrm{maxbias}(\mathtt{opt}(X, Y)) = \frac{1}{K^2} \max_{\mathbf{0} \neq v \in \{0,1\}^m} \left\{ \left| \sum_{(i,j) \in R} (-1)^{\langle A_{ij}, v \rangle} \right| \right\} \leq \frac{1}{dM}.$$

Do lema 2.2.20 segue que $\mathtt{opt}(X, Y)$ é $1/d$-próximo de uniforme.

Na subseção 7.1.1 mostraremos como construir o espaço de probabilidade usado nesta demonstração, ou seja, construiremos um espaço de tamanho $(D^2m)^{K^2m}$ que nos forneça bits com as propriedades desejadas. Para fazer uma busca exaustiva no espaço que gera todas as possíveis matrizes A, todos os retângulos R e todos os vetores $\mathbf{0} \neq v \in \{0,1\}^m$ e ainda checar se o evento de (7.1) é satisfeito para (A, R, v) gastamos tempo proporcional a

$$(D^2m)^{K^2m} \binom{D}{K}^2 MK^2m \leq 2^{5d^{14}},$$

como queríamos. □

7.1.1 Simulando Espaços de Probabilidade k-a-k Independentes

Seguindo a construção dada em [AS00], podemos usar códigos BCH para construir um espaço de probabilidade Ω de tamanho n^k e n variáveis aleatórias $X_i \colon \Omega \to \{0,1\}$, com $i = 0, \ldots, n-1$ tais que para qualquer conjunto $S \subseteq [n]$ com $|S| \leq k$, temos

$$\mathbf{P}\Big[\bigwedge_{i \in S} X_i = a_i\Big] = \prod_{i \in S} \mathbf{P}[X_i = a_i].$$

Teorema 7.1.2. *[AS00, p. 255] Suponha que $n = 2^r - 1$ e $k = 2t + 1$. Então existe um espaço de probabilidade Ω de tamanho $2(n+1)^t$ e variáveis aleatórias k-a-k independentes $y_1, \ldots, y_n$ sobre Ω onde cada uma delas é uniforme sobre $\{0,1\}$. Ademais, Ω e $y_1, \ldots, y_n$ podem ser construídos explicitamente dada uma representação do corpo $\mathrm{GF}(2^k)$ como um espaço vetorial sobre $\mathrm{GF}(2)$.*

Observação 7.1.3. *Existe uma noção mais fraca de independência k-a-k. Em [AGHP90] são exibidas construções de espaços de probabilidade $\mathcal{S}$ onde são necessários apenas $O\big(\log\log n + k + \log \varepsilon^{-1}\big)$ bits para representar pontos de $\mathcal{S}$, e n variáveis aleatórias $Z_1, \ldots, Z_n \in \{0,1\}$ são definidas de forma que, para qualquer $I \subseteq [n]$ com $|I| \leq k$, temos $\mathrm{dist}\big((Z_i)_{i \in I}, U_{|I|}\big) \leq \varepsilon$.*

Raz [Raz04] utilizou diretamente a construção de [AGHP90] para obter poderosos extratores.

7.2 A Construção do Extrator de 3 Fontes

Agora construiremos um extrator de 3 fontes conforme o teorema 7.0.1. Sejam l, β, η as constantes do teorema 6.0.13 para os parâmetros δ e $\gamma = 1$. Tome $d = d(\delta, \varepsilon, m)$ satisfazendo

- $d \geq d_0$ (onde d_0 é a constante no lema 7.1.1);
- $d\varepsilon > 2$;
- $\log d \geq m$;
- $d/l - \log l \geq 6 \log d$.

Tomando n suficientemente grande, $\beta n > d/l$ e então podemos tomar $m = d/l$ no teorema 6.0.13, obtendo $\texttt{s\_ext}_{d/l}$, que tem como saída l blocos de tamanho d/l que são $2^{-\eta n}$-próximos de uma distribuição uniforme-em-algum-lugar.

Seja `opt` o extrator ótimo obtido na seção 7.1 com comprimento de entrada d. Como d é constante, `opt` pode ser computado em tempo constante. Observe que a saída desse extrator tem tamanho $\lfloor \log d \rfloor \geq m$. Será conveniente truncar a saída até m bits. Note que ao fazermos isso, pela proposição 2.2.5 e o fato 2, a saída ainda será $1/d$-próximo de uniforme.

A construção. O extrator é dado por:

$$\texttt{3ext}(x_1, x_2, x_3) = \texttt{opt}\Big(\texttt{s\_ext}_{d/l}(x_1, x_2), \texttt{s\_ext}_{d/l}(x_2, x_3)\Big).$$

Prova do Teorema 7.0.1. Sejam X_1, X_2, X_3 três δ-fontes independentes de n-bits. Lembre-se que `s_ext` é um extrator-em-algum-lugar forte e como "boa entrada" foi definida em 6.0.12. Para $i = 1, 3$, defina $B_i = \{x \mid x \text{ é ruim para } X_i\}$ e seja $B = B_1 \cup B_3$. Pela hipótese, temos

$$\mathbf{P}_{x \in_R X_2}[x \in B] \leq 2 \cdot 2^{-\eta n}.$$

Observe que, fixado x_2, as distribuições $\texttt{s\_ext}(X_1, x_2)$ e $\texttt{s\_ext}(x_2, X_3)$ são independentes. Se $x_2 \notin B$ então, além de independentes, elas são $2^{-\eta n}$-próximas de uma distribuição uniforme-em-algum-lugar com l blocos de tamanho d/l.

Pelo fato 1, tais saídas (condicionadas em $x_2 \notin B$) são $2^{-\eta n}$-próximas de terem entropia pelo menos $d/l - \log l$ que, por definição, é pelo menos $6 \log d$. Portanto (usando a proposição 2.2.6), ao aplicar `opt` a essas saídas obtemos uma distribuição $(2 \cdot 2^{-\eta n} + 1/d)$-próxima de uniforme. Isso implica (novamente usando a proposição 2.2.6) que a saída de `3ext` é $(4 \cdot 2^{-\eta n} + 1/d)$-próxima de uniforme. Como $(4 \cdot 2^{-\eta n} + 1/d) < \varepsilon$ para n grande o suficiente, mostramos que `3ext` satisfaz as propriedades enunciadas. □

7.3 O Extrator de Raz

O extrator definido por Raz [Raz04] atinge parâmetros melhores dos que os apresentados neste capítulo. Observamos que é possível aplicar a técnica de Shaltiel [Sha05] para aumentar o número de bits na saída do extrator de Raz pagando um leve preço em termos aumento no erro final.

Teorema 7.3.1. *Sejam $\delta > 0$ e $\varepsilon > 0$ constantes fixadas. Então existem constantes $c, \alpha > 0$ tais que se X_1, X_2 e X_3 fontes de $n_1, n_2, n_3 > c$ bits com min-entropias k_1, k_2 e k_3 satisfazendo $k_1 = \delta n_1$, $k_2, k_3 \geq 5 \log n_1$ e $n_1 \geq c \log n_2, c \log n_3$ então existe um extrator explícito que extrai $m \geq \min\{\alpha n_1, k_2\}/200$ bits e tem erro ε.*

Capítulo 8

Dispersores de 2 fontes

Nesta seção utilizaremos objetos como os extratores-em-algum-lugar do capítulo 6 e o extrator ótimo da seção 7.1 para obter um dispersor de 2 fontes. A parte mais substancial desta seção é o desenvolvimento do chamado *mecanismo desafio-resposta*, apresentado em [BKS+05]. O seguinte dispersor será obtido.

Teorema 8.0.2 (Dispersor de 2 fontes)**.** *Para todo $\delta > 0$ e $m \geq 1$ inteiro, existe um dispersor sem-erro, polinomialmente computável, de 2 fontes* `disp`$: \{0,1\}^{n\times 2} \to \{0,1\}^m$ *com requerimento de entropia pelo menos δ. Ademais, existe uma constante $\vartheta = \vartheta(\delta, m) > 0$ tal que, para quaisquer duas δ-fontes independentes X e Y e para todo $z \in \{0,1\}^m$, temos*

$$\mathbf{P}[\texttt{disp}(X,Y) = z] \geq \vartheta.$$

Para obter o dispersor prometido, iremos particionar os bits da entrada de diversas maneiras e será conveniente definir uma notação para tais partições.

Definição 8.0.3 (Partições)**.** *As palavras da entrada (que tem comprimento n) serão divididas em t (onde t é uma constante) partes de tamanho igual a n/t (podemos supor que t divide n adicionando bits sem informação a palavra, pagando o custo de deixar δ levemente menor). Se x é uma palavra de n-bits então $x[i]$ $(1 \leq i \leq t)$ é o i-ésimo bloco de x da direita para a esquerda. Em outras palavras, $x = x[t]x[t-1]\cdots x[2]x[1]$.*

Para todo $i \in [t]$ consideramos a partição de x em 3 blocos. O primeiro bloco é denotado por $x_1^i = x[t]x[t-1]\cdots x[i+1]$, o segundo bloco é $x_2^i \equiv x[i]$ e o terceiro bloco é $x_3^i \equiv x[i-1]x[i-2]\cdots x[1]$. Observe que a partição permite blocos vazios. Omitiremos o índice i quando este for claro no contexto.

Uma partição *I é um par (i_x, i_y) usada para particionar duas palavras $x, y \in \{0,1\}^n$. Mais precisamente, consideramos os blocos $x_j^I = x_j^{i_x}$ e $y_j^I = y_j^{i_y}$ para $j = 1,2,3$. Omitiremos I quando este for claro no contexto.*

Definiremos a seguinte *ordem parcial* $\prec$ sobre as partições. Temos $(i_x, i_y) \prec (i'_x, i'_y)$ se e somente se $i_x \leq i'_x$ e $i_y \leq i'_y$.

Preliminares e parâmetros. Lembre-se que os valores $\delta > 0$ e $m \geq 1$ são constantes fixadas e $n = |x| = |y|$ é maior que uma função $n_0(\delta, m)$ para que a construção do dispersor funcione.

Pelo teorema 6.0.13, podemos obter um extrator-em-algum-lugar de 2 fontes cuja razão dos tamanhos das entradas é uma constante e cuja razão de entropia é pelo menos $\delta > 0$. Abusaremos da notação e usaremos `s_ext` para entradas de tamanho variando de $\delta^3 n$ a n e requerimento de entropia variando de $\delta^5 n$ a n. Observe que existe uma constante l que é um limitante superior

para o número de blocos na saída de todos os extratores da família considerada acima. Podemos assumir que o número de blocos é sempre l simplesmente adicionando blocos arbitrários. Também podemos supor que o erro é $2^{-\eta n}$, para uma constante $\eta > 0$, em todas as aplicações de $\texttt{s\_ext}$. Denotaremos por $\texttt{s\_ext}_b$ um extrator-em-algum-lugar cuja saída consiste de blocos de tamanho b.

Usaremos o extrator ótimo do lema 7.1.1 $\texttt{opt}\colon \{0,1\}^{d\times 2} \to \{0,1\}^m$ tal que a constante d é grande o suficiente para satisfazer $\log d > 10m$ e $d/l - \log l > 10 \log d$. Observe que esses são requerimentos similares aos da seção 7.2. Ademais, tais requerimentos garantem que $\texttt{opt}$ pode ser aplicado a saída de $\texttt{s\_ext}_{d/l}$.

Podemos então definir um dispersor relativo a uma partição fixada.

Definição 8.0.4. *Dada uma partição I, com blocos x_1^I e y_1^I não vazios, definimos*

$$\texttt{disp}_I(x,y) = \texttt{opt}\Big(\texttt{s\_ext}_{d/l}(x_1, y_2), \texttt{s\_ext}_{d/l}(y_1, x_2)\Big). \tag{8.1}$$

Grosseiramente falando, a partição I que escolheremos no final será tal que, possivelmente tomando sub-fontes $X' \subseteq X$ e $Y' \subseteq Y$, teremos alta min-entropia em $(X'_1 \mid X'_2 = x_2)$ e $(Y'_1 \mid Y'_2 = y_2)$ para quase todo par $(x_2, y_2) \in \mathrm{supp}(X'_2) \times \mathrm{supp}(Y'_2)$. Ademais, como $\texttt{s\_ext}_{d/l}$ é forte, quase todo par (x_2, y_2) faz com que x_2 seja bom com relação a $(Y'_1 \mid Y'_2 = y_2)$ e que y_2 seja bom com relação a $(X'_1 \mid X'_2 = x_2)$. Dessa forma, temos que, fixando um par bom, a saída dos $\texttt{s\_ext}_{d/l}$ é uniforme-em-algum-lugar e, portanto, podemos aplicar o extrator ótimo assim como no caso dos extratores de 3 fontes do capítulo anterior. Observe que, como podemos precisar passar para uma sub-fonte, não temos um extrator de 2 fontes, apenas um dispersor.

8.1 O Mecanismo *Desafio-Resposta*

Apresentaremos um procedimento $\texttt{teste}_I\colon \{0,1\}^{n\times 2} \to \{0,1\}$ que indicará quando uma partição I é boa para tomarmos $\texttt{disp}(x,y) = \texttt{disp}_I(x,y)$. Observe que a entrada do procedimento $\texttt{teste}_I$ é a mesma do dispersor. O fato de que esse teste é possível é uma grande contribuição de [BKS+05].

O procedimento $\texttt{teste}_I(x,y)$. Dada uma partição I, definimos $\texttt{teste}_I(x,y)$ como segue.

1. Seja k uma constante suficientemente grande que escolheremos *a posteriori*. Seja $c_1 = \texttt{s\_ext}_k(x_3, y)$, $c_2 = \texttt{s\_ext}_k(y_3, x)$. Se x_3 é um bloco vazio, então definimos c_1 como uma palavra constante arbitrária e fazemos o mesmo para c_2 se y_3 é um bloco vazio. Seja $c = c_1c_2$ e observe que $|c| = 2lk$, já que o número de blocos em todos os extratores-em-algum-lugar que estamos usando é l. Chamaremos c de *desafio* de x, y para a partição I. Se ambos x_3 e y_3 forem vazios, ou seja $I = (1,1)$, teremos sempre $\texttt{teste}_I(x,y) = 1$ independente de x e y.

2. Para uma sub-palavra $\hat{x} = x_{j+1}x_{j+2}\ldots x_{j+\delta^3 n}$ de $x = x_1 \ldots x_n$ tal que j é múltiplo de $\delta^3 n$ e uma sub-palavra $\hat{y}$ de y com as mesmas propriedades, defina $c_{\hat{x},\hat{y}} = \texttt{s\_ext}_{2lk}(\hat{x}, \hat{y})$ (observe que $c_{\hat{x},\hat{y}}$ consiste de l blocos de comprimento $|c| = 2lk$). Nos referimos a $c_{\hat{x},\hat{y}}$ como o *chute*. Note que o número de chutes possíveis a partir de x e y fixados é uma constante.

3. Se existem sub-palavras $\hat{x}$ e $\hat{y}$ da forma acima tais que c é um dos blocos do chute $c_{\hat{x},\hat{y}}$, diremos que o desafio foi respondido e definimos $\texttt{teste}_I(x,y) = 1$. Caso contrário, definimos $\texttt{teste}_I(x,y) = 0$.

