Concentration bounds for quantum states and limitations on the QAOA from polynomial approximations

Anurag Anshu[1] and Tony Metger[2]

[1]School of Engineering and Applied Sciences, Harvard University
[2]Institute for Theoretical Physics, ETH Zurich

Abstract

We prove concentration bounds for the following classes of quantum states: (i) output states of shallow quantum circuits, answering an open question from [DPMRF23]; (ii) injective matrix product states; (iii) output states of dense Hamiltonian evolution, i.e. states of the form $e^{\iota H^{(p)}} \cdots e^{\iota H^{(1)}} |\psi_0\rangle$ for any n-qubit product state $|\psi_0\rangle$, where each $H^{(i)}$ can be any local commuting Hamiltonian satisfying a norm constraint, including dense Hamiltonians with interactions between any qubits. Our proofs use polynomial approximations to show that these states are close to local operators. This implies that the distribution of the Hamming weight of a computational basis measurement (and of other related observables) concentrates.

An example of (iii) are the states produced by the quantum approximate optimisation algorithm (QAOA). Using our concentration results for these states, we show that for a random spin model, the QAOA can only succeed with negligible probability even at super-constant level $p = o(\log \log n)$, assuming a strengthened version of the so-called overlap gap property. This gives the first limitations on the QAOA on dense instances at super-constant level, improving upon the recent result [BGMZ22].

Anurag Anshu: anuraganshu@fas.harvard.edu
Tony Metger: tmetger@ethz.ch

Note: a short version of this work has appeared in the Proceedings of the 14th Innovations in Theoretical Computer Science Conference (ITCS 2023).

1 Introduction

Concentration bounds deal with the deviation of random variables from their expectation. Such bounds describe important structural properties of probability distributions that carry low correlation, and as a result have become ubiquitous tools in mathematics, computer science, and physics. A series of recent works has extended concentration bounds to the quantum many-body setting, where weakly correlated quantum states hold physical and computational relevance and their concentration properties explain important physical effects. For example, the concentration of expectation values in various quantum states such as product states [GV89, HMH04, Abr20], Gibbs quantum states [DPR22, KS20b], and finitely correlated states [Ans16] explains the equivalence of ensembles [BC15, BCG15, Tas18, Alh22] and eigenstate thermalisation [KS20a] in quantum statistical mechanics. Concentration bounds are also an important proof technique in quantum complexity theory: for example, concentration properties of quantum states generated by low-depth quantum circuits have been used to prove circuit lower bounds on low-energy states of quantum Hamiltonians [EH17], leading to the recent proof of the NLTS conjecture [AN22, AB22, ABN22]. Furthermore, concentration results play an important role in analysing and bounding the performance of variational quantum algorithms for classical constraint optimisation problems, e.g. the quantum approximate optimisation algorithm (QAOA) [FGG14, FGG20, CLSS22, BGMZ22].

The standard method to prove concentration inequalities is the moment method. The slow growth of the moments of weakly correlated probability distributions or quantum states serves as a signature of concentration. Arguably the simplest example is the classical Chernoff-Hoeffding bound, which shows that the probability that the sum of n independent random variables deviates from its expectation value by more than k is at most $e^{-\Omega(k^2/n)}$. We call this *Gaussian concentration*. The moment method also extends to non-commuting observables with product quantum states [Kuw16] or quantum states with exponential decay of correlation [Ans16], albeit sometimes with weaker bounds of the form $e^{-\Omega(k^\alpha/n^\beta)}$ for some $\alpha, \beta > 0$. We call these weaker bounds *exponential concentration*.[1]

An alternative to the moment method was introduced in [KAAV17]. The idea of this approach is to approximate the quantum state of interest by a local operator and use this approximation to prove concentration bounds. Because the local approximation is usually constructed from low-degree polynomials, we call this the *polynomial-based method*. The strength of the resulting concentration bounds depends on the locality of the approximation; see Section 2.4 for details. [KAAV17] used this method to show exponential concentration of the form $e^{-\Omega(k/\sqrt{n})}$ for local classical observables on the ground states of gapped local Hamiltonians. However, their method was not able to produce Gaussian concentration even for the simplest case of local observables on an i.i.d. distribution (e.g. the sum of i.i.d. random variables), a case in which the Chernoff-Hoeffding bound does give Gaussian concentration results.

