





A Pseudorandom Generator for Functions of Low-Degree Polynomial Threshold Functions

Penghui Yao  

State Key Laboratory for Novel Software Technology, New Cornerstone Science Laboratory,
Nanjing University, Nanjing 210023, China
Hefei National Laboratory, Hefei 230088, China

Mingnan Zhao  

State Key Laboratory for Novel Software Technology, New Cornerstone Science Laboratory,
Nanjing University, Nanjing 210023, China

Abstract

Developing explicit *pseudorandom generators* (PRGs) for prominent categories of Boolean functions is a key focus in computational complexity theory. In this paper, we investigate the PRGs against the functions of degree- d *polynomial threshold functions* (PTFs) over Gaussian space. Our main result is an explicit construction of PRG with seed length $\text{poly}(k, d, 1/\epsilon) \cdot \log n$ that can fool *any* function of k degree- d PTFs with probability at least $1 - \epsilon$. More specifically, we show that the summation of L independent R -moment-matching Gaussian vectors ϵ -fools functions of k degree- d PTFs, where $L = \text{poly}(k, d, \frac{1}{\epsilon})$ and $R = O(\log \frac{kd}{\epsilon})$. The PRG is then obtained by applying an appropriate discretization to Gaussian vectors with bounded independence.

2012 ACM Subject Classification Theory of computation \rightarrow Pseudorandomness and derandomization

Keywords and phrases Pseudorandom generators, polynomial threshold functions

Digital Object Identifier 10.4230/LIPIcs.ICALP.2025.134

Category Track A: Algorithms, Complexity and Games

Related Version *Previous Version*: <https://arxiv.org/abs/2504.10904>

Funding PY and MZ were supported by National Natural Science Foundation of China (Grant Nos. 62332009 and 12347104), Innovation Program for Quantum Science and Technology (Grant No. 2021ZD0302901), NSFC/RGC Joint Research Scheme (Grant No. 12461160276), Natural Science Foundation of Jiangsu Province (Grant No. BK20243060), and New Cornerstone Science Foundation.

1 Introduction

In computational complexity theory, derandomization is a powerful technique that aims to reduce randomness in algorithms without sacrificing efficiency or accuracy. A versatile approach for derandomization is to design explicit *pseudorandom generators* (PRGs) for notable families of Boolean functions. A PRG for a family of Boolean functions is able to consume few random bits and produce a distribution over high-dimensional vectors, which is indistinguishable from a target distribution, such as the uniform distribution over Boolean cube, by any function in the family. In this paper, we concern ourselves with the Gaussian distribution over \mathbb{R}^n . Formally,

► **Definition 1.** Let $\mathcal{F} \subseteq \{f : \mathbb{R}^n \rightarrow \{0, 1\}\}$ be a family of Boolean functions. A function $G : \{0, 1\}^r \rightarrow \mathbb{R}^n$ is a pseudorandom generator for \mathcal{F} with error ϵ over Gaussian distribution $\mathcal{N}(0, 1)^n$ if for any $f \in \mathcal{F}$,

$$\left| \mathbb{E}_{s \sim_u \{0, 1\}^r} [f(G(s))] - \mathbb{E}_{x \sim \mathcal{N}(0, 1)^n} [f(x)] \right| \leq \epsilon .$$

We call r the seed length of G . We also say G ϵ -fools \mathcal{F} over the Gaussian distribution.



© Penghui Yao and Mingnan Zhao;

licensed under Creative Commons License CC-BY 4.0

52nd International Colloquium on Automata, Languages, and Programming (ICALP 2025).

Editors: Keren Censor-Hillel, Fabrizio Grandoni, Joël Ouaknine, and Gabriele Puppis

Article No. 134; pp. 134:1–134:20



Leibniz International Proceedings in Informatics

LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



There has been a considerable amount of research developing PRGs for various Boolean function families, including halfspaces, polynomial threshold functions and intersections of halfspaces. Let $\text{sign} : \mathbb{R} \rightarrow \{0, 1\}$ be the function such that $\text{sign}(x) = 1$ iff $x \geq 0$. A *halfspace* is a Boolean function of the form $f(x) = \text{sign}(a_1x_1 + \dots + a_nx_n - b)$ for some $a_1, \dots, a_n, b \in \mathbb{R}$. Halfspaces are a fundamental class of Boolean functions which have found significant applications in machine learning, complexity theory, theory of approximation and more. A very successful series of work produced PRGs that ϵ -fool halfspaces with seed length poly-logarithmic in n and ϵ^{-1} over both Boolean space [28, 5, 21, 7] and Gaussian space [19]. *Polynomial threshold functions* (PTFs) are functions of the form $f(x) = \text{sign}(p(x))$ where p is a polynomial. We call f is a degree- d PTF if p is a degree- d polynomial. PTFs are natural generalization for halfspaces since a halfspace is a degree-1 PTF. An explicit PRG that ϵ -fools PTFs over Boolean space has been achieved with seed length $(d/\epsilon)^{O(d)} \cdot \log n$ [21]. As for Gaussian space, a sequence of work [6, 12, 13, 14, 21, 15, 16, 24, 17] succeeds in giving a PRG with seed length polynomial in d , ϵ^{-1} and $\log n$ [24, 17]. Another extension for halfspaces is *intersections* of k halfspaces which are polytopes with k facets. A line of work [8, 9, 27, 4, 25] results in PRGs with seed length polynomial in $\log k$, $\log n$ and $1/\epsilon$ over Boolean space [25] and over Gaussian space [4].

Considering the prosperity of PRGs for these functions families, we commence designing PRGs for *functions of degree- d polynomial threshold functions*.

► **Definition 2.** We say a function $F : \mathbb{R}^n \rightarrow \{0, 1\}$ is a function of k degree- d PTFs if there exist k polynomials $p_1, \dots, p_k : \mathbb{R}^n \rightarrow \mathbb{R}$ of degree d and a Boolean function $f : \{0, 1\}^k \rightarrow \{0, 1\}$ such that $F(x) = f(\text{sign}(p_1(x)), \dots, \text{sign}(p_k(x)))$.

This family consumes all three function families we discussed above. For example, it includes intersections of halfspaces by setting $d = 1$ and $f(x) = x_1 \cdots x_k$. The research on PRGs for functions of PTFs is driven by several motivations beyond its fundamental role in derandomization tasks. For instance, the collection of satisfying assignments of an intersection of k degree-2 PTFs corresponds to the feasible solutions set of an $\{0, 1\}$ -integer quadratic programming [22] with k constraints. The investigation into the structure of these sets has been a central focus of extensive research in areas including learning theory, counting, optimization, and combinatorics.

In this work, we consider building explicit PRGs for functions of degree- d PTFs over Gaussian space. Before presenting our main result, we briefly revisit relevant prior work on fooling functions of halfspaces.

1.1 Prior Work

The related work is summarized in Table 1. Gopalan, O'Donnell, Wu and Zuckerman [8] constructed PRGs for *monotone functions of halfspaces*. They modified the PRG for halfspaces in [21] and showed the modified PRG ϵ -fools any monotone function of k halfspaces over a broad class of *product distributions* with seed length $O((k \log(k/\epsilon) + \log n) \cdot \log(k/\epsilon))$. When $k/\epsilon \leq \log^c n$ any $c > 0$, the seed length can be further improved to $O(k \log(k/\epsilon) + \log n)$.

Harsha, Klivans and Meka [9] considered designing PRGs for intersections of *regular* halfspaces (i.e., halfspaces with low influence). A halfspace $f(x) = \text{sign}(a_1x_1 + \dots + a_nx_n - b)$ is δ -*regular* if $\sum_i a_i^4 \leq \delta^2 \sum_i a_i^2$. They gave an explicit PRG construction for intersections of k δ -regular halfspaces over *proper* and *hypercontractive* distributions with seed length $O(\log n \log k/\epsilon)$ when δ is no more than a threshold. Their proof is based on developing an invariance principle for intersections of regular halfspaces via a generalization of the well-known Lindeberg method [20] and an anti-concentration result of polytopes in Gaussian space from [18].

■ **Table 1** Related Work on PRGs for Intersections of PTFs.

Reference	Function Family	Seed Length
[8]	Monotone functions of k halfspaces	$O((k \log(k/\epsilon) + \log n) \cdot \log(k/\epsilon))$
[9]	Intersections of k δ -regular halfspaces	$O(\log n \log k/\epsilon)$ for $\delta \leq \epsilon^5 / (\log^{8.1} k \log(1/\epsilon))$
[27]	Intersections of k weight- t halfspaces	$\text{poly}(\log n, \log k, t, 1/\epsilon)$
[25]	Intersections of k halfspaces	$\text{polylog } m \cdot \epsilon^{-(2+\delta)} \cdot \log n$ for any absolute constant $\delta \in (0, 1)$
[4]	Intersections of k halfspaces Arbitrary functions of k halfspaces	$O(\log n + \text{poly}(\log k, 1/\epsilon))$ $O(\log n + \text{poly}(k, 1/\epsilon))$
[6]	Intersections of k degree-2 PTFs	$O(\log n \cdot \text{poly}(k, 1/\epsilon))$

By extending the approach of [9] and combing the results on bounded independence fooling CNF formulas [1, 26], Servedio and Tan [27] designed an explicit PRG that ϵ -fools intersections of k *weight- t* halfspaces over Boolean space with $\text{poly}(\log n, \log k, t, 1/\epsilon)$ seed length. A halfspace $f(x) = \text{sign}(a_1x_1 + \dots + a_nx_n - b)$ is said to be *weight- t* if each a_i is an integer in $[-t, t]$.

As for intersections of k general halfspaces, O’Donnell, Servedio and Tan [25] gave a PRG construction over Boolean space with a polylogarithmic seed length dependence on k and n . Their proof involves a novel invariance principle for intersections of arbitrary halfspaces and a Littlewood–Offord style anticoncentration inequality for polytopes over Boolean space.

Concurrently, Chattopadhyay, De and Servedio [4] proposed a simple PRG that ϵ -fools intersections of k general halfspaces over Gaussian space, building upon the concept of *Johnson-Lindenstrauss transform* [10, 11]. The seed length is $O(\log n + \text{poly}(\log k, 1/\epsilon))$. Additionally, they show that the same PRG with seed length $O(\log n + \text{poly}(k, 1/\epsilon))$ is able to fool arbitrary functions of k halfspaces.

Speaking of fooling functions of PTFs, the study by Diakonikolas, Kane and Nelson [6] stands out as the sole work that constructs a PRG for intersections of k degree-2 PTFs. Their PRG is specific to degree $d \leq 2$ with a $O(\log n \cdot \text{poly}(k, 1/\epsilon))$ seed length.

1.2 Main Result

In this work, we investigate the PRGs fooling any function of low-degree PTFs. The main result is the following.

► **Theorem 3** (Informal version of Theorem 20). *There exists an explicit PRG ϵ -fools any function of k degree- d PTFs over Gaussian space with seed length $\text{poly}(k, d, 1/\epsilon) \cdot \log n$.*

The proof is inspired by the PRG proposed in [13] and the work [17]. This theorem follows from two components.

(1) Bounded independence fools functions of k degree- d PTFs

Consider the continuous random vector $Y = \frac{1}{\sqrt{L}} \sum_{i=1}^L Y_i$ where Y_i is a R -wise independent standard Gaussian vector of length n . Every $Y_{i,j}$ is a standard Gaussian variable and for any degree- R polynomial f , $\mathbb{E}[f(Y_i)] = \mathbb{E}_{y \sim \mathcal{N}(0,1)^n}[f(y)]$. We will prove that

► **Theorem 4.** (Informal version of Theorem 16) With $R = O(\log \frac{kd}{\epsilon})$ and $L = \text{poly}(k, d, \frac{1}{\epsilon})$, the distribution of Y ϵ -fools any function of k degree- d PTFs over Gaussian space.