O dispersor final $\texttt{disp}(x, y)$ será $\texttt{disp}_I(x, y)$ para $I = I(x, y)$. De fato, a escolha de I será uma partição maximal (com relação a $\prec$) que satisfaz $\texttt{teste}_I(x, y) = 1$. Da seguinte asserção seguirá a construção explícita do dispersor almejado.

Asserção 8.1.1. *Para todo $\delta > 0$ e $m \geq 1$ inteiro, existe uma constante $\nu = \nu(\delta, m) > 0$ tal que para todo par X, Y de δ-fontes independentes existem sub-fontes (definição 2.2.11) independentes $\underline{X} \subseteq X$ e $\underline{Y} \subseteq Y$ de densidade pelo menos ν. Ademais, $\underline{X}$ e $\underline{Y}$ podem ser expressas como combinação convexa de fontes $\tilde{X}$ e $\tilde{Y}$, respectivamente, para as quais existe uma partição $\tilde{I}$ satisfazendo:*

(i) $\texttt{disp}_{\tilde{I}}(\tilde{X}, \tilde{Y})$ *é 2^{-2m}-próximo de uniforme;*

(ii) A partição $\tilde{I} = \tilde{I}(\tilde{x}, \tilde{y})$ é selecionada com probabilidade $1 - 2^{-2m}$ no espaço de probabilidade $(\tilde{x}, \tilde{y}) \in_R (\tilde{X}, \tilde{Y})$.

Prova do Teorema 8.0.2 a partir da Asserção 8.1.1. Observe que o item (i) implica que, para todo $z \in \{0,1\}^m$, temos $\mathbf{P}[\texttt{disp}_{\tilde{I}}(\tilde{X}, \tilde{Y}) = z] \geq 2^{-m} - 2^{-2m}$. Juntando isso ao item (ii), concluímos que

$$\begin{aligned}\mathbf{P}[\texttt{disp}(\tilde{X}, \tilde{Y}) = z] &\geq \mathbf{P}[\texttt{disp}_{\tilde{I}}(\tilde{X}, \tilde{Y}) = z]\mathbf{P}[\tilde{I} \text{ é selecionada}] \\ &\geq (2^{-m} - 2^{-2m})(1 - 2^{-2m}) \\ &= 2^{-m} - 2^{-2m} - 2^{-3m} + 2^{-4m}.\end{aligned}$$

Observe que a densidade de $(X', Y') \subseteq (X, Y)$ é pelo menos ν^2 e que a cota acima implica que $\mathbf{P}[\texttt{disp}(X', Y') = z] \geq 2^{-m} - 2^{-2m} - 2^{-3m} + 2^{-4m}$ para todo $z \in \{0,1\}^m$. Isso mostra o teorema 8.0.2 para

$$\vartheta = \vartheta(\delta, m) = \nu(\delta, m)^2(2^{-m} - 2^{-2m} - 2^{-3m} + 2^{-4m}) > 0.$$

□

Primeiramente, vamos estabelecer condições suficientes para que duas fontes independentes $\tilde{X}$ e $\tilde{Y}$ e uma partição $\tilde{I}$ satisfaçam $\text{dist}\left(\texttt{disp}_{\tilde{I}}(\tilde{X}, \tilde{Y}), U_m\right) \leq 2^{-2m}$.

Asserção 8.1.2. *Sejam $\tilde{X}$ e $\tilde{Y}$ fontes independentes e $\tilde{I}$ uma partição satisfazendo as seguintes propriedades.*

1. $H^\infty(\tilde{X}_1), H^\infty(\tilde{Y}_1) \geq \delta n/100$;
2. $H^\infty(\tilde{X}_2), H^\infty(\tilde{Y}_2) \geq \delta^2 n/100$;
3. $\texttt{teste}_{\tilde{I}}(\tilde{X}, \tilde{Y}) = 1$ *com probabilidade* 1;
4. *com probabilidade pelo menos* $1 - 2^{-2m}$, $\texttt{teste}_{I'}(\tilde{X}, \tilde{Y}) = 0$ *para todo* $I' \not\prec \tilde{I}$.

Então (8.1.1.i) $\texttt{disp}_{\tilde{I}}(\tilde{X}, \tilde{Y})$ *é 2^{-2m}-próximo de uniforme e (8.1.1.ii) a partição $\tilde{I} = \tilde{I}(\tilde{x}, \tilde{y})$ é selecionada com probabilidade $1 - 2^{-2m}$ no espaço de probabilidade $(\tilde{x}, \tilde{y}) \in_R (\tilde{X}, \tilde{Y})$.*

Observe que a propriedade 4 nos garante que a partição selecionada é $\tilde{I}$ com probabilidade pelo menos $1 - 2^{-2m}$ no espaço $(\tilde{X}, \tilde{Y})$, pois se $I' \prec \tilde{I}$, então, como escolhemos uma partição maximal cujo teste seja aceito, não importa o valor de $\texttt{teste}_{I'}(\tilde{X}, \tilde{Y})$. Mostraremos que $\texttt{disp}_{\tilde{I}}(\tilde{X}, \tilde{Y})$ é 2^{-2m}-próximo de uniforme na seção 8.5.

8.2 Resultados Técnicos

Lema 8.2.1. *Sejam Z uma δ-fonte de n-bits, $0 < \alpha < 1$ e $t \in \mathbb{N}$ constantes. Existe uma sub-fonte $Z' \subseteq Z$ de densidade pelo menos $1 - \alpha$ tal que Z' é combinação convexa $Z' = \sum_i \beta_i X_i$ onde $\beta_i \geq \alpha/(4t)$. Ademais, para todo $X = X_i$ existe um índice j tal que para todo $x \in \text{supp}(X)$, temos (usando a notação de partições da definição 8.0.3)*

(a) $H^\infty(X_2^j \mid X_3^j = x_3^j) > \frac{\delta n}{4t} - \log t + \log \alpha - 2$;

(b) $H^\infty(X_1^j \mid X_2^j = x_2^j, X_3^j = x_3^j) > \frac{3\delta n}{4} - \frac{n}{t} - \log t + \log \alpha - 2$.

Demonstração. Seja $p_k(x) = \mathbf{P}[Z_2^k = x_2^k \mid Z_3^k = x_3^k]$. Observe que, para todo x, temos $2^{-\delta n} \geq \mathbf{P}[Z = x] = \prod_{k=1}^t p_k(x)$ (um simples produtório "telescópico" da fórmula de Bayes). Disso segue que existe um índice $l \in [t]$ para o qual $p_l(x) \leq 2^{-\delta n/t}$. Seja $v = \delta n/(4t)$ e

$$A_i = \left\{ x \in \text{supp}(Z) \;\middle|\; i = \min_l \{ p_l(x) \leq 2^{-v} \} \right\}.$$

Pelas observações acima, é claro que $\{A_i\}_{i=1}^t$ é uma partição de $\text{supp}(Z)$. Dizemos que A_i é *grande* se $\mathbf{P}[Z \in A_i] \geq \alpha/(2t)$ e, neste caso, i é um *bom índice*. Para um A_i grande, dizemos que $x \in A_i$ é *pequeno* em A_i se $p_i(x) < \alpha 2^{-n/t}/(4t)$. Temos, para todo A_i grande,

$$\begin{aligned} &\mathbf{P}\Big[Z \text{ peq. em } A_i\Big] \\ &= \sum_{(x_2^i, x_3^i)} \mathbf{P}\Big[Z \text{ peq. em } A_i, Z_2^i = x_2^i \mid Z_3^i = x_3^i\Big] \mathbf{P}\Big[Z_3^i = x_3^i\Big] \\ &< \sum_{(x_2^i, x_3^i)} \Big\{ \alpha 2^{-n/t}/(4t) \Big\} \mathbf{P}\Big[Z_3^i = x_3^i\Big] = \frac{\alpha 2^{-n/t}}{4t} 2^{n/t} = \frac{\alpha}{4t}. \end{aligned} \tag{8.2}$$

Definimos o conjunto B como sendo a união de todos os A_i's que não são grandes e todos os elementos x que são pequenos. Usando a cota da união, obtemos $\mathbf{P}[Z \in B] \leq t\alpha/(2t) + t\alpha/(4t) < \alpha$. Para um índice bom i, seja $B_i = A_i \setminus B$. Segue que $\mathbf{P}[Z \in B_i] \geq \alpha/(4t)$. Fixe algum índice bom i e tome $X_i = Z \mid B_i$. Observe que a fonte $Z' = Z \mid \bigcup_{i \text{ bom}} B_i$ é idêntica a $Z \mid \text{supp}(Z) \setminus B$. Dessa forma, temos que $Z' \subseteq Z$ tem densidade pelo menos $1 - \alpha$. Ademais, Z' pode ser expresso como combinação convexa dos X_i's. Resta mostrarmos que cada X_i satisfaz as duas propriedades enunciadas. Seja $X = X_i$. Para o item (a), temos, para todo $x \in B_i$,

$$\begin{aligned} &\mathbf{P}\Big[X_2^i = x_2^i \;\Big|\; X_3^i = x_3^i\Big] = \mathbf{P}\Big[Z_2^i = x_2^i \;\Big|\; Z \in B_i, Z_3^i = x_3^i\Big] \\ &= \frac{\mathbf{P}\Big[Z_2^i = x_2^i, Z \in B_i \;\Big|\; Z_3^i = x_3^i\Big]}{\mathbf{P}[Z \in B_i]} \leq \frac{\mathbf{P}\Big[Z_2^i = x_2^i \;\Big|\; Z_3^i = x_3^i\Big]}{\mathbf{P}[Z \in B_i]} \\ &\leq p_i(x) \cdot 4t/\alpha \leq 2^{-\delta n/(4t) - \log t + \log \alpha + 2}. \end{aligned} \tag{8.3}$$

Para o item (b), observe que, para todo $x \in B_i$,

$$\begin{aligned} &\mathbf{P}\Big[X_1^i = x_1^i \;\Big|\; X_2^i = x_2^i, X_3^i = x_3^i\Big] \\ &= \mathbf{P}\Big[Z_1^i = x_1^i \;\Big|\; Z \in B_i, Z_2^i = x_2^i, Z_3^i = x_3^i\Big] \\ &= \frac{\mathbf{P}\Big[Z_1^i = x_1^i, Z_2^i = x_2^i, Z_3^i = x_3^i, Z \in B_i\Big]}{\mathbf{P}\Big[Z \in B_i, Z_2^i = x_2^i, Z_3^i = x_3^i\Big]} \\ &= \frac{\mathbf{P}[Z = x]}{\mathbf{P}\Big[Z \in B_i, Z_2^i = x_2^i, Z_3^i = x_3^i\Big]}. \end{aligned} \tag{8.4}$$

Como $\mathbf{P}[Z = x] \leq 2^{-\delta n}$, resta mostrar que $\mathbf{P}\Big[Z \in B_i, Z_2^i = x_2^i, Z_3^i = x_3^i\Big]$ é grande. Note que se $Z_2^i = x_2^i$ e $Z_3^i = x_3^i$ então $Z \in B_i$ (pois os únicos blocos que são relevantes para B_i são os blocos 2 e 3). Portanto,

$$\begin{aligned} \mathbf{P}\Big[Z \in B_i, Z_2^i = x_2^i, Z_3^i = x_3^i\Big] &= \mathbf{P}\Big[Z_2^i = x_2^i, Z_3^i = x_3^i\Big] \\ &= \prod_{k=1}^{i} p_k(x) > (2^{-v})^{i-1} \frac{\alpha}{4t} 2^{-n/t}. \end{aligned} \tag{8.5}$$

Como $(i-1)v \leq tv = \delta n/4$, segue o item (b). □

Corolário 8.2.2. *Seja Z uma δ-fonte de n-bits. Fixe $t = 10/\delta$. Existe uma sub-fonte $Z' \subseteq Z$ de densidade pelo menos $1/2$ em Z e tal que Z' pode ser expresso como combinação convexa onde cada componente X satisfaz, para algum índice $j \in [t]$:*

- X_3^j *é constante;*
- $H^\infty(X_2^j) \geq \frac{\delta^2 n}{50}$;
- $H^\infty(X_1^j) \geq \frac{\delta n}{2}$.

Demonstração. Seja Z' a fonte obtida pelo lema 8.2.1 para $\alpha = 1/2$. Temos que Z' é combinação convexa de fontes X' satisfazendo as propriedades (a) e (b) para algum $j \in [t]$. Observando que cada X' é combinação convexa das fontes $X = (X' \mid X_3'^j = x_3'^j)$, onde $x \in \mathrm{supp}(X')$, concluímos a demonstração do corolário. □

Também vamos precisar do seguinte corolário.

Corolário 8.2.3. *Seja Z uma δ-fonte de n-bits e $t \in \mathbb{N}$. Existe uma sub-fonte $X \subseteq Z$ com densidade pelo menos $1/(8t)$ e um índice $j \in [t]$ tal que $H^\infty(X[j]) \geq \delta n/(5t)$.*

Demonstração. Basta tomar $\alpha = 1/2$ no lema 8.2.1 e obter uma sub-fonte $X \subseteq Z$ de densidade $\alpha/(4t) = 1/(8t)$ tal que $H^\infty(X_2^j \mid X_3^j = x_3^j)$ para todo $x \in \mathrm{supp}(X)$. Segue que $X[j] = X_2^j$ satisfaz $H^\infty(X[j]) \geq \delta n/(5t)$ para n suficientemente grande. □

Lema 8.2.4. *Seja X uma fonte com $H^\infty(X) \geq k$ e seja E tal que $\mathbf{P}[X \in E] \geq p$. Então $H^\infty(X \mid X \in E) \geq k - \log(1/p)$.*

Demonstração. Temos, para todo $x \in E$,

$$\begin{aligned}\mathbf{P}[X = x \mid X \in E] &= \mathbf{P}[X = x, X \in E]/\mathbf{P}[X \in E] \\ &= \mathbf{P}[X = x]/\mathbf{P}[X \in E] \leq 2^{-k}/p.\end{aligned}$$

Como $\mathbf{P}[X = x \mid X \in E] = 0$ para todo $x \notin E$, o lema segue. □

Lema 8.2.5. *Sejam X_1 e X_2 fontes aleatórias tais que $H^\infty(X_1) \geq k$ e X_2 tem r-bits. Então, com probabilidade pelo menos $1 - 2^{-r}$ sobre o espaço $x_2 \in_R X_2$,*

$$H^\infty(X_1 \mid X_2 = x_2) \geq k - 2r.$$

Demonstração. Evidentemente, só precisamos analisar o caso em que $k > 2r$. Para todo $x \in \text{supp}(X_1)$, temos

$$2^{-k} \geq \mathbf{P}[X_1 = x] = \sum_{x_2 \in \text{supp}(X_2)} \mathbf{P}[X_1 = x \mid X_2 = x_2]\,\mathbf{P}[X_2 = x_2].$$

Seja $S_x = \left\{x_2 \in \text{supp}(X_2) \,\middle|\, \mathbf{P}[X_1 = x \mid X_2 = x_2] > 2^{2r-k}\right\}$ para $x \in \text{supp}(X_1)$. Como

$$2^{-k} \geq \mathbf{P}[X_1 = x] \geq 2^{2r-k}\mathbf{P}[X_2 \in S_x],$$

segue que $\mathbf{P}[X_2 \in S_x] \leq 2^{-2r}$. Se $S_x \neq \emptyset$, para todo $x_2 \in S_x$, devemos ter $\mathbf{P}[X_2 = x_2] \leq 2^{-2r}$, donde concluímos que

$$\mathbf{P}\Big[X_2 \in \bigcup_{x \in \text{supp}(X_1)} S_x\Big] \leq 2^r \cdot 2^{-2r} = 2^{-r}.$$

Por definição, se $x_2 \notin \bigcup_{x \in \text{supp}(X_1)} S_x$ então $H^\infty(X_1 \mid X_2 = x_2) \geq k - 2r$, e o lema segue. □

8.3 Obtendo as Sub-fontes da Asserção 8.1.1

Começamos aplicando o corolário 8.2.2 às fontes independentes X e Y da entrada do dispersor. Sejam

$$\bar{X} \subseteq X \text{ e } \bar{Y} \subseteq Y, \tag{8.6}$$

subfontes que podem ser decompostas como combinações convexas

$$\bar{X} = \sum_j \alpha_i X_j \text{ e } \bar{Y} = \sum_k \beta_k Y_k, \tag{8.7}$$

onde cada X_j e Y_k satisfaz as propriedades do corolário 8.2.2. Seja $X' = X_j$ para algum uma das componentes de $\bar{X}$ que e i'_x um índice tal que os blocos $X_1'^{i'_x}, X_2'^{i'_x}, X_3'^{i'_x}$ satisfazem as propriedades do corolário 8.2.2. Da mesma forma, seja $Y' = Y_k$ para algum k e i'_y definido de forma análoga a i'_x. A partição $I' = (i'_x, i'_y)$ será assumida implicitamente, ou seja, $X'_j = X'^{I'}_j$ e $Y'_j = Y'^{I'}_j$ para $j = 1, 2, 3$.

