Electronic Colloquium on Computational Complexity, Report No. 202 (2018)

A stochastic calculus approach to the oracle separation of BQP and PH

Xinyu Wu [*]

October 19, 2018

Abstract

After presentations of the oracle separation of BQP and PH result [RT18], several people (e.g. Ryan O'Donnell, James Lee, Avishay Tal) suggested that the proof may be simplified by stochastic calculus. In this short note, we describe such a simplification.

1 Reduction to a Fourier bound

The main technical part of [RT18] shows that, for a Boolean function $f : \{-1,1\}^N \to \{-1,1\}$ computable by an AC^0 circuit, and a multivariate Gaussian distribution $\boldsymbol{Z} \in \mathbb{R}^N$,

$$|\,\mathbf{E}[f(\mathsf{trnc}(\boldsymbol{Z}))] - \mathbf{E}[f(\boldsymbol{U}_N)]| \leq O(\gamma \cdot \mathrm{polylog}(n)),$$

where γ is a bound on the (pairwise) covariance of the coordinates of $\boldsymbol{Z}$, trnc truncates $\boldsymbol{Z}$ so that the resulting random variable is within $[-1,1]^N$, and $\boldsymbol{U}_N$ is the uniform distribution over $\{-1,1\}^N$. The important condition here is that AC^0 has second level Fourier coefficients bounded by $\mathrm{polylog}(n)$, and that this holds under any restriction of the function.

Another natural way of viewing a multivariate Gaussian distribution is as the result of an N-dimensional Brownian motion stopped at a fixed time. We can also build the truncation into the stopping time. This allows us to use tools from stochastic calculus to analyze the distribution.

We first recall the definition of restrictions of Boolean functions.

Definition 1. Let $f : \{-1,1\}^N \to \mathbb{R}$ and let $\rho \in \{-1,1,*\}^N$. Let $\mathrm{free}(\rho)$ be the set of coordinates with $*$'s. We define the restriction of f by ρ as $f_\rho : \{-1,1\}^N \to \mathbb{R}$, and $f_\rho(x)$ is f evaluated at ρ with x replacing the $*$'s in ρ.[1]

Henceforth, we also identify Boolean functions $f : \{-1,1\}^N \to \mathbb{R}$ with their multilinear polynomial representations (or Fourier expansions)

$$f(x) = \sum_{|S| \subseteq [N]} \hat{f}(S) \prod_{i \in S} x_i.$$

[*] Computer Science Department, Carnegie Mellon University. xinyuw1@andrew.cmu.edu

[1] Although f_ρ's domain is $\{-1,1\}^N$, it only depends on the coordinates in $\mathrm{free}(\rho)$.

ISSN 1433-8092

We make some observations about Fourier coefficients. First, the Fourier coefficients of f_ρ satisfy $\widehat{f_\rho}(S) = 0$ for all $S \not\subseteq \mathrm{free}(\rho)$. We also have that

$$\hat{f}(S) = \partial_S f(0), \tag{1}$$

where $\partial_S = \prod_{i \in S} \partial_i$ and ∂_i is the usual calculus derivative. Further, because f is multilinear, for any $h \in \mathbb{R} \setminus \{0\}$ and any standard basis vector e_i we have

$$\partial_i f(x) = \frac{f(x + he_i) - f(x)}{h}. \tag{2}$$

The following lemma is similar to [CHLT18, Claim A.5], which first appeared in [BB18] and [CHHL18, Claim 3.3].

Lemma 1. *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a multilinear polynomial. For any $x \in [-1/2, 1/2]^N$, there exists a distribution $\mathcal{R}_x$ over restrictions $\boldsymbol{\rho} \in \{-1, 1, *\}^N$, such that for any $i, j \in [N]$,*

$$\partial_{ij} f(x) = 4 \mathop{\mathbf{E}}_{\boldsymbol{\rho} \sim \mathcal{R}_x} \left[\partial_{ij} f_{\boldsymbol{\rho}}(0) \right].$$

Proof. We define $\mathcal{R}_x$ as such: for each coordinate $i \in [N]$ we independently set $\boldsymbol{\rho}_i$ to be 1 with probability $\frac{1}{4} + \frac{x_i}{2}$, to be -1 with probability $\frac{1}{4} - \frac{x_i}{2}$, and to be $*$ with probability $\frac{1}{2}$.

Using that f is a multilinear polynomial, and that the coordinates are independent, we deduce that for any $y \in \mathbb{R}^N$, $f(x + y) = \mathbf{E}_{\boldsymbol{\rho} \sim \mathcal{R}_x}[f(2y)]$. Then, using Equation (2),

$$\begin{aligned} \partial_{ij} f(x) &= f(x + e_i + e_j) - f(x + e_i) - f(x + e_j) + f(x) \\ &= \mathop{\mathbf{E}}_{\boldsymbol{\rho} \sim \mathcal{R}_x} \left[f(2e_i + 2e_j) - f(2e_j) - f(2e_i) + f(0) \right] = 4 \mathop{\mathbf{E}}_{\boldsymbol{\rho} \sim \mathcal{R}_x} \left[\partial_{ij} f_{\boldsymbol{\rho}}(0) \right]. \end{aligned}$$ □

We now show the main result, which is a restatement of [CHLT18, Therorem A.7] and [RT18, Theorem 2.4].

