The Quantum and Classical Streaming Complexity of Quantum and Classical Max-Cut

John Kallaugher
Sandia National Labs
Albuquerque, USA
jmkall@sandia.gov

Ojas Parekh
Sandia National Labs
Albuquerque, USA
odparek@sandia.gov

***Abstract*—We investigate the space complexity of two graph streaming problems: MAX-CUT and its quantum analogue, QUANTUM MAX-CUT. Previous work by Kapralov and Krachun [STOC '19] resolved the *classical* complexity of the *classical* problem, showing that any $(2-\varepsilon)$-approximation requires $\Omega(n)$ space (a 2-approximation is trivial with $\mathbf{O}(\log n)$ space). We generalize both of these qualifiers, demonstrating $\Omega(n)$ space lower bounds for $(2-\varepsilon)$-approximating MAX-CUT and QUANTUM MAX-CUT, even if the algorithm is allowed to maintain a quantum state. As the trivial approximation algorithm for QUANTUM MAX-CUT only gives a 4-approximation, we show tightness with an algorithm that returns a $(2+\varepsilon)$-approximation to the QUANTUM MAX-CUT value of a graph in $\mathbf{O}(\log n)$ space. Our work resolves the quantum and classical approximability of quantum and classical Max-Cut using $\mathbf{o}(n)$ space.**

We prove our lower bounds through the techniques of *Boolean Fourier analysis*. We give the first application of these methods to *sequential one-way quantum communication*, in which each player receives a quantum message from the previous player, and can then perform arbitrary quantum operations on it before sending it to the next. To this end, we show how Fourier-analytic techniques may be used to understand the application of a quantum channel.

***Index Terms*—quantum algorithms, graph algorithms, streaming algorithms, fourier analysis**

I. INTRODUCTION

Quantum approaches for discrete optimization, such as the Quantum Approximate Optimization Algorithm (QAOA) have received significant attention. The seminal work of Farhi, Goldstone, and Gutmann [1] showed that QAOA applied to an NP-hard classical constraint satisfaction problem (CSP) gave a better worst-case approximation than the best known classical approximation algorithm at the time. An improved classical approximation algorithm subsequently followed [2]; however, this seeded the question of whether a quantum approximation algorithm might offer a provably better approximation guarantee than the best classical approximation for some CSP or discrete optimization problem, which still remains open. One potential barrier is that classical hardness of approximation results may also restrict quantum approximation algorithms. For example, it is generally not expected that $\mathsf{NP} \subseteq \mathsf{BQP}$, so a quantum approximation is not expected to overcome NP-hardness of approximation results. Even possibly weaker hardness assumptions such as Unique-Games-hardness may impede quantum approximations. It would be surprising if a quantum approximation were able to achieve a $(1/0.878\ldots-\varepsilon)$-approximation[1] for the Maximum Cut Problem (MAX-CUT), which is Unique-Games-hard [3].

Although the prospects for quantum approximations for classical CSPs may seem limited, a natural question is whether quantum approximations can offer provably better guarantees for *quantum versions* of CSPs. The k-Local Hamiltonian Problem (k-LH) serves as the canonical QMA-hard quantum generalization of k-CSP. A recent line of work has enjoyed success in devising nontrivial classical approximations for 2-LH [4], [5], [6], [7], [8], [9], [10], [11]; however, truly quantum approximations for LH remain elusive. Hardness of approximation results with respect to QMA are even more elusive, as the existence of a quantum analogue of the classical PCP theorem, a cornerstone for hardness of approximation, remains a major open question [12], [13].

We seek to understand the power of quantum versus classical approximations for 2-CSP and 2-LH in the streaming setting, where *space* is the computational quantity of interest. In particular we consider the MAX-CUT (MC) and QUANTUM MAX-CUT (QMC) problems. MAX-CUT is a prototypical CSP in the sense that approximation and hardness results are typically devised for MAX-CUT and then generalized to other CSPs; QUANTUM MAX-CUT has emerged to serve a similar role in approximating 2-LH. QUANTUM MAX-CUT is

[1]This result is more typically stated as $0.878\ldots+\varepsilon$, where an α-approximation is held to mean returning a value in $[\alpha \cdot \mathrm{OPT}, \mathrm{OPT}]$, for OPT the correct value. However, we follow previous work on streaming MAX-CUT by instead using a K-approximation to mean returning a value in $[\mathrm{OPT}, K \cdot \mathrm{Opt}]$.

2575-8454/22/$31.00 ©2022 IEEE
DOI 10.1109/FOCS54457.2022.00054

	MAX-CUT		QUANTUM MAX-CUT	
Approximation	$2+\varepsilon$	$2-\varepsilon$	$2+\varepsilon$	$2-\varepsilon$
Classical	$\mathrm{O}(\log n)$	$\Omega(n)$	$\mathbf{O(\log n)}$	$\mathbf{\Omega(n)}$
Quantum	$\mathrm{O}(\log n)$	$\mathbf{\Omega(n)}$	$\mathbf{O(\log n)}$	$\mathbf{\Omega(n)}$

TABLE I: The space needed by quantum and classical algorithms for quantum and classical Max-Cut. Results from this paper are shown in bold.

also closely related to the quantum Heisenberg model (see [7]), which is a well-studied model of quantum magnetism introduced in the late 1920s.

For *classical* algorithms applied to *classical* MAX-CUT, tight bounds[2] for the space complexity in terms of the approximation factor are known [14]—our work generalizes these results in both ways, giving tight bounds on the approximation factor attainable in $\mathrm{o}(n)$ space by quantum streaming algorithms for classical MAX-CUT and by quantum and classical algorithms for QUANTUM MAX-CUT.