Fixando o desafio. Lembre-se que $c_1 = \texttt{s\_ext}_k(x_3, y)$ e $c_2 = \texttt{s\_ext}_k(y_3, x)$. Vamos restringir nossa atenção a sub-fontes para as quais $c = c_1 c_2$ é constante. Em particular, estamos lidando com o par

de sub-fontes independentes (X', Y'). Como Y'_3 é constante, o valor de c_2 só depende X'. Logo, existe uma palavra c'_2 tal que $\mathbf{P}[\texttt{s\_ext}_k(Y'_3, X') = c'_2] \geq 2^{-|c'_2|} = 2^{-lk}$, que é constante. Seja

$$X'' = (X' \mid \texttt{s\_ext}_k(Y'_3, X') = c'_2). \tag{8.8}$$

Pelo lema 8.2.4, segue que $H^\infty(X'') \geq \delta n/2 - lk$, ou seja, perdemos apenas alguns bits de entropia. Repetimos o mesmo processo para Y' com algum c'_1 e obtemos Y'' com alta min-entropia. É simples verificar que as seguintes propriedades valem para as fontes X'' e Y'':

- $H^\infty(X''_2), H^\infty(Y''_2) \geq \frac{\delta^2 n}{60}$;
- $H^\infty(X''_1), H^\infty(Y''_1) \geq \frac{\delta n}{4}$;
- nas fontes X'', Y'', o valor do desafio é fixo.

(Note que a partição I' é omitida da notação nas fontes X'' e Y''.)

Encontrando um bloco com aleatoriedade suficiente. Aplicamos o corolário 8.2.3 na distribuição X'' dividindo-a em $t = 1/\delta^3$ blocos de tamanho $\delta^3 n$. Como $H^\infty(X'') \geq \delta n/4$, segue que há uma sub-fonte

$$X''' \subseteq X'' \tag{8.9}$$

de densidade pelo menos $1/(8t)$ e um índice $j_x \in [t]$ tal que $H^\infty(X'''[j_x]) \geq \delta^4 n/20$. Verifica-se na demonstração do corolário 8.2.3 e do lema 8.2.1 que $X''' = X'' \mid X'' \in E$ para algum E tal que $\mathbf{P}[X'' \in E] \geq 1/(8t)$, donde segue, pelo lema 8.2.4, que

- $H^\infty(X'''_2) > \frac{\delta^2 n}{70}$;
- $H^\infty(X'''_1) > \frac{\delta n}{8}$
- $H^\infty(X'''[j_x]) \geq \delta^4 n/20$
- nas fontes X''', Y''', o valor do desafio é fixo.

(Novamente estamos omitindo I' na notação.)

Fazemos o mesmo para Y'' de forma a obter Y''' e j_y com as mesmas características.

Respondendo o desafio. Relembrando a seção 8.1, para cada entrada $x = x_1 x_2 \dots x_n, y = y_1 \dots y_n$ são computados os chutes $\texttt{s\_ext}_{2lk}(\hat{x}, \hat{y})$ para todo $\hat{x}$ e $\hat{y}$ da forma $\hat{x} = x_{a+1} \dots x_{a+\delta^3 n}$ e $\hat{y} = y_{b+1} \dots y_{b+\delta^3 n}$ com a e b múltiplos de $\delta^3 n$. Para (X''', Y'''), o desafio é uma palavra binária fixada $c' = c'_1 c'_2$. Como $H^\infty(X'''[j_x]) \geq \delta^4 n/20$ e $H^\infty(Y'''[j_y]) \geq \delta^4 n/20$, temos que $\mathcal{R} = \texttt{s\_ext}_{2lk}(X'''[j_x], Y'''[j_y])$ é $2^{-\eta n}$-próximo de ser uniforme-em-algum-lugar. Sendo assim, existe uma variável aleatória J sobre $[l]$ (que depende de $(X'''[j_x], Y'''[j_y])$) tal que $\mathcal{R}_J$ é $2^{-\eta n}$-próximo de ser uniforme. Logo, o desafio c' é igual a $\mathcal{R}_J$ com probabilidade pelo menos $2^{-|c'|} - 2^{-\eta n} = 2^{-2lk} - 2^{-\eta n}$ (que é maior que uma constante positiva como 2^{-2lk-1}, por exemplo). Isso significa que, com probabilidade positiva, o desafio é respondido. Agora considere todos os possíveis valores $\hat{x}$ e $\hat{y}$ tais que

- $\texttt{s\_ext}_{2lk}(\hat{x}, \hat{y})$ contém o desafio c' como sub-bloco;
- $\mathbf{P}\left[X'''[j_x] = \hat{x}\right], \mathbf{P}\left[Y'''[j_y] = \hat{y}\right] > 2^{-2\delta^3 n}$.

Chamaremos um par $(\hat{x}, \hat{y})$ de *bom* se este satisfaz as duas condições acima.

Seja $B_X = \left\{\hat{x} \;\middle|\; \mathbf{P}\left[X'''[j_x] = \hat{x}\right] \leq 2^{-2\delta^3 n}\right\}$. É simples verificar que

$$\mathbf{P}\left[X'''[j_x] \in B_X\right] = \sum_{\hat{x} \in B_X} \mathbf{P}\left[X'''[j_x] = \hat{x}\right] \leq 2^{\delta^3 n} 2^{-2\delta^3 n} = 2^{-\delta^3 n}.$$

Sendo assim, a probabilidade da segunda condição falhar para um par $(\hat{x}, \hat{y}) \in_R (X'''[j_x], Y'''[j_y])$ é no máximo $2^{1-\delta^3 n}$. Pelo que vimos no parágrafo anterior, a probabilidade de satisfazer a primeira condição é maior que uma constante positiva. Sendo assim, a probabilidade de um par ser bom é uma constante positiva.

Para cada par bom $(\hat{x}, \hat{y})$ definimos sub-fontes

$$\tilde{X} = (X''' \mid X'''[j_x] = \hat{x}) \text{ e } \tilde{Y} = (Y''' \mid X'''[j_y] = \hat{y}). \tag{8.10}$$

Pelo lema 8.2.4 e a segunda condição dos pares bons, concluímos que a entropia de cada bloco de $\tilde{X}$ e $\tilde{Y}$ é no mínimo $2\delta^3 n$ menor que a do mesmo bloco em X''' e Y'''. Juntando isso ao fato que em $(\tilde{X}, \tilde{Y})$ o desafio é sempre respondido, obtemos, para $\tilde{I} = I'$,

- $H^\infty(\tilde{X}_1), H^\infty(\tilde{Y}_1) \geq \frac{\delta n}{100}$;
- $H^\infty(\tilde{X}_2), H^\infty(\tilde{Y}_2) \geq \frac{\delta^2 n}{100}$;
- $\texttt{teste}_{\tilde{I}}(\tilde{X}, \tilde{Y}) = 1$ com probabilidade 1.

Observe que essas são as condições 1–3 da asserção 8.1.2. A condição 4 será verificada a seguir.

Evitando partições "incorretas". Fixe sub-fontes $\tilde{X}$ e $\tilde{Y}$ como acima e sua partição associada $\tilde{I}$. Seja I uma partição tal que $I \not\prec \tilde{I}$. Queremos mostrar que $\texttt{teste}_I(\tilde{X}, \tilde{Y}) = 0$ com probabilidade próxima a 1. Sejam $I = (i_x, i_y)$ e $\tilde{I} = (\tilde{i}_x, \tilde{i}_y)$. Suponha que $i_x > \tilde{i}_x$ (o caso em que $i_y > \tilde{i}_y$ é análogo). Observe que o bloco $\tilde{X}_3^I = \tilde{X}[i_x - 1]\tilde{X}[i_x - 2] \cdots \tilde{X}[1]$ contém o bloco $\tilde{X}_2^{\tilde{I}} = \tilde{X}[\tilde{i}_x]$. Portanto, $H^\infty(\tilde{X}_3^I) \geq H^\infty(\tilde{X}_2^{\tilde{I}}) \geq \delta^2 n/100$.

Fixe a e b múltiplos de $\delta^3 n$ e seja $\hat{X} = \tilde{X}_{a+1} \cdots \tilde{X}_{b+\delta^3 n}$ e $\hat{Y} = \tilde{Y}_{b+1} \cdots \tilde{Y}_{b+\delta^3 n}$. Vamos mostrar que, com alta probabilidade, o chute $c_{\hat{X},\hat{Y}}$ não responde ao desafio.

Aplicando o lema 8.2.5 às variáveis $\tilde{X}_3^I$ e $\hat{X}$, verificamos que com probabilidade $1 - 2^{-\delta^3 n}$ na escolha $\hat{x} \in_R \hat{X}$, temos

$$H^\infty(\tilde{X}_3^I \mid \hat{X} = \hat{x}) \geq \delta^2 n/100 - 2\delta^3 n \geq \delta^2 n/200.$$

O mesmo vale para $\tilde{Y}_3^I$ e $\hat{Y}$. Para todo par $(\hat{x}, \hat{y})$ tal que $H^\infty(\tilde{X}_3^I \mid \hat{X} = \hat{x}), H^\infty(\tilde{Y}_3^I \mid \hat{Y} = \hat{y}) \geq \delta^2 n/200$, considere a distribuição

$$\mathcal{D} = \left((\tilde{X}_3^I, \tilde{Y}) \mid \hat{X} = \hat{x}, \hat{Y} = \hat{y}\right).$$

Segue que os dois blocos de $\mathcal{D}$ são independentes e tem alta min-entropia. Logo, a saída de $\texttt{s\_ext}_k(\mathcal{D})$ é $2^{-\eta n}$-próxima de uniforme-em-algum-lugar. Isso significa que existe uma variável aleatória J sobre $[l]$ (que depende de $\mathcal{D}$) tal que, ao calcularmos $c_1 = \texttt{s\_ext}_k(\tilde{x}, \tilde{y})$ com $(\tilde{x}, \tilde{y}) \in_R \mathcal{D}$, temos $(c_1)_J$ é $2^{-\eta n}$-próximo de uniforme. Observe que o chute $c_{\hat{X},\hat{Y}}$ está fixado na distribuição $\mathcal{D}$. Considere a subdivisão de $c_{\hat{X},\hat{Y}}$ em blocos de tamanho k. Evidentemente, o desafio é respondido pelo chute $c_{\hat{X},\hat{Y}}$ somente se $(c_1)_J$ aparece como um dos blocos da subdivisão. Como há $2l^2$ blocos na subdivisão, a probabilidade de $(c_1)_J$ assumir um dos $2l^2$ valores desses blocos é no máximo $2l^2 \cdot$

$2^{-k} + 2^{-\eta n}$. Considerando ainda o caso em que a escolha $x \in_R \hat{X}$ não foi boa, a probabilidade do desafio ser respondido no espaço $(\hat{x}, \hat{y}) \in_R (\hat{X}, \hat{Y})$ é no máximo $2l^2 \cdot 2^{-k} + 2^{-\eta n} + 2^{-\delta^3 n} < 3l^2 \cdot 2^{-k}$, para n suficientemente grande.

Estamos livres para escolher a constante k, logo, podemos escolher um valor de forma que a cota da união para todas as possíveis escolhas de a, b e toda partição I dos eventos em que o desafio é respondido é $\leq 2^{-2m}$. Com isso garantimos que com probabilidade $\geq 1 - 2^{-2m}$, todos os desafios *não* são respondidos simultaneamente.

8.4 Prova da Asserção 8.1.1

Uma esquematização do processo de obtenção das fontes descritas na subseção anterior é dada a seguir. No diagrama, abreviamos s.f.d. α = "sub-fonte de densidade $\geq \alpha$" e c.c. = "combinação convexa"

$$\begin{aligned} X \xrightarrow[(8.6)]{\text{s.f.d. } 1/2} \bar{X} \xrightarrow[(8.7)]{\text{c.c.}} X' \xrightarrow[(8.8)]{\text{s.f.d. } 2^{-lk}} X'' \\ \xrightarrow[(8.9)]{\text{s.f.d. } \delta^3/8} X''' \xrightarrow[(8.10)]{\text{s.f.d. } 2^{-2lk-2}} \sum_{\hat{x} \text{ bom}} \alpha_{\hat{x}} \tilde{X}_{\hat{x}} \xrightarrow{\text{c.c.}} \tilde{X}. \end{aligned} \tag{8.11}$$

Suponha que $\bar{X} = \sum_{j=1}^{r} \alpha_j X'_j$, onde cada X'_j possui as mesmas propriedades de X' no diagrama. Para cada X'_j há uma fonte X'''_j associada que possui uma sub-fonte da forma $\sum_{k=1}^{s_j} \beta_{j,k} \tilde{X}_{j,k}$. Defina

$$\underline{X} = \sum_{j=1}^{r} \sum_{k=1}^{s_j} \alpha_j \beta_{j,k} \tilde{X}_{j,k}. \tag{8.12}$$

Pela proposição 2.2.12, temos que $\underline{X} \subseteq X$ tem densidade pelo menos

$$\frac{1}{2} \cdot 2^{-lk} \cdot \frac{\delta^3}{8} \cdot 2^{-2lk-2} = 2^{-3lk-6} \delta^3.$$

Definimos $\underline{Y} \subseteq Y$ de forma análoga. Observe que $\underline{X}$ e $\underline{Y}$ são combinações convexas de fontes satisfazendo as propriedades enunciadas na asserção 8.1.1.