The prior work [17] shows that bounded independence fools a single low-degree polynomial threshold function. This generalizes their work to the case of functions of k low-degree PTFs.

(2) Discretization of bounded independence Gaussians

An explicit PRG construction requires a discrete approximation to Gaussian vectors with bounded independence. The idea is to use a finite entropy random variable X to approximate Y . Previous work [13] uses the idea that a single Gaussian variable can be produced by two uniform random variables in $[0, 1]$ through the Box–Muller transform [2]. Therefore bounded independence Gaussian variables Y_i can be generated by using bounded independence uniform random variables. Then by truncating these uniform $[0, 1]$ random variables to a sufficient precision, we obtain vectors X_i that serve as a discrete approximation of Y_i . We prove that X also fools functions of k degree- d PTFs as long as X is a good approximation to Y .

► **Lemma 5** (Informal version of Lemma 19). If $X_{i,j}$ and $Y_{i,j}$ are sufficiently close with high probability, then X also fools functions of k degree- d PTFs.

2 Preliminary

Basic Notation

For $n \in \mathbb{N}$, $[n]$ denotes the set $\{1, 2, \dots, n\}$. For $\alpha \in \mathbb{R}^n$ and $i \in [n]$, α_i denotes the i -th coordinate of α , $|\alpha| = \sum_{i=1}^n |\alpha_i|$ and $\|\alpha\|_\infty = \max_{1 \leq i \leq n} |\alpha_i|$. For $\alpha, \beta \in \mathbb{R}^n$, $\alpha - \beta$ denotes the vector v such that $v_i = \alpha_i - \beta_i$ for all $i \in [n]$, and $\alpha^\beta = \prod_{i=1}^n \alpha_i^{\beta_i}$. For $\alpha \in \mathbb{N}^n$, $\alpha! = \prod_{i=1}^n \alpha_i!$. When it is clear from the context, we will use both subscript and superscript as indices.

Derivatives and Multidimensional Taylor Expansion

For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $\alpha \in \mathbb{N}^n$, we use $\partial^\alpha f$ to denote the partial derivative taken α_i times in the i -th coordinate and define $\|\nabla^t f(x)\| = \sqrt{\sum_{\alpha \in \mathbb{N}^n, |\alpha|=t} (\partial^\alpha f(x))^2}$. For $f(a, b) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ and $\alpha, \beta \in \mathbb{N}^n$, we use $\partial_a^\alpha \partial_b^\beta f$ to denote the partial derivative taken α_i times in a_i and β_i times in b_i . Using these notations, one has:

► **Theorem 6** (Multidimensional Taylor's Theorem). Let $d \in \mathbb{N}$ and $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a \mathcal{C}^{d+1} function. Then for all $x, y \in \mathbb{R}^n$,

$$f(y) = \sum_{\alpha \in \mathbb{N}^n, |\alpha| \leq d} \frac{\partial^\alpha f(x)}{\alpha!} (y - x)^\alpha + \sum_{\alpha \in \mathbb{N}^n, |\alpha|=d+1} \frac{\partial^\alpha f(z)}{\alpha!} (y - x)^\alpha$$

where $z = cx + (1 - c)y$ for some $c \in (0, 1)$.

Bump Function

Consider the bump function $\Psi : \mathbb{R} \rightarrow \mathbb{R}$ defined by $\Psi(x) = \begin{cases} e^{\frac{1}{x^2-1}}, & \text{if } |x| < 1, \\ 0, & \text{if } |x| \geq 1. \end{cases}$ It is well known that this function is infinitely differentiable and the derivatives are bounded.

► **Fact 7.** For all $t \in \mathbb{N}$, $|\Psi^{(t)}(x)| \leq t^{(3+o(1))t}$.

Let ρ be the smooth univariate function defined by $\rho(x) = \begin{cases} 1, & \text{for } x \geq 1, \\ e \cdot e^{\frac{1}{(\epsilon-1)^{2-1}}} & \text{for } 0 < x < 1, \\ 0, & \text{for } x \leq 0. \end{cases}$ It

is easy to see ρ is obtained from Ψ via translation, stretch, and concatenation. We have

► **Fact 8.** For all $t \in \mathbb{N}$, $|\rho^{(t)}(x)| \leq t^{(3+o(1))t}$.

► **Fact 9.** Let $r(u, v) := \rho(\log u - \log v + c)$ for some constant c . Then we have that for all $n, m \in \mathbb{N}$, $\left| \frac{\partial^n \partial^m r(u, v)}{\partial u^n \partial v^m} \right| \leq \frac{(n+m)^{6(n+m)}}{|u|^n |v|^m}$.

We include the proof for the above three facts in Appendix A for self-containment.

Gaussian Space and the Gaussian Noise Operator

We denote by $y \sim \mathcal{N}(0, 1)^n$ that $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ is a random vector whose components are independent standard Gaussian variables (i.e., with mean 0 and variance 1). We say a random vector $Y \in \mathbb{R}^n$ is a k -wise independent standard Gaussian vector if every component of Y is a standard Gaussian variable and $\mathbb{E}[p(Y)] = \mathbb{E}_{y \sim \mathcal{N}(0,1)^n}[p(y)]$ for all polynomials $p : \mathbb{R}^n \rightarrow \mathbb{R}$ with degree at most k . For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ on Gaussian space and $1 \leq p \leq \infty$, the p -norm is denoted by $\|f\|_p = (\mathbb{E}_{y \sim \mathcal{N}(0,1)^n}[|f(y)|^p])^{1/p}$. For $\rho \in [0, 1]$, the Gaussian noise operator U_ρ is the operator on the space of functions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by $U_\rho f(x) = \mathbb{E}_{y \sim \mathcal{N}(0,1)^n} [f(\rho x + \sqrt{1-\rho^2}y)]$.

The probabilists' Hermite polynomials [23, Section 11] $\{H_j\}_{j \in \mathbb{N}}$ are defined by $H_j(y) = \frac{(-1)^j}{\varphi(y)} \cdot \frac{d^j \varphi(y)}{d y^j}$ where $\varphi(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$. The univariate Hermite polynomials $\{h_j\}_{j \in \mathbb{N}}$ are defined by normalization: $h_j = \frac{1}{\sqrt{j!}} H_j$. For a multi-index $\alpha \in \mathbb{N}^n$, the (multivariate) Hermite polynomial $h_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ is $h_\alpha(y) = \prod_{j=1}^n h_{\alpha_j}(y_j)$. The degree of h_α is $|\alpha|$. The Hermite polynomials $\{h_\alpha\}_{\alpha \in \mathbb{N}^n}$ form an orthonormal basis for the functions over Gaussian space: $\mathbb{E}_{y \sim \mathcal{N}(0,1)^n}[h_\alpha(y)h_\beta(y)] = 1$ iff $\alpha = \beta$, and every degree- d polynomial $f : \mathbb{R}^n \rightarrow \mathbb{R}$ can be uniquely expanded as $f(y) = \sum_{\alpha \in \mathbb{N}^n, |\alpha| \leq d} \hat{f}(\alpha) h_\alpha(y)$. We can also expand the function $f(x + \sqrt{\lambda}y)$ in the Hermite basis in a manner similar to Taylor expansion.

► **Lemma 10** (Lemma 16 in [17]). Suppose $f(y) = \sum_{\alpha \in \mathbb{N}^n} \hat{f}(\alpha) h_\alpha(y)$, we have $f(x + \sqrt{\lambda}y) = \sum_{\alpha \in \mathbb{N}^n} \frac{\partial^\alpha \phi(x)}{\sqrt{\alpha!}} \lambda^{|\alpha|/2} h_\alpha(y)$, where $\phi(x) = U_{\sqrt{1-\lambda}} f\left(\frac{x}{\sqrt{1-\lambda}}\right)$.

The function $U_\rho f$ has the following expansion: $U_\rho f(y) = \sum_{\alpha \in \mathbb{N}^n, |\alpha| \leq d} \rho^{|\alpha|} \hat{f}(\alpha) h_\alpha(y)$. The definition of U_ρ can be extended to $\rho > 1$ by its action on the Hermite polynomials: $U_\rho h_\alpha(y) = \rho^{|\alpha|} h_\alpha(y)$. We will use the following hypercontractive inequality:

► **Theorem 11.** Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $2 \leq p \leq \infty$, $\|f\|_p \leq \|U_{\sqrt{p-1}} f\|_2$.

For more details on analysis over Gaussian space, readers may refer to [23].

Low-Degree Polynomials

Low-degree polynomials are extensively studied in the literature. We list some results used in this paper. It is well-know that low-degree polynomials have the following anti-concentration property:

► **Lemma 12** (Theorem 8 in [3]). Let $p : \mathbb{R}^n \rightarrow \mathbb{R}$ be a polynomial of degree d with $\|p\|_2 = 1$. Then we have $\Pr_{x \sim \mathcal{N}(0,1)^n}[|p(x)| \leq \epsilon] = O(d\epsilon^{1/d})$.

134:6 A PRG for Functions of Low-Degree PTFs

Suppose p is a low-degree polynomial, the following gives an estimation on the deviation of $p(x)$ caused by a small perturbation.

► **Lemma 13** (Lemma 22 in [13]). *Let $p : \mathbb{R}^n \rightarrow \mathbb{R}$ be a polynomial of degree d with $\|p\|_2 = 1$. Suppose $x \in \mathbb{R}^n$ be a vector with $\|x\|_\infty \leq B$ ($B > 1$). Let x' be another vector such that $\|x - x'\|_\infty \leq \delta < 1$. Then we have $|p(x) - p(x')| \leq \delta n^{d/2} O(B)^d$.*

The magnitudes of the derivatives of a low-degree polynomial are likely to grow at a moderate rate with high probability. Formally,

► **Lemma 14** (Lemma 6 in [17]). *Let $p : \mathbb{R}^n \rightarrow \mathbb{R}$ be an arbitrary polynomial of degree d and $y \sim \mathcal{N}(0, 1)^n$, the following holds with probability at least $1 - \epsilon d^3$:*

$$\|\nabla^t p(y)\| \leq O(\epsilon^{-1}) \|\nabla^{t-1} p(y)\| \text{ for all } 1 \leq t \leq d.$$

The following lemma gives quantitative bounds on how much the derivatives $\nabla^t p(x + \sqrt{\lambda}y)$ are concentrated around those of $\phi(x) = \mathbb{E}_{y \sim \mathcal{N}(0,1)^n} [p(x + \sqrt{\lambda}y)]$ when $y \sim \mathcal{N}(0, 1)^n$.