We find, perhaps surprisingly, that quantum streaming algorithms offer no advantage over classical ones in approximating MAX-CUT or QUANTUM MAX-CUT. Although our main contribution is a quantum hardness result, the matching upper bound for approximating QUANTUM MAX-CUT in the stream requires analyzing a nontrivial streaming algorithm, which is a departure from the case of MAX-CUT.

A. Our Contributions

Ours is the first work to consider streaming versions of 2-LH or any kind of quantum optimization problem. Just as the results of [14] have been expanded for more general CSPs, we expect that our results for QUANTUM MAX-CUT will apply to more general instances of 2-LH. Indeed there is precedent for this in the standard approximation setting [15], [8].

We give tight (up to an arbitrarily small additive constant in the approximation factor) characterizations of the best possible approximation factor achievable in $\mathrm{o}(n)$ space for quantum and classical algorithms for quantum and classical Max-Cut. Our results are laid out in Table I.

Our lower bounds are encompassed in the following theorem (the classical lower bound for QUANTUM MAX-CUT is a special case of this, as any classical streaming algorithm can be implemented as a quantum streaming algorithm).

Theorem 1. *For any $\varepsilon > 0$, any quantum streaming algorithm for* MAX-CUT *or* QUANTUM MAX-CUT *that returns a $(2-\varepsilon)$-approximation with probability $2/3$ requires*

$$n/2^{\mathrm{O}(1/\varepsilon^2)}$$

qubits of storage.

For MAX-CUT the upper bound for $(2+\varepsilon)$-approximation (in fact, even 2-approximation) is trivial, as a graph on m edges always has MAX-CUT value between $m/2$ and m. However, for QUANTUM MAX-CUT the trivial approximation is only a 4-approximation, so we give an algorithm that returns a $(2+\varepsilon)$-approximation using $\mathrm{O}(\log n)$ space. We also give an algorithm for *weighted* graphs, but in this case we are only able to attain a $(5/2+\varepsilon)$-approximation (the lower bound remains a 2-approximation).

Theorem 2. *Let G be a weighted graph on n vertices with weights that are multiples of $1/\operatorname{poly}(n)$. Then for any $\varepsilon, \delta \in (0,1)$ there is a streaming algorithm that returns a $(5/2+\varepsilon)$-approximation to the* QUANTUM MAX-CUT *value of G with probability at least $1-\delta$ using $\mathrm{O}\big(\frac{1}{\varepsilon^2}\log\frac{1}{\delta}\log n\big)$ space. If all the weights in the graph are 1, it returns a $(2+\varepsilon)$-approximation instead.*

We note here two lacunae in our results for (unweighted) graphs. Firstly, for classical MAX-CUT it is possible to achieve a $(1+\varepsilon)$-approximation in $\widetilde{\mathrm{O}}(n)$ space, through the use of cut-preserving sparsifiers [16]. However, analogous results on sparsifiers for general 2-local Hamiltonians are not known, and indeed there are results pointing in the opposite direction [17]. So while we can characterize the approximation factors possible in sublinear space the semi-streaming complexity remains open. Secondly, our $\mathrm{O}(\log n)$-space upper bound for QUANTUM MAX-CUT only gives a $(2+\varepsilon)$-approximation instead of a 2-approximation. This is a consequence of the fact that it is based on graph parameters that must themselves be approximated rather than just the number of edges, which can be calculated exactly.

a) Fourier Analysis for Quantum Channels: The technical core of our lower bound is a quantum communication complexity bound for a sequential one-way communication problem (originally introduced in [14] in the classical setting), in which the first player sends a message to the second player, the second to the third, and so on. Our bound for this problem is based on a novel application of *Boolean Fourier analysis*[3]—in particular, we prove that a key inequality associated with this problem, analyzed in [14] for the classical case, is preserved even in the presence of quantum communication.

[2]Up to log factors in the space complexity, as is typical for streaming algorithms.

[3]For a general overview of Boolean Fourier analysis, see [18].

The application of Boolean Fourier Analysis to two-player one-way communication problems in the classical setting goes back to [19], in which it was used to prove lower bounds for the Boolean Hidden Matching problem (and its application to communication complexity more generally goes back further, e.g. [20], [21]). This problem, and its generalization in the Boolean Hidden Hypermatching problem (analyzed in [22]), are the main route by which Fourier analysis has contributed to lower bounds for streaming algorithms.

However, in later years these techniques have been extended to communication lower bounds (and corresponding streaming lower bounds) with different configurations of players. In [23] they were applied to problems where many players communicate with a single referee, while in [24] they were extended to problems where players communicate in a line, as is the case in the Distributed Implicit Partition Problem (from [14]) we make use of in this paper. In [25], a generalization of this problem was studied through the use of Fourier analysis on $\mathbb{Z}_q^n$.

The core ingredient of most of these lower bounds is a hypercontractive Fourier coefficients lemma from [26], that can be seen as generalizing facts about sampling protocols, in which a player chooses some subset of their input to send to protocols where players send arbitrary (classical) messages. This lemma was generalized to *matrix*-valued functions in [27], opening the door to the application of Fourier-analytic methods to lower bounds for *quantum* communication protocols, as these can be seen as functions from inputs to density matrices.

This was first used to prove quantum lower bounds on the complexity of the Boolean Hidden Hypermatching problem [28]. This result was further generalized in [29], while [30] generalized the Fourier coefficients lemma further, in order to obtain quantum lower bounds for MAX-CUT and more general hypergraph problems. However, these are all two-player one-way communication problems. We give the first application of Fourier-analytic techniques to *sequential quantum one-way communication*. The key technical challenge is in finding methods for applying Fourier analysis to the application of *quantum channels*.