8.5 Prova da Asserção 8.1.2

Lembre-se que

$$\texttt{disp}_I(x, y) = \texttt{opt}\Big(\texttt{s\_ext}_{d/l}(x_1, y_2), \texttt{s\_ext}_{d/l}(y_1, x_2)\Big).$$

Nosso objetivo é mostrar que `opt` é aplicado a duas fontes independentes com requerimento de min-entropia adequado.

Aplicando o lema 8.2.5 a $\tilde{X}_1$ e $\tilde{X}_2$ concluímos que, com probabilidade pelo menos $1 - 2^{-\delta^2 n/100}$ na escolha de $x_2 \in_R \tilde{X}_2$, temos

$$H^\infty(\tilde{X}_1 \mid \tilde{X}_2 = x_2) \geq \delta n/100 - \delta^2 n/50 > \delta n/200. \tag{8.13}$$

Diremos que x_2 é *útil* se satisfaz a equação (8.13). Evidentemente, podemos fazer o mesmo para $\tilde{Y}_1$ e $\tilde{Y}_2$. Dado que $\texttt{s\_ext}_{d/l}$ é um extrator-em-algum-lugar forte com erro $\varepsilon = 2^{-\eta n}$ (definição 6.0.12), verifica-se que para todo x_2 útil,

$$\mathbf{P}[\tilde{Y}_2 \text{ é bom para } (\tilde{X}_1 \mid \tilde{X}_2 = x_2)] \geq 1 - 2^{-\eta n}.$$

É simples observar que também vale, para todo y_2 útil,

$$\mathbf{P}[\tilde{X}_2 \text{ é bom para } (\tilde{Y}_1 \mid \tilde{Y}_2 = y_2)] \geq 1 - 2^{-\eta n}.$$

Segue que, com probabilidade pelo menos $1 - 2 \cdot 2^{-\eta n} - 2 \cdot 2^{-\delta^2 n/100}$, a escolha $(x_2, y_2) \in_R (\tilde{X}_2, \tilde{Y}_2)$ é um *bom casamento*, ou seja, o par é tal que x_2 é bom para $(\tilde{Y}_1 \mid \tilde{Y}_2 = y_2)$ e y_2 é bom para $(\tilde{X}_1 \mid \tilde{X}_2 = x_2)$. Isso significa que, para todo bom casamento (x_2, y_2) fixado, $\mathcal{D}_x = \texttt{s\_ext}\Big((\tilde{X}_1 \mid \tilde{X}_2 = x_2), y_2\Big)$ é $2^{-\eta n}$-próximo de uma distribuição uniforme-em-algum-lugar com l blocos de tamanho d/l. O mesmo vale para $\mathcal{D}_y = \texttt{s\_ext}\Big((\tilde{Y}_1 \mid \tilde{Y}_2 = y_2), x_2\Big)$.

Pelo fato 1, segue que $\mathcal{D}_x$ (e $\mathcal{D}_y$) tem min-entropia $d/l - \log l$ e este valor é pelo menos $10 \log d$ pela nossa escolha de d. Portanto, $\texttt{opt}(\mathcal{D}_x, \mathcal{D}_y)$ é $1/d$-próximo de uniforme. Como exigimos $\log d \geq 10m$, temos que $1/d \leq 2^{-10m}$. Segue da proposição 2.2.7 que $\texttt{disp}_I(\tilde{X}, \tilde{Y})$ é 2^{-2m}-próximo de uniforme, pois $2^{-10m} + 2 \cdot 2^{-\eta n} + 2 \cdot 2^{-\delta^2 n/100} \leq 2^{-2m}$.

8.6 Construções Explícitas de Grafos de Ramsey

Um dispersor $\texttt{disp}\colon \{0,1\}^{n \times 2} \to \{0,1\}$ pode ser encarado como um grafo bipartido G da seguinte forma. Sejam $N = 2^n$ e $M = 2^{\delta n}$ para algum $\delta > 0$. As classes do grafo G serão U e W, consistindo de cópias de $\{0,1\}^n$. Dado $(u, w) \in U \times W$, temos a aresta $uw \in E(G)$ se e somente se $\texttt{disp}(u, w) = 1$. É simples verificar que se $\texttt{disp}$ é um dispersor com requerimento de min-entropia δn então G não possui nenhum subgrafo bipartido completo $K_{M,M}$ nem um subgrafo bipartido (induzido) vazio $\overline{K}_{M,M}$.

De fato, se houvesse um desses subgrafos então poderíamos tomar conjuntos $U' \subseteq U$ e $W' \subseteq W$ tais que $|U'|, |W'| \geq M$ e $G[U', W']$ é *homogêneo* (ou seja, é completo ou vazio). Considere então as δ-fontes planas sobre U' e W' respectivamente. Pelas hipóteses sobre o dispersor, devemos ter

$$\mathbf{P}_{u \in_R U', w \in_R W'}[\texttt{disp}(u, w) = i] > 0,$$

para $i = 0, 1$. Mas isso contradiz o fato de que $G[U', W']$ é homogêneo.

Os parágrafos anteriores talvez não façam jus à dificuldade do problema de obter construções *explícitas* de grafos de Ramsey. Por explícita, entendemos que há um algoritmo polinomial (em $n = \log N$) que, ao receber um par de vértices, indica se há ou não aresta ligando os dois. Na versão de *coloração*, o grafo bipartido é sempre completo e o que interessa é a cor dada a uma aresta (é evidente que a primeira versão é idêntica a versão bi-colorida). Na versão de coloração, desejamos que para qualquer subgrafo $G' = G[U, W]$ com $|U|, |W|$ razoavelmente grandes, *todas* as cores apareçam em G'. A tradução de dispersores com saída em $\{0,1\}^m$ para grafos de Ramsey 2^m-coloridos é análoga a descrita acima.

Diremos que G é um grafo (N, K)-*Ramsey* se G tem N vértices e para qualquer $G[S]$ com $S \subseteq V(G)$ e $|S| \geq K$, o grafo $G[S]$ não é homogêneo. O caso bipartido é de fato mais difícil do que o não bipartido. A redução é simples: suponha que H é um grafo bipartido $N \times N$ que não

contém $K_{M,M}$ ou $\overline{K}_{M,M}$; forme o grafo G de N vértices com uma ordem total em seus vértices. Para todo $x < y \in V(G)$, coloca-se a aresta $\{x, y\} \in E(G)$ se e somente se $(x, y) \in E(H)$. O grafo G não possui conjuntos homogêneos de tamanho $2M$. De fato, seja $S \subseteq V(G)$ com $|S| = 2M$ e $G' = G[S]$. Podemos definir uma partição $S = S_1 \cup S_2$ na qual $\max(S_1) < \min(S_2)$ e $|S_1| = |S_2| = M$. O grafo $H' = H[S_1, S_2]$ define completamente as arestas entre S_1 e S_2 em G': como H' não é homogêneo, temos que G' também não é.

O recorde de melhor cota construtiva para grafos de Ramsey bipartidos devida a Frankl e Wilson [FW81] ($K = \exp\left\{\sqrt{\log N \log\log N}\right\}$ para o caso não-bipartido e $K = \sqrt{N}$ para o caso bipartido) perdurou por 25 anos, sendo quebrada por Pudlák e Rödl [PR04] ($K = o(\sqrt{N})$). As construções discutidas aqui, mostram cotas construtivas com $K = O(N^{\delta})$ para qualquer constante $\delta > 0$, ou seja, $K = 2^{o(n)}$. Os dispersores mencionados na seção 8.7 vão ainda mais longe e obtêm $K = 2^{n^{o(1)}}$.

Em contraste, uma das primeiras aplicações do Método Probabilístico, introduzido por Erdős [Erd47], demonstra que quase todo grafo aleatório é $\big(N, O(\log N)\big)$-Ramsey (e a mesma cota de $O(\log N)$ vale para grafos bipartidos). Essa disparidade entre o que se consegue fazer não-construtivamente através do método probabilístico e o que se sabe construir explicitamente é tema recorrente: de fato, todas as construções obtidas até o momento são trivialmente obtidas através do método probabilístico e, em geral[1], os parâmetros alcançados são substancialmente melhores do que os das construções.

8.7 Dispersores para Min-entropia $n^{o(1)}$

Em um trabalho recente [BRSW06], os dispersores obtidos neste capítulo puderam ser melhorados, passando a necessitar de muito menos min-entropia. O mecanismo desafio-resposta continua sendo chave nessa nova construção, porém ela é substancialmente mais envolvida e utiliza construções de extratores recentes de [Raz04, Bou05, Rao05]. O enunciado do teorema segue.

Teorema 8.7.1. *Existe dispersor computável em tempo polinomial* $D\colon \{0,1\}^{n\times 2} \to \{0,1\}$ *que funciona para qualquer par de fontes independentes com min-entropia* $n^{o(1)}$.

[1] Talvez a única exceção aqui sejam os extratores com semente (veja o capítulo 9), que atingem assintoticamente os parâmetros do extrator ótimo (que por sinal, pode ser obtido probabilisticamente).

Capítulo 9

Extratores com Semente

Historicamente, os primeiros extratores estudados foram aqueles que utilizavam uma pequena quantidade de bits genuinamente aleatórios (isto é uma amostra de bits que seguem uma distribuição uniforme) para extrair aleatoriedade de δ-fontes. Shaltiel [Sha02] escreveu um interessante *survey* sobre o assunto.

O estado da arte na construção de extratores com semente [LRVW03, GUV06] atingiu seu objetivo, de obter uma construção explícita que atinge (assintoticamente) os parâmetros ótimos estabelecidos em [RTS00]. Não discutiremos tais construções neste capítulo.

O conteúdo desenvolvido nessa seção será utilizado na seção seguinte, onde provamos uma melhora substancial dos teoremas 7.0.1 e 8.0.2 que exibem construções explícitas de extratores e dispersores.

Definição 9.0.2 (Extrator com semente)**.** *Uma função*

$$E\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$$

é chamada de um (k,ε)-extrator com semente de tamanho d e saída de m-bits *se, para toda fonte X com n-bits e min-entropia k, a saída de $E(X,U_d)$ é ε-próxima de uniforme (onde U_d é independente de X).*

Diremos que E é um (k,ε)*-extrator* forte *se, para qualquer fonte X de min-entropia k, tivermos*

$$\mathbf{P}_{s\in_R\{0,1\}^d}[E(X,s) \text{ é } \varepsilon\text{-próximo de uniforme}] > 1-\varepsilon.$$

O estudo de extratores com semente (um conceito relacionado a teoria da informação) está, surpreendentemente, fortemente relacionado a existência de geradores pseudoaleatórios (um conceito relacionado a computação).

É importante frisar que todo este trabalho tem como finalidade, a construção *explícita* de extratores de aleatoriedade. Para considerarmos um extrator explícito, basta que este seja um algoritmo polinomialmente computável.[1]

9.1 *Leftover-hash-lemma*

Uma construção simples baseada em técnicas de *hashing* nos permite a construção de extratores, como veremos a seguir. Primeiro, vamos relembrar alguns conceitos básicos sobre funções de hash. Uma visão mais detalhada deste tópico pode ser consultada em [CSRL01].

[1] Alguns autores fazem distinção de extratores (fracamente explícitos) dados por algoritmos polinomiais e aqueles (fortemente explícitos) dados através de uma fórmula algébrica fechada.

Definição 9.1.1. *Uma família $\mathcal{H} = \{h_a\}_{a \in A}$ de funções $h_a \colon \{0,1\}^n \to \{0,1\}^m$ é uma* família de hash universal *se para todos $x, x' \in \{0,1\}^n$ distintos e todos $y, y' \in \{0,1\}^m$ fixados, temos*

$$\mathbf{P}_{a \in_R A}\Big[\big(h_a(x), h_a(x')\big) = (y, y')\Big] = 2^{-2m}. \tag{9.1}$$

Como exemplo de uma família de hash universal simples, identifique palavras binárias de n-bits com elementos do corpo $\mathrm{GF}(2^n)$. Seja $A = \mathrm{GF}(2^n)^2$ e defina $h_{a,b}(x) = a \cdot x + b$ onde a aritmética é sobre o corpo $\mathrm{GF}(2^n)$. Para verificar que tal família é universal, observe que dados x e x' distintos, temos

$$\begin{bmatrix} h_{a,b}(x) \\ h_{a,b}(x') \end{bmatrix} = \begin{bmatrix} x & 1 \\ x' & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

Como a matriz 2×2 acima é não-singular, a transformação dada por ela é uma bijeção $(a, b) \leftrightarrow \big(h_{a,b}(x), h_{a,b}(x')\big)$. Segue imediatamente a propriedade (9.1).

Lema 9.1.2 (Leftover-hash-lemma)**.** *Seja $X \subseteq \{0,1\}^n$ com $|X| \geq 2^r$. Seja $k > 0$ e $\mathcal{H}$ uma família de hashing universal mapeando n bits a $r - 2k$ bits. Então, a distribuição $\big(h, h(x)\big)$ onde $h \in_R \mathcal{H}$ e $x \in_R X$ é 2^{-k}-próxima da distribuição uniforme em $\mathcal{H} \times \{0,1\}^{r-2k}$*

Demonstração. Seja A o conjunto que indexa a família $\mathcal{H}$. Defina $\mathcal{D}$ como a distribuição $\big(a, h_a(x)\big)$ onde $a \in_R A$ e $x \in_R X$. Vamos provar que

$$\mathrm{cp}(\mathcal{D}) \leq \frac{1 + 2^{-2k}}{2^{r-2k}|A|}.$$

Pelo lema 2.2.19, isso basta para esta demonstração.

Temos

$$\begin{aligned}
\mathrm{cp}(\mathcal{D}) &= \mathbf{P}_{a,a' \leftarrow_R A, x, x' \leftarrow_R S}[(a, h_a(x)) = (a', h_{a'}(x))] \\
&= \mathbf{P}_{a,a' \leftarrow_R A}[a = a']\, \mathbf{P}_{x,x' \leftarrow_R S}[h_a(x) = h_a(x')] \\
&= \frac{1}{|A|}\Big(\mathbf{P}_{x,x' \leftarrow_R S}[x = x'] + \mathbf{P}[h_a(x) = h_a(x') \mid x \neq x']\Big) \\
&= \frac{1}{|A|}\Big(2^{-r} + 2^{-r+2k}\Big) \\
&= \frac{1}{2^{r-2k}|A|}\Big(1 + 2^{-2k}\Big),
\end{aligned}$$

como queríamos. □

Observe que o *leftover hash lemma* exibe explicitamente um extrator de aleatoriedade. O problema de tal construção é que gastamos mais bits aleatórios do que efetivamente extraímos (isso é compensado pelo fato que todos os bits inicialmente "investidos" podem ser concatenados a saída do extrator). É interessante procurar obter construções cujo tamanho da semente seja consideravelmente menor do que o número de bits fornecidos pela fonte defeituosa. De fato, o Método Probabilístico nos garante a *existência* de (k, ε)-extratores cujas sementes têm tamanho $d = \log(n - k) + \log(1/\varepsilon) + O(1)$. O extrator obtido pelo método probabilístico tem parâmetros ótimos que atingem os limites inferiores assintóticos provados em [RTS00].