1.1 Main results

We extend the polynomial-based method in two ways: we show that in cases where the moment method produces Gaussian concentration, the polynomial-based method can do so, too; and we

[1] Note that with this definition, technically Gaussian concentration is a special case of exponential concentration for $\alpha = 2$ and $\beta = 1$. However, when we say "exponential concentration" it is usually implicit that α and β are such that the bounds are generally weaker than for Gaussian concentration. A typical case is $\alpha = 1$ and $\beta = 1/2$.

show that the polynomial-based method can be applied to a much wider class of states, including the output states of shallow quantum circuits, injective matrix product states, and the output states of dense Hamiltonian evolutions (explained below). An example of a dense Hamiltonian evolution is the QAOA for solving classical constraint optimisation problems (COPs). We therefore obtain concentration bounds for the QAOA, and combining this with the so-called *overlap gap property* first introduced in the classical literature [GL18, GJ21], we prove strong limitations on the performance of the QAOA even at (admittedly only slightly) super-constant level $p = o(\log \log n)$. Crucially, our method works for *dense* COPs, which may have constraints between any variables, and our proofs are fairly straightforward. This improves upon the recent work [BGMZ22], which proved similar results for constant-level QAOA on dense instances using a highly technical proof.

We now describe our main results in more detail. In all cases, concentration bounds are obtained by approximating the quantum state of interest by a local operator, so in sketching the proof ideas, we only focus on the construction of such a local approximation. Below, we only give concentration bounds for the Hamming weight distribution W_ρ of an n-qubit quantum state ρ, i.e. the probability distribution over $\{1, \ldots, n\}$ describing the Hamming weight of a computational basis measurement of ρ. More formally, $\Pr[W_\rho = i] = \sum_{x \in \{0,1\}^n : |x| = i} \langle x|\rho|x\rangle$. However, it is straightforward to extend these bounds to other observables that only change slowly as the Hamming weight changes; for an example, see Corollary 4.8.

	Prior work	This work
Depth t quantum states	$e^{-\Omega(k/\sqrt{2^{2t}n})}$ (implicit in [KAAV17])	$e^{-\Omega(k^2/2^{2t}n)}$
Injective matrix product states	$e^{-\Omega(k/\sqrt{n})}$ ([KAAV17, Ans16])	$e^{-\Omega(k^2/n)}$
Dense Hamiltonian evolution	$o_n(1)$ for the special case of QAOA with level $p = O(1)$ [BGMZ22]	$e^{-\Omega(n^{1/8})}$ for $k = o(n)$ and level $p = o(\log \log n)$

Table 1: Main results and comparison with prior work. For each case, we consider a state on n qubits and bound the probability that the Hamming weight of a computational basis measurements deviates by more than $k \in [0, n]$ from its median (or mean). Note that [BGMZ22] were able to show concentration over both the choice of instance and the randomness of the QAOA output, whereas our bounds, while stronger, only deal with concentration over the latter.

Shallow quantum circuits. Consider a depth-t quantum circuit, i.e. a circuit comprised of t layers of arbitrary 2-qubit gates applied to the initial state $|0\rangle^{\otimes n}$. We denote the unitary implemented by this circuit by U. The output state of this circuit is the unique maximum-energy eigenstate of a 2^t-local Hamiltonian that is a sum of commuting projectors. To see this, observe that $U|0\rangle^{\otimes n}$ is the unique joint $(+1)$-eigenstate of the operators $H_i = U|0\rangle\langle 0|_i U^\dagger$, where $|0\rangle\langle 0|_i$ acts as identity on all qubits except i. By a standard lightcone argument, H_i only acts non-trivially on 2^t qubits, so $H = \frac{1}{n}\sum H_i$ is a 2^t-local Hamiltonian with $U|0\rangle^{\otimes n}$ as its unique $(+1)$-eigenstate. Therefore, the output of the circuit can be written as $U|0\rangle\langle 0|^{\otimes n} U^\dagger = \delta_1(H)$, where $\delta_1(1) = 1$ and $\delta_1(x) = 0$ for $x \neq 1$. We can now approximate $\delta_1(x)$ using a degree-d polynomial P_d constructed