► **Lemma 15** (Lemma 23 in [17]). *Let $0 \leq \lambda < 1$ and $p : \mathbb{R}^n \rightarrow \mathbb{R}$ be an arbitrary polynomial of degree d and $\phi(x) = U_{\sqrt{1-\lambda}} p\left(\frac{x}{\sqrt{1-\lambda}}\right) = \mathbb{E}_{y \sim \mathcal{N}(0,1)^n} [p(x + \sqrt{\lambda}y)]$. For $0 \leq t \leq d$ and $y \sim \mathcal{N}(0, 1)^n$,*

$$\left(\mathbb{E}_{y \sim \mathcal{N}(0,1)^n} \left[\left\| \nabla^t p(x + \sqrt{\lambda}y) - \nabla^t \phi(x) \right\|^R \right] \right)^{\frac{1}{R}} \leq \sqrt{\sum_{j=t+1}^d (\lambda d R)^{j-t} \|\nabla^j \phi(x)\|^2}.$$

3 Fooling the Functions of PTFs via Bounded Independence

In this section, we show that a random Gaussian vector matching certain moments fools *any* function of low-degree polynomial threshold functions. Formally, we prove

► **Theorem 16.** *Fix a small constant $0 < \epsilon < 1$ and let $R \in \mathbb{N}$ be an integer. Let $p_1, \dots, p_k : \mathbb{R}^n \rightarrow \mathbb{R}$ be arbitrary polynomials of degree d and $f : \{0, 1\}^k \rightarrow \{0, 1\}$ be an arbitrary Boolean function. Define function*

$$F(x) := f(\text{sign}(p_1(x)), \dots, \text{sign}(p_k(x))) .$$

Let $Y = \frac{1}{\sqrt{L}} \sum_{i=1}^L Y_i$ where Y_i is a $2dR$ -wise independent standard Gaussian vector of length n and $L = \Omega\left(\frac{k^2 d^3 R^{15}}{\epsilon^2}\right)$. Then, we have

$$\left| \mathbb{E}_Y [F(Y)] - \mathbb{E}_{y \sim \mathcal{N}(0,1)^n} [F(y)] \right| = O(\epsilon k d^3) + kdL \cdot 2^{-\Omega(R)} .$$

The key idea in the proof of Theorem 16 is to analyze the derivatives of the disturbed function $\phi_i(x) = \mathbb{E}_{y \sim \mathcal{N}(0,1)^n} [p_i(x + \sqrt{\lambda}y)]$. We will see that once the derivatives of ϕ_i are well-controlled by its preceding order derivative at x , $\nabla^t p_i(x + \sqrt{\lambda}y)$ is concentrated around $\nabla^t \phi_i(x)$ for a random y , and $p_i(x + \sqrt{\lambda}y)$ and $\phi_i(x)$ share the same sign with high probability. Starting from this point, we use the mollifier introduced in [17]

$$G(x) := \prod_{i=1}^k \prod_{t=0}^{d-1} \rho \left(\log \left(\frac{\|\nabla^t p_i(x)\|^2}{16\epsilon^2 \|\nabla^{t+1} p_i(x)\|^2} \right) \right) \quad (1)$$

to judge whether the derivatives are all well-controlled for all k polynomials. $G(x) = 0$ as long as a certain order of derivative that is not controlled by its preceding order derivative. Our proof consists of following steps:

- **Approximation using the mollifier G :** We first establish that

$$\left| \mathbb{E}_Y[F(Y)] - \mathbb{E}_y[F(y)] \right| \approx \left| \mathbb{E}_Y[F(Y)G(Y)] - \mathbb{E}_y[F(y)G(y)] \right| .$$

This approximation enables us to focus primarily on the analysis of $F(y)G(y)$ in the subsequent steps.

- **Hybrid argument:** Let $\lambda = L^{-1}$, $y = \sqrt{\lambda} \sum_{i=1}^L y_i$ where $y_i \sim \mathcal{N}(0, 1)^n$ and $Z^i = \sqrt{\lambda}(y_1 + \dots + y_{i-1} + Y_{i+1} + \dots + Y_L)$. We will show

$$\mathbb{E} \left[F(Z^i + \sqrt{\lambda}Y_i)G(Z^i + \sqrt{\lambda}Y_i) \right] \approx \mathbb{E} \left[F(Z^i + \sqrt{\lambda}y_i)G(Z^i + \sqrt{\lambda}y_i) \right] . \quad (2)$$

Therefore by the triangle inequality, we have

$$\begin{aligned} \mathbb{E}_Y[F(Y)G(Y)] &= \mathbb{E} \left[F(Z^1 + \sqrt{\lambda}Y_1)G(Z^1 + \sqrt{\lambda}Y_1) \right] \\ &\approx \mathbb{E} \left[F(Z^1 + \sqrt{\lambda}y_1)G(Z^1 + \sqrt{\lambda}y_1) \right] = \mathbb{E} \left[F(Z^2 + \sqrt{\lambda}Y_2)G(Z^2 + \sqrt{\lambda}Y_2) \right] \\ &\approx \dots \approx \mathbb{E} \left[F(Z^L + \sqrt{\lambda}y_L)G(Z^L + \sqrt{\lambda}y_L) \right] = \mathbb{E}[F(y)G(y)] . \end{aligned}$$

To prove (2), we show for any fixed x ,

$$\mathbb{E} \left[F(x + \sqrt{\lambda}Y_i)G(x + \sqrt{\lambda}Y_i) \right] \approx \mathbb{E} \left[F(x + \sqrt{\lambda}y_i)G(x + \sqrt{\lambda}y_i) \right] .$$

This is done by a case analysis:

- The derivatives of all k polynomials $\phi_j(x)$ are well-controlled at point x . In this case, all $p_j(x + \sqrt{\lambda}Y_i)$ and $p_j(x + \sqrt{\lambda}y_i)$ share the same sign with high probability. Thus, it is highly likely that $F(x + \sqrt{\lambda}Y_i)$ and $F(x + \sqrt{\lambda}y_i)$ are nearly the same constant. It suffices to show Y_i fools the mollifier function $G(x + \sqrt{\lambda}y_i)$.
- At least one derivative is not controlled. In this case, we will show that $G(x + \sqrt{\lambda}Y_i)$ and $G(x + \sqrt{\lambda}y_i)$ are 0 with high probability. This implies that $F(x + \sqrt{\lambda}Y_i)G(x + \sqrt{\lambda}Y_i) = F(x + \sqrt{\lambda}y_i)G(x + \sqrt{\lambda}y_i) = 0$ with overwhelming probability.

In the subsequent sections, Section 3.1 first demonstrates that Y_i is able to fool the mollifier function G when x is a well-behaved point. Section 3.2 shows the closeness of a single step in the hybrid argument. Lastly, we prove Theorem 16 using approximation and the hybrid argument in Section 3.3.

3.1 Fooling the Mollifier G

We begin with proving that a $2dR$ -wise independent standard Gaussian vector Y fools the mollifier function $G(x + \sqrt{\lambda}y)$. To achieve this, we utilize the Taylor expansion to expand the mollifier function $G(x + \sqrt{\lambda}y)$ up to a specified order. As a result, $G(x + \sqrt{\lambda}y)$ is decomposed into two parts: a degree- $d(R - 1)$ polynomial $l(y)$ and a remainder term $\Delta(y)$. We mainly show that $\mathbb{E}[\Delta]$ is negligible under both pseudorandom distribution and true Gaussian distribution. This leads us to the conclusion that $\mathbb{E}[G(x + \sqrt{\lambda}y)] \approx \mathbb{E}[l(y)]$ and $\mathbb{E}[G(x + \sqrt{\lambda}Y)] \approx \mathbb{E}[l(Y)]$. Furthermore, since $l(y)$ has degree at most dR , it follows that $\mathbb{E}[l(y)] = \mathbb{E}[l(Y)]$. Thus, we conclude that $\mathbb{E}[G(x + \sqrt{\lambda}y)] \approx \mathbb{E}[G(x + \sqrt{\lambda}Y)]$.

- **Lemma 17.** Fix a small constant $0 < \epsilon < 1$ and let $R \in \mathbb{N}$ be an integer. Let $p_1, \dots, p_k : \mathbb{R}^n \rightarrow \mathbb{R}$ be arbitrary polynomials of degree d . Define $\phi_i(x) := U_{\sqrt{1-\lambda}p_i} \left(\frac{x}{\sqrt{1-\lambda}} \right) = \mathbb{E}_{y \sim \mathcal{N}(0, 1)^n} [p_i(x + \sqrt{\lambda}y)]$ for all p_i . Suppose that a fix point $x \in \mathbb{R}^n$ satisfies $\|\nabla^{t+1} \phi_i(x)\| \leq$

134:8 A PRG for Functions of Low-Degree PTFs

$\frac{1}{\epsilon} \|\nabla^t \phi_i(x)\|$ for any $1 \leq i \leq k$ and $0 \leq t \leq d-1$. Let Y be a $2dR$ -wise independent standard Gaussian vector of length n . For $\lambda = O(k^{-2}d^{-3}R^{-15}\epsilon^2)$, we have

$$\left| \mathbb{E}_Y \left[G(x + \sqrt{\lambda}Y) \right] - \mathbb{E}_{y \sim \mathcal{N}(0,1)^n} \left[G(x + \sqrt{\lambda}y) \right] \right| = kd \cdot 2^{-\Omega(R)} ,$$

where G is defined in (1).

Proof. Let $\sigma(z) := \rho(z - \log 16\epsilon^2)$ and by the definition of function $G(\cdot)$ defined in (1),

$$G(z) = \prod_{i=1}^k \prod_{t=0}^{d-1} \sigma \left(\log \|\nabla^t p_i(z)\|^2 - \log \|\nabla^{t+1} p_i(z)\|^2 \right) .$$

Define variables $\{s_i^t\}_{1 \leq i \leq k, 0 \leq t \leq d-1}$ and $\{r_i^t\}_{1 \leq i \leq k, 0 \leq t \leq d-1}$ by letting $s_i^t = \|\nabla^t p_i(x + \sqrt{\lambda}y)\|^2$ and $r_i^t = \|\nabla^{t+1} p_i(x + \sqrt{\lambda}y)\|^2$ as functions of y . Apparently, we have

$$G(x + \sqrt{\lambda}y) = g(s, r) := \prod_{i=1}^k \prod_{t=0}^{d-1} \sigma \left(\log s_i^t - \log r_i^t \right) .$$

Therefore, it is equivalent to prove

$$\left| \mathbb{E}_{y \sim \mathcal{N}(0,1)^n} [g(s, r)] - \mathbb{E}_Y [g(s, r)] \right| = kd \cdot 2^{-\Omega(R)} .$$

To this end, we expand $g(s, r)$ into R -th order using the Taylor expansion at some point (a, b) : $g(s, r) = l(s, r) + \Delta$, where $l(s, r)$ is a polynomial of y of degree at most dR and Δ is the remainder. Since Y is $2dR$ -wise independent, we know $\mathbb{E}_{y \sim \mathcal{N}(0,1)^n} [l(s, r)] = \mathbb{E}_Y [l(s, r)]$. Therefore, it suffices to show $\mathbb{E}_{y \sim \mathcal{N}(0,1)^n} [\Delta]$ and $\mathbb{E}_Y [\Delta]$ are bounded by $kd \cdot 2^{-\Omega(R)}$.

More specifically, we choose to expand $g(s, r)$ at points $a_i^t = \|\nabla^t \phi_i(x)\|^2$ and $b_i^t = \|\nabla^{t+1} \phi_i(x)\|^2$ via Theorem 6: $g(s, r) = l(s, r) + \Delta$, where

$$l(s, r) = \sum_{\substack{(\alpha_i^t)_{i \in [k], 0 \leq t \leq d-1} \in \mathbb{N}^{kd} \\ (\beta_i^t)_{i \in [k], 0 \leq t \leq d-1} \in \mathbb{N}^{kd} \\ |\alpha| + |\beta| < R}} \frac{\partial_s^\alpha \partial_r^\beta g(a, b)}{\alpha! \beta!} (s-a)^\alpha (r-b)^\beta$$

and

$$\Delta = \sum_{\substack{(\alpha_i^t)_{i \in [k], 0 \leq t \leq d-1} \in \mathbb{N}^{kd} \\ (\beta_i^t)_{i \in [k], 0 \leq t \leq d-1} \in \mathbb{N}^{kd} \\ |\alpha| + |\beta| = R}} \frac{\partial_s^\alpha \partial_r^\beta g(s^*, r^*)}{\alpha! \beta!} (s-a)^\alpha (r-b)^\beta .$$

for some (s^*, r^*) on the line segment joining (s, r) and (a, b) . It is not hard to see that $l(s, r)$ is a polynomial of y of degree at most $d(R-1)$. For the remainder term Δ , we will prove the following bound: $\mathbb{E}_{y \sim \mathcal{N}(0,1)^n} [\Delta] \leq kd \cdot 2^{-\Omega(R)}$.