In classical communication, as long as we consider a single "hard" input distribution, a player's message can be without loss of generality assumed to be a deterministic function of the message they received and their input. In sequential quantum communication, however, the player may apply an arbitrary quantum channel to the message they receive. Our key insight is that, as quantum channels are linear operators, many of the techniques of Fourier analysis, including the convolution lemma for Fourier coefficients, may be applied to them.

B. Other Related Work

a) Streaming bounds for MAX-CUT*:* The fact that that a 2-approximation for MAX-CUT is possible in $\mathrm{O}(\log n)$ space is an immediate consequence of the fact that the MAX-CUT value is always at least $m/2$. Less immediately, but still a consequence of standard results in streaming algorithms, is the fact that it can be $(1+\varepsilon)$-approximated in $\widetilde{\mathrm{O}}(n)$ space, through sparsifiers that preserve cut values [16]. The question, then, was whether a better approximation than the first could be attained in less space than the second.

In [31], [32], it was shown that any $(2-\varepsilon)$-approximation would require at least polynomial space in n, while [24] showed that $(1+\varepsilon)$-approximation would require $\Omega(n)$ space. This left open the possibility of intermediate results, but [14] closed the door on this possibility, proving that $(2-\varepsilon)$-approximation would require $\Omega(n)$ space for any constant $\varepsilon > 0$.

However, the above results are only for classical algorithms. In [30], a polynomial lower bound was shown that applies even to quantum streaming algorithms, but this left open the possibility that a $(2-\varepsilon)$ approximation was possible in $\mathrm{o}(n)$ space for quantum algorithms.

b) Quantum streaming algorithms: The first work on quantum streaming was [33], which showed that there are problems that that are exponentially easier for quantum streaming algorithms than classical ones. In [34], it was shown that this is true even for a function that does not depend on the order of the stream (the more "standard" streaming model).

Later work has investigated the question of whether quantum streaming can obtain advantages over classical for problems of independent classical interest (as the aforementioned work is for problems constructed for the purpose of proving separations). The problem of recognizing $\texttt{Dyck}(2)$ in the stream was considered as a candidate problem in [35], [36], but only negative results were found. For problems where $\omega(1)$ passes are allowed over the stream, [37] and [38] showed an advantage for the well-studied moment estimation problem. Later, [39] showed that an advantage exists in the one-pass setting for the problem of counting triangles in graph streams.

c) Approximating QUANTUM MAX-CUT*:* QUANTUM MAX-CUT was introduced in [7], where a classical $1/0.498$-approximation algorithm was given, akin to the Goemans-Williamson algorithm for MAX-CUT, that produces an unentangled product state. Since the gap between the best product state and best entangled quantum state on a single edge is two, at best a 2-approximation

is possible for algorithms that return product states, and so the $1/0.498$-approximation is nearly optimal among such algorithms. By rounding to entangled states, [9] gave the first approximation with guarantee better than 2. Subsequently [10] showed how to use higher levels of the quantum Lasserre hierarchy of semidefinite programs to obtain a slight improvement over [9].

II. Proof Overview

A. Lower Bounds

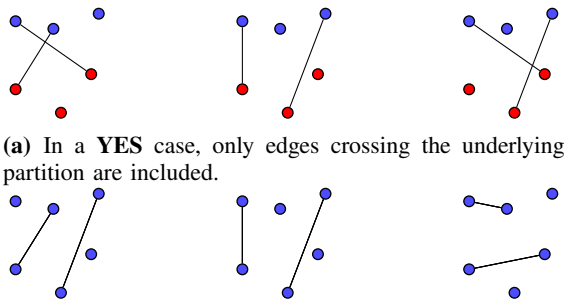
Our lower bounds for the quantum streaming complexity of MAX-CUT and QUANTUM MAX-CUT are derived from a new analysis of the Distributional Implicit Hidden Partition (**DIHP**) problem introduced in [14] to prove lower bounds for the streaming complexity of approximating classical MAX-CUT. We restate this problem here.

1) The Distributed Implicit Hidden Partition Problem: In an instance of **DIHP**(n, α, T), T players are each given a partial matching M_t of αn edges on n vertices, with each edge labelled with a bit. Either these bit labels are generated by choosing a random partition of $[n]$ and assigning 1 to the edges crossing the partition (a **YES** case) or they are chosen uniformly at random (a **NO** case).

The players are allowed one-way communication, from player i to player $i+1$ for each i, and are additionally given the matching edges (but not the edge labels) of every previous player for free. Their goal is to determine whether their inputs were drawn from a **YES** case or a **NO** case with probability at least $2/3$ over the random draw and any internal randomness they may use.

a) Reduction to Classical Max-Cut: If each player t creates the graph G_t consisting of edges labelled 1 in M_t, $G = \bigcup_t G_t$ will be bipartite in a **YES** case, and close to random in a **NO** case. This means it is possible to cut every edge in the first case, and not much more than half of them in the second. Therefore, an algorithm that returns a $(2-\varepsilon)$-approximation to MAX-CUT can distinguish them if ε is large enough (by making α small enough and T large enough, we can make the necessary ε arbitrarily small). Therefore, a MAX-CUT algorithm using S space gives a protocol in which each player sends a size-S message, by having each player run the algorithm on their input and then send their algorithm's state to the next player. The graphs the players get with this reduction are illustrated in Figure 1.

One problem with this is that, if the players' matchings are randomly chosen they may share edges. Our approach to this differs somewhat from that of [14]. Instead of considering multigraphs, we take advantage of the fact that player t is allowed to know the matching (but not the edge labels) of players $s < t$. This means we can have them decline to add edges that are present in previous matchings, guaranteeing that the final graph is simple. We show that, as the number of edges thus removed is small, it has little effect on the reduction.



(a) In a **YES** case, only edges crossing the underlying partition are included.

(b) In a **NO** case, each edge received is included with probability $\frac{1}{2}$.