in [KLS96, BCDZ99, AAG22]. Polynomials spread the locality of operators in a controllable way. We can therefore show that $P_d(H)$ is a $(d \cdot 2^t)$-local operator that approximates $U|0\rangle\langle 0|^{\otimes n}U^\dagger$. Hence, we have constructed a local operator approximation to $U|0\rangle\langle 0|^{\otimes n}U^\dagger$ and can use this approximation to show that for any depth-t circuit, the output state $|\psi\rangle\langle\psi| = U|0\rangle\langle 0|^{\otimes n}U^\dagger$ has the following concentration property for $k \in (2^t\sqrt{n}, 2^t n)$ (see Corollary 3.2 for the formal statement):

$$\Pr[|W_\psi - \mathrm{median}(W_\psi)| \geq k] \leq e^{-\Omega\left(\frac{k^2}{2^{2t}n}\right)}.$$

This generalises the Chernoff-Hoeffding bound for product distributions (which corresponds to the case of a single layer ($t = 1$), since then the measurement distribution of the output state is a product of distributions on one or two bits) and shows Gaussian concentration for any constant-depth quantum circuit, answering an open question from [DPMRF23]. Similar statements have also appeared in [AN22, AB22, ABN22]. Note that the same argument applies to any ground state of a Hamiltonian that is a sum of commuting projectors, not just output states of shallow circuits.

Injective matrix product states. Matrix product states (MPSs) are a widely used tensor network representation of quantum states. Injective MPSs have an additional property that ensures that they are the unique ground state of a local "parent Hamiltonian" with a constant spectral gap. We can therefore approximate an injective MPS as a polynomial of its parent Hamiltonian. Using near-optimal polynomial approximations constructed in [AAG22], we obtain Gaussian concentration bounds for injective MPSs (Lemma 3.3). Our bounds are stronger than previous ones [Ans16, KAAV17], which only showed exponential concentration. We also note that conditionally independent probability distributions can be encoded into injective MPSs, and that in that case our concentration bounds reproduce a (version of) Azuma's inequality.

Dense Hamiltonian evolution. Concentration bounds are natural for quantum states that have weak long-range correlations such as the output states of shallow quantum circuits. A priori, one would not expect similar bounds to hold for quantum states with long-range correlations. Recently, [BGMZ22] considered the output distribution of the QAOA (explained below) on random dense COPs, i.e. local COPs that can have constraints between any variables. For dense COPs, the operations implemented by the QAOA can include interactions between any qubits, and as a result the output distribution can have long-range correlations. Remarkably, [BGMZ22] showed that the variance of the average energy density (averaged over the randomness of the QAOA as well as the choice of random instance) vanishes asymptotically. This means that the energy density of the output, which corresponds to the quality of the COP solution produced by the QAOA, concentrates about the average. However, [BGMZ22] were only able to prove an asymptotic statement without explicit tail bounds and the proof was highly non-trivial.

We consider a more general class of states that includes the output states of the QAOA as a special case. Specifically, we define the output of a *dense Hamiltonian evolution* as a state of the form $e^{\iota H^{(p)}} \cdots e^{\iota H^{(1)}}|\psi_0\rangle$ for any n-qubit product state $|\psi_0\rangle$. Here, each $H^{(i)}$ can be any commuting local Hamiltonian (though the different $H^{(i)}$ themselves are of course not required to commute). Importantly, $H^{(i)}$ are allowed to be *dense* Hamiltonians, i.e. Hamiltonians with interactions between any qubits. As explained below, the QAOA applied to a dense COP is a special case of dense Hamiltonian evolution.