We now give bounds on $\mathbb{E}_{y \sim \mathcal{N}(0,1)^n} [\Delta]$. The same argument applies to Y as well. Fix a small constant $0 < \delta < \frac{1}{kdR^7}$. Let \mathcal{E}_i^t be the event that

$$\left| \|\nabla^t p_i(x + \sqrt{\lambda}y)\| - \|\nabla^t \phi_i(x)\| \right| \leq \delta \|\nabla^t \phi_i(x)\| .$$

Now let $\mathcal{E} = \bigwedge_{1 \leq i \leq k, 0 \leq t \leq d} \mathcal{E}_i^t$. Note that

$$\Delta = \Delta \cdot \mathbf{1}_{\mathcal{E}} + \Delta \cdot \mathbf{1}_{\bar{\mathcal{E}}} = \Delta \cdot \mathbf{1}_{\mathcal{E}} + g(s, r) \cdot \mathbf{1}_{\bar{\mathcal{E}}} - l(s, r) \cdot \mathbf{1}_{\bar{\mathcal{E}}}.$$

Here, $\mathbf{1}_{\mathcal{A}} = 1$ when event \mathcal{A} occurs and $\mathbf{1}_{\mathcal{A}} = 0$ otherwise. Therefore, by the triangle inequality, we have

$$\begin{aligned} \mathbb{E}_y[|\Delta|] &\leq \mathbb{E}_y[\Delta \cdot \mathbf{1}_{\mathcal{E}}] + \mathbb{E}_y[g(s, r) \cdot \mathbf{1}_{\bar{\mathcal{E}}}] + \mathbb{E}_y[|l(s, r)| \cdot \mathbf{1}_{\bar{\mathcal{E}}}] \\ &\leq \mathbb{E}_y[\Delta \cdot \mathbf{1}_{\mathcal{E}}] + \mathbb{E}_y[\mathbf{1}_{\bar{\mathcal{E}}}] + \mathbb{E}_y[|l(s, r)| \cdot \mathbf{1}_{\bar{\mathcal{E}}}] && (g(s, r) \in [0, 1]) \\ &\leq \underbrace{\mathbb{E}_y[\Delta \cdot \mathbf{1}_{\mathcal{E}}]}_{\text{(Term 1)}} + \underbrace{\mathbb{E}_y[\mathbf{1}_{\bar{\mathcal{E}}}]}_{\text{(Term 2)}} + \underbrace{\sqrt{\mathbb{E}_y[l^2(s, r)]}}_{\text{(Term 3)}} \cdot \sqrt{\mathbb{E}_y[\mathbf{1}_{\bar{\mathcal{E}}}]}. && \text{(Cauchy-Schwarz)} \end{aligned}$$

We are next to bound Term 1 \sim 3.

Bounding Term 1. If event \mathcal{E} occurs, we have

$$\begin{aligned} |\Delta| &\leq \sum_{\substack{(\alpha_i^t) \in \mathbb{N}^{kd}, (\beta_i^t) \in \mathbb{N}^{kd} \\ |\alpha| + |\beta| = R}} \frac{|\partial_s^\alpha \partial_r^\beta g(s^*, r^*)|}{\alpha! \beta!} \prod_{i=1}^k \prod_{t=0}^{d-1} |s_i^t - a_i^t|^{\alpha_i^t} |r_i^t - b_i^t|^{\beta_i^t} \\ &\leq \sum_{\substack{(\alpha_i^t) \in \mathbb{N}^{kd}, (\beta_i^t) \in \mathbb{N}^{kd} \\ |\alpha| + |\beta| = R}} R^{6R} \cdot \prod_{i=1}^k \prod_{t=0}^{d-1} \frac{|s_i^t - a_i^t|^{\alpha_i^t} |r_i^t - b_i^t|^{\beta_i^t}}{|s_i^{*t}|^{\alpha_i^t} |r_i^{*t}|^{\beta_i^t}} \\ &\leq \sum_{\substack{(\alpha_i^t) \in \mathbb{N}^{kd}, (\beta_i^t) \in \mathbb{N}^{kd} \\ |\alpha| + |\beta| = R}} R^{6R} \cdot \left(\frac{2\delta + \delta^2}{(1 - \delta)^2} \right)^R = \binom{R + 2kd - 1}{R} \cdot R^{6R} \cdot (4\delta)^R \leq 2^{-R}, \end{aligned}$$

where the second inequality is from Fact 9, and the third one is true by the following facts:

- $(1 - \delta)^2 a_i^t \leq s_i^t \leq (1 + \delta)^2 a_i^t$, $(1 - \delta)^2 b_i^t \leq r_i^t \leq (1 + \delta)^2 b_i^t$,
- $s_i^{*t} \geq \min\{s_i^t, a_i^t\} \geq (1 - \delta)^2 a_i^t$ and $r_i^{*t} \geq \min\{r_i^t, b_i^t\} \geq (1 - \delta)^2 b_i^t$, since (s_i^{*t}, r_i^{*t}) lies between (s_i^t, r_i^t) and (a_i^t, b_i^t) .

This gives us $\mathbb{E}_y[\Delta \cdot \mathbf{1}_{\mathcal{E}}] \leq 2^{-R}$.

Bounding Term 2. Since

$$\left\| \nabla^t p_i(x + \sqrt{\lambda}y) \right\| - \left\| \nabla^t \phi_i(x) \right\| \leq \left\| \nabla^t p_i(x + \sqrt{\lambda}y) - \nabla^t \phi_i(x) \right\|$$

and

$$\left\| \nabla^t \phi_i(x) \right\| - \left\| \nabla^t p_i(x + \sqrt{\lambda}y) \right\| \leq \left\| \nabla^t p_i(x + \sqrt{\lambda}y) - \nabla^t \phi_i(x) \right\|,$$

we have

$$\left\| \left\| \nabla^t p_i(x + \sqrt{\lambda}y) \right\| - \left\| \nabla^t \phi_i(x) \right\| \right\| \leq \left\| \nabla^t p_i(x + \sqrt{\lambda}y) - \nabla^t \phi_i(x) \right\| \quad (3)$$

134:10 A PRG for Functions of Low-Degree PTFs

This gives us

$$\begin{aligned}
\Pr_y [\mathcal{E}_i^t] &\geq \Pr_y \left[\left\| \nabla^t p_i(x + \sqrt{\lambda}y) - \nabla^t \phi_i(x) \right\| \leq \delta \left\| \nabla^t \phi_i(x) \right\| \right] \\
&\geq 1 - \frac{\mathbb{E}_y \left[\left\| \nabla^t p_i(x + \sqrt{\lambda}y) - \nabla^t \phi_i(x) \right\|^R \right]}{\delta^R \left\| \nabla^t \phi_i(x) \right\|^R} \\
&\geq 1 - \frac{\left(\sum_{j=t+1}^d (\lambda d R)^{j-t} \left\| \nabla^j \phi(x) \right\|^2 \right)^{\frac{R}{2}}}{\delta^R \left\| \nabla^t \phi_i(x) \right\|^R} \\
&\geq 1 - \left(\frac{1}{\delta} \right)^R \left(\sum_{j=1}^{d-t} \left(\frac{\lambda d R}{\epsilon^2} \right)^j \right)^{\frac{R}{2}} \geq 1 - \left(\sqrt{\frac{2\lambda d R}{\delta^2 \epsilon^2}} \right)^R
\end{aligned}$$

where the second inequality is from Markov's inequality, the third one is from Lemma 15 and the fourth one holds since $x \in \mathbb{R}^n$ satisfies $\left\| \nabla^{t+1} \phi_i(x) \right\| \leq \frac{1}{\epsilon} \left\| \nabla^t \phi_i(x) \right\|$ for any $1 \leq i \leq k$ and $0 \leq t \leq d-1$. For $\lambda = O(k^{-2} d^{-3} R^{-15} \epsilon^2)$, we have $\Pr_y [\mathcal{E}_i^t] \geq 1 - 2^{-R}$. Consequently, $\Pr_y [\mathcal{E}] \geq 1 - kd2^{-R}$. Therefore, we have $\mathbb{E}_y [\mathbb{1}_{\bar{\mathcal{E}}}] \leq kd2^{-R}$.

Bounding Term 3. We next to upper bound $\mathbb{E}_y [l^2(s, r)]$. Note that

$$\begin{aligned}
&\mathbb{E}_y [l^2(s, r)] \\
&\leq \sum_{\substack{|\alpha|+|\beta| < R \\ |\alpha'|+|\beta'| < R}} \frac{|\partial_s^\alpha \partial_r^\beta g(a, b)|}{\alpha! \beta!} \frac{|\partial_s^{\alpha'} \partial_r^{\beta'} g(a, b)|}{\alpha'! \beta'!} \mathbb{E}_y \left[\left| (s-a)^\alpha (r-b)^\beta (s-a)^{\alpha'} (r-b)^{\beta'} \right| \right] \\
&\leq \sum_{\substack{q_1 < R \\ q_2 < R}} \sum_{\substack{|\alpha|+|\beta|=q_1 \\ |\alpha'|+|\beta'|=q_2}} R^{6(q_1+q_2)} \cdot \underbrace{\mathbb{E}_y \left[\prod_{i=1}^k \prod_{t=0}^{d-1} \frac{|s_i^t - a_i^t|^{\alpha_i^t} |r_i^t - b_i^t|^{\beta_i^t} |s_i^t - a_i^t|^{\alpha_i'^t} |r_i^t - b_i^t|^{\beta_i'^t}}{|a_i^t|^{\alpha_i^t} |b_i^t|^{\beta_i^t} |a_i^t|^{\alpha_i'^t} |b_i^t|^{\beta_i'^t}} \right]}_{(\star)}
\end{aligned}$$

By generalized Hölder's inequality, for $q_1 + q_2 \neq 0$

$$\begin{aligned}
(\star) &\leq \prod_{i,t} \frac{\left(\mathbb{E}_y \left[|s_i^t - a_i^t|^{q_1+q_2} \right] \right)^{\frac{\alpha_i^t}{q_1+q_2}} \left(\mathbb{E}_y \left[|r_i^t - b_i^t|^{q_1+q_2} \right] \right)^{\frac{\beta_i^t}{q_1+q_2}}}{|a_i^t|^{\alpha_i^t} |b_i^t|^{\beta_i^t}} \\
&\quad \cdot \frac{\left(\mathbb{E}_y \left[|s_i^t - a_i^t|^{q_1+q_2} \right] \right)^{\frac{\alpha_i'^t}{q_1+q_2}} \left(\mathbb{E}_y \left[|r_i^t - b_i^t|^{q_1+q_2} \right] \right)^{\frac{\beta_i'^t}{q_1+q_2}}}{|a_i^t|^{\alpha_i'^t} |b_i^t|^{\beta_i'^t}}.
\end{aligned}$$