Fig. 1: The graphs each player receives when reducing **DIHP** to MAX-CUT.

b) Extending the Reduction to Quantum Max-Cut: The reduction to Quantum Max-Cut uses exactly the same mapping from **DIHP** instances to graphs. We consider the following SDP,

$$\max_{f:V\to S^{n-1}} \sum_{uv\in E} -\langle f(u), f(v)\rangle$$

which is a shifted version of the standard Goemans-Williamson SDP for MAX-CUT. In particular, its optimal value is an upper bound on $2K - m$, where K is the MAX-CUT value of a graph. Usefully, when $2K - m$ is small, a converse property holds, as the optimal value of this SDP is at most a constant factor times larger than $2K - m$ [40]. This means the graphs generated by **NO** instances of **DIHP** will have small values of this SDP.

This gives us a QUANTUM MAX-CUT lower bound, because this SDP *also* upper bounds $\frac{4}{3}Q - \frac{m}{3}$, where Q is the QUANTUM MAX-CUT value of the graph[4]. So **NO** instances will create graphs with QUANTUM MAX-CUT value approximately $m/4$. Conversely **YES** instances will create graphs with QUANTUM MAX-CUT value at least $m/2$, as they are bipartite and the QUANTUM MAX-CUT value is always at least half the MAX-CUT value. So a $(2-\varepsilon)$ approximation algorithm would suffice to distinguish between the two.

2) Quantum Communication Lower Bounds for DIHP: In [14] it was shown that **DIHP** is hard when the

[4]See Section 2.3 of [41]. Note that in the cited work, both QUANTUM MAX-CUT and the SDP are scaled by $\frac{1}{m}$ relative to our usage.

players are only allowed to send classical messages, requiring $\Omega(n)$ space when α and T are constant. The majority of the technical difficulty of our lower bounds is in proving that **DIHP** is hard even if the players are allowed to send *quantum* messages. This immediately implies that quantum algorithms must use $\Omega(n)$ space to $(2-\varepsilon)$-approximate MAX-CUT or QUANTUM MAX-CUT, and so no quantum advantage for either problem is possible.

a) Reduction to Boolean Fourier Analysis: As with the classical lower bound of [14], our proof depends on applying Fourier analysis to functions on the Boolean cube. In particular, we will show that a bound on Fourier coefficients used in the classical proof is maintained even in the presence of quantum communication. We start by providing an intuition for the significance of this bound.

Suppose the game is in a **YES** case, and so player t's input depends only on the matching M_t and the randomly chosen partition (which we may write $x \in \{0,1\}^n$, with the bit of vertex i determining which side of the partition it is on). Then, fixing $(M_s)_{s=1}^t$, we can write a function

$$f_t : \{0,1\}^n \to \mathbb{C}^{2^\beta \times 2^\beta}$$

where $f(x)$ is the density matrix sent by player t if the partition is x, and β is the number of qubits used to represent that state.

Now suppose player $t+1$ would like to determine whether they are in a **YES** or a **NO** case. They have received $f_t(x)$ if they are in a **YES** case, and they want to determine if it is consistent with being in a **YES** case. In addition, they have the bit labels of the edges in M_{t+1}. Therefore, for any odd-cardinality set of edges in M_{t+1}, they know the parity of the set of vertices in x matched by these edges. We write such sets of vertices as $M_{t+1}^{\text{tr}} s$ for a string $s \in \{0,1\}^{\alpha n}$ indexing a subset of the edges in M_{t+1}.

Now suppose the player looked at only one of these sets s, and so knew the parity of the vertices $M_{t+1}^{\text{tr}} s$ alone. To tell whether $f_t(x)$ could come from a **YES** instance, they need[5] its average value when the parity of $M_{t+1}^{\text{tr}} s$ is 0 to be distinguishable from its average value when the parity of $M_{t+1}^{\text{tr}} s$ is 1.

The distinguishability of two distributions over quantum states is given by the trace norm of the difference between their density matrices, so the quantity the player would need to be large is

$$\frac{1}{2}\left\| \frac{1}{2^{n-1}} \sum_{\substack{x\in\{0,1\}^n: \\ x\cdot M_{t+1}^{\text{tr}} s=0}} f_t(x) - \frac{1}{2^{n-1}} \sum_{\substack{x\in\{0,1\}^n: \\ x\cdot M_{t+1}^{\text{tr}} s=1}} f_t(x) \right\|_1$$

which equals

$$\frac{1}{2^n}\left\| \sum_{x\in\{0,1\}^n} f_t(x)(-1)^{x\cdot M_{t+1}^{\text{tr}} s} \right\|_1 = \|\widehat{f_t}(M_{t+1}^{\text{tr}} s)\|_1$$

where we now introduce $\widehat{f_t}$, the *Fourier transform* of f_t, given by

$$\widehat{f}(S) = \frac{1}{2^n} \sum_{x\in\{0,1\}^n} f(x)(-1)^{S\cdot x}.$$

It turns out that this sums nicely—it can be shown that player T's ability to distinguish between a **YES** and a **NO** case is bounded by

$$\sum_{t=1}^{T} \sum_{s\in\{0,1\}^{\alpha n}\setminus\{\emptyset\}} \|\widehat{f_t}(M_{t+1}^{\text{tr}} s)\|_1$$

and so our goal will be to prove that this sum is small in expectation over $(M_t)_{t=1}^T$.