For our concentration bounds, we further require each $H^{(i)}$ to satisfy a norm constraint, which limits the norm of the Hamiltonian restricted to a subset of the qubits (see Equation (4.1) for the

formal statement). In particular, this condition is satisfied with overwhelming probability for the random dense model from [BGMZ22]. Under this condition, we can prove (Theorem 4.1) that the output state of a dense Hamiltonian evolution is ϵ-close in operator norm to a k_p-local operator for

$$k_p \leq c_1^p n^{1-(1-\alpha)^p/4}, \quad \epsilon \leq e^{-\Omega(n^{1/8})}.$$

Here, c_1 and $0 \leq \alpha < 1$ are constants and p is *level* of the dense evolution, i.e. the number of unitaries $e^{\iota H^{(i)}}$ that have been applied. In particular, if we choose $p = o(\log\log n)$, then $k = o(n)$. This implies the following exponential concentration result (Corollary 4.7): for ρ_p the output of a dense Hamiltonian evolution with level $p = o(\log\log n)$ satisfying the above conditions,

$$\Pr[|W_{\rho_p} - \mathrm{median}(W_{\rho_p})| > o(n)] \leq e^{-\Omega(n^{1/8})}.$$

Here, we have only stated the asymptotic result, but in Corollary 4.7 we give explicit bounds for any choice of p. This concentration result can also be extended beyond just the Hamming weight of a computational basis measurement: for example, it also holds for the energy density of ρ_p with respect to any classical Hamiltonian satisfying a similar norm constraint to the one mentioned above (Corollary 4.8).

To prove that the output of a dense Hamiltonian evolution can be approximated by a local operator, we again make use of polynomial approximations. Recall that we are interested in states of the form $\rho_p = e^{\iota H^{(p)}} \cdots e^{\iota H^{(1)}} |\psi_0\rangle\langle\psi_0| e^{-\iota H^{(1)}} \cdots e^{-\iota H^{(p)}}$ for a pure product state $|\psi_0\rangle$. As a first step, we approximate $|\psi_0\rangle\langle\psi_0|$ by a local operator (Lemma 4.5). For this, we observe that since $|\psi_0\rangle = \otimes_i |\psi_0\rangle_i$ is a product state, it is the unique ground state of the 1-local Hamiltonian $H = \frac{1}{n}\sum |\psi_0\rangle\langle\psi_0|_i$. If we apply a linear combination of Chebyshev polynomials to this Hamiltonian, we obtain a good local approximation to $|\psi_0\rangle\langle\psi_0|$. Then, for each unitary $e^{\iota H^{(i)}}$ in the dense Hamiltonian evolution, we approximate the exponential function by its truncated Taylor series. The fact that the Hamiltonian evolution is applied to an approximately local operator allows us to use the norm constraint mentioned above to obtain an improved error bound for the truncated Taylor series (see Lemma A.2 for details). Therefore, the truncated Taylor series spreads the locality of the state in a controllable way without degrading the quality of the approximation too much. Applying this argument recursively for each layer of the dense Hamiltonian evolution, we obtain a local approximation to the output state ρ_p.

Limitations on the QAOA from concentration bounds. The QAOA [FGG14] is an algorithm for solving local COPs (i.e. COPs consisting of any number of clauses, each with at most $q = O(1)$ variables) on a quantum computer. We can associate a q-local Hamiltonian H with every q-local COP C by replacing the variables in C with Pauli-Z matrices acting on different qubits. The resulting Hamiltonian is diagonal in the computational basis and has the property that for any string $x \in \{0,1\}^n$, $C(x) = \langle x|H|x\rangle$. The QAOA attempts to find a "good" solution x (i.e. one for which $C(x)$ is as large as possible) by starting from the state $|+\rangle^{\otimes n}$ and then applying p layers of unitaries of the form $e^{\iota\beta_i \sigma_X^{\otimes n}} e^{\iota\gamma_i H}$. Here, β_i and γ_i are real parameters that can be tuned to the problem instance. It is clear that this is a special case of the dense Hamiltonian evolution we have described earlier. Our results will apply for any choice of β_i and γ_i and we will always implicitly consider a family of COPs, one for each number n of input bits, in order to make asymptotic statements.