Note that for $0 < q \leq 2R$,

$$\begin{aligned}
&\mathbb{E}_y \left[\left(\frac{\left\| \nabla^t p_i(x + \sqrt{\lambda}y) \right\|^2 - \left\| \nabla^t \phi_i(x) \right\|^2}{\left\| \nabla^t \phi_i(x) \right\|^2} \right)^q \right] \\
&\leq \mathbb{E}_y \left[\left(\frac{2 \left\| \nabla^t p_i(x + \sqrt{\lambda}y) \right\| - \left\| \nabla^t \phi_i(x) \right\| \cdot \left\| \nabla^t \phi_i(x) \right\| + \left\| \nabla^t p_i(x + \sqrt{\lambda}y) \right\| - \left\| \nabla^t \phi_i(x) \right\|^2}{\left\| \nabla^t \phi_i(x) \right\|^2} \right)^q \right] \\
&\leq \mathbb{E}_y \left[\left(\frac{2 \left\| \nabla^t p_i(x + \sqrt{\lambda}y) - \nabla^t \phi_i(x) \right\|}{\left\| \nabla^t \phi_i(x) \right\|} + \frac{\left\| \nabla^t p_i(x + \sqrt{\lambda}y) - \nabla^t \phi_i(x) \right\|^2}{\left\| \nabla^t \phi_i(x) \right\|^2} \right)^q \right]
\end{aligned}$$

$$\begin{aligned}
 &= \sum_{j=0}^q 2^j \mathbb{E}_y \left[\frac{\|\nabla^t p_i(x + \sqrt{\lambda}y) - \nabla^t \phi_i(x)\|^{2q-j}}{\|\nabla^t \phi_i(x)\|^{2q-j}} \right] \\
 &\leq \sum_{j=0}^q 2^j \cdot \left(\sqrt{\frac{4\lambda dR}{\epsilon^2}} \right)^{2q-j} = 4^q \cdot \sum_{j=0}^q \left(\sqrt{\frac{\lambda dR}{\epsilon^2}} \right)^{2q-j} \\
 &\leq 2 \cdot 4^q \cdot \left(\sqrt{\frac{\lambda dR}{\epsilon^2}} \right)^q \leq \left(\sqrt{\frac{17\lambda dR}{\epsilon^2}} \right)^q
 \end{aligned}$$

where the first inequality is from $|a^2 - b^2| = |a - b| |a + b| \leq |a - b| (|a| + |b|) \leq |a - b| (2|a| + |a - b|)$, the second one is from Eq.(3) and the third inequality is from Lemma 15. Therefore,

$$(\star) \leq \prod_{i,t} \left(\sqrt{\frac{17\lambda dR}{\epsilon^2}} \right)^{\alpha_i^t + \beta_i^t + \alpha_i^t + \beta_i^t} = \left(\sqrt{\frac{17\lambda dR}{\epsilon^2}} \right)^{q_1 + q_2} .$$

Consequently,

$$\begin{aligned}
 \mathbb{E}_y[l^2(s, r)] &\leq \sum_{\substack{q_1 < R \\ q_2 < R}} \sum_{\substack{|\alpha| + |\beta| = q_1 \\ |\alpha'| + |\beta'| = q_2}} R^{6(q_1 + q_2)} \cdot \left(\sqrt{\frac{17\lambda dR}{\epsilon^2}} \right)^{q_1 + q_2} \\
 &\leq \sum_{\substack{q_1 < R \\ q_2 < R}} (R + 2kd - 1)^{q_1 + q_2} \cdot R^{6(q_1 + q_2)} \cdot \left(\sqrt{\frac{17\lambda dR}{\epsilon^2}} \right)^{q_1 + q_2} \\
 &= \sum_{q=0}^{2R-2} (q + 1) \cdot \left((R + 2kd - 1) R^6 \sqrt{\frac{17\lambda dR}{\epsilon^2}} \right)^q .
 \end{aligned}$$

For $\lambda = O(k^{-2}d^{-3}R^{-15}\epsilon^2)$ sufficiently small, we have $\mathbb{E}_y[l^2(s, r)] \leq \sum_{q=0}^{2R-2} (q+1) \cdot 4^{-q} < 2R$.

Thus, putting everything together, we have

$$\begin{aligned}
 \mathbb{E}_y[|\Delta|] &\leq \mathbb{E}_y[\Delta \cdot \mathbf{1}_{\mathcal{E}}] + \mathbb{E}_y[\mathbf{1}_{\bar{\mathcal{E}}}] + \sqrt{\mathbb{E}_y[l^2(s, r)]} \sqrt{\mathbb{E}_y[\mathbf{1}_{\bar{\mathcal{E}}}] } \\
 &\leq 2^{-R} + kd2^{-R} + \sqrt{2Rkd2^{-R}} \\
 &\leq kd \cdot 2^{-\Omega(R)} .
 \end{aligned}$$

We now regard $l(s, r)$ as a function of y , and from the above inequality we have

$$\left| \mathbb{E}_y[G(x + \sqrt{\lambda}y)] - \mathbb{E}_y[l(s, r)] \right| \leq \mathbb{E}_y[|\Delta|] \leq kd \cdot 2^{-\Omega(R)} .$$

Similarly, the same argument applying on $2dR$ -wise independent Gaussian vector Y gives us

$$\left| \mathbb{E}_Y[G(x + \sqrt{\lambda}Y)] - \mathbb{E}_Y[l(s, r)] \right| \leq kd \cdot 2^{-\Omega(R)} .$$

The lemma then follows from the fact that $\mathbb{E}_y[l(s, r)] = \mathbb{E}_Y[l(s, r)]$, since $l(s, r)$ is a polynomial of y of degree at most $d(R - 1)$. \blacktriangleleft

3.2 A Single Step in the Hybrids

In this section, we analyze one single step in the entire hybrid argument. We will show that for any x , we have that $\mathbb{E}_Y[F(x + \sqrt{\lambda}Y)G(x + \sqrt{\lambda}Y)] \approx \mathbb{E}_y[F(x + \sqrt{\lambda}y)G(x + \sqrt{\lambda}y)]$ for $2dR$ -wise independent Gaussian Y and true Gaussian y .

134:12 A PRG for Functions of Low-Degree PTFs

Let $\phi_i(x) = U_{\sqrt{1-\lambda}} p_i\left(\frac{x}{\sqrt{1-\lambda}}\right) = \mathbb{E}_y[p_i(x + \sqrt{\lambda}y)]$. The proof proceeds through a case analysis based on the behavior of ϕ_i at the fixed point x . Specifically, we define x as well-behaved if $\|\nabla^{t+1}\phi_i(x)\| \leq \frac{1}{\epsilon} \|\nabla^t\phi_i(x)\|$ for all $t \in [d]$ and $i \in [k]$. In other words, for each function ϕ_i , its t -th order derivatives are controlled by its $(t-1)$ -th order derivatives.

- In the scenario where x is not well-behaved, we can identify an i_0 and a t_0 such that with at least probability $1 - 2^{-R+1}$,

$$\left\| \nabla^{t_0+1} p_{i_0}(x + \sqrt{\lambda}y) \right\| > \frac{1}{4\epsilon} \left\| \nabla^{t_0} p_{i_0}(x + \sqrt{\lambda}y) \right\| .$$

Thus, it is highly probable that the mollifier function $G(x + \sqrt{\lambda}y) = 0$. So, the expectation of $F(x + \sqrt{\lambda}y)G(x + \sqrt{\lambda}y)$ is no more than 2^{-R+1} . The same argument works for Y as well.

- For the case that x is well-behaved, we will show that for all p_i , $p_i(x + \sqrt{\lambda}y)$ and $p_i(x + \sqrt{\lambda}Y)$ are nearly the same constant. This implies $F(x + \sqrt{\lambda}y)$ and $F(x + \sqrt{\lambda}Y)$ are equal in most situations. Then it suffices to show Y fools the mollifier, as discussed in the previous section.

► **Lemma 18.** Fix a small constant $0 < \epsilon < 1$ and let $R \in \mathbb{N}$ be an integer. Let $p_1, \dots, p_k : \mathbb{R}^n \rightarrow \mathbb{R}$ be arbitrary polynomials of degree d and $f : \{0, 1\}^k \rightarrow \{0, 1\}$ be an arbitrary Boolean function. Define function

$$F(x) := f(\text{sign}(p_1(x)), \dots, \text{sign}(p_k(x))) .$$

Let Y be a $2dR$ -wise independent standard Gaussian vector of length n . For any $x \in \mathbb{R}^n$ and $\lambda = O(k^{-2}d^{-3}R^{-15}\epsilon^2)$

$$\left| \mathbb{E}_Y \left[F(x + \sqrt{\lambda}Y)G(x + \sqrt{\lambda}Y) \right] - \mathbb{E}_{y \sim \mathcal{N}(0,1)^n} \left[F(x + \sqrt{\lambda}y)G(x + \sqrt{\lambda}y) \right] \right| = kd2^{-\Omega(R)} ,$$

where G is defined in (1).

Proof. Let $\phi_i(x) = U_{\sqrt{1-\lambda}} p_i\left(\frac{x}{\sqrt{1-\lambda}}\right)$. Define x is good if for any $1 \leq i \leq k$ and $0 \leq t \leq d-1$, $\|\nabla^{t+1}\phi_i(x)\| \leq \frac{1}{\epsilon} \|\nabla^t\phi_i(x)\|$. We prove this lemma by considering x is good or not.

We first consider that x is not good. In this case, we will show $G(x + \sqrt{\lambda}y) = 0$ holds with high probability. Consequently, $F(x + \sqrt{\lambda}y)G(x + \sqrt{\lambda}y)$ is zero with high probability. To this end, it suffices to find an i_0 and a t_0 such that

$$\left\| \nabla^{t_0+1} p_{i_0}(x + \sqrt{\lambda}y) \right\| > \frac{1}{4\epsilon} \left\| \nabla^{t_0} p_{i_0}(x + \sqrt{\lambda}y) \right\| .$$

We choose an arbitrary i_0 satisfying that there exists $0 \leq t \leq d-1$ such that $\|\nabla^{t+1}\phi_{i_0}(x)\| > \frac{1}{\epsilon} \|\nabla^t\phi_{i_0}(x)\|$. Since x is not good, we know such i_0 exists. And let t_0 be the largest t such that $\|\nabla^{t+1}\phi_{i_0}(x)\| > \frac{1}{\epsilon} \|\nabla^t\phi_{i_0}(x)\|$ holds. It is not hard to check

- $\|\nabla^{t_0}\phi_{i_0}(x)\| < \epsilon \|\nabla^{t_0+1}\phi_{i_0}(x)\|$,
- $\|\nabla^{t+1}\phi_{i_0}(x)\| \leq \frac{1}{\epsilon} \|\nabla^t\phi_{i_0}(x)\|$ for $t \geq t_0 + 1$.

We are next to prove the following inequalities hold with high probability,

$$(a) \left\| \nabla^{t_0} p_{i_0}(x + \sqrt{\lambda}y) \right\| < 2\epsilon \|\nabla^{t_0+1}\phi_{i_0}(x)\|$$

$$(b) \left\| \nabla^{t_0+1} p_{i_0}(x + \sqrt{\lambda}y) \right\| > \frac{1}{2} \|\nabla^{t_0+1}\phi_{i_0}(x)\|$$

It is easy to see (a) and (b) give us $\left\| \nabla^{t_0+1} p_{i_0}(x + \sqrt{\lambda}y) \right\| > \frac{1}{4\epsilon} \left\| \nabla^{t_0} p_{i_0}(x + \sqrt{\lambda}y) \right\|$.