Diversos trabalhos surgiram desde o início dos anos 1990 com o objetivo de obter um extrator com parâmetros ótimos. Tal objetivo foi alcançado em [LRVW03], onde são obtidos extratores com parâmetros ótimos a menos de constante multiplicativa.

9.2 O extrator de Nisan-Zuckerman

Nisan e Zuckerman [NZ96] utilizaram uma idéia engenhosa para poderem utilizar o extrator obtido pelo *leftover hash lemma* na construção de um extrator com parâmetros melhores. De uma maneira grosseira, o extrator construído por eles utiliza uma parte da semente aleatória para obter blocos $X_1, X_2, \ldots, X_s$ da fonte original de forma que X_{i+1} tem uma certa min-entropia mesmo quando os valores de $X_1, \ldots, X_i$ são conhecidos. Chamaremos tal estrutura de blocos de uma *bloco-fonte* e postergamos sua definição formal.

Informalmente, o extrator de Nisan-Wigderson utiliza uma coleção de famílias de hash universal $\{\mathcal{H}_i\}_{i=0}^{s}$, com tamanhos de entrada e saída apropriados de forma que o extrator final é definido recursivamente da seguinte forma. Seja y a semente aleatória fornecida ao extrator e $x_1, \ldots, x_s$ os valores da bloco-fonte. Usamos y para escolher uma função de hash $h_s \in \mathcal{H}_s$ de maneira uniforme (cada família de hash é indexada por um conjunto de palavras binárias de tamanho apropriado). Definimos então $h_{i-1} = h_i(x_i) \circ h_i$ para todo $i \geq 1$. O extrator devolve o valor h_0 excluindo os bits correspondentes a h_s (ou seja, a cópia da semente). Observe que estamos abusando da notação e identificando funções de hash com seus índices (palavras binárias). É evidente que todos os parâmetros devem ser ajustados para que a definição indutiva seja válida. Um problema mais substancial é o de obter as bloco-fontes a partir de uma única δ-fonte.

A técnica pioneira de [NZ96] permite obter $(\delta n, \varepsilon)$-extratores com $\delta \geq 1/n$ e $\varepsilon \geq 2^{-\delta n}$ e semente de tamanho $O\left(\log \varepsilon^{-1} \log^2 n \frac{\log \delta^{-1}}{\delta}\right)$.

9.3 Extrator de Trevisan

Em um artigo revolucionário, Trevisan [Tre99] demonstrou uma conexão entre aleatoriedade sobre o ponto de vista *computacional* e aleatoriedade sobre o ponto de vista de *teoria da informação*.

Compare a seguinte definição de *distância computacional* com a definição 2.2.4 de distância estatística.

Definição 9.3.1. *Sejam X e Y distribuições sobre $\{0,1\}^n$. Diremos que X é* ε-indistinguível *da distribuição Y, se para todo algoritmo de decisão A, computável em tempo polinomial sobre n, temos*

$$\left|\mathbf{P}[A(X) = 1] - \mathbf{P}[A(Y) = 1]\right| \leq \varepsilon.$$

Observe que para todos X e Y que são ε-próximos (distância estatística) vale que X e Y são ε-indistinguíveis (distância computacional), ou seja, distância estatística é um conceito mais forte.

Definição 9.3.2. *Um* gerador de pseudoaleatoriedade $G\colon \{0,1\}^m \to \{0,1\}^n$ *de erro ε é uma função tal que $G(U_m)$ é ε-indistinguível de uniforme. (Evidentemente, só é interessante termos $m < n$.)*

Observe que, a partir de um gerador G que roda em tempo $q(m)$, qualquer algoritmo polinomial A com tempo de execução $p(l)$ sobre uma entrada de l bits e que dependa de $n = n(l) = \text{poly}(l)$ bits aleatórios pode ser simulado em tempo $2^m q(m)p(l)$. Basta executar o gerador para cada possível semente $s \in \{0,1\}^m$ e rodar o algoritmo A com os bits gerados. Tomando-se uma maioria de votos sobre a saída de A (ou seja, contamos para quantas sementes s a resposta de A é 0 ou 1 e decidimos pela maioria), obtemos um algoritmo determinístico que tem praticamente a mesma probabilidade de acerto que A.

Observe que não poderíamos trocar ε-indistinguível por ε-próximo na definição acima pois $\{0,1\}^n \setminus G(\{0,1\}^m)$ tem pelo menos $2^n - 2^m$ elementos e todos eles têm probabilidade 0 na distribuição $G(U_m)$, o que implica $\text{dist}\Big(G(U_m), U_n\Big) \geq (2^n - 2^m)2^{-n} = 1 - 2^{m-n}$.

Infelizmente todas as construções de geradores pseudoaleatórios dependem de hipóteses não provadas. De fato, as construções de geradores seguem o paradigma *dificuldade vs. aleatoriedade*. Em linhas gerais, esse paradigma é fundamentado no fato que a existência de funções *difíceis de ser computadas* implica a existência de geradores de pseudoaleatoriedade. Mais detalhes sobre esta relação na seção 9.5.

9.4 Códigos corretores de erros

O uso de códigos corretores de erro na construção de extratores é mais uma das inúmeras aplicações da bela Teoria de Códigos [MS77]. Vamos começar com algumas definições e um aquecimento.

Definição 9.4.1. *Seja Σ uma coleção finita de símbolos. Chamaremos Σ de* alfabeto. *Nesta dissertação, lidaremos com o caso $\Sigma = \{0,1\}$.*[*] *Uma* codificação *é uma função injetora $EC\colon \Sigma^n \to C \subseteq \Sigma^{\hat{n}}$. Chamamos de* palavra-código *um elemento do* código *C. Quando queremos comunicar uma palavra x de tamanho n computamos $y = EC(x)$ (portanto, desejamos que EC seja computável de maneira eficiente) e enviamos y através de um canal onde erros podem ser introduzidos. Do outro lado, uma palavra y' (que pode ser diferente de y) é recebida. A* distância (de Hamming) *entre y e y'—dada por $dH(y,y') = |\{i \mid y_i \neq y'_i\}|$—é o parâmetro usado para avaliar a quantidade de erros que o código permite corrigir. Dizemos que o código EC permite a correção de e erros se, para quaisquer palavras-código $z \neq z' \in C$, temos $dH(z,z') > e$. Uma* decodificação *para EC consiste de uma função (que desejamos ser computável eficientemente) que, para todo $z \in \Sigma^{\hat{n}}$, calcula $EC^{-1}(z')$ com $z' \in C$ sendo a palavra-código mais próxima de z.*

Uma construção explícita de códigos que servem para nosso propósito encontra-se no lema abaixo.

Lema 9.4.2. *Para todo $n \in \mathbb{N}$ e $\varepsilon > 0$ existe $EC = EC_{n,\varepsilon}\colon \{0,1\}^n \to \{0,1\}^{\hat{n}}$, onde $\hat{n} = \text{poly}(n, 1/\varepsilon)$, tal que toda bola de Hamming de raio $(1/2 - \varepsilon)\hat{n}$ contém no máximo $1/\varepsilon^2$ palavras-código. Ademais, EC pode ser computado em tempo $\text{poly}(n, 1/\varepsilon)$.*

Na literatura de códigos, a função EC do lema 9.4.2 é classificada como um $(\varepsilon, \varepsilon^{-2})$-*código lista-decodificável*, pois, para cada $p \in \{0,1\}^{\hat{n}}$, o tamanho da lista $L_p = \Big\{x \mid dH\Big(EC(x), p\Big) \leq (1/2 - \varepsilon)\hat{n}\Big\}$ é limitado pelo parâmetro ε^{-2}.

9.4.1 Um Aquecimento

Vamos definir um extrator com base no código do lema 9.4.2. Defina

$$E(x,y)\colon \{0,1\}^n \times \{0,1\}^{\log \hat{n}} \to \{0,1\}$$

como $E(x,y) = EC(x)_y$, ou seja, codificamos x e tomamos a y-ésima coordenada do código.

Seja $d = \log \hat{n}$. Provaremos que, para qualquer δ-fonte X, a distribuição $\mathcal{D} = E(X,y) \circ y$, com $y \in_R U_d$ é próxima da distribuição uniforme em U_{d+1}. Para cada $y \in \{0,1\}^{\log \hat{n}}$, defina o

[*] É comum Σ possuir estrutura algébrica, como por exemplo um corpo finito ou um anel de polinômios. No caso $\Sigma = \{0,1\}$, identificamos Σ com os elementos de GF(2).

profeta[2] $A(y) \in \{0,1\}$ como sendo a saída mais provável de $EC(X,y)$. Observe que a distância estatística está relacionada com a *vantagem* do profeta A:

$$[*] \equiv \mathbf{P}_{x\in_R X, y\in_R U_d}\Big[A(y) = E(x,y)\Big] = \frac{1}{2}(1 + \mathrm{dist}(\mathcal{D}, U_{d+1})). \tag{9.2}$$

Note também que nenhum profeta pode ter vantagem superior a de A; isso será importante para a análise abaixo.

Defina $\mathbf{a} = \Big(A(y)\Big)_y$ e $\mathbf{v}(x) = \Big(E(x,y)\Big)_y$. Observe que $\mathbf{a}, \mathbf{v}(x) \in \{0,1\}^{\hat{n}}$. Temos

$$\begin{aligned}[*] &= \sum_{x\in X} \mathbf{P}_y\Big[A(y) = E(x,y)\Big]\mathbf{P}\Big[X = x\Big] \\ &= \sum_{x\in X} \mathbf{P}[X = x]\Big\{1 - \frac{dH\Big(\mathbf{v}(x), \mathbf{a}\Big)}{\hat{n}}\Big\}.\end{aligned} \tag{9.3}$$

Seja $B \subseteq \mathrm{supp}(X)$ o conjunto dos x tais que $dH\Big(\mathbf{v}(x), \mathbf{a}\Big) \leq (1/2 - \varepsilon)\hat{n}$. Pelo lema 9.4.2 temos $|B| \leq 1/\varepsilon^2$. Como X é δ-fonte, segue que

$$\begin{aligned}[*] &\leq \sum_{x\in B} P[X = x] \times 1 + \sum_{x\in X\setminus B} P[X = x]\Big\{1 - \frac{dH\Big(\mathbf{v}(x), \mathbf{a}\Big)}{\hat{n}}\Big\} \\ &\leq 2^{-\delta n}\,|B| + \frac{1}{2} + \varepsilon \leq \frac{1}{2} + \varepsilon + \frac{2^{-\delta n}}{\varepsilon^2}.\end{aligned} \tag{9.4}$$

Pelas equações (9.2) e (9.4) temos $\mathrm{dist}(\mathcal{D}, U_{d+1}) \leq 2\varepsilon + 2^{1-\delta n}\varepsilon^{-2}$.

9.4.2 Paradigma da Reconstrução

Para a construção do extrator de Trevisan, usaremos de uma forma mais sofisticada as idéias contidas no aquecimento prévio.

Lema 9.4.3 ([Yao82])**.** *Suponha que (Y,Z) seja uma distribuição em $\{0,1\}^{d+m}$ tal que $Y \sim U_d$ porém $\mathrm{dist}\Big((Y,Z), U_{d+m}\Big) > \varepsilon$. Então existe $i \in [m]$ e uma função $A\colon \{0,1\}^d \times \{0,1\}^{i-1} \to \{0,1\}$ chamada de* profeta do próximo-bit *satisfazendo*

$$\mathbf{P}_{(y,z)\in_R(Y,Z)}\Big[A(y, z_{1,2\ldots i-1}) = z_i\Big] > \frac{1}{2} + \frac{\varepsilon}{m}. \tag{9.5}$$

Demonstração. Primeiramente, observe que o melhor profeta toma $A(y, z_1, \ldots, z_{i-1})$ como sendo o elemento de maior probabilidade na distribuição $(Z_i \mid Y = y, Z_{1\ldots i-1} = z_{1\ldots i-1})$. A melhor vantagem possível para um profeta do i-ésimo bit é, portanto,

$$v^i(Y,Z) = \frac{1}{2}\sum_{y,z_1,\ldots,z_i} \mathbf{P}\Big[Y = y, Z_{1\ldots i-1} = z_{1\ldots i-1}\Big] \times \Big|\mathbf{P}\Big[Z_i = z_i \mid Y = y, Z_{1\ldots i-1} = z_{1\ldots i-1}\Big] - \frac{1}{2}\Big|.$$

[2] Um profeta prevê o valor de uma parte da saída de uma função dado acesso a semente usada e, possivelmente, alguma outra parte da saída. Observe que, ao contrário da definição 9.3.1, não exigimos para esta construção que o profeta tenha poder computacional limitado.

Vamos provar o lema mostrando que se (9.5) não vale para nenhum i então, para todo $j = 0, \ldots, m$, temos $\operatorname{dist}\big((Y, Z_{1\ldots j}), U_{d+j}\big) \leq j\varepsilon/m$. No caso $j = 0$, nada temos a fazer. Seja $j > 0$ e suponha que a hipótese valha para todo valor menor que j. Temos

$$\begin{aligned}
[*] &\equiv 2\operatorname{dist}\big((Y, Z_{1\ldots j}), U_{d+j}\big) \\
&= \sum_{y, z_1, \ldots, z_j} \left|\mathbf{P}\big[Y = y, Z_{1\ldots j} = z_{1\ldots j}\big] - 2^{-d-j}\right| \\
&\leq 2v^j(Y, Z) + \sum_{y, z_1, \ldots, z_j} \left|\mathbf{P}\big[Y = y, Z_{1\ldots j-1} = z_{1\ldots j-1}\big]/2 - 2^{-d-j}\right| \\
&= 2v^j(Y, Z) + \sum_{y, z_1, \ldots, z_{j-1}} \left|\mathbf{P}\big[Y = y, Z_{1\ldots j-1} = z_{1\ldots j-1}\big] - 2^{-d-(j-1)}\right| \\
&= 2v^j(Y, Z) + 2\operatorname{dist}\big((Y, Z_{1\ldots j-1}), U_{d+j-1}\big).
\end{aligned} \tag{9.6}$$

Segue que $\operatorname{dist}\big((Y, Z_{1\ldots j}), U_{d+j}\big) \leq j\varepsilon/m$, como queríamos. □

Também precisamos introduzir a noção de *designs*. Para esta demonstração, utilizaremos uma versão refinada, denominada design fraco (*weak design*), definida em [RRV99] com o propósito de melhorar a análise do extrator de Trevisan.

Definição 9.4.4. *Uma coleção de conjuntos $S_1, \ldots, S_m \subseteq [d]$ é um* (l, ρ)-design fraco *se valem, para todo i,*

$$\begin{aligned}
|S_i| &= l; \\
\sum_{j<i} 2^{|S_i \cap S_j|} &\leq \rho(m-1).
\end{aligned} \tag{9.7}$$

O extrator de Trevisan é definido em relação a um design fraco $\mathcal{S} = \{S_1, \ldots, S_m\}$ da seguinte forma. Dado $y \in \{0,1\}^d$ seja $y|_{S_i}$ a projeção de y nos índices em S_i (isso dá um vetor com l coordenadas). Dado $u \in \{0,1\}^n$, utilize o código $EC = EC_{n,\varepsilon/(2m)}$ do lema 9.4.2 para obter $\bar{u} = EC(u)$, e interprete $\bar{u}$ como uma função $\bar{u} \colon \{0,1\}^l \times \{0,1\}$ (devemos fixar $l = \log \hat{n}$ para tal fim). Defina

$$\mathtt{Ext}_{\mathcal{S}}(u, y) = \bar{u}(y|_{S_1}) \cdots \bar{u}(y|_{S_m}). \tag{9.8}$$

Observe que u é usado para escolher um "predicado" e y é usado para escolher m elementos onde o predicado é avaliado.