To prove this, we bound the total value of weight-2ℓ Fourier coefficients for every ℓ. As a $\sim\binom{\alpha n}{\ell}/\binom{n}{2\ell}$ fraction of these will end up being matched by a set of ℓ matching edges, it suffices to prove that the value is bounded by[6]

$$\left(\frac{\sqrt{\beta n}}{\ell}\right)^\ell$$

where we have dropped some constants exponential in T and ℓ. Then if $\beta \ll n$, this expression will be small enough for the final states to be hard to distinguish.

b) The Evolution of Fourier Coefficients: We will bound the expression above by induction on t, considering how these coefficients evolve based on the message sent from player t to player $t+1$. This is where the *quantum* difficulty of the proof will arise, and is the most important novel element in our analysis—the combinatorial aspects of the evolution are similar to those in the classical case but now player t may apply a quantum channel to generate f_t rather than sending a deterministic[7] message based on their input and the message f_{t-1}.

[5]We are eliding the possibility that, for instance, the state player $t+1$ receives is impossible or unlikely in a **YES** case due to, for instance, only arising if a triangle in previously arrived edges has every edge labelled 1. However it turns out this possibility is already accounted for by considering what a previous player would've seen on receiving the third edge of that triangle.

[6]This expression changes somewhat when $\ell \geq \beta$, but we will disregard those highest-order terms in this overview.

[7]When proving a lower bound for a classical communication problem with a known input distribution, one may without loss of generality assume the players act deterministically.

The base case of the induction is straightforward (for simplicity we can think of player 1 as receiving 0^β from a player 0, and consider only an inductive step). For the inductive step, we need to understand the effect of player t applying a quantum channel $\mathcal{A}$ to $f_{t-1}(x)$. This quantum channel itself is determined by player t's input, and therefore (again fixing $(M_s)_{s=1}^t$) we can write $\mathcal{A}_x$ for its value when the underlying partition is x. As quantum channels are linear operators, we can define a Fourier transform

$$\widehat{\mathcal{A}}_S = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} \mathcal{A}_x (-1)^{S \cdot x}$$

that in particular obeys the convolution lemma for the Boolean Fourier transform, which tells us that

$$\widehat{\mathcal{A}_x f_{t-1}(x)}(S) = \sum_U \widehat{\mathcal{A}}_U \widehat{f}_{t-1}(U \oplus S).$$

Using the fact that $\widehat{\mathcal{A}}_U$ is 0 whenever U is not $M_t s$ for some $s \in \{0,1\}^{\alpha n}$ (intuitively, this is because $\mathcal{A}_x$ only depends on the edge labels of the edges in M_t), we can write down "mass transfer" lemmas describing how coefficients of weight $2\ell_2$ of f_t are formed from coefficients of weight $2\ell_1$ of f_{t-1}. We want to know how much weight can be contributed to $\widehat{\mathcal{A}_x f_{t-1}(x)}(S)$ from $\widehat{\mathcal{A}}_U \widehat{f}_{t-1}(M_t^{\mathrm{tr}} s \oplus S)$ where $|S| = \ell_2$ and $M_t^{\mathrm{tr}} s \oplus S = \ell_1$.

We can think of this in terms of three more parameters, a the number of edges from $M_t^{\mathrm{tr}} s$ that are entirely contained in S, b the number of edges that each have one endpoint in S, c the number that are entirely outside of S (so $\ell_2 = \ell_1 - a + c$). We end up with the amount of "mass" transferred from ℓ_1-weight coefficients via $M_t^{\mathrm{tr}} s$ with this property being bounded by

$$\sum_{\substack{S \in \{0,1\}^n \\ |S| = 2\ell_1}} \sum_{\substack{u \in \{0,1\}^{\alpha n} \\ |u| = a}} \sum_{\substack{v \in \{0,1\}^{\alpha n} \\ |v| = b}} I_S(u) B_S(v) \times \sum_{\substack{w \in \{0,1\}^{\alpha n} \\ |w| = c}} \| \widehat{\mathcal{A}}^t_{M_t^{\mathrm{tr}}(u \oplus v \oplus w)} \widehat{f}_{t-1}(S) \|_1$$

where $I_S(u)$ and $B_S(v)$ are indicator variables on whether $M_t^{\mathrm{tr}} u$ is entirely contained in S and $M_t^{\mathrm{tr}} v$ has one endpoint of each edge in S. See Figure 2 for an illustration.

The final tool we need to bound this is an extension of the matrix-valued Fourier coefficients inequality, a consequence of Theorem 1 of [27] (itself a generalization of a lemma of [26]) that has previously been used for two-player quantum lower bounds [28], [29], [30]. This will tell us that[8]

$$\sum_{\substack{w \in \{0,1\}^{\alpha n} \\ |w| = c}} \| \widehat{\mathcal{A}}^t_{M_t^{\mathrm{tr}}(u \oplus v \oplus w)} \widehat{f}_{t-1}(S) \|_1$$

is

$$\binom{\mathrm{O}(\beta)}{c} \left\| \widehat{f}_{t-1}(S) \right\|_1.$$

With this in place, and using the fact that

$$\mathbb{P}[I_S(u) B_S(v)] \sim \frac{\binom{\ell_1}{a}}{\binom{n}{a}} \cdot \frac{\binom{\ell_1}{b}}{\binom{n}{b}}$$

we can bound the above in expectation over M_t, and from then the proof becomes an exercise in carefully evaluating sums.

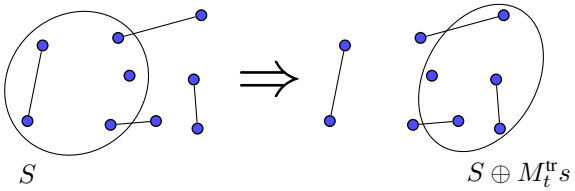


Fig. 2: When player t receives the matching M_t, each subset s of the edges in M_t and each Fourier coefficient $\widehat{f}_{t-1}(S)$ corresponds to a new Fourier coefficient $\widehat{f}_t(S \oplus M_t^{\mathrm{tr}} s)$. In this example s includes $a = 1$ edge internal to S, $b = 2$ edges with one endpoint in S, and $c = 1$ edge outside, so the resulting coefficient $S \oplus M_t^{\mathrm{tr}} s$ has weight $|S| + a - c = 5$.