Showing (a). By Markov's inequality, we have

$$\begin{aligned} & \Pr_y \left[\left\| \nabla^{t_0} p_{i_0}(x + \sqrt{\lambda}y) - \nabla^{t_0} \phi_{i_0}(x) \right\| \geq \epsilon \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\| \right] \\ & \leq \frac{\mathbb{E}_y \left[\left\| \nabla^{t_0} p_{i_0}(x + \sqrt{\lambda}y) - \nabla^{t_0} \phi_{i_0}(x) \right\|^R \right]}{\epsilon^R \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\|^R} \leq \frac{\left(\sum_{t=t_0+1}^d (\lambda dR)^{t-t_0} \left\| \nabla^t \phi(x) \right\|^2 \right)^{\frac{R}{2}}}{\epsilon^R \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\|^R} \\ & \leq \frac{\left(\sum_{t=t_0+1}^d (\lambda dR)^{t-t_0} \left(\frac{1}{\epsilon^2} \right)^{t-t_0-1} \left\| \nabla^{t_0+1} \phi(x) \right\|^2 \right)^{\frac{R}{2}}}{\epsilon^R \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\|^R} \leq \left(\sum_{t=1}^{d-t_0} \left(\frac{\lambda dR}{\epsilon^2} \right)^t \right)^{\frac{R}{2}}. \end{aligned}$$

Here the second inequality is from Lemma 15. The third inequality uses the condition that $\left\| \nabla^{t+1} \phi_{i_0}(x) \right\| \leq \frac{1}{\epsilon} \left\| \nabla^t \phi_{i_0}(x) \right\|$ for $t \geq t_0 + 1$. Since $\lambda \leq \frac{\epsilon^2}{100dR}$, this probability is bounded by 2^{-R} . Therefore, with probability at least $1 - 2^{-R}$,

$$\left\| \nabla^{t_0} p_{i_0}(x + \sqrt{\lambda}y) - \nabla^{t_0} \phi_{i_0}(x) \right\| < \epsilon \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\|.$$

Moreover, we know $\left\| \nabla^{t_0} \phi_{i_0}(x) \right\| < \epsilon \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\|$. So, we have with probability at least $1 - 2^{-R}$,

$$\left\| \nabla^{t_0} p_{i_0}(x + \sqrt{\lambda}y) \right\| \leq \left\| \nabla^{t_0} \phi_{i_0}(x) \right\| + \left\| \nabla^{t_0} p_{i_0}(x + \sqrt{\lambda}y) - \nabla^{t_0} \phi_{i_0}(x) \right\| < 2\epsilon \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\|.$$

Showing (b). Similarly, we have

$$\begin{aligned} & \Pr_y \left[\left\| \nabla^{t_0+1} p_{i_0}(x + \sqrt{\lambda}y) - \nabla^{t_0+1} \phi_{i_0}(x) \right\| \geq \frac{1}{2} \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\| \right] \\ & \leq \frac{\mathbb{E}_y \left[\left\| \nabla^{t_0+1} p_{i_0}(x + \sqrt{\lambda}y) - \nabla^{t_0+1} \phi_{i_0}(x) \right\|^R \right]}{2^{-R} \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\|^R} \leq \frac{\left(\sum_{t=t_0+2}^d (\lambda dR)^{t-t_0-1} \left\| \nabla^t \phi(x) \right\|^2 \right)^{\frac{R}{2}}}{2^{-R} \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\|^R} \\ & \leq \frac{\left(\sum_{t=t_0+2}^d \left(\frac{\lambda dR}{\epsilon^2} \right)^{t-t_0-1} \left\| \nabla^{t_0+1} \phi(x) \right\|^2 \right)^{\frac{R}{2}}}{2^{-R} \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\|^R} \leq \left(4 \cdot \sum_{t=1}^{d-t_0-1} \left(\frac{\lambda dR}{\epsilon^2} \right)^t \right)^{\frac{R}{2}}. \end{aligned}$$

Here the second inequality is from Lemma 15. The third inequality uses the condition that $\left\| \nabla^{t+1} \phi_{i_0}(x) \right\| \leq \frac{1}{\epsilon} \left\| \nabla^t \phi_{i_0}(x) \right\|$ for $t \geq t_0 + 1$. Since $\lambda \leq \frac{\epsilon^2}{100dR}$, this probability is bounded by 2^{-R} . Therefore, with probability at least $1 - 2^{-R}$,

$$\left\| \nabla^{t_0+1} p_{i_0}(x + \sqrt{\lambda}y) \right\| \geq \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\| - \left\| \nabla^{t_0+1} p_{i_0}(x + \sqrt{\lambda}y) - \nabla^{t_0+1} \phi_{i_0}(x) \right\| > \frac{1}{2} \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\|.$$

Thus, combining (a) and (b), we have that with probability at least $1 - 2 \cdot 2^{-R}$,

$$\left\| \nabla^{t_0+1} p_{i_0}(x + \sqrt{\lambda}y) \right\| > \frac{1}{2} \left\| \nabla^{t_0+1} \phi_{i_0}(x) \right\| > \frac{1}{4\epsilon} \left\| \nabla^{t_0} p_{i_0}(x + \sqrt{\lambda}y) \right\|,$$

and consequently $G(x + \sqrt{\lambda}y) = 0$. This gives us the bound on the following expectation

$$\mathbb{E}_{y \sim \mathcal{N}(0,1)^n} \left[F(x + \sqrt{\lambda}y) G(x + \sqrt{\lambda}y) \right] \leq 2 \cdot 2^{-R}.$$

Since Y is dR -wise independent, the above argument still holds for Y . So,

$$\left| \mathbb{E}_Y \left[F(x + \sqrt{\lambda}Y) G(x + \sqrt{\lambda}Y) \right] - \mathbb{E}_{y \sim \mathcal{N}(0,1)^n} \left[F(x + \sqrt{\lambda}y) G(x + \sqrt{\lambda}y) \right] \right| \leq 4 \cdot 2^{-R}.$$

134:14 A PRG for Functions of Low-Degree PTFs

Now suppose that x is good. That is, for any $1 \leq i \leq k$ and $0 \leq t \leq d-1$, $\|\nabla^{t+1}\phi_i(x)\| \leq \frac{1}{\epsilon} \|\nabla^t\phi_i(x)\|$. In this case, we will prove that the sign of $p_i(x + \sqrt{\lambda}y)$ is the same as the sign of $\phi_i(x)$ with high probability over the random variable y . Therefore, $F(x + \sqrt{\lambda}y)$ is almost like a constant, since the value of $F(x + \sqrt{\lambda}y)$ only depends on the signs of all $p_i(x + \sqrt{\lambda}y)$. To show the signs of $p_i(x + \sqrt{\lambda}y)$ and $\phi_i(x)$ are the same, it suffices to show $\frac{p_i(x + \sqrt{\lambda}y)}{\phi_i(x)} > 0$. Let

$$q_i(y) := \frac{p_i(x + \sqrt{\lambda}y)}{\phi_i(x)} - 1 = \frac{1}{\phi_i(x)} \sum_{0 < |\alpha| \leq d} \frac{\partial^\alpha \phi_i(x)}{\sqrt{\alpha!}} \lambda^{|\alpha|/2} h_\alpha(y) .$$

Here, we expand $p_i(x + \sqrt{\lambda}y) = \phi_i(x) + \sum_{0 < |\alpha| \leq d} \frac{\partial^\alpha \phi_i(x)}{\sqrt{\alpha!}} \lambda^{|\alpha|/2} h_\alpha(y)$ according to Lemma 10. We have by applying hypercontractive inequality in Theorem 11,

$$\begin{aligned} \|q_i(y)\|_R &\leq \|U_{\sqrt{R}} q_i(y)\|_2 = \left\| \frac{1}{\phi_i(x)} \sum_{0 < |\alpha| \leq d} \frac{\partial^\alpha \phi_i(x)}{\sqrt{\alpha!}} R^{|\alpha|/2} \lambda^{|\alpha|/2} h_\alpha(y) \right\|_2 \\ &\leq \sqrt{\sum_{0 < t \leq d} \frac{\|\nabla^t \phi_i(x)\|^2}{\phi_i(x)^2} (\lambda R)^t} \leq \sqrt{\sum_{0 < t \leq d} \left(\frac{\lambda R}{\epsilon^2}\right)^t} . \end{aligned}$$

In the last inequality, we use our assumption that $\|\nabla^t \phi_i(x)\| \leq \frac{1}{\epsilon} \|\nabla^{t-1} \phi_i(x)\| \leq \dots \leq \frac{|\phi_i(x)|}{\epsilon^t}$. Since $\frac{\lambda R}{\epsilon^2}$ is sufficiently small, we have $\|q_i(y)\|_R \leq \frac{1}{4}$. Therefore, by Markov's inequality, we have $\Pr_y [|q_i(y)| \geq \frac{1}{2}] \leq 2^R \cdot \|q_i(y)\|_R^R \leq 2^{-R}$. This means that with probability at least $1 - 2^{-R}$, we have $\left| \frac{p_i(x + \sqrt{\lambda}y)}{\phi_i(x)} - 1 \right| \leq \frac{1}{2}$, and therefore $\frac{p_i(x + \sqrt{\lambda}y)}{\phi_i(x)} \geq \frac{1}{2}$. Thus, with probability at least $1 - 2^{-R}$, the sign of $p_i(x + \sqrt{\lambda}y)$ is the same as the sign of $\phi_i(x)$. Then by a union bound,

$$\Pr_y \left[\forall 1 \leq i \leq k, \text{sign}(p_i(x + \sqrt{\lambda}y)) = \text{sign}(\phi_i(x)) \right] \geq 1 - k \cdot 2^{-R} .$$

Let $c = f(\text{sign}(\phi_1(x)), \dots, \text{sign}(\phi_k(x)))$ be a constant. We have

$$\left| \mathbb{E}_y [F(x + \sqrt{\lambda}y)G(x + \sqrt{\lambda}y)] - \mathbb{E}_y [c \cdot G(x + \sqrt{\lambda}y)] \right| \leq k \cdot 2^{-R} .$$

The same argument applying on $2dR$ -wise independent Gaussian vector Y gives us

$$\left| \mathbb{E}_Y [F(x + \sqrt{\lambda}Y)G(x + \sqrt{\lambda}Y)] - \mathbb{E}_Y [c \cdot G(x + \sqrt{\lambda}Y)] \right| \leq k \cdot 2^{-R} .$$

By Lemma 17, we know $\left| \mathbb{E}_Y [G(x + \sqrt{\lambda}Y)] - \mathbb{E}_{y \sim \mathcal{N}(0,1)^n} [G(x + \sqrt{\lambda}y)] \right| = kd \cdot 2^{-\Omega(R)}$. Therefore, we have

$$\left| \mathbb{E}_Y [F(x + \sqrt{\lambda}Y)G(x + \sqrt{\lambda}Y)] - \mathbb{E}_{y \sim \mathcal{N}(0,1)^n} [F(x + \sqrt{\lambda}y)G(x + \sqrt{\lambda}y)] \right| = kd 2^{-\Omega(R)} . \quad \blacktriangleleft$$

3.3 Proof of Theorem 16

Proof of Theorem 16. By Lemma 14, the following holds with probability at least $1 - \epsilon kd^3$:

$$\|\nabla^t p_i(y)\| \leq O\left(\frac{1}{\epsilon}\right) \|\nabla^{t-1} p_i(y)\| \text{ for all } 1 \leq t \leq d \text{ and } 1 \leq i \leq k .$$

Recall the function $G(x) = \prod_{i=1}^k \prod_{t=0}^{d-1} \rho \left(\log \left(\frac{\|\nabla^t p_i(x)\|^2}{16\epsilon^2 \|\nabla^{t+1} p_i(x)\|^2} \right) \right)$ as defined in (1). Note that $G(x) = 0$ if there exists some i and t such that $\|\nabla^t p_i(y)\| > O\left(\frac{1}{\epsilon}\right) \|\nabla^{t-1} p_i(y)\|$. Thus, $\Pr_y[G(y) = 1] \geq 1 - O(\epsilon k d^3)$. We have

$$\begin{aligned} \mathbb{E}_y[F(y)] &= \mathbb{E}_y[F(y)(1 - G(y))] + \mathbb{E}_y[F(y)G(y)] \\ &\leq \mathbb{E}_y[1 - G(y)] + \mathbb{E}_Y[F(Y)G(Y)] + \left| \mathbb{E}_y[F(y)G(y)] - \mathbb{E}_Y[F(Y)G(Y)] \right| \\ &\leq O(\epsilon k d^3) + \mathbb{E}_Y[F(Y)] + \left| \mathbb{E}_y[F(y)G(y)] - \mathbb{E}_Y[F(Y)G(Y)] \right|. \end{aligned}$$