Theorem 1. *Let $f : \{-1, 1\}^N \to \{-1, 1\}$ be a Boolean function, and let $t > 0$ such that for any restriction ρ,*

$$\sum_{\substack{S \subseteq [N] \\ |S| = 2}} |\widehat{f_\rho}(S)| \le t.$$

Let $\gamma > 0$ and let $\mathbf{X}$ be an N-dimensional Brownian motion with mean 0 and covariance matrix Σ, in the sense that $\mathbf{E}[(\mathbf{X}_t)_i] = 0$ for all $i \in [N]$, and $\mathbf{Cov}((\mathbf{X}_t - \mathbf{X}_s)_i, (\mathbf{X}_t - \mathbf{X}_s)_j) = (t - s)\Sigma_{ij}$. Further assume that $|\Sigma_{ij}| \le \gamma$ for $i \ne j$.

Let $\varepsilon > 0$ and define the stopping time

$$\tau := \min\{\varepsilon, \textit{first time that } \mathbf{X}_t \textit{ exits } [-1/2, 1/2]^N\}.$$

Then, identifying f with its multilinear expansion, we have

$$|\mathbf{E}[f(\mathbf{X}_\tau)] - \mathbf{E}[f(\boldsymbol{U}_n)]| \le 2\varepsilon\gamma t.$$

Proof. First, we note that $\mathbf{E}[f(\boldsymbol{U}_N)] = f(0)$. Next, let $\sigma = \Sigma^{1/2}$.[2] $\boldsymbol{X}$ satisfies the stochastic differential equation

$$d\boldsymbol{X}_t = \sigma d\boldsymbol{B}_t.$$

Note that $\boldsymbol{X}_\tau$ is always within $[-1/2, 1/2]^N$. We can apply Dynkin's formula[3][4]

$$\mathbf{E}[f(\boldsymbol{X}_\tau)] - f(0) = \mathbf{E}\left[\int_0^\tau \frac{1}{2}\sum_{i,j\in[N]} \Sigma_{ij}\partial_{ij} f(\boldsymbol{X}_s)\, ds\right].$$

Then, we upper bound $\tau \le \varepsilon$, and use that $\partial_{ii} f = 0$ for all $i \in [N]$ because f is multilinear, to get

$$\begin{aligned}
|\,\mathbf{E}[f(\boldsymbol{X}_\tau)] - f(0)| &\le \varepsilon\, \mathbf{E}\left[\sup_{s\in[0,\tau]}\left|\frac{1}{2}\sum_{i,j\in[N]}\Sigma_{ij}\partial_{ij} f(\boldsymbol{X}_s)\right|\right] \\
&\le \frac{\varepsilon\gamma}{2} \sup_{x\in[-1/2,1/2]^N} \sum_{i\ne j} |\partial_{ij} f(x)| \\
&= 2\varepsilon\gamma \sup_{x\in[-1/2,1/2]^N} \sum_{i\ne j} \left|\mathop{\mathbf{E}}_{\rho\sim\mathcal{R}_x}[\partial_{ij} f_\rho(0)]\right| && \text{(Lemma 1)} \\
&\le 2\varepsilon\gamma \sup_{x\in[-1/2,1/2]^N} \mathop{\mathbf{E}}_{\rho\sim\mathcal{R}_x}\left[\sum_{i\ne j} |\partial_{ij} f_\rho(0)|\right] \\
&\le 2\varepsilon\gamma \sup_{x\in[-1/2,1/2]^N} \mathop{\mathbf{E}}_{\rho\sim\mathcal{R}_x}\left[\sum_{\substack{S\subseteq \mathrm{free}(\rho)\\ |S|=2}} \left|\hat{f}_\rho(S)\right|\right] && \text{(Equation (1))} \\
&\le 2\varepsilon\gamma t.
\end{aligned}$$

□

2 Application to the oracle separation of BQP and PH

The distribution $\mathcal{D}$. Let $N = 2n$ where n is a power of 2 and

$$\Sigma := \begin{pmatrix} I_n & H_n \\ H_n & I_n \end{pmatrix},$$

where H_n is the Walsh–Hadamard matrix. Now we define $\boldsymbol{X}$ and τ as in Theorem 1, with $\varepsilon = 1/(8\ln N)$, and our distribution $\mathcal{D}$ will be the distribution defined by $\boldsymbol{X}_\tau$. At each time t, we can also look at $\boldsymbol{X}_t$ as a pair of random variables in $\mathbb{R}^n$, $(\boldsymbol{x}, \boldsymbol{y})$ such that $\boldsymbol{y}$ is the Hadamard transform of $\boldsymbol{x}$.

AC0 lower bound. From [Tal17, Theorem 37] there exists a universal constant c such that every function $f : \{-1,1\}^N \to \{-1,1\}$ computable by an AC^0 circuit with at most $(\ln N)^\ell$ gates and

[2] This exists since Σ is symmetric and positive definite.