[BGMZ22] considered the performance of the QAOA on a random spin model on n qubits, described by the q-local Hamiltonian

$$H_n^q(J) = \frac{1}{n^{(q-1)/2}} \sum_{i_1,\dots i_q=1}^{n} J_{i_1,\dots i_q} \sigma_{i_1}^Z \dots \sigma_{i_q}^Z \,, \tag{1.1}$$

where $J_{i_1,\dots i_q} \sim \mathcal{N}(0,1)$ are sampled from i.i.d. standard Gaussians. For this model, [BGMZ22] were able to show that for constant even $q \geq 4$ and level $p = O(1)$, the value achieved by the QAOA (for fixed β_i, γ_i) in expectation over J and the internal randomness of the QAOA is bounded away from the optimal value by a constant as $n \to \infty$. They were also able to show the asymptotic concentration property described above.

Here, we use our concentration results to show limitations on the QAOA for a class of COPs that includes the random spin model above. For this, we consider local COPs that have the so-called overlap gap property (OGP) [Gam21], which roughly says that "good" solutions to the COP are clustered in the sense that two good solutions are either close or far in Hamming distance. Combining this with our concentration results for dense Hamiltonian evolution, we can show that for any COP with a sufficiently strong OGP whose associated Hamiltonian satisfies Equation (4.1), if the QAOA produced a good solution with noticeable probability, then the probability distribution over good solutions produced by the QAOA would have to be concentrated on one such cluster. This allows us to show that the QAOA cannot succeed with noticeable probability on *symmetric* COPs (i.e. COPs that are invariant under flipping all the input bits) that have a strong OGP. This is because the symmetry of the COP is in contradiction with the existence of a single cluster on which most of the probability distribution is concentrated: if such a cluster existed, we could take the strings in that cluster and flip all their bits to produce another cluster which, by symmetry, must have the same probability weight, a contradiction. This argument is similar to [BKKT]. As a result, we obtain the following limitation on the QAOA (see Lemma 5.5).

Theorem (informal)**.** *Consider a local symmetric COP $C(x)$ with a sufficiently strong OGP and suppose that the associated Hamiltonian H satisfies the norm constraint in Equation (4.1). Then, the value of the solution to $C(x)$ produced by the QAOA with level $p = o(\log\log n)$ is bounded away from the optimal value by at least a constant except with probability $e^{-\Omega(n^{1/8})}$.*

For even q, the random spin model from Equation (1.1) is symmetric and satisfies the norm constraint in Equation (4.1) with overwhelming probability. Furthermore, it was shown in [CGPR19] that it satisfies the OGP with overwhelming probability. However, we note that here we need a stronger version of the OGP than was shown in [CGPR19]. This stronger version appears to be implicit in their proof, too, although we leave its formal proof for future work and assume it here as a conjecture. Assuming this stronger OGP, we can show the following (see Section 5.3).

Theorem (informal)**.** *With probability $1 - O(e^{-n})$ over the choice of J (with i.i.d. Gaussian entries), the value of the solution to the random spin model (Equation (1.1)) produced by the QAOA with level $p = o(\log\log n)$ is bounded away from the optimal value by at least a constant except with probability $e^{-\Omega(n^{1/8})}$.*

This places strong limitations on the performance of the QAOA because it does not just bound the *expectation* value away from the optimal value as in [BGMZ22], but instead asserts that the QAOA output is bounded away from the optimal value with overwhelming probability, even at super-constant level $p = o(\log\log n)$.

1.2 Discussion and open questions

We have shown that polynomial approximations can be used to derive Gaussian concentration bounds for the output states of constant-depth quantum circuits and injective matrix product states, and exponential concentration bounds for the output states of dense Hamiltonian evolution. The latter can be used to derive strong limitations on the performance of the QAOA at super-constant level $p = o(\log \log n)$ even on dense instances such as random spin models.

At first sight, it is surprising that the (provably optimal) polynomial approximations [KLS96, BCDZ99] we use for shallow quantum circuits are able to reproduce the (likewise provably optimal) Chernoff-Hoeffding bound in the classical case. It would be interesting to explore whether there is a deeper conceptual connection between optimal polynomial approximations and optimal concentration bounds.