Let $y = \frac{1}{\sqrt{L}} \sum_{i=1}^L y_i$ where $y_i \sim \mathcal{N}(0, 1)^n$ and denote $Z^i = \frac{1}{\sqrt{L}}(y_1 + \dots + y_{i-1} + Y_{i+1} + \dots + Y_L)$. We have

$$\begin{aligned} &\left| \mathbb{E}_y[F(y)G(y)] - \mathbb{E}_Y[F(Y)G(Y)] \right| \\ &\leq \sum_{i=1}^L \left| \mathbb{E}_{Z^i, Y_i} \left[F \left(Z^i + \frac{1}{\sqrt{L}} Y_i \right) G \left(Z^i + \frac{1}{\sqrt{L}} Y_i \right) \right] - \mathbb{E}_{Z^i, y_i} \left[F \left(Z^i + \frac{1}{\sqrt{L}} y_i \right) G \left(Z^i + \frac{1}{\sqrt{L}} y_i \right) \right] \right| \\ &\leq k d L \cdot 2^{-\Omega(R)} \end{aligned}$$

where the last inequality is from Lemma 18. Thus,

$$\mathbb{E}_y[F(y)] \leq \mathbb{E}_Y[F(Y)] + O(\epsilon k d^3) + k d L \cdot 2^{-\Omega(R)}.$$

And the other side follows from considering $1 - F(x)$. ◀

4 Discretization

To give an explicit construction of a PRG, we need a discretization of R -wise independent Gaussian distributions. In this section, we show an algorithm which outputs L vectors $\{X_i\}_{1 \leq i \leq L}$ approximating Y_i , that is, $|X_{i,j} - Y_{i,j}|$ is sufficiently small. Before that, we first prove that if X and Y are close enough, then X also fools any function of low-degree polynomial threshold functions.

► **Lemma 19.** *Let $0 < \epsilon, \delta < 1$, and $R \in \mathbb{N}$ be an integer. Let $Y = \frac{1}{\sqrt{L}} \sum_{i=1}^L Y_i$ where Y_i is an R -wise independent Gaussian vector of length n for $1 \leq i \leq L$. Let $p_1, \dots, p_k : \mathbb{R}^n \rightarrow \mathbb{R}$ be arbitrary polynomials of degree d and $f : \{0, 1\}^k \rightarrow \{0, 1\}$ be an arbitrary Boolean function. Define functions*

$$F(x) := f(\text{sign}(p_1(x)), \dots, \text{sign}(p_k(x)))$$

Suppose that for any such function F ,

$$\left| \mathbb{E}_Y[F(Y)] - \mathbb{E}_{y \sim \mathcal{N}(0,1)^n}[F(y)] \right| \leq \epsilon.$$

Suppose that $\{X_i\}_{1 \leq i \leq L}$ are random vectors of length n and there is a joint distribution over X and Y such that for each $1 \leq i \leq L, 1 \leq j \leq n$, $\Pr[|X_{i,j} - Y_{i,j}| \leq \delta] \geq 1 - \delta$.

Let $X = \frac{1}{\sqrt{L}} \sum_{i=1}^L X_i$ and we have that for any such function F

$$\left| \mathbb{E}_X[F(X)] - \mathbb{E}_{y \sim \mathcal{N}(0,1)^n}[F(y)] \right| \leq \epsilon + k 2^{2k} d \delta^{1/d} \sqrt{nL} \log \frac{1}{\delta} + O(2^{2k} n L \delta).$$

134:16 A PRG for Functions of Low-Degree PTFs

Proof. Let $q_i(x) = p_i(x) + \delta(nL)^{d/2} (\log \frac{1}{\delta})^d$ and $\tilde{F}(x) = f(\text{sign}(q_1(x)), \dots, \text{sign}(q_k(x)))$. We will prove that

$$(a) \mathbb{E}[F(X)] \leq \mathbb{E}[\tilde{F}(y)] + \epsilon + O(2^{2k}nL\delta),$$

$$(b) \mathbb{E}[\tilde{F}(y)] \leq \mathbb{E}[F(y)] + k2^{2k}d\delta^{1/d}\sqrt{nL}\log \frac{1}{\delta}.$$

Combing (a) and (b), we have

$$\mathbb{E}[F(X)] \leq \mathbb{E}[F(y)] + k2^{2k}d\delta^{1/d}\sqrt{nL}\log \frac{1}{\delta} + \epsilon + O(2^{2k}nL\delta) .$$

The other side can be obtained in a similar way by considering $1 - F(x)$.

Proving (a). Since \tilde{F} is a function of degree- d PTFs and Y fools such a function, we have $\mathbb{E}[\tilde{F}(Y)] \leq \mathbb{E}[\tilde{F}(y)] + \epsilon$. Therefore, it suffices to prove that $\mathbb{E}[F(X)] \leq \mathbb{E}[\tilde{F}(Y)] + O(2^{2k}nL\delta)$.

Fix a set $S \subseteq [k]$. Let $P_S(x) = \prod_{i \in S} \text{sign}(p_i(x))$ and $Q_S(x) = \prod_{i \in S} \text{sign}(q_i(x))$. We first show that

$$\mathbb{E}[P_S(X)] \leq \mathbb{E}[Q_S(Y)] + O(nL\delta) .$$

Let event \mathcal{E} denote that for all $i \in [L], j \in [n], |Y_{i,j}| \leq \log \frac{1}{\delta}$ and $|X_{i,j} - Y_{i,j}| \leq \delta$. By the tail bound of the standard Gaussian distribution, we have $\Pr[\mathcal{E}] \geq 1 - O(nL\delta)$. We have

$$\begin{aligned} \mathbb{E}[P_S(X)] &= \Pr \left[\bigwedge_{i \in S} p_i(X) \geq 0 \right] \leq \Pr \left[\bigwedge_{i \in S} p_i(X) \geq 0 \wedge \mathcal{E} \right] + \Pr [\bar{\mathcal{E}}] \\ &\leq \Pr \left[\bigwedge_{i \in S} p_i(Y) \geq -\delta(nL)^{d/2} \left(\log \frac{1}{\delta} \right)^d \right] + O(nL\delta) \\ &= \mathbb{E}[Q_S(Y)] + O(nL\delta) , \end{aligned}$$

where the second inequality is by Lemma 13 and viewing $p_i\left(\frac{1}{\sqrt{L}} \sum_{i=1}^L X_i\right)$ as a degree d function of nL variables.

Note that $f(x) = \sum_{S \subseteq [k]} c_S \mathbb{1}_S \prod_{i \in S} x_i \prod_{i \notin S} (1 - x_i)$ where $\mathbb{1}_S$ denotes the length- k string with 1's only at coordinates in S . This can be further simplified in the form

$$f(x) = \sum_{S \subseteq [k]} c_S \prod_{i \in S} x_i ,$$

with $\sum_S |c_S| \leq 2^{2k}$. So, we have

$$\begin{aligned} \mathbb{E}[F(X)] &= \sum_{S \subseteq [k]} c_S \mathbb{E}[P_S(X)] = \sum_{S \subseteq [k]} c_S \mathbb{E}[Q_S(Y)] + \sum_{S \subseteq [k]} c_S (\mathbb{E}[P_S(X)] - \mathbb{E}[Q_S(Y)]) \\ &\leq \sum_{S \subseteq [k]} c_S \mathbb{E}[Q_S(Y)] + O(2^{2k}nL\delta) = \mathbb{E}[\tilde{F}(Y)] + O(2^{2k}nL\delta) . \end{aligned}$$

Proving (b). Next we show $\mathbb{E}[\tilde{F}(y)]$ and $\mathbb{E}[F(y)]$ are close. Note that

$$\mathbb{E}[Q_S(y)] = \Pr \left[\bigwedge_{i \in S} p_i(y) \geq -\delta(nL)^{d/2} \left(\log \frac{1}{\delta} \right)^d \right] \leq \Pr \left[\bigwedge_{i \in S} p_i(y) \geq 0 \right] + kd\delta^{1/d}\sqrt{nL}\log \frac{1}{\delta} ,$$

where the inequality is from Lemma 12. Thus, we have

$$\mathbb{E}[Q_S(y)] \leq \mathbb{E}[P_S(y)] + kd\delta^{1/d}\sqrt{nL} \log \frac{1}{\delta} .$$

Furthermore, we know that

$$\begin{aligned} \mathbb{E}[\tilde{F}(y)] &= \sum_{S \subseteq [k]} c_S \mathbb{E}[Q_S(y)] = \sum_{S \subseteq [k]} c_S \mathbb{E}[P_S(y)] + \sum_{S \subseteq [k]} c_S \left(\mathbb{E}[Q_S(y)] - P_S(y) \right) \\ &\leq \mathbb{E}[F(y)] + k2^{2k}d\delta^{1/d}\sqrt{nL} \log \frac{1}{\delta} . \end{aligned} \blacktriangleleft$$

We now prove the main theorem for constructing an explicit pseudorandom generator. The idea is that a standard Gaussian variable can be generated using two uniform $[0, 1]$ random variables through the Box–Muller transform [2]. Let $Y_{i,j} = \sqrt{-2 \log u_{i,j}} \cos(2\pi v_{i,j})$ where $u_{i,j}$ and $v_{i,j}$ are uniform in $[0, 1]$. Then $Y_{i,j}$ is a Gaussian variable. Thus, if we truncate $u_{i,j}$ and $v_{i,j}$ to a certain precision and produce $X_{i,j}$ in a similar manner, X approximates Y with high probability.

► **Theorem 20.** *There exists an explicit PRG which ϵ -fools any functions of any k degree- d polynomial threshold functions over $\mathcal{N}(0, 1)^n$ with seed length $O\left(\frac{k^5 d^{11}}{\epsilon^2} \log \frac{kdn}{\epsilon}\right)$.*

Proof. In Theorem 16, set parameter ϵ as $\frac{\epsilon}{kd^3}$ and set R as $C \log \frac{kd}{\epsilon}$ for some large constant C . Then for $L = C' \cdot \frac{k^4 d^9}{\epsilon^2} \cdot \text{polylog} \frac{kd}{\epsilon}$ where C' is a large constant, we have

$$\left| \mathbb{E}_Y[F(Y)] - \mathbb{E}_{y \sim \mathcal{N}(0,1)^n}[F(y)] \right| \leq O(\epsilon) .$$

Similar to the proof of Corollary 2 in [13], we can let $Y_{i,j}$ generated by

$$Y_{i,j} = \sqrt{-2 \log u_{i,j}} \cos(2\pi v_{i,j})$$

where $u_i = (u_{i,1}, \dots, u_{i,n})$ and $v_i = (v_{i,1}, \dots, v_{i,n})$ are $2dR$ -wise independent uniform $[0, 1]$ random vectors. Then let $u'_{i,j}$ and $v'_{i,j}$ be $M = C''kd \log \frac{kdn}{\epsilon}$ -bit approximations to $u_{i,j}$ and $v_{i,j}$ (i.e., round $u_{i,j}$ and $v_{i,j}$ to multiples of 2^{-M}), where C'' is a large constant, and let

$$X_{i,j} = \sqrt{-2 \log u'_{i,j}} \cos(2\pi v'_{i,j}) .$$

Letting $\delta = \Omega(2^{-M/2})$, for the same reason as in the proof of Corollary 2 in [13], we have $|X_{i,j} - Y_{i,j}| < \delta$ with probability at least $1 - \delta$. Then, by Lemma 19,

$$\left| \mathbb{E}_X[F(X)] - \mathbb{E}_{y \sim \mathcal{N}(0,1)^n}[F(y)] \right| \leq O(\epsilon) + k2^{2k}d\delta^{1/d}\sqrt{nL} \log \frac{1}{\delta} + O(2^{2k}nL\delta) = O(\epsilon) .$$

Note that X_i can be generated by $2dR$ -wise independent random variables $u'_{i,j}$ and $v'_{i,j}$ taken uniformly from $\{2^{-M}, 2 \cdot 2^{-M}, 3 \cdot 2^{-M}, \dots, 1\}$ using $O(dRM)$ randomness. Thus generating X uses $O(LdRM) = O\left(\frac{k^5 d^{11}}{\epsilon^2} \log \frac{kdn}{\epsilon}\right)$ randomness. \blacktriangleleft

References

- 1 Louay M. J. Bazzi. Polylogarithmic independence can fool DNF formulas. *SIAM Journal on Computing*, 38(6):2220–2272, 2009. doi:10.1137/070691954.
- 2 G. E. P. Box and Mervin E. Muller. A note on the generation of random normal deviates. *The Annals of Mathematical Statistics*, 29(2):610–611, 1958. doi:10.1214/aoms/1177706645.