A principal tecnicalidade—e o que de fato define o paradigma introduzido por Trevisan—aparece nos lemas a seguir. Abstratamente, o paradigma da reconstrução consiste da seguinte técnica. Deseja-se provar que uma dada função $f \colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ é um extrator. Dado um profeta A, defina B_A como o conjunto de elementos de $\{0,1\}^n$ para o qual o profeta possui uma vantagem razoável, digamos, maior que $\varepsilon/(2m)$.

Prova-se que para todo profeta A, uma fração constante α dos elementos em B_A pode ser codificada com poucos bits (digamos, r) de forma que existe um algoritmo que, a partir desses bits e com acesso a um oráculo computando A, reconstrói exatamente o elemento (daí o nome *paradigma da reconstrução*).

Dada uma distribuição X, do lema 9.4.3, obtemos

$$\begin{aligned}
&\mathbf{P}_{x\in_R X, y\in_R U_d}\Big[A(y, f(x)_{1,\ldots,i-1}) = f(x)_i\Big] \\
&= \mathbf{P}_{x\in_R (X|B_A), y}\Big[A(y, f(x)_{1,\ldots,i-1}) = f(x)_i\Big]\mathbf{P}[X \in B_A] \\
&+ \mathbf{P}_{x\in_R (X|X\setminus B_A), y}\Big[A(y, f(x)_{1,\ldots,i-1}) = f(x)_i\Big]\mathbf{P}[X \notin B_A] \\
&\leq \mathbf{P}[X \in B_A] + \Big(\frac{1}{2} + \frac{\varepsilon}{2m}\Big)\mathbf{P}[X \notin B_A] \\
&\leq \frac{1}{2} + \frac{\varepsilon}{2m} + \max_A\Big\{\mathbf{P}\Big[X \in B_A\Big]\Big\}.
\end{aligned} \tag{9.9}$$

Do lema 9.4.3 e da equação (9.9), obtemos

$$\operatorname{dist}(f(X), U_m) \leq \varepsilon/2 + m \cdot \max_A\Big\{\mathbf{P}\Big[X \in B_A\Big]\Big\}. \tag{9.10}$$

Como estamos trabalhando com δ-fontes, a probabilidade de um conjunto $|B_A|$ é no máximo $2^{-\delta n}|B_A|$. Isso implica que o limite sobre a cardinalidade de B_A, obtido implicitamente através da codificação (ou seja, $\alpha|B_A| \leq 2^r$), é suficiente para provar que f é extrator.

Lema 9.4.5. *Fixe um design fraco $\mathcal{S} = \{S_1, \ldots, S_m\}$. Para todo i, existe uma família $\mathcal{F}_i$ de funções $f\colon \{0,1\}^l \to \{0,1\}^{d+i-1}$ (dependendo apenas de $\mathcal{S}$ e i) tal que*

1. *Para toda função $P\colon \{0,1\}^l \to \{0,1\}$ e todo profeta $A\colon \{0,1\}^{d+i-1} \to \{0,1\}$ existe uma função $f \in \mathcal{F}_i$ tal que*
$$\mathbf{P}_{x\in_R U_l}\Big[A\Big(f(x)\Big) = P(x)\Big] \geq \mathbf{P}_{y\in_R U_d}\Big[A\Big(y, P(y|_{S_1})\cdots P(y|_{S_{i-1}})\Big) = P(y|_{S_i})\Big];$$
2. $\log|\mathcal{F}_i| \leq d + \sum_{j<i} 2^{|S_i\cap S_j|}$.

Para reforçar os parágrafos informais acima, observe que a idéia do lema 9.4.5 é mostrar que existe uma maneira compacta de representar uma coleção de funções que serve de boas aproximações para funções que são previsíveis de acordo com algum profeta. Em particular, dado um elemento $u \in \{0,1\}^n$ para o qual existe um profeta A tal que $A\Big(y, \bar{u}(y|_{S_1}), \ldots, \bar{u}(y|_{S_{i-1}})\Big)$ acerta $\bar{u}(y|_{S_i})$ com probabilidade $\geq 1/2 + \varepsilon'$, existe $f \in \mathcal{F}_i$ tal que $A \circ f$ é uma boa aproximação para $\bar{u}$. Isso significa que o vetor $\Big(A(f(x))\Big)_{x\in\{0,1\}^l}$ tem pelo menos uma fração $\geq 1/2 + \varepsilon'$ de coordenadas iguais a $\bar{u}$.

Passo I: a codificação. Para provar que $\texttt{Ext}_{\mathcal{S}}$ é extrator usando o paradigma da reconstrução, primeiro definimos como codificar os elementos $u \in \{0,1\}^n$ para os quais o predicado $\bar{u}$ pode ser previsto com vantagem $\varepsilon/(2m)$ por algum profeta do i-ésimo bit A. Para o par (u, A) existe uma função $f \in \mathcal{F}_i$ tal que $A \circ f$ é boa aproximação de $\bar{u}$. Pela construção do código EC, sabemos que há no máximo $(2m/\varepsilon)^2$ palavras-código tão próximas de $A \circ f$. Suponha que $\bar{u}$ seja a k-ésima palavra-código (em ordem lexicográfica) na bola de raio $(1/2 + \varepsilon/(2m))\hat{n}$ e centro $A \circ f$. Codificamos u através do par (f, k). Pela equação (9.7), podemos representar f usando $\log m + d + \sum_{j<i} 2^{|S_i\cap S_j|} \leq \log m + d + \rho(m-1)$ bits: usamos $\log m$ bits para descrever o índice $i \in [m]$ e os demais bits indexam funções de $\mathcal{F}_i$. Já k pode ser representado usando-se $\leq 2\log(2m/\varepsilon)$ bits.

Passo II: a reconstrução. Dado o par (f,k), calculamos $A \circ f$,* enumeramos todas as palavras código que estejam a distância $\leq (1/2 + \varepsilon/(2m))\hat{n}$ de $A \circ f$ e tomamos a k-ésima delas (em ordem lexicográfica). Com isso obtemos $\bar{u}$. Usando o mapa inverso EC^{-1} determinamos u. Observe que A não faz parte da codificação. Em essência, usamos A como um oráculo para uma máquina de Turing que computa u a partir de (f,k).

Vamos agora terminar a demonstração de que $\texttt{Ext}$ é um extrator para δ-fontes com erro ε quando $\rho = \Big(\delta n - 3\log(2m/\varepsilon) - d\Big)/m$.† Suponha por contradição que existe uma δ-fonte X para o qual $\texttt{Ext}$ falha. Seja A um profeta do i-ésimo bit para $\texttt{Ext}(X, U_d)$. Pelo paradigma da reconstrução, o conjunto B_A dos x tais que A tem vantagem $\varepsilon/(2m)$ na previsão de $\texttt{Ext}(x, U_d)$ é limitado pelo tamanho de sua codificação, a saber,

$$\log|B_A| \leq 2\log(m/\varepsilon) + d + \rho(m-1) < \delta n - \log(2m/\varepsilon).$$

Segue que $\mathbf{P}[X \in B_A] \leq 2^{-\delta n}|B_A| < \varepsilon/(2m)$. Donde concluímos da equação (9.10) que a saída do extrator é ε-próxima de uniforme mesmo quando a semente usada é concatenada a saída.

9.5 Dificuldade vs. Aleatoriedade

Nesta seção discutiremos a interessante relação entre dificuldade computacional e pseudoaleatoriedade. Em termos grosseiros, se dispomos de uma[3] função $f\colon \{0,1\}^n \to \{0,1\}$ difícil de ser computada então podemos obter um gerador de pseudoaleatoriedade (veja a definição 9.3.2).

Uma importante descoberta em relação a PRGs (do inglês, *pseudorandom generator*) é devida a Nisan e Wigderson [NW94] que introduziram seu gerador. Dada uma função $f\colon \{0,1\}^l \to \{0,1\}$ e uma família de conjuntos $\mathcal{S} = \{S_1, \ldots, S_m\} \subseteq \binom{[d]}{l}$, definimos $NW^f_{\mathcal{S}}\colon \{0,1\}^d \to \{0,1\}^m$, tal que

$$NW^f_{\mathcal{S}}(x) = f(x|_{S_1}) \circ \cdots \circ f(x|_{S_m}). \tag{9.11}$$

Observe que o extrator (9.8) é definido em termos do gerador de Nisan Wigderson. De fato, temos $\texttt{Ext}_{\mathcal{S}}(u,y) = NW^{\bar{u}}_{\mathcal{S}}(y)$.

Diremos que a função $f = f_l\colon \{0,1\}^l \to \{0,1\}$ é $k(l)$-*difícil* se o menor *circuito* que computa f tem tamanho $k(l)$. A *complexidade de circuito* para f neste caso é $k = k(l)$. É interessante observar que há uma relação entre a complexidade de circuito de f e o tempo que uma máquina de Turing leva para computar f. Essa perspectiva de complexidade pareceu promissora na missão de provar que P $\neq$ NP, porém, até o momento os esforços nesse sentido não chegaram a uma solução. Observamos também que não se conhece nenhuma função explícita $f_l \in \mathrm{E} = \mathrm{DTIME}\Big(2^{O(l)}\Big)$ tal que $l = o\Big(k(l)\Big)$.[4]

O trabalho [NW94] foi melhorado ao longo dos anos por diversos autores, *e.g.* [IW97, STV98, ISW06]. Em particular, em [IW97] prova-se que se E necessita de circuitos exponenciais então

*Não impomos qualquer restrição sobre o tempo que tais algoritmos levam para computar a codificação/reconstrução.

†Evidentemente, deve-se verificar que tal valor de ρ é atingível por construções explícitas. Felizmente, isso é possível desaleatorizando uma construção probabilística (usando o Método das Esperanças Condicionais).

[3]Como é usual, abusamos da notação e exigimos, na verdade, uma família de funções $\Big\{f_n\colon \{0,1\}^n \to \{0,1\}\Big\}_{n \geq n_0}$ onde cada uma delas é difícil de ser computada.

[4]Isso significa que os PRGs exibidos dependem de funções explícitas cuja complexidade de circuito é muito maior do que o máximo conhecido.

existem bons PRGs, bons o suficiente para desaleatorizar qualquer algoritmo em BPP e, portanto, P = BPP em tal cenário. Onde BPP é a classe de problemas de decisão para os quais há um algoritmo polinomial probabilístico que acerta a resposta de toda instância com probabilidade $\geq 2/3$.

Capítulo 10

Melhorando as Construções: Extratores e Dispersores com Saída Linear

Recentemente, Anup Rao [Rao05] exibiu construções explícitas de extratores com parâmetros substancialmente melhores que os de [BKS+05]. Um fato interessante é que as técnicas usadas não empregam Teoria Aditiva dos Números. Inspirado em suas idéias e utilizando descobertas recentes por parte de Bourgain [Bou05], propomos uma alteração no método desenvolvido por Barak et al. [BKS+05] para obtermos extratores e dispersores cujo número de bits devolvidos é linear no tamanho da entrada.[1]

O leitor atento pode ter percebido que, em ambas as construções do extrator e do dispersor descritas acima, as δ-fontes são, numa primeira fase (a saída de `s_ext`), convertidas em fontes quebradas em blocos onde sabemos algo muito forte em termos estruturais; num segundo momento, no entanto, tal informação estrutural é praticamente descartada. De fato, sabemos que as fontes obtidas na primeira fase do tipo $Z = (Z_i)_{i=1}^{l} \subseteq \{0,1\}^{m \times l}$ e que existe uma função determinística $I\colon \subseteq \{0,1\}^{m \times l} \to [l]$ que só depende da distribuição de Z tal que $Z_{I(Z)}$ é uniforme. No entanto, tal estrutura é completamente ignorada, usando-se somente uma cota para a min-entropia de Z.

Primeiramente, vamos mostrar que podemos interpretar Z de uma forma ainda mais estruturada.

Lema 10.0.1. *Seja Z uma fonte consistindo de l blocos de m-bits cada para o qual existe uma função determinística $I\colon \subseteq \{0,1\}^{m \times l} \to [l]$ (um seletor) que só depende da distribuição de Z tal que $Z_{I(Z)}$ tem min-entropia k. Então Z é $l\,2^{1-a}$ próximo de ser combinação convexa de fontes $\{Z^i\}_{i=1}^{l}$ onde Z_i^i tem min-entropia pelo menos $k - a$.*

Demonstração. Por definição, para todo $z \in \{0,1\}^m$, temos

$$2^{-k} \geq \mathbf{P}[Z_{I(Z)} = z] = \sum_{j=1}^{l} \mathbf{P}[Z_j = z \mid I(Z) = j]\,\mathbf{P}[I(Z) = j].$$

Segue que, para todo j e todo z,

$$\mathbf{P}[Z_j = z \mid I(Z) = j] \leq 2^{-k} \Big/ \mathbf{P}[I(Z) = j].$$

[1] Soubemos através de Avi Wigderson que, independentemente, Rao e outros também observaram este fato. Observamos também que Raz [Raz04] obteve, de uma maneira diferente dos métodos aqui empregados, extratores de três fontes mais poderosos do que nossa técnica obteve (porém seu trabalho não parece se estender a dispersores).

Portanto, se $H^\infty(Z_j \mid I(Z) = j) < k - a$ então $\mathbf{P}[I(Z) = j] < 2^{-a}$. Seja $J \subseteq [l]$ o conjunto dos índices j para o qual $H^\infty(Z_j \mid I(Z) = j) < k - a$. Estamos interessados no caso em que $l\,2^{-a} < 1/2$ (caso contrário, o lema é trivial). Segue que J é não-vazio. Defina a distribuição

$$Z' = \mathbf{P}[I(Z) \in J]^{-1} \sum_{j \in J} \mathbf{P}[I(Z) = j]\,(Z \mid I(Z) = j).$$

É simples verificar que Z' é sub-fonte de Z com densidade $\mathbf{P}[I(Z) \in J] \geq l\,2^{-a}$. Disso segue que $\mathrm{dist}(Z, Z') \leq 2l\,2^{-a}$. Ademais, como cada fonte $Z^j = (Z \mid I(Z) = j)$ com $j \in J$, por construção é tal que $H^\infty(Z^j_j) \geq k - a$, concluímos a prova do lema. □

Diremos que uma fonte $X = (X_j)_{j=1}^l$ tem min-entropia k em algum bloco se existe um índice $j \in [l]$ tal que $H^\infty(X_j) \geq k$. Se $Y = (Y_j)_{j=1}^l$ é tal que existe $j \in [l]$ com $H^\infty(X_j), H^\infty(Y_j) \geq k$, dizemos que X e Y são alinhados.

Usando o lema acima podemos aplicar a saída da primeira fase (saída `s_ext`) ao seguinte condensador.