B. Space Upper Bounds for Quantum Max-Cut

For classical MAX-CUT a trivial classical algorithm achieving a 2-approximation in logarithmic space is already known—count the number of edges (or total weight for a weighted graph) m and report m, which is at most twice the true value. As our lower bound for quantum algorithms for classical MAX-CUT is the same as the classical one, nothing more is needed here. However, for QUANTUM MAX-CUT the story is a bit different. The trivial lower bound in this case is $m/4$, and so the aforementioned algorithm would only guarantee a 4-approximation.

We give a simple algorithm that achieves a $(2+\varepsilon)$-approximation in the unweighted case, and a $(5/2+\varepsilon)$-approximation in the weighted case. The basic idea will be the same in both cases, so for ease of exposition the rest of the discussion in this section will assume a weighted graph, and we will point out where every edge having unit weight allows a better approximation.

[8] Again, this has to be somewhat changed when c is particularly large.

a) Upper Bounding the QMC Value: Let m be the total weight of the graph, and let $W = \sum_{u \in V} \max_{v \in N(u)} w_{uv}$, the sum of the max-weight edges incident to each vertex (so W is just the number of non-isolated vertices in the unweighted case). It is known [9] that

$$\frac{m}{2} + \frac{W}{4}$$

is an upper bound for QUANTUM MAX-CUT. So we want lower bounds in terms of m and W.

b) Lower Bounding the QMC Value in General Weighted Graphs: We use a modified version of an argument of [9]. Consider the subgraph formed by taking the highest-weight edge incident to every vertex. We can decompose this into a matching M consisting of every edge "chosen" by two vertices, and a forest F of all the other edges (note that the two together are also a forest). Abusing notation to use the names of the objects to also denote their total weights, we have $2M + F = W$.

Now, for any edge, it is possible to earn QUANTUM MAX-CUT energy equal to its weight by assigning its vertices the singlet. Secondly, when we have a collection of vertex-disjoint graphs it is possible to maximize each of their QUANTUM MAX-CUT energies separately and still earn energy $w_e/4$ for each edge e *between* distinct pairs of graphs. So there is a solution earning $M + (m - M)/4$.

Secondly, as $M \cup F$ is a forest, cutting it clasically earns energy $(M + F)/2$, as any classical cut gives a quantum cut earning at least half as much energy. By minimizing these two expressions subject to $2M + F = W$ it can be shown that the QUANTUM MAX-CUT value is at least

$$\frac{m}{5} + \frac{W}{10}$$

giving a $(5/2)$-approximation determined only by m and W. We illustrate this construction in Figure 3.

c) Lower Bounding the QMC Value in Unweighted Graphs: In the unweighted case we have the advantage that any method for choosing a maximal tree chooses one of optimal weight, and so (inspired by a method of [42]) we consider *depth-first search trees*. We will assume the graph is connected—note that m, W, and the QUANTUM MAX-CUT value all sum up over components, so as long as the lower bound we show is linear in m and W this will immediately generalize.

In the weighted case, trying to optimize the energy we earned from our tree meant potentially earning nothing from edges outside the tree, as we had no control over how they might cross the tree. However, with a DFS tree, we have the following useful property: for any node in the tree, the subtrees rooted at its children are

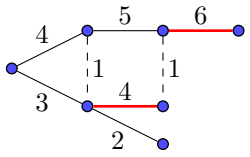


Fig. 3: Proving a lower bound for the optimal QUANTUM MAX-CUT value of a weighted graph. The edges "chosen" by one vertex (in solid black) form a tree T, while the edges "chosen" by two (in solid red) form a matching M. There is an assignment earning energy $\frac{T+M}{2} = 12$ from assigning a perfect classical cut to $T \cup M$, and one earning energy $M + \frac{M-m}{4} = 14$ from assigning the singlet to every edge in M and earning $w_e/4$ on every other edge.

disconnected from each other (because otherwise those connecting edges would have been explored before both subtrees were). This means we can do the following: choose either the even or odd levels (with level i being edges from depth-i vertices to depth-$(i+1)$ vertices) of the tree, one of which will contain at least half the edges; call this set of edges H. Now, H consists of disjoint bipartite subgraphs, and no edge outside the tree connects two edges in the same level of H. Thus, as noted above for the weighted case, there is an optimal QUANTUM MAX-CUT solution for H that still earns $1/4$ from every edge outside the tree, and from the edges in the unchosen levels. An optimal classical solution of this kind on H earns a QUANTUM MAX-CUT value of $1/2$ on each edge in H (by randomly selecting either a fixed assignment that cuts all the edges or the "bit-flipped" assignment, independently for each component of H).

Now, as the tree contains $W - 1$ edges, merely using the optimal classical solution would only earn us at least $\frac{W-1}{4} + \frac{m-(W-1)/2}{4}$, which is not quite as strong as we want. But each level of the tree is a disjoint union of stars, and the optimal QUANTUM MAX-CUT assignment for a star with d leaves earns $\frac{d+1}{2}$. So we can earn at least $1/2$ more energy, giving us

$$\frac{W-1}{4} + \frac{m-(W-1)/2}{4} + \frac{1}{2} > \frac{m}{4} + \frac{W}{8}$$

for a 2-approximation determined only by m and W. We illustrate this construction in Figure 4.