- 3 Anthony Carbery and James Wright. Distributional and L^q norm inequalities for polynomials over convex bodies in \mathbb{R}^n . *Mathematical Research Letters*, 8:233–248, 2001. doi:10.4310/MRL.2001.v8.n3.a1.
- 4 Eshan Chattopadhyay, Anindya De, and Rocco A. Servedio. Simple and efficient pseudorandom generators from Gaussian processes. In *Proceedings of the 34th Computational Complexity Conference*, Dagstuhl, DEU, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CCC.2019.4.
- 5 Ilias Diakonikolas, Parikshit Gopalan, Ragesh Jaiswal, Rocco A. Servedio, and Emanuele Viola. Bounded independence fools halfspaces. *SIAM Journal on Computing*, 39(8):3441–3462, 2010. doi:10.1137/100783030.
- 6 Ilias Diakonikolas, Daniel M. Kane, and Jelani Nelson. Bounded independence fools degree-2 threshold functions. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 11–20, 2010. doi:10.1109/FOCS.2010.8.
- 7 Parikshit Gopalan, Daniel M. Kane, and Raghu Meka. Pseudorandomness via the discrete fourier transform. *SIAM Journal on Computing*, 47(6):2451–2487, 2018. doi:10.1137/16M1062132.
- 8 Parikshit Gopalan, Ryan O’Donnell, Yi Wu, and David Zuckerman. Fooling functions of half-spaces under product distributions. In *2010 IEEE 25th Annual Conference on Computational Complexity*, pages 223–234, 2010. doi:10.1109/CCC.2010.29.
- 9 Prahladh Harsha, Adam Klivans, and Raghu Meka. An invariance principle for polytopes. *J. ACM*, 59(6), 2013. doi:10.1145/2395116.2395118.
- 10 William B Johnson, Joram Lindenstrauss, and Gideon Schechtman. Extensions of Lipschitz maps into Banach spaces. *Israel Journal of Mathematics*, 54(2):129–138, 1986. doi:10.1007/BF02764938.
- 11 Daniel Kane, Raghu Meka, and Jelani Nelson. Almost optimal explicit johnson-lindenstrauss families. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 628–639, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg. doi:10.1007/978-3-642-22935-0_53.
- 12 Daniel M. Kane. k -independent Gaussians fool polynomial threshold functions. In *2011 IEEE 26th Annual Conference on Computational Complexity*, pages 252–261, 2011. doi:10.1109/CCC.2011.13.
- 13 Daniel M. Kane. A small PRG for polynomial threshold functions of Gaussians. In *Proceedings of the 2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 257–266, USA, 2011. IEEE Computer Society. doi:10.1109/FOCS.2011.16.
- 14 Daniel M. Kane. A structure theorem for poorly anticoncentrated Gaussian chaos and applications to the study of polynomial threshold functions. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 91–100, 2012. doi:10.1109/FOCS.2012.52.
- 15 Daniel M. Kane. A pseudorandom generator for polynomial threshold functions of Gaussian with subpolynomial seed length. In *2014 IEEE 29th Conference on Computational Complexity*, pages 217–228, 2014. doi:10.1109/CCC.2014.30.
- 16 Daniel M. Kane. A Polylogarithmic PRG for Degree 2 Threshold Functions in the Gaussian Setting. In David Zuckerman, editor, *30th Conference on Computational Complexity (CCC 2015)*, volume 33 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 567–581, Dagstuhl, Germany, 2015. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CCC.2015.567.
- 17 Zander Kelley and Raghu Meka. Random restrictions and PRGs for PTFs in Gaussian space. In Shachar Lovett, editor, *37th Computational Complexity Conference (CCC 2022)*, volume 234 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 21:1–21:24, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CCC.2022.21.

- 18 Adam R. Klivans, Ryan O'Donnell, and Rocco A. Servedio. Learning geometric concepts via Gaussian surface area. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 541–550, 2008. doi:10.1109/FOCS.2008.64.
- 19 Pravesh K. Kothari and Raghu Meka. Almost optimal pseudorandom generators for spherical caps: extended abstract. In *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, pages 247–256, New York, NY, USA, 2015. Association for Computing Machinery. doi:10.1145/2746539.2746611.
- 20 J. W. Lindeberg. Eine neue herleitung des exponentialgesetzes in der wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 15:211–225, 1922. doi:10.1007/BF01494395.
- 21 Raghu Meka and David Zuckerman. Pseudorandom generators for polynomial threshold functions. *SIAM Journal on Computing*, 42(3):1275–1301, 2013. doi:10.1137/100811623.
- 22 Jorge Nocedal and Stephen J. Wright, editors. *Quadratic Programming*, pages 438–486. Springer New York, New York, NY, 1999. doi:10.1007/0-387-22742-3_16.
- 23 Ryan O'Donnell. *Analysis of boolean functions*. Cambridge University Press, 2014. doi:10.1017/CB09781139814782.
- 24 Ryan O'Donnell, Rocco A. Servedio, and Li-Yang Tan. Fooling Gaussian PTFs via local hyperconcentration. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1170–1183, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3357713.3384281.
- 25 Ryan O'Donnell, Rocco A. Servedio, and Li-Yang Tan. Fooling polytopes. *J. ACM*, 69(2), January 2022. doi:10.1145/3460532.
- 26 Alexander Razborov. A simple proof of Bazzi's theorem. *ACM Trans. Comput. Theory*, 1(1), February 2009. doi:10.1145/1490270.1490273.
- 27 R. A. Servedio and L. Tan. Fooling intersections of low-weight halfspaces. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 824–835, Los Alamitos, CA, USA, October 2017. IEEE Computer Society. doi:10.1109/FOCS.2017.81.
- 28 Rocco A. Servedio. Every linear threshold function has a low-weight approximator. In *21st Annual IEEE Conference on Computational Complexity (CCC'06)*, pages 18–32, 2006. doi:10.1109/CCC.2006.18.

A Facts about Bump Function

► **Fact 7** For all $t \in \mathbb{N}$, $|\Psi^{(t)}(x)| \leq t^{(3+o(1))t}$.

Proof. It is easy to check there exists a series of polynomials $\{P_t\}_{t \in \mathbb{N}}$ such that for $x \in (-1, 1)$

$$\Psi^{(t)}(x) = \frac{P_t(x)}{(1-x^2)^{2t}} \cdot \Psi(x) .$$

Besides, $\{P_t\}_{t \in \mathbb{N}}$ has the following recursion:

$$P_0(x) = 1, \quad P_1(x) = -2x, \quad P_t = (1-x^2)^2 P'_{t-1}(x) + 4(t-1)x(1-x^2)P_{t-1}(x) - 2xP_{t-1}(x) .$$

The degree of P_t is at most $3t$. Therefore we have

$$\left| \Psi^{(t)}(x) \right| \leq \left(\max_{x \in (-1,1)} |P_t(x)| \right) \cdot \left(\max_{x \in (-1,1)} \frac{\Psi(x)}{(1-x^2)^{2t}} \right) .$$

Let $f(x) = x \cdot e^{-\frac{x}{2t}}$ for $x \in (1, +\infty)$. We have $f(x) \leq f(2t) = \frac{2t}{e}$ by a simple calculation. Thus, we have that $\frac{\Psi(x)}{(1-x^2)^{2t}} \leq \left(\frac{2t}{e}\right)^{2t}$. We are left to bound $\max_{x \in (-1,1)} |P_t(x)|$.

Define $\|P\|_1$ be the sum of absolute values of all coefficients of the polynomial P . Since $x \in (-1, 1)$, we know $\max_{x \in (-1,1)} |P_t(x)| \leq \|P_t\|_1$. By the recursion, we have

$$\|P_t\|_1 \leq 4 \|P'_{t-1}\|_1 + 8t \|P_{t-1}\|_1 \leq 20t \|P_{t-1}\|_1$$

134:20 A PRG for Functions of Low-Degree PTFs

where the last inequality is from $\|P'_{t-1}\|_1 \leq 3t \|P_{t-1}\|_1$ since the degree of P_{t-1} is at most $3t - 3$. So we know $\max_{x \in (-1,1)} \leq 20^t \cdot t!$. Therefore,

$$\left| \Psi^{(t)}(x) \right| \leq \left(\frac{2t}{e} \right)^{2t} \cdot 20^t \cdot t! = t^{(3+o(1))t} . \quad \blacktriangleleft$$

► **Fact 8** For all $t \in \mathbb{N}$, $|\rho^{(t)}(x)| \leq t^{(3+o(1))t}$.

Proof. Directly from Fact 7 by observing that $\rho(x) = e \cdot \Psi(1 - x)$ for $0 < x < 1$ and is constant elsewhere. ◀

► **Fact 9** Let $r(u, v) := \rho(\log u - \log v + c)$ for some constant c . Then we have that for all $n, m \in \mathbb{N}$, $\left| \frac{\partial^n \partial^m r(u, v)}{\partial u^n \partial v^m} \right| \leq \frac{(n+m)^{6(n+m)}}{|u|^n |v|^m}$.

Proof. Let $g(u, v) = \log u - \log v + c$. Then by the generalized chain rule for the derivative of the composition of two functions (also known as Faà di Bruno's formula), we have

$$\begin{aligned} \frac{\partial^n \partial^m r(u, v)}{\partial u^n \partial v^m} &= \sum_{\substack{(a_1, \dots, a_n) \in \mathbb{N}^n \\ a_1 + 2 \cdot a_2 + \dots + n \cdot a_n = n}} \sum_{\substack{(b_1, \dots, b_m) \in \mathbb{N}^m \\ b_1 + 2 \cdot b_2 + \dots + m \cdot b_m = m}} \frac{n! m!}{\prod_{i=1}^n (i!)^{a_i} a_i! \prod_{i=1}^m (i!)^{b_i} b_i!} \\ &\quad \cdot \rho^{(a_1 + \dots + a_n + b_1 + \dots + b_m)}(g(u, v)) \cdot \prod_{i=1}^n \left(\frac{(-1)^i i!}{u^i} \right)^{a_i} \prod_{i=1}^m \left(\frac{(-1)^{i+1} i!}{v^i} \right)^{b_i} . \end{aligned}$$

Therefore

$$\left| \frac{\partial^n \partial^m r(u, v)}{\partial u^n \partial v^m} \right| \leq n^n \cdot m^m \cdot n! \cdot m! \cdot (m+n)^{4(m+n)} \cdot \frac{1}{|u|^n |v|^m} \leq \frac{(n+m)^{6(n+m)}}{|u|^n |v|^m} . \quad \blacktriangleleft$$