Teorema 10.0.2. *Sejam $a > b$ constantes positivas tais que $a + b < 1$. Suponha que X e Y sejam fontes independentes, alinhadas, com $l \leq m^b$ blocos de tamanho m cada, com razão de min-entropia $1 - m^{-a}$ em algum bloco. Existe uma função polinomialmente computável*

$$\mathtt{cond}\colon \{0,1\}^{l \times m} \times \colon \{0,1\}^{l \times m} \to \{0,1\}^{\lceil l/2 \rceil \times m'} \times \{0,1\}^{\lceil l/2 \rceil \times m'},$$

com $m' = m - \Theta(m^{1+b-a})$, tal que $\mathtt{cond}(X, Y)$ é ε-próximo de ser combinação convexa de fontes independentes alinhadas com razão de min-entropia $1 - m'^{-a}$ em algum bloco, onde $\varepsilon = 2^{-m^{\Omega(1)}}$.

Para demonstrarmos o teorema 10.0.2 precisaremos das seguintes construções recentes de extratores.[2]

Teorema 10.0.3 (Extrator de Bourgain [Bou05])**.** *Existem uma função explícita polinomialmente computável*

$$\mathtt{Bou}\colon \{0,1\}^{n \times 2} \to \{0,1\}^m,$$

com $m = \Omega(n)$, e uma constante universal $\alpha_0 > 0$ tais que se X e Y são $(1/2 - \alpha_0)$-fontes independentes, então, com probabilidade $1 - 2^{-\Omega(n)}$ sobre a escolha de $y \in_R Y$, vale

$$\mathrm{dist}\Big(\mathtt{Bou}(X, y), U_m\Big) \leq 2^{-\Omega(n)}.$$

(O mesmo vale para $x \in_R X$ e $\mathtt{Bou}(x, Y)$.)

Teorema 10.0.4. *Para todo $n \in \mathbb{N}$, $\delta > 0$ constante e $\varepsilon = \varepsilon(\delta, n) > 0$ existe um extrator-com-semente computável em tempo* $\mathrm{poly}(n)$

$$\mathtt{Ext}\colon \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{\delta n - O(d)},$$

com $d = O\Big(\log^3(n/\varepsilon)\Big)$, que funciona para qualquer δ-fonte e cuja saída é ε-próxima de uniforme. Ademais, `Ext` é extrator forte.

[2] Bourgain obteve o primeiro extrator para duas fontes independentes cuja entropia é menor do que 1/2. Conjectura-se que grafos de Paley (cujas arestas são determinadas por $(x, y) \in \mathbb{F}_q^2$ tais que $\chi(x - y) = 1$, onde χ é o caractere quadrático de $\mathbb{F}_q$ e $q \equiv 1 \pmod 4$) fornecem extratores para fontes de n-bits com entropia $\mathrm{polylog}(n)$, porém, provar tal conjectura parece ser extremamente difícil.

Também precisamos do seguinte lema.

Lema 10.0.5. *Seja X uma fonte com n-bits indexados por $[n]$ tal que $H^\infty(X) \geq k$. Dado $S \subseteq [n]$ de cardinalidade s, a fonte X_S (isto é, a projeção de X nas coordenadas de S) satisfaz*

$$H^\infty(X_S) \geq k + s - n.$$

Em termos da razão de min-entropia, o resultado é $r(X_S) = H^\infty(X_S)/s \geq 1 + (k-n)/s$.

Demonstração. Seja $x^* \in \mathrm{supp}(X_S)$ qualquer. Seja x o elemento da distribuição $(X \mid X_S = x^*)$ de maior probabilidade. Como $|\mathrm{supp}(X \mid X_S = x^*)| \leq 2^{n-s}$, claramente temos $\mathbf{P}[X = x \mid X_S = x^*] \geq 2^{s-n}$. Segue que

$$2^{-k} \geq \mathbf{P}[X = x] = \mathbf{P}[X = x \mid X_S = x^*]\,\mathbf{P}[X_S = x^*]$$

e, portanto, $\mathbf{P}[X_S = x^*] \leq 2^{-k-s+n}$. □

Com os resultados acima podemos provar o teorema 10.0.2.

Prova do Teorema 10.0.2. Suponha sem perda de generalidade que l é par (se não for, adicione um bloco arbitrário a entrada). Para $i = 1, \ldots, l/2$ denote por X_i os dois blocos X^{2i-1} e X^{2i} de X, onde os bits desses dois blocos estão intercalados (por exemplo, todo índice ímpar de X_i corresponde a um bit em X^{2i-1} e todo índice par corresponde a um bit em X^{2i}). Defina Y_i's da mesma maneira.

Suponha que j é tal que ambos X_h e Y_h contêm um bloco com razão de min-entropia $1 - m^{-a}$. Denote por $X_{h,1}$ e $Y_{h,1}$ os primeiros $s \equiv 2\alpha_0^{-1} m^{1-a}$ bits de X_h (onde α_0 é a constante do teorema 10.0.3). Como $X_{h,1}$ contém $s/2$ bits de uma fonte de m-bits com min-entropia pelo menos $m - m^{1-a}$, pelo lema 10.0.5, a razão de entropia de $X_{h,1}$ é pelo menos $(1 - \alpha_0)/2$. O mesmo vale para $Y_{h,1}$.

Denote por $\mathtt{Bou}$ o extrator de Bourgain com tamanho de entrada s e saída truncada a $d = m^{1-a-b}/2 \ll s$ bits (truncar não afeta a uniformidade da saída, vide proposição 2.2.5). Seja $\mathtt{Ext}$ o extrator-com-semente do teorema 10.0.4, com entrada de tamanho $2m$, requerimento de min-entropia $m - 4\alpha_0^{-1} m^{1+b-a}$, semente de tamanho d e erro ε_{ext}^2. Observe que podemos tomar $\varepsilon_{ext} = 2^{-m^{\Omega(1)}}$ e o extrator terá saída de tamanho $m' = m - \Theta(m^{1+b-a})$.

Seja $\overline{X}$ a distribuição $(X_{i,1})_{i=1}^{l/2}$ e $\overline{Y} = (Y_{i,1})_{i=1}^{l/2}$. Considere todas as maneiras de fixar $\overline{x} \in \mathrm{supp}(\overline{X})$ tais que

(a) $H^\infty(X_h \mid \overline{X} = \overline{x}) \geq m - m^{1-a}(2\alpha_0^{-1} l + 1) \geq m - 4\alpha_0^{-1} m^{1+b-a}$;

(b) $\mathtt{Bou}(\overline{x}_h, Y_{h,1})$ é $2^{-\Omega(s)}$-próximo de uniforme.

Pelo lema 8.2.5, a probabilidade da condição (a) falhar é no máximo $2^{-sl/2}$. Como $\mathtt{Bou}$ é forte e $r(X_{h,1}), r(Y_{h,1}) \geq 1/2 - \alpha_0$, a probabilidade de (b) falhar é no máximo $2^{-\Omega(s)}$. A cota da união nos garante que as maneiras boas de fixar $\overline{X}$ têm probabilidade pelo menos $1 - 2^{-\Omega(s)}$. Em particular, X está a distância $2^{-\Omega(s)}$ de ser combinação convexa de fontes X' tais que $\overline{X}'$ é constante e as condições (a) e (b) valem para X'.

Seja $X' = (X \mid \overline{X} = \overline{x})$ uma dessas fontes e considere todas as maneiras de fixar $\overline{y} \in \mathrm{supp}(Y)$ tais que

(c) $H^\infty(Y_h \mid \overline{Y} = \overline{y}) \geq m - 4\alpha_0^{-1} m^{1+b-a}$;

(d) $z = \mathtt{Bou}(\overline{x}'_h, \overline{y}_h)$ é uma boa semente com relação a $\mathtt{Ext}\left(X'_h, z\right)$.

A condição (c) é similar a (a) e o mesmo raciocínio usado acima se aplica. Por construção, temos $H^{\infty}(X'_h) \geq m - 4\alpha_0^{-1}m^{1+b-a}$ e $\texttt{Bou}(\overline{x}', Y_{h,1})$ é $2^{-\Omega(s)}$-próximo de uniforme, o que implica que a probabilidade de z ser uma semente ruim é no máximo $2^{-\Omega(s)} + \varepsilon_{ext}$.

Segue que Y está a distância $2^{-n^{\Omega(1)}}$ de ser uma combinação convexa de fontes Y' com $\overline{Y}'$ constante (para algum valor $\overline{y}'$) tal que

$$Z_h = \texttt{Ext}\Big(X'_h, \texttt{Bou}(\overline{x}'_h, \overline{y}'_h)\Big)$$

é ε_{ext}-próximo de uniforme.

Defina Z_i, para todo i, da mesma forma que Z_h foi definido. A distribuição $Z = (Z_i)_{i=1}^{l/2}$ é independente de Y' já que ela depende apenas de X'_h. Denote por $Z_{i,1}$ os primeiros d bits de Z_i. Pela proposição 2.2.5, segue que $Z_{h,1}$ é ε_{ext}-próximo de uniforme. Repetindo o argumento anterior, Z é $O(\varepsilon_{ext})$-próximo de uma combinação convexa de fontes Z', com $Z'_{i,1}$ fixado para todo i, tais que

(e) $Z'_{h,1}$ é boa semente em $\texttt{Ext}(Y'_h, Z'_{h,1})$;

(f) $Z'_h = (Z_h \mid Z'_{i,1} = \overline{z}'_i)$ tem min-entropia pelo menos $m' - dl = m' - m^{1-a}/2 \geq m' - m'^{1-a}$.

Seja

$$W = (W_i)_{i=1}^{l/2} = \Big(\texttt{Ext}(Y'_i, \overline{z}'_i)\Big)_{i=1}^{l/2}.$$

Observe que W não depende de Z' e que W_h é ε_{ext}-próximo de uniforme.

Defina `cond` como o algoritmo implicitamente definido acima. Mais especificamente, dada uma entrada (x, y), computamos

$$z = \Big(\texttt{Ext}(x_i, \texttt{Bou}(x_{i,1}, y_{i,1}))\Big)_{i=1}^{l/2}$$

e

$$w = \Big(\texttt{Ext}(y_i, z_{i,1})\Big)_{i=1}^{l/2},$$

onde $z_{i,1}$ denota os primeiros d bits de z_i. A saída do algoritmo é o par (z, w). □

Aplicando o teorema 10.0.2 recursivamente as fontes obtidas na primeira fase obtemos o seguinte resultado.

Teorema 10.0.6 (Melhora do Teorema 7.0.1)**.** *Para toda constante $\delta > 0$ existe uma função polinomialmente computável*

$$\texttt{3ext}\colon \{0,1\}^{n\times 3} \to \{0,1\}^m,$$

com $m = \Omega(n)$, tal que para quaisquer três δ-fontes independentes X, Y e Z temos que a saída de $\texttt{3ext}(X, Y, Z)$ é $2^{-n^{\Omega(1)}}$-próxima de uniforme.

O mesmo tipo de melhora pode ser obtido no caso de dispersores, também substituindo a função `opt` por aplicações sucessivas do condensador do Teorema 10.0.2.

Teorema 10.0.7 (Melhora do Teorema 8.0.2)**.** *Para toda constante $\delta > 0$ existe uma função polinomialmente computável*

$$\texttt{2disp}\colon \{0,1\}^{n\times 2} \to \{0,1\}^m,$$

com $m = \Omega(n)$, tal que para quaisquer duas δ-fontes independentes X e Y, vale que[3]

$$\operatorname{supp}\Big(\texttt{2disp}(X, Y)\Big) = \{0,1\}^m.$$

10.1 Condensadores assimétricos

Encerraremos a seção fazendo uma breve menção a outro resultado original desta pesquisa. Na linha do capítulo 5, podemos obter um condensador de semente constante cujos parâmetros são essencialmente ótimos.

Teorema 10.1.1. *Sejam $5/9 < \delta \leq 1$, $\alpha \in (0,1)$ e $\gamma \in [1,2]$ constantes tais que*

$$\delta > \frac{3+\gamma}{(\gamma+1)^2}, \text{ and } \alpha > \frac{1+\gamma}{2+\gamma}. \tag{10.1}$$

Sejam $n_1 = n/(1+\gamma)$ e $n_2 = n - n_1 = \gamma n_1$.[†] *Existe um algoritmo computável em tempo* $\operatorname{poly}(n)$

$$\texttt{acond}\colon \{0,1\}^n \to \{0,1\}^{n_1} \times \{0,1\}^{n_2} \times \{0,1\}^{n_1}$$

e $\varepsilon = 2^{-n^{\Omega(1)}}$ tais que $\texttt{acond}(X)$ é ε-próximo de ser combinação convexa de fontes com razão de min-entropia $\alpha \cdot \delta + (1-\alpha)\cdot 1$ em algum bloco, desde que X seja uma δ-fonte de n bits.

A assimetria do condensador acima pode ser desfeita utilizando-se o *merger* de [Raz04].

[3] Algumas modificações sutis devem ser feitas para demonstrar que a construção de fato funciona. Em particular, como $m \to \infty$, não podemos esperar que a probabilidade de cada elemento $z \in \{0,1\}^m$ seja limitada inferiormente por uma constante positiva como no teorema 8.0.2.

[†] Por simplicidade, vamos supor que n_1 e n_2 são inteiros.

Apêndice A

Construção Explícita de Had

A.1 Um pouco de álgebra linear

Lema A.1.1. *Suponha que existem matrizes $A_1, \dots, A_k$, de dimensões $n \times n$ (com $k \leq 0.1n$), tais que, para todo $\emptyset \neq S \subseteq [k]$, a matriz $\sum_{i \in S} A_i$ tem posto cheio. Defina a função* $\texttt{Had}\colon \{0,1\}^n \to \{0,1\}^k$ *como* $\texttt{Had}(x, y) = (x^T A_1 y, \dots, x^T A_k y)$. *Então* Had *é um extrator forte para qualquer par de fontes independentes cuja razão de entropia seja pelo menos* 0.6 *e sua saída é* $2^{-\Omega(n)}$*-próxima de uniforme.*

Demonstração. Por conveniência, denotaremos $p_x = \mathbf{P}[X = x]$ e $q_y = \mathbf{P}[Y = y]$. Fica subentendido que os índices são elementos de $\{0,1\}^n$. Primeiramente, observamos que

$$\operatorname{dist}(\texttt{Had}(X, Y), U_k) \leq \sum_y q_y \operatorname{dist}(\texttt{Had}(X, y), U_k).$$

A prova da desigualdade acima segue os moldes da prova da proposição 2.2.7.

Do lema 2.2.20 e da desigualdade de Jensen segue que

$$\begin{aligned} \sum_y q_y \operatorname{dist}(\texttt{Had}(X, y), U_k) &\leq \sum_y q_y \Big(\sum_{0 \neq v \in \mathbb{Z}_2^k} \operatorname{dist}\Big(\langle \texttt{Had}(X, y), v\rangle, U_1\Big)^2\Big)^{1/2} \\ &\leq \Big(\sum_y q_y \sum_{0 \neq v \in \mathbb{Z}_2^k} \operatorname{dist}\Big(\langle \texttt{Had}(X, y), v\rangle, U_1\Big)^2\Big)^{1/2} \end{aligned} \tag{A.1}$$

onde o produto interno é calculado $\bmod 2$.