d) Estimating W in the Stream: To obtain an actual algorithm we will need a $(1+\varepsilon)$-multiplicative approximation to $m/2 + W/4$. Counting m is trivial, and in the unweighted case W can be approximated with cardinality estimation algorithms. So the problem we need to resolve (ideally in $O(\log n)$ space) is estimating

$$\sum_{u \in V} \max_{v \in N(u)} w_{uv}$$

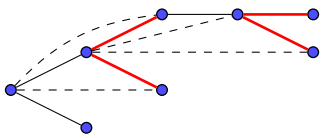


Fig. 4: Proving a lower bound for the optimal QUANTUM MAX-CUT value of an unweighted graph based on a DFS tree (the solid edges in the graph). The heavier half of the levels in the DFS (colored in red) are given an optimal assignment, and then every other edge (solid and dashed) earns 1/4. The total energy earned in this example is $\frac{3}{2} + \frac{3}{2} + \frac{7}{4} = 4.75$.

in the stream. Our approach is to use reservoir sampling to sample edges e with probability proportion to w_e, choose an endpoint at random, and then check whether they are higher-weight than every edge that arrives *after* them in the stream (since we can't check edges that arrive earlier). If we defined an estimator that is 1 whenever this happened and 0 otherwise, we would get a contribution of $w_e/2$ for every vertex u and $v \in N(u)$ such that w_{uv} was a "scenic viewpoint", an edge of higher weight than all subsequent edges incident to u.

To correct for this, we also check the weight w' of the highest-weight edge to arrive incident to u after uv (calling it 0 if uv is the last edge) and then subtract w'/w_{uv}. This gives us an estimator with expectation

$$W/2m$$

and constant variance, that we can compute in logarithmic space. So we could have trouble getting a multiplicative estimate of W if $m \gg W$, but this isn't a problem—we only want an estimate of $m/2 + W/4$, and so a εm-approximation of W suffices. This then gives us our full streaming algorithm, obtaining a $(2+\varepsilon)$- and $(5/2+\varepsilon)$-approximation in the unweighted and weighted case, respectively, using $\mathrm{O}(\log n)$ space if ε is constant.

FULL VERSION

Detailed proofs of the above claims are deferred to the full version of the paper, available at https://arxiv.org/abs/2206.00213.

ACKNOWLEDGEMENTS

The authors were supported by the Laboratory Directed Research and Development program at Sandia National Laboratories, a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. Also supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Accelerated Research in Quantum Computing program.

NOTICE

This article has been authored by an employee of National Technology & Engineering Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The employee owns all right, title and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan https://www.energy.gov/downloads/doe-public-access-plan.

REFERENCES

[1] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithm," *arXiv preprint arXiv:1411.4028*, 2014.

[2] B. Barak, A. Moitra, R. O'Donnell, P. Raghavendra, O. Regev, D. Steurer, L. Trevisan, A. Vijayaraghavan, D. Witmer, and J. Wright, "Beating the Random Assignment on Constraint Satisfaction Problems of Bounded Degree," in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2015)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), N. Garg, K. Jansen, A. Rao, and J. D. P. Rolim, Eds., vol. 40. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2015, pp. 110–123. [Online]. Available: http://drops.dagstuhl.de/opus/volltexte/2015/5298

[3] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell, "Optimal inapproximability results for max-cut and other 2-variable csps?" *SIAM J. Comput.*, vol. 37, no. 1, pp. 319–357, 2007.

[4] S. Gharibian and J. Kempe, "Approximation algorithms for qma-complete problems," *SIAM Journal on Computing*, vol. 41, no. 4, pp. 1028–1050, 2012.

[5] F. G. Brandão and A. W. Harrow, "Product-state approximations to quantum states," *Communications in Mathematical Physics*, vol. 342, no. 1, pp. 47–80, 2016.

[6] S. Bravyi, D. Gosset, R. König, and K. Temme, "Approximation algorithms for quantum many-body problems," *Journal of Mathematical Physics*, vol. 60, no. 3, p. 032203, 2019.

[7] S. Gharibian and O. Parekh, "Almost Optimal Classical Approximation Algorithms for a Quantum Generalization of Max-Cut," in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), D. Achlioptas and L. A. Végh, Eds., vol. 145. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2019, pp. 31:1–31:17. [Online]. Available: http://drops.dagstuhl.de/opus/volltexte/2019/11246

[8] O. Parekh and K. Thompson, "Beating Random Assignment for Approximating Quantum 2-Local Hamiltonian Problems," in *29th Annual European Symposium on Algorithms (ESA 2021)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), vol. 204, 2021, pp. 74:1–74:18. [Online]. Available: https://drops.dagstuhl.de/opus/volltexte/2021/14655

[9] A. Anshu, D. Gosset, and K. Morenz, "Beyond Product State Approximations for a Quantum Analogue of Max Cut," in *15th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2020)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), S. T. Flammia, Ed., vol. 158. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2020, pp. 7:1–7:15. [Online]. Available: https://drops.dagstuhl.de/opus/volltexte/2020/12066

[10] O. Parekh and K. Thompson, "Application of the Level-2 Quantum Lasserre Hierarchy in Quantum Approximation Algorithms," in *48th International Colloquium on Automata, Languages, and Programming (ICALP 2021)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), vol. 198, 2021, pp. 102:1–102:20. [Online]. Available: https://drops.dagstuhl.de/opus/volltexte/2021/14171

[11] A. Anshu, D. Gosset, K. J. Morenz Korol, and M. Soleimanifar, "Improved approximation algorithms for bounded-degree local hamiltonians," *arXiv preprint arXiv:2105.01193*, 2021.

[12] D. Aharonov, I. Arad, and T. Vidick, "Guest column: the quantum pcp conjecture," *Acm sigact news*, vol. 44, no. 2, pp. 47–79, 2013.

[13] A. Natarajan and T. Vidick, "Low-degree testing for quantum states, and a quantum entangled games pcp for qma," in *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2018, pp. 731–742.