On a more technical level, there are a number of interesting improvements one could hope to make to our bounds. Firstly, our bounds for MPSs (Lemma 3.3) can only deal with sub-linear deviations $k = O(n^{1-\delta})$ for any $\delta > 0$. It would be desirable to extend this result to arbitrary values of k. Additionally, one could hope to prove similar concentration bounds for PEPSs, the two-dimensional analogue of MPSs.

Secondly, we only achieve exponential, not Gaussian, concentration bounds for dense Hamiltonian evolutions with level $p = o(\log \log n)$. Can one improve these results to Gaussian concentration and also extend them to higher levels, e.g. $p = O(\log n)$ or even $p = O(n^{\delta})$ for a small $\delta > 0$? Furthermore, we show concentration for the output states of dense Hamiltonian evolution for a fixed instance, but we cannot show that for random COPs, the output states also have concentration properties over the choice of random instance, e.g. over the choice of $J_{i_1 \ldots i_q} \sim \mathcal{N}(0, 1)$ in the case of the random spin model introduced earlier. [BGMZ22] do show such a concentration property, albeit only in the asymptotic regime without explicit bounds. Can our polynomial approximation techniques also be used to prove explicit concentration bounds over the choice of random instance? If so, it might be possible to extend the limitations on the performance of dense evolutions for COPs we prove in Section 5.2 beyond symmetric COPs and optimisers.

Finally, our techniques may also be useful for problems in condensed matter physics. As an example, consider the Lieb-Schultz-Mattis theorem [LSM61] and its higher-dimensional generalisation [Has04], seminal results in condensed matter physics. Their main idea is that sufficient symmetry and non-degeneracy of the ground space prevents a Hamiltonian from being gapped. Inspired by our application to symmetric QAOA (Section 5.3), we can ask whether an alternative proof of this result can be obtained using concentration bounds and polynomial approximations, e.g. by showing that the concentration properties of unique gapped ground states are in conflict with the symmetry requirements.

Acknowledgements. We thank Joao Basso, David Gamarnik, Song Mei, and Leo Zhou for very helpful discussions and especially for suggesting the application to symmetric QAOA (Section 5.3). AA also thanks Daniel Stilck França, Tomotaka Kuwahara, Cambyse Rouzé, and Juspreet Singh Sandhu for helpful discussions. This work was done in part while the authors were visiting the Simons Institute for the Theory of Computing. AA acknowledges support through the NSF CAREER Award No. 2238836 and NSF award QCIS-FF: Quantum Computing & Information Science Faculty Fellow at Harvard University (NSF 2013303). TM acknowledges support from the ETH Zurich Quantum Center.

2 Preliminaries

2.1 Notation

For a bitstring $x \in \{0,1\}^n$, we denote by $(-1)^x$ the string $((-1)^{x_1}, \ldots, (-1)^{x_n})$. For $n \in \mathbb{N}$, $[n] \coloneqq \{1, \ldots, n\}$. We consider the n-qubit Hilbert space $\mathcal{H} = (\mathbb{C}^2)^n$ throughout. $L(\mathcal{H})$ is the set of linear operators on $\mathcal{H}$, and $\mathcal{D}(\mathcal{H}) = \{A \in \mathrm{Pos}(\mathcal{H}) \mid \mathrm{Tr}[A] = 1\}$ is the set of density matrices on $\mathcal{H}$. For a subset $S \in \{0,1\}^n$, we use Π_S to denote the projector onto S, i.e. $\Pi_S = \sum_{x \in S} |x\rangle\langle x|$. The single qubit Pauli operators are $\sigma_X = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\sigma_Z = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$.

For a vector $|\psi\rangle \in \mathcal{H}$ with entries ψ_i and $p \in [1, \infty]$, the p-norm is defined as $\||\psi\rangle\|_p = \left(\sum_i |\psi_i|^p\right)^{1/p}$. For an operator $A \in L(\mathcal{H})$ and $p \in [1, \infty]$, the Schatten p-norm $\|A\|_p$ is the (vector) p-norm of the vector of singular values of A. For $p = \infty$, this norm is called the operator norm and can also be written as $\|A\|_\infty = \sup_{\||\psi\rangle\|_2 = 1} \|A|\psi\rangle\|_2$. For convenience, we usually drop the subscript for the operator norm, i.e. $\|A\| \coloneqq \|A\|_\infty$.