Seja $0 \neq v \in \mathbb{Z}_2^k$ um vetor qualquer. Observe que $\langle \texttt{Had}(x, y), v\rangle = \sum_{i\colon v_i = 1} x^T A_i y$. Tome $B = B_v = \sum_{i\colon v_i = 1} A_i$, de forma que $\langle \texttt{Had}(x, y), v\rangle = x^T B y$. Para esta demonstração, defina a matriz $M = M_v$, cujas colunas e linhas são indexadas por $\{0,1\}^n$ e $M_{x,y} = (-1)^{x^T B y}$. Denotaremos por m_x a x-ésima linha de M.

Asserção A.1.2. *As linhas de M são duas-a-duas ortogonais (no sentido Euclidiano).*

A asserção segue diretamente do lema 2.2.21. De fato, se $x, x' \in \{0,1\}^n$ são elementos distintos, então $x + x' \neq \mathbf{0}$. Ademais, como B tem posto cheio, obtemos

$$\langle m_x, m_{x'}\rangle = \sum_y (-1)^{(x+x')^T B y} = \sum_z (-1)^{\langle x+x', z\rangle} = 0.$$

Seja $\mathbf{d} \in \mathbb{R}^{2^n}$ definido por $\mathbf{d} = \frac{1}{2}\sum_x p_x m_x$. Observe que

$$\begin{aligned} |d_y| &= \frac{1}{2}\left|\mathbf{P}_{x\in_R X}[M_{x,y} = 1] - \mathbf{P}_{x\in_R X}[M_{x,y} = -1]\right| \\ &= \operatorname{dist}(\langle \mathtt{Had}(X, y), v\rangle, U_1). \end{aligned}$$

Queremos então uma estimativa para $\sum_y q_y |d_y|^2$. Suponha que $H^\infty(X) = \delta_X n$ ($\geq 0.6n$) e $H^\infty(Y) = \delta_Y n$ ($\geq 0.6n$). Temos $\sum_y q_y |d_y|^2 \leq 2^{-\delta_Y n}\|\mathbf{d}\|_2^2$. Dada a ortogonalidade das linhas de M (asserção A.1.2), temos

$$\|\mathbf{d}\|_2^2 = \frac{1}{4}\sum_x p_x^2 \|m_x\|_2^2 = 2^{n-2}\sum_x p_x^2 \leq 2^{n-2}2^{-\delta_X n}\sum_x p_x = 2^{n-2-\delta_X n}.$$

Da desigualdade (A.1), obtemos

$$\begin{aligned} \sum_y q_y \operatorname{dist}(\mathtt{Had}(X, y), U_k) &\leq \Big(\sum_{0\neq v\in\mathbb{Z}_2^k}\sum_y q_y |d_y|^2\Big)^{1/2} \\ &\leq 2^{(k+n-2-\delta_X n-\delta_Y n)/2}. \end{aligned} \tag{A.2}$$

Seja $\alpha = (\delta_X n + \delta_Y n + 2 - k - n)/(2n)$ (observe que $\alpha > 0.05$). Vamos agora mostrar que Had é extrator forte. Seja

$$\mathcal{B} = \left\{y \in \{0,1\}^n \;\middle|\; \operatorname{dist}(\mathtt{Had}(X, y), U_k) > 2^{-\alpha n/2}\right\}.$$

Temos, pela desigualdade (A.2),

$$2^{-\alpha n} \geq \sum_y q_y \operatorname{dist}(\mathtt{Had}(X, y), U_k) > 2^{-\alpha n/2}\sum_{y\in\mathcal{B}} q_y,$$

donde podemos concluir que, para $\varepsilon = 2^{-\alpha n/2}$,

$$\mathbf{P}_{y\in_R Y}\Big[\operatorname{dist}\Big(\mathtt{Had}(X, y), U_k\Big) \leq \varepsilon\Big] > 1 - \varepsilon.$$

Por um argumento simétrico, temos $\mathbf{P}_{x\in_R X}\Big[\operatorname{dist}\Big(\mathtt{Had}(x, Y), U_k\Big) \leq \varepsilon\Big] > 1-\varepsilon$ e, portanto, o extrator Had é forte. □

A.2 Corpos finitos e espaços vetoriais

Para obter explicitamente Had, devemos exibir uma família de matrizes com a propriedade enunciada no lema A.1.1. Isto é feito no lema A.2.1.

Lema A.2.1. *Podemos encontrar n matrizes $A_1, \ldots, A_n \in \mathrm{GF}(2)^{n\times n}$ tais que, para todo $\emptyset \neq S \subseteq [n]$, a matriz $A_S \equiv \sum_{j\in S} A_j$ tem posto cheio. Ademais, isso pode ser feito em tempo* poly(n).

Demonstração. Primeiramente observamos que é possível encontrar (deterministicamente) um polinômio de grau n, irredutível sobre GF(2), em tempo $O(n^4)$ (veja [Sho88, BFSS]). Se P é o polinômio encontrado então

$$\mathbb{F} \equiv \mathrm{GF}(2)[X]/\langle P(X)\rangle$$

é uma representação concreta do corpo GF(2^n) na qual sabemos realizar operações aritméticas.

Tome $B = \{1, X, X^2, \ldots, X^{n-1}\}$ como base de $\mathbb{F}$. Para $i = 0, \ldots, n-1$, seja $T_i \colon \mathbb{F} \to \mathbb{F}$ uma transformação tal que $T_i \colon y \mapsto X^i y$. Observe que T_i é uma transformação linear. Seja A_{i+1} a matriz cuja j-ésima coluna é a representação de $T_i(X^{j-1}) = X^{i+j-1} \mod P(X)$ na base B. Então A_{i+1} é uma matriz que representa a transformação T_i tendo B como base fixada para $\mathbb{F}$ (ou seja, dado qualquer $Q \in \mathbb{F}$ representado como um vetor q na base B, o produto $A_{i+1}q$ é X^iQ na base B).

Afirmamos que as matrizes $A_1, \ldots, A_n$ satisfazem o enunciado do teorema. Já verificamos que elas podem ser obtidas em tempo poly(n). Resta apenas mostrar que, para todo $\emptyset \neq S \subseteq [n]$, a matriz A_S tem posto cheio. Seja $X_S = \sum_{j \in S} X^{j-1}$. Então X_S é um elemento não-nulo de $\mathbb{F}$ e, portanto, possui um inverso X_S^{-1} no corpo. Usando os mesmos argumentos do parágrafo anterior, podemos obter uma matriz U_S que representa a transformação linear $y \in \mathbb{F} \mapsto X_S^{-1} y$ de acordo com a base fixada B. Por construção $U_S A_S$ é a identidade. Isso implica que A_S tem posto cheio. □

Prova do teorema 6.0.14. Tome $k = \lfloor 0.1n \rfloor$ matrizes das n obtidas pelo lema A.2.1 e aplique o lema A.1.1. □

Referências Bibliográficas

[AGHP90] Noga Alon, Oded Goldreich, Johan Hastad, and Rene Peralta. Simple constructions of almost k-wise independent random variables. In *IEEE Symposium on Foundations of Computer Science*, pages 544–553, 1990. 7.1.3

[AS00] Noga Alon and Joel Spencer. *The Probabilistic Method*. John Wiley and Sons, New York, 2nd edition, 2000. 7.1.1, 7.1.2

[BFSS] Alin Bostan, Philippe Flajolet, Bruno Salvy, and Eric Schost. Fast computation with two algebraic numbers. A.2

[BIW04] Boaz Barak, Russell Impagliazzo, and Avi Wigderson. Extracting randomness using few independent sources. In *FOCS*, pages 384–393. IEEE Computer Society, 2004. 1.1, 3, 3, 4

[BKS+05] Boaz Barak, Guy Kindler, Ronen Shaltiel, Benny Sudakov, and Avi Wigderson. Simulating independence: new constructions of condensers, Ramsey graphs, dispersers, and extractors. In *STOC '05: Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 1–10, New York, NY, USA, 2005. ACM Press. 1.1, 1.2, 2.2, 4, 8, 8.1, 10

[BKT04] Jean Bourgain, Nets Katz, and Terence Tao. A sum-product estimate in finite fields, and applications. *Geometric And Functional Analysis*, 14:27–57, 2004. 4.1, 4.1

[Blu86] M Blum. Independent unbiased coin flips from a correlated biased source: a finite state markov chain. *Combinatorica*, 6(2):97–108, 1986. 1

[Bou99] Jean Bourgain. On the dimension of Kakeya sets and related maximal inequalities. *Geometric And Functional Analysis*, 9:256–282, 1999. 4.1

[Bou05] J. Bourgain. More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory*, 1:1–32, 2005. 1.2, 8.7, 10, 10.0.3

[BRSW06] Boaz Barak, Anup Rao, Ronen Shaltiel, and Avi Wigderson. 2-source dispersers for $n^{o(1)}$ entropy and Ramsey graphs beating the Frankl-Wilson construction. In *STOC '06: Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, New York, NY, USA, 2006. ACM Press. 1.2, 8.7

[CG85] Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity (extended abstract). In *IEEE Symposium on Foundations of Computer Science*, pages 429–442, 1985. 1

[CSRL01] Thomas H. Cormen, Clifford Stein, Ronald L. Rivest, and Charles E. Leiserson. *Introduction to Algorithms*. McGraw-Hill Higher Education, 2001. 9.1

[Erd47] P. Erdös. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.*, 53:292–294, 1947. 8.6

[FW81] Peter Frankl and Richard M. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, 1(4):357–368, 1981. 8.6

[Gol95] Oded Goldreich. Three XOR-lemmas—an exposition. *Electronic Colloquium on Computational Complexity (ECCC)*, 2(056), 1995. 2.2.1, 2.2.1

[Gow98] W. T. Gowers. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.*, 8(3):529–551, 1998. 4.1

[GRS04] Ariel Gabizon, Ran Raz, and Ronen Shaltiel. Deterministic extractors for bit-fixing sources by obtaining an independent seed. In *FOCS '04: Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science (FOCS'04)*, pages 394–403, Washington, DC, USA, 2004. IEEE Computer Society. 1.1

[GUV06] Venkatesan Guruswami, Christopher Umans, and Salil Vadhan. Extractors and condensers from univariate polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, 6(134), 2006. 9

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc.*, 43:439–561, 2006. 1.1

[ISW06] Russell Impagliazzo, Ronen Shaltiel, and Avi Wigderson. Reducing the seed length in the nisan-wigderson generator. *Combinatorica*, 26:647–681, 2006. 9.5

[IW97] Russell Impagliazzo and Avi Wigderson. P=BPP unless E has subexponential circuits: derandomizing the XOR lemma. In *Proceedings of the 29th STOC*, pages 220–229, 1997. 9.5

[KRVZ06] Jesse Kamp, Anup Rao, Salil Vadhan, and David Zuckerman. Deterministic extractors for small-space sources. In *STOC '06: Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 691–700, New York, NY, USA, 2006. ACM Press. 1.1

[KZ03] Jesse Kamp and David Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient cryptography. In *FOCS '03: Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, page 92, Washington, DC, USA, 2003. IEEE Computer Society. 1.1

[LRVW03] C. Lu, O. Reingold, S. Vadhan, and A. Wigderson. Extractors: Optimal up to constant factors, 2003. 9, 9.1

[MS77] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. North Holland, 1977. 9.4

[Nat96] Melvyn B. Nathanson. *Additive Number Theory: Inverse Problems and the Geometry of Sumsets*. Number 165 in Graduate Texts in Mathematics. Springer, New York, 1996. 4.1

[NW94] N. Nisan and A. Wigderson. Hardness vs randomness. *J. Comput. Syst. Sci.*, 2(42):148–167, 1994. 9.5, 9.5

[NZ96] Noam Nisan and David Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–52, 1996. 9.2

[PR04] Pavel Pudlák and Vojtěch Rödl. Pseudorandom sets and explicit constructions of Ramsey graphs. In *Complexity of computations and proofs*, volume 13 of *Quad. Mat.*, pages 327–346. Dept. Math., Seconda Univ. Napoli, Caserta, 2004. 8.6

[Rao05] Anup Rao. Extractors for a constant number of polynomial min-entropy independent sources. *Electronic Colloquium on Computational Complexity (ECCC)*, 5(106), 2005. 1.1, 1.2, 3, 8.7, 10

[Raz04] Ran Raz. Extractors with weak random seeds. *Electronic Colloquium on Computational Complexity (ECCC)*, 4(099), 2004. 1.1, 1.2, †, 1, 7.1.3, 7.3, 8.7, 1, 10.1

[RRV99] R. Raz, O. Reingold, and S. Vadhan. Extracting all the randomness and reducing the error in trevisan's extractors, 1999. 9.4.2

[RTS00] Jaikumar Radhakrishnan and Amnon Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM J. Discret. Math.*, 13(1):2–24, 2000. 9, 9.1

[Ruz99] Imre Ruzsa. An analog of Freiman's theorem in groups, Structure theory of set adition. *Astérisque*, 258:323–326, 1999. 4.2

[Sha02] R. Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the EATCS*, pages 67–95, 2002. 9

[Sha05] Ronen Shaltiel. How to get more mileage from randomness extractors. *Electronic Colloquium on Computational Complexity (ECCC)*, 5(145), 2005. 7.3

[Sho88] Victor Shoup. New algorithms for finding irreducible polynomials over finite fields. In *IEEE Symposium on Foundations of Computer Science*, pages 283–290, 1988. A.2

[SS96] M. Sipser and D. A. Spielman. Expander codes. *IEEE Transactions on Information Theory*, 42(6):1710–1722, 1996. 1.1

[SSV05] B. Sudakov, E. Szemerédi, and V. H. Vu. On a question of Erdős and Moser. *Duke Math. J.*, 129:129–155, 2005. 4.1

[STV98] Madhu Sudan, Luca Trevisan, and Salil P. Vadhan. Pseudorandom generators without the XOR lemma. *Electronic Colloquium on Computational Complexity (ECCC)*, 5(074), 1998. 9.5

[SV86] Miklos Santha and Umesh V Vazirani. Generating quasi-random sequences from semi-random sources. *J. Comput. Syst. Sci.*, 33(1):75–87, 1986. 1

[Tre99] Luca Trevisan. Construction of extractors using pseudo-random generators (extended abstract). In *STOC '99: Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 141–148, New York, NY, USA, 1999. ACM Press. 9.3

[TSZ04] A. Ta-Shma and D. Zuckerman. Extractor codes. *Information Theory, IEEE Transactions on*, 50(12):3015–3025, 2004. 2

[TV06] Terence Tao and Van H. Vu. *Additive Combinatorics*. Cambridge Studies in Advanced Mathematics, 2006. 4

[vN51] John von Neumann. Various techniques used in connection with random digits. *Applied Math Series*, 12:36–38, 1951. 1

[WZ99] Avi Wigderson and David Zuckerman. Expanders that beat the eigenvalue bound: Explicit construction and applications. *Combinatorica*, 19(1):125–138, 1999. 1.1

[Yao82] Andrew C. Yao. Theory and applications of trapdoor functions. In *FOCS '82: Proc. 22nd IEEE Symp. on Foundations of Comp. Science*, pages 80–91, 1982. 9.4.3

[Zuc90] Zuckerman. General weak random sources. In *FOCS: IEEE Symposium on Foundations of Computer Science (FOCS)*, 1990. 1

Índice Remissivo