Quantum Learning of Concentrated Boolean Functions

Authors: Krishna Palem, Duc H. Pham, M. V. Panduranga Rao

Paper Review:
Manish Kumar

Quantum Technology
M.Tech. IISc Bengaluru

December 20, 2023

Table of Contents

1. Main Results
   - Context & Key Results
2. Brief Review of key concepts
   - PAC learning: Brief Review
   - Boolean Function & Fourier Series: Brief Review
3. Algorithm for Quantum PAC learning
   - CBF learning: A two stage process
   - Stage-I: Learning the concentration of a CBF
   - Satge-II: Learning the function from its concentration
4. Benchmarking
   - Complexity: Quantum and Classical PAC learning
   - Theoretical limits on Quantum learning
5. Conclusion

Table of Contents

1. Main Results
   - Context & Key Results
2. Brief Review of key concepts
   - PAC learning: Brief Review
   - Boolean Function & Fourier Series: Brief Review
3. Algorithm for Quantum PAC learning
   - CBF learning: A two stage process
   - Stage-I: Learning the concentration of a CBF
   - Satge-II: Learning the function from its concentration
4. Benchmarking
   - Complexity: Quantum and Classical PAC learning
   - Theoretical limits on Quantum learning
5. Conclusion

Learning Concentrated Boolean Function (CBF)

Objective: Is quantum speedup possible in learning Concentrated Boolean function via (i) PAC or (ii) Exact learning methods?

Learning CBF via Quantum PAC learner: **Results**

- There exists a poly-time quantum algorithms to **PAC learn** a CBF with (asymptotically) fewer query than the best-known classical algorithm. (due to Theorem-3)
- Source of Speedup: Sparse fourier sampling, an essential subroutinue, requires fewer query if done quantumly (via Quantum fourier sampling)

Learning CBF via Quantum Exact learner: **Results**

- Any **Exact learning** algorithm must make exponential number of query to learn CBF with high success probability. Thus, the problem remain intractable in Quantum regime too.
- Reason: A tight (exponential) lower bound on query complexity exist due to quantum information theoretic argument. (due to Theorem-10)

Table of Contents

1. Main Results
   - Context & Key Results
2. Brief Review of key concepts
   - PAC learning: Brief Review
   - Boolean Function & Fourier Series: Brief Review
3. Algorithm for Quantum PAC learning
   - CBF learning: A two stage process
   - Stage-I: Learning the concentration of a CBF
   - Satge-II: Learning the function from its concentration
4. Benchmarking
   - Complexity: Quantum and Classical PAC learning
   - Theoretical limits on Quantum learning
5. Conclusion

PAC learning: Brief Review

Probably Approximately Correct (PAC) learning: Key terms

- Task: To learn a (Boolean) function f promised to belong to the set of function $\mathcal{C} = \{f_1, ..., f_r\}$ [a]
- Learning paradigm: Supervised Learning [an access to $(x,\ f(x))$]
- Learning Algorithm: A randamized algorithm that output a hypothesis function h. [with success probability $\geq (1-\delta)$]
- α-approximate learning: Function h must agree to f on at least (1-α) fraction of the input. i.e., $Pr[h(x_i) = f(x_i)] \geq (1-\alpha)\ \forall\ i$
- Bounds on parameters: $0 < \delta, \alpha < 0.5$

[a] $\mathcal{C}$ is also called concept class (Hypothesis space) in VC-dimension.

Definition: (δ, α)-PAC learner

- An algorithm is called (δ, α)-learner, if for a given $f \in \mathcal{C}$, $x \in D$(a distribution over input sapce) and query access to $(x, f(x))$, it output a α-approximate function h with success prob $\geq (1-\delta)$.

Boolean Function & its Fourier Series: Brief Review

Boolean Function: Notations & Terms

- Boolean Function [in $(-1, 1)$ basis [a]]:
  $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$. $(Let,\ 0 \rightarrow 1; 1 \rightarrow -1)$
- Fourier (Series) Expansion:

$$f(x) = \sum_{S \subseteq [n]} \hat{f}(S) x^S$$

- Notations: $S \subseteq [n]$ is an element of the powerset;

$$x^S = \prod_{i \in S} x_i$$

  where x^S is fourier term; $\hat{f}(S)$ is fourier coefficient.
- Normalization: $\sum_{S \subseteq [n]} \hat{f}(S)^2 = 1$
- Perseval's identity: $\hat{f}(S) = \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} f(x) x^S$

[a] Hadamard-Walsh basis

Concentrated Boolean Functions: Brief Review

Concentarted Boolean Function(CBF)

- Definition-1: A ϵ-concentrated boolean function f in a set $\mathcal{M} \subseteq [n]$ if:

$$\sum_{S \in \mathcal{M}} \hat{f}(S)^2 \geq (1-\epsilon) \Leftrightarrow \sum_{S \notin \mathcal{M}} \hat{f}(S)^2 \leq \epsilon$$

- Informally, a concentrated Boolean function on a set $\mathcal{M}$ has all the crucial information about the function concentrated in a few important Fourier terms in the Fourier expansion.