[3] https://en.wikipedia.org/wiki/Dynkin%27s_formula

[4] This works for functions which are C^2, and not merely C_c^2, since $\boldsymbol{X}_\tau$ is bounded.

depth d satisfies

$$\sum_{\substack{S\subseteq[N]\\|S|=k}} |\hat{f}(S)| \le (c\cdot \ln^{\ell} N)^{(d-1)k}.$$

Since AC^0 is closed under restrictions, we can apply Theorem 1 with $\varepsilon = 1/(8\ln N)$ and $\gamma = \frac{1}{\sqrt{n}}$, to deduce that

$$|\,\mathbf{E}[f(\boldsymbol{X}_\tau)] - f(0)| \le \frac{\mathrm{polylog}\, N}{\sqrt{N}}.$$

Quantum algorithm. Finally, we show that a quantum algorithm can distinguish $\mathcal{D}$ from the uniform distribution. This is virtually identical to the argument in [RT18, Section 6], but we can again use some stochastic calculus tools on the stopping time built into the distribution. Using the Forrelation query algorithm, there is a quantum algorithm Q with inputs $x, y \in \{-1,1\}^n$ which accepts with probability $(1+\varphi(x,y))/2$, where

$$\varphi(x,y) := \frac{1}{n}\sum_{i,j\in[n]} x_i\cdot H_{ij}\cdot y_j.$$

We show the following, which is Claim 6.3 in [RT18], from which one can deduce the existence of a quantum algorithm distinguishing $\mathcal{D}$ from uniform with 1 query and running in time $O(\log N)$.

Proposition 1. $\mathbf{E}_{(\boldsymbol{x},\boldsymbol{y})\sim\mathcal{D}}[\varphi(\boldsymbol{x},\boldsymbol{y})] \ge \frac{\varepsilon}{4}$.

Proof. By the linearity of expectation and optional sampling theorem,

$$\begin{aligned}\mathop{\mathbf{E}}_{(\boldsymbol{x},\boldsymbol{y})\sim\mathcal{D}}[\varphi(\boldsymbol{x},\boldsymbol{y})] &= \frac{1}{n}\sum_{i,j\in[n]} H_{ij}\cdot \mathbf{E}[\boldsymbol{x}_i\cdot\boldsymbol{y}_j]\\ &= \frac{1}{n}\sum_{i,j\in[n]} H_{ij}\cdot\mathbf{E}[\tau]\cdot H_{ij} = \mathbf{E}[\tau].\end{aligned}$$

Since τ is bounded by ε,

$$\mathbf{E}[\tau] = \int_0^{\varepsilon} \mathbf{Pr}[\tau > t]\,dt \ge \frac{\varepsilon}{2}\mathbf{Pr}[\tau > \tfrac{\varepsilon}{2}].$$

If $\tau < \varepsilon$, it must be the case that the path exits $[-1/2,1/2]^N$ earlier than ε. Hence, we can upper bound

$$\mathbf{Pr}\big[\tau \le \tfrac{\varepsilon}{2}\big] \le N\cdot\mathbf{Pr}\big[\text{1st coordinate exits } \big[-\tfrac{1}{2},\tfrac{1}{2}\big] \text{ earlier than } \tfrac{\varepsilon}{2}\big].$$

Each coordinate of $\boldsymbol{X}$ is a standard 1D Brownian motion since $\Sigma_{ii} = 1$ for all i. An application of Doob's submartingale inequality tells us that, for a standard 1D Brownian motion $\boldsymbol{B}_t$,

$$\mathbf{Pr}\left[\sup_{0\le t\le\varepsilon/2} |\boldsymbol{B}_t| \ge \frac{1}{2}\right] \le 2e^{-1/4\varepsilon} = 2e^{-2\ln N} \le \frac{1}{2N} \quad \text{for } N \ge 4.$$

Therefore, $\mathbf{Pr}[\tau \le \frac{\varepsilon}{2}] \le \frac{1}{2}$, so $\mathbf{E}[\tau] \ge \frac{\varepsilon}{4}$. □

3 Acknowledgments

I would like to thank Ryan O'Donnell and Avishay Tal for helpful discussions and their suggestions concerning this work.

References

[BB18] Boaz Barack and Jarosław Błasiok. On the Raz-Tal oracle separation of BQP and PH. https://windowsontheory.org/2018/06/17/on-the-raz-tal-oracle-separation-of-bqp-and-ph/, 2018.

[CHHL18] Eshan Chattopadhyay, Pooya Hatami, Kaave Hosseini, and Shachar Lovett. Pseudorandom generators from polarizing random walks. 2018.

[CHLT18] Eshan Chattopadhyay, Pooya Hatami, Shachar Lovett, and Avishay Tal. Pseudorandom generators from the second Fourier level and applications to AC0 with parity gates. Technical Report TR18-155, Electronic Colloquium on Computational Complexity, 2018.

[RT18] Ran Raz and Avishay Tal. Oracle separation of BQP and PH. Technical Report 107, Electronic Colloquium on Computational Complexity, 2018.

[Tal17] Avishay Tal. Tight bounds on the Fourier spectrum of AC0. 2017.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il