[14] M. Kapralov and D. Krachun, "An optimal space lower bound for approximating max-cut," in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, ser. STOC 2019. New York, NY, USA: Association for Computing Machinery, 2019, pp. 277–288. [Online]. Available: https://doi.org/10.1145/3313276.3316364

[15] S. Hallgren, E. Lee, and O. Parekh, "An approximation algorithm for the MAX-2-Local Hamiltonian problem," in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2020)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2020.

[16] K. Ahn and S. Guha, "Graph sparsification in the semi-streaming model," *ICALP*, pp. 328–338, 2009.

[17] D. Aharonov and L. Zhou, "Hamiltonian sparsification and gap-simulation," in *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, ser. LIPIcs, vol. 124, 2019, pp. 2:1–2:21.

[18] R. O'Donnell, *Analysis of Boolean functions*. Cambridge University Press, 2014.

[19] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf, "Exponential separations for one-way quantum communication complexity, with applications to cryptography," in *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing*, ser. STOC '07. New York, NY, USA: ACM, 2007, pp. 516–525. [Online]. Available: http://doi.acm.org/10.1145/1250790.1250866

[20] R. Raz, "Fourier analysis for probabilistic communication complexity," *Computational Complexity*, vol. 5, no. 3, pp. 205–221, 1995.

[21] H. Klauck, "Lower bounds for quantum communication complexity," *SIAM Journal on Computing*, vol. 37, no. 1, pp. 20–46, 2007.

[22] E. Verbin and W. Yu, "The streaming complexity of cycle counting, sorting by reversals, and other problems," *SODA*, pp. 11–25, 2011.

[23] J. Kallaugher, M. Kapralov, and E. Price, "The sketching complexity of graph and hypergraph counting," in *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2018, pp. 556–567.

[24] M. Kapralov, S. Khanna, M. Sudan, and A. Velingker, "$(1 + \Omega(1))$-Approximation to MAX-CUT requires linear space," in *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19*, 2017, pp. 1703–1722. [Online]. Available: https://doi.org/10.1137/1.9781611974782.112

[25] C.-N. Chou, A. Golovnev, M. Sudan, A. Velingker, and S. Velusamy, "Linear space streaming lower bounds for approximating CSPs," in *STOC'22*, 2022. [Online]. Available: https://eccc.weizmann.ac.il/report/2021/086/

[26] J. Kahn, G. Kalai, and N. Linial, "The influence of variables on boolean functions," in *Proceedings of the 29th Annual Symposium on Foundations of Computer Science*, ser. SFCS '88. Washington, DC, USA: IEEE Computer Society, 1988, pp. 68–80. [Online]. Available: https://doi.org/10.1109/SFCS.1988.21923

[27] A. Ben-Aroya, O. Regev, and R. de Wolf, "A hypercontractive inequality for matrix-valued functions with applications to quantum computing and ldcs," in *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*, 2008.

[28] Y. Shi and X. Wu, "Limits of quantum one-way communication by matrix hypercontractive inequality," 2012.

[29] J. F. Doriguello and A. Montanaro, "Exponential Quantum Communication Reductions from Generalizations of the Boolean Hidden Matching Problem," in *15th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2020)*, vol. 158, 2020, pp. 1:1–1:16.

[30] S. Arunachalam and J. F. Doriguello, "Matrix hypercontractivity, streaming algorithms and ldcs: the large alphabet case," *arXiv preprint arXiv:2109.02600*, 2021.

[31] D. Kogan and R. Krauthgamer, "Sketching cuts in graphs and hypergraphs," in *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science*, 2015, pp. 367–376.

[32] M. Kapralov, S. Khanna, and M. Sudan, "Streaming lower bounds for approximating MAX-CUT," in *26th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2015.

[33] F. Le Gall, "Exponential separation of quantum and classical online space complexity," in *Proceedings of the eighteenth annual ACM symposium on Parallelism in algorithms and architectures*, ser. SPAA '06. New York, NY, USA: Association for Computing Machinery, Jul. 2006, pp. 67–73. [Online]. Available: https://doi.org/10.1145/1148109.1148119

[34] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf, "Exponential separation for one-way quantum communication complexity, with applications to cryptography," *SIAM J. Comput.*, vol. 38, no. 5, pp. 1695–1708, 2008.

[35] R. Jain and A. Nayak, "The Space Complexity of Recognizing Well-Parenthesized Expressions in the Streaming Model: The Index Function Revisited," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6646–6668, Oct. 2014.

[36] A. Nayak and D. Touchette, "Augmented index and quantum streaming algorithms for dyck(2)," in *Proceedings of the 32nd Computational Complexity Conference*, ser. CCC '17. Dagstuhl, DEU: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2017.

[37] A. Montanaro, "The quantum complexity of approximating the frequency moments," *Quantum Info. Comput.*, vol. 16, no. 13–14, pp. 1169–1190, Oct. 2016.

[38] Y. Hamoudi and F. Magniez, "Quantum Chebyshev's Inequality and Applications," in *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), vol. 132. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2019, pp. 69:1–69:16.

[39] J. Kallaugher, "A quantum advantage for a natural streaming problem," in *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, 2022, pp. 897–908.

[40] M. Charikar and A. Wirth, "Maximizing quadratic programs: extending grothendieck's inequality," in *45th Annual IEEE Symposium on Foundations of Computer Science*, 2004, pp. 54–60.

[41] Y. Hwang, J. Neeman, O. Parekh, K. Thompson, and J. Wright, "Unique games hardness of quantum max-cut, and a vector-valued borell's inequality," *arXiv preprint arXiv:2111.01254*, 2021.

[42] G. Gutin and A. Yeo, "Lower bounds for maximum weighted cut," *arXiv preprint arXiv:2104.05536*, 2021.