Intuitive connection of CBF and Quantum Speedups:

Shor's (factoring) algorithm: Factor and non-factor integer of the input N corresponds to important and non-important fourier terms.

HHL algorithm: Sparsity requirement on the matrix corresponds to concentration of information idea.

Table of Contents

1 Main Results
- Context & Key Results

2 Brief Review of key concepts
- PAC learning: Brief Review
- Boolean Function & Fourier Series: Brief Review

3 Algorithm for Quantum PAC learning
- CBF learning: A two stage process
- Stage-I: Learning the concentration of a CBF
- Satge-II: Learning the function from its concentration

4 Benchmarking
- Complexity: Quantum and Classical PAC learning
- Theoretical limits on Quantum learning

5 Conclusion

CBF learning: A two stage process

Structure of the algorithm

- Input: Partial (Fourier domain) information about a CBF to be learned is supplied. Mainly, a promise of CBF to be $\epsilon-$concentarted on an unknown set of size atmost $\mathcal{M}$.
- Resource available: Quantum/Classical Query access to function input-output pairs, e.g, $(x, f(x))$
- Output: A function h that α-approximate CBF f.

Two stage process of learning CBF

- Stage-I: Learning the concentration of the given CBF
- Stage-II: Learning the function from the concentration estimated in stage-I

Stage-I of Quantum PAC learn a CBF

Learning the concentration of a CBF

- Definition-2: (***Sparse fourier sampling (SFS) problem***) Given f to be $\epsilon-$concentrated Boolean function on an unknown set of size at most $\mathcal{M}$; sample a set L of fourier terms such that L contains all the fourier terms with coefficients greater than or equal to a threshold η.
- Set $\eta = O(\sqrt{\frac{\epsilon}{\mathcal{M}}})$. This relates the SFS problem to estimating the concentration of a CBF. (See definition-1)
- Use **Quantum fourier sampling (QFS)** to solve SFS problem, and hence, learn the concentration. [Algorithm on next page]

Complexity of Quantum Fourier Sampling: Results

- Query complexity: $O(\frac{\mathcal{M}}{\epsilon})$; (due to Theorem-2)
- Time complexity: $O(\frac{\mathcal{M}}{\epsilon})$; (due to Theorem-2)

Algorithm for **Quantum Fourier Sampling (QFS)**

- Step-I: Satrt with $\frac{1}{2^{n/2}} \sum_{x \in \{0,1\}^n} |x\rangle|1\rangle$; (uniform superposition)
- Step-II: Apply H gate to the last qubit: $\frac{1}{2^{n/2}} \sum_{x \in \{0,1\}^n} |x\rangle|-\rangle$
- Step-III: Query the oracle with the above state as input to get $\frac{1}{2^{n/2}} \sum_{x \in \{0,1\}^n} |x\rangle| - \oplus f'(x)\rangle \implies \frac{1}{2^{n/2}} \sum_{x \in \{0,1\}^n} (-1)^{f'(x)} |x\rangle|-\rangle$
- Step-IV: Apply H gate to the last qubit: $\frac{1}{2^{n/2}} \sum_{x \in \{0,1\}^n} f'(x)|x\rangle|1\rangle$
- Step-V: Apply $H^{\otimes n}$ gate to the first n-qubits: $\frac{1}{2^{n/2}} \sum_{x \in \{0,1\}^n} f'(x)(\frac{1}{2^n/2} \sum_S (-1)^{x \cdot S}|S\rangle)|1\rangle$ $\implies \frac{1}{2^n} \sum_x \sum_S (-1)^{x \cdot S} f'(x)|S\rangle)|1\rangle = \sum_S \hat{f}(S)|S\rangle$.
- Step-VI: Measure the resulting state to obtain the desired result.

Stage-I: Learning the concentration (continue...)

Learning the concentration via QFS subroutinue

- Call the QFS subroutinue for $O(\mathcal{M}/\epsilon)$ times; store the result of each call in a list $L = \{S_1, ..., S_{\mathcal{M}'}\}$, where $\mathcal{M}' = O(\mathcal{M}/\epsilon)$. Output list L as the result of SFS problem. [due to theorem-2]
- With high probability, the list L contains mainly the dominant terms of the fourier series. Thus, we learned the concentration of the CBF f.

Proof of correctness

- Step-V of the QFS algorithms produce state $|\psi\rangle = \sum_S \hat{f}(S)|S\rangle$. Thus the measurement output the state $|S\rangle$ with probability $|\hat{f}(S)|^2$. Or, dominant fourier terms are more likey to be the measurement outcome.
- Let 'Good event-1' be a fourier tems with coefficient $> \sqrt{\epsilon/\mathcal{M}}$ is output in a single QFS call. Then Pr[Good-1] $\geq 2/3\mathcal{M}$. [Note: We want to sample such dominent terms. See definition-2]
- Let 'Good event-M' be a fourier tems with coefficient $> \sqrt{\epsilon/\mathcal{M}}$ is output in all $\sim O(\mathcal{M})$ QFS calls. Then Pr[Good-M] $\geq (2/3\mathcal{M}) \cdot (\mathcal{M}) = 2/3$ (due to Union Bound).

Stage-II: Learning f from the estimated concentration

- Let the output from the stage-I be the list $L = \{S_1, ..., S_{\mathcal{M}'}\}$; the indices for dominant fourier terms
- Step-I: Make a list of t random queries: $\{(x_i, f(x_i))\}\ \forall\ i = \{1, ..., t\}$. Where, queries are made uniformly at random: $x_i \in_{u.a.r} \{-1, 1\}^n$. [This is classical query. We will reuse it whenever required.]
- Step-II: Use Perseval's identity to estimate the fourier coefficients as: $\tilde{f}(S) = \frac{1}{t} \cdot \sum_{i=1}^{t} f(x_i) x_i^S$.
- The **Lemma-1** guarantees $\tilde{f}(S)$ well approximate the actual value $\hat{f}(S)$ if the value of t (i.e., # of queries) is sufficiently large: $|\hat{f}(S) - \tilde{f}(S)| < \gamma$; if, $t = O(\frac{1}{\gamma^2})$. [via Chernoff bound]

Theorem 1 *Let $f : \{-1, 1\}^n \to \{-1, 1\}$ be an ϵ-concentrated Boolean function on an unknown set of size at most M. Then, if given a set L of Fourier terms such that all the Fourier terms with coefficient greater than or equal to $\sqrt{\epsilon/M}$ are included in L, there is a learning algorithm that can PAC learn the target function, with error $\alpha = O(\epsilon)$ and high success probability, in $O(\frac{|L|^2}{\epsilon})$ time using $O(\frac{|L|}{\epsilon})$ classical uniform random queries.*

Figure: From the paper

(Continue...) Stage-II: Learning f from the concentration

- **Remark-II**: Theorem-I prescribe $\gamma = \sqrt{\frac{\epsilon}{L}}$. (We use it in later analysis)
- Step-III: Using the strategy similar to Kushilevitz-Monsour Algorithm (a classical algorithm); define $g(x) = \sum_{S \in L} \tilde{f}(S) x^S$ and $h(x) = sign(g(x))$.
- Step-IV: Due to theorem-I, the above estimated $h(x)$ is proven to α-approximate $f(x)$. Output $h(x)$ as PAC learned function.

Note: The Stage-II is classical in nature.

Theorem 2 *Given $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ be an ϵ-concentrated Boolean function on an unknown set of size at most M, there exists a quantum algorithm that, using $O(\frac{M}{\epsilon})$ quantum uniform queries and time, recovers a list L of at size $O(\frac{M}{\epsilon})$ containing all Fourier terms of f with Fourier coefficient greater than $\sqrt{\epsilon/M}$ with high success probability.*

Figure: From the paper

Time and Query complexity: Combining Stage-I and II

- The Stage-I requires $O(\frac{\mathcal{M}}{\epsilon})$ calls to subroutinue QFS. Assume each QFS call take $O(1)$ time. Thus, stage-I time $= O(\frac{\mathcal{M}}{\epsilon})$.
- The Stage-II has two requirements:
- (II.a) Creating the list of t random queries: $\{(x_i, f(x_i))\}\ \forall\ i = \{1, ..., t\}$. It requires $t = O(\frac{1}{\gamma^2}) = O(\frac{|L|}{\epsilon})$ queries. (due to Remark-II)
- (II.b) Estimating each of $\tilde{f}(S)$, where $S \in L = \{S_1, ..., S_{\mathcal{M}'}\}$. For each $\tilde{f}(S)$, we use the list L and apply Perseval's identity. It need $O(|L| \cdot |L|/\epsilon)$ time.
- Replace $|L| = O(\frac{\mathcal{M}}{\epsilon})$ [due to theorem-2] to get stage-II Query: $O(\frac{\mathcal{M}}{\epsilon^2})$ and Time: $O(\frac{\mathcal{M}^2}{\epsilon^3})$.

Theorem 3 *Let $\mathcal{C}$ be a concept class such that every $f : \{-1, 1\}^n \to \{-1, 1\}$ in $\mathcal{C}$ has its Fourier spectrum ϵ-concentrated on a collection of at most M Fourier terms. Then, $\mathcal{C}$ can be PAC learned using $O(\frac{M}{\epsilon^2})$ quantum uniform queries in $O(\frac{M^2}{\epsilon^3})$ time.*

Figure: From the paper

Table of Contents

1. Main Results
   - Context & Key Results
2. Brief Review of key concepts
   - PAC learning: Brief Review
   - Boolean Function & Fourier Series: Brief Review
3. Algorithm for Quantum PAC learning
   - CBF learning: A two stage process
   - Stage-I: Learning the concentration of a CBF
   - Satge-II: Learning the function from its concentration
4. Benchmarking
   - Complexity: Quantum and Classical PAC learning
   - Theoretical limits on Quantum learning
5. Conclusion

Complexity table: Classical and Quantum PAC learning

Table 2 Different algorithms solving the learning of concentrated Boolean functions problem using different types of query

Algorithm	Time complexity	Query complexity	Query type
Hassanieh2012 [23]	$O(M\log(2^n/M)n/\epsilon)$	$O(M\log(2^n/M)n/\epsilon)$	Uniform random query
Indyk2014 [25]	$\tilde{O}(2^n n^c)$	$\tilde{O}(Mn)$	Uniform random query
Kushilevitz–Mansour [20, 26]	$O(nM^3/\epsilon^3)$	$O(nM^3/\epsilon^3)$	Membership query
Our algorithm	$O(M^2/\epsilon^3)$	$O(M/\epsilon^2)$	Quantum uniform random query

Figure: (From the paper) **Hassanieh2012 algorithm** is the best-known classical algorithmic for CBF PAC learning (as per Query×time metric). The **Quantum PAC** learning of the paper perform relatively better in terms of query complexity. For better time complexity, $\mathcal{M} = \tilde{O}(n^2)$ is essential. **(Note:** First three algorithms are classical.)

- Notation: $\tilde{O}(f(n)) = O(f(n) \cdot log^k(n))$

Theoretical bounds on query complexity for Quantum learning of a CBF

- Any **Quantum PAC** learning algorithm for a ϵ-concentrated Boolean function on an unknown set sized $\mathcal{M}$ **must make** $\Omega(\mathcal{M})$ **queries** to be 'decently' successful.
- Reason: **(Theorem-9)** Let $\mathcal{C} = \{f_1, ..., f_r\}$ be the class of boolean function that are ϵ-concentrated on an unknown set of size at most $\mathcal{M}$. Then, the number of quantum queries necessary to PAC learn $\mathcal{C}$ with approx. error $\alpha \leq 1/10$ under arbitary distribution is $\Omega(M)$.
- Any **Quantum Exact**[1] learning algorithm for a ϵ-concentrated Boolean function on an unknown set sized $\mathcal{M}$ **must require exponential number** of queries to be decently successful.
- Reason: **(Theorem-10)** Let $\mathcal{C} = \{f_1, ..., f_r\}$ be the class of boolean function that are ϵ-concentrated on an unknown set of size at most $\mathcal{M}$. Then, the number of uniform quantum queries examples to exact learn $\mathcal{C}$ is $\Omega(\epsilon \cdot 2^n \cdot log(\mathcal{M})/n)$

[1] i.e., approximation error $\alpha = 0$

Table of Contents

1. Main Results
   - Context & Key Results
2. Brief Review of key concepts
   - PAC learning: Brief Review
   - Boolean Function & Fourier Series: Brief Review
3. Algorithm for Quantum PAC learning
   - CBF learning: A two stage process
   - Stage-I: Learning the concentration of a CBF
   - Satge-II: Learning the function from its concentration
4. Benchmarking
   - Complexity: Quantum and Classical PAC learning
   - Theoretical limits on Quantum learning
5. Conclusion

Conclusions

- Asymptotically, certain Concentarted Boolean function can be quantumly PAC learned using lesser query and time resources.
- Exact learning remain interactable both in the classical and the quantum way.

Open Problem: Can Satge-II be optimized?

- Stage-II estimate the hypothesis function h based on the concentration learned in stage-I. In this paper, it is done using Kushilevitz-Mansour algorithms.
- Can Hassanieh (2012) algorithm, a highly optimized classical PAC algorithm that stress on finding high-level structure for important fourier terms, be used in this context to reduce time/query complexity?

Thanks!