Definition 2.1 (Hamming weight distribution)**.** For a quantum state $\rho \in \mathcal{D}(\mathcal{H})$ we denote by W_ρ the random variable indicating the Hamming weight of the outcome of measuring ρ in the computational basis, i.e.

$$\Pr[W_\rho = i] = \sum_{x \in \{0,1\}^n : |x| = i} \langle x|\rho|x\rangle \,.$$

2.2 Local operators

Definition 2.2 (Local operators)**.** Let $k \in \mathbb{N}$ and $\epsilon > 0$. An operator $R \in L(\mathcal{H})$ is called k-local if it can be written as $R = \sum_i R_i$ for operators R_i that only act non-trivially on k subsystems. Whenever we write $R = \sum_i R_i$ for a k-local operator R, this is understood to be such a local decomposition. An operator $Q \in L(\mathcal{H})$ is called (k, ϵ)-local if there exists a k-local operator R such that $\|Q - R\| \leq \epsilon$.

We will frequently consider local operators with additional properties and extend the above definition in the obvious way: for example, a (k, ϵ)-local state is a quantum state that is ϵ-close in operator norm to a k-local operator R. Note that the operator R does not need to be a quantum state itself.

Definition 2.3 (Total local norm)**.** A k-local operator R has total local norm $\textsc{tln}(R) \leq r$ if there exists a local decomposition $R = \sum R_i$ such that $\sum \|R_i\| \leq r$. Similarly, a (k, ϵ)-local operator Q has approximate total local norm $\textsc{tln}_\epsilon(Q) \leq q$ if there exists a k-local operator R within ϵ-distance (in operator norm) from Q with $\textsc{tln}(R) \leq q$.

Note that in the definition of $\textsc{tln}(R)$ and $\textsc{tln}_\epsilon(Q)$, it always has to be clear from the context which locality k we are considering, since for different choices of k, different values of the total local norm can be achieved. Therefore we will always make statements of the form: R is a k-local operator with $\textsc{tln}(R) = \ldots$ (and likewise for the approximate case). Hence, strictly speaking the subscript in $\textsc{tln}_\epsilon(Q)$ is unnecessary as it must anyway be specified that Q is a (k, ϵ)-local operator for some values of k and ϵ, but we find it useful to include the subscript as a reminder of this nonetheless.

Example 2.4. The locality of an operator increases in a controllable way when we apply a polynomial. Specifically, consider a k-local operator $R = \sum R_i$ and a degree-d polynomial P. Then, $P(R)$ is a $(d \cdot k)$-local operator because we can expand it into a sum of terms, each of which contains a product of at most d different local terms R_i.

Example 2.5. The state $|+\rangle\langle+|^{\otimes n}$ is a non-local operator. However, it can be approximated by a local operator. For this, consider the Hamiltonian $H_0 = \frac{1}{n}\sum_{i=1}^{n} |+\rangle\langle+|_i$, where $|+\rangle\langle+|_i$ acts as identity everywhere except on the i-th qubit. H_0 is a 1-local operator, with ground state $|+\rangle\langle+|^{\otimes n}$ and spectral gap $1/n$. It is easy to see that $\||+\rangle\langle+|^{\otimes n} - H_0^k\| \leq (1-\frac{1}{n})^k \leq e^{-k/n}$, and by Example 2.4 H_0^k is a k-local operator. We therefore have shown that $|+\rangle\langle+|^{\otimes n}$ is a $\left(k, (1-\frac{1}{n})^k\right)$-local operator. Expanding $H_0^k = \frac{1}{n^k}\sum_{i_1,i_2,\ldots i_k} |+\rangle\langle+|_{i_1} \cdots |+\rangle\langle+|_{i_k}$ into n^k terms each with operator norm at most 1, we also see that $\textsc{tln}(H_0^k) \leq 1$ for any k, and consequently $\textsc{tln}_\epsilon(|+\rangle\langle+|^{\otimes n}) \leq 1$ for any ϵ.

Definition 2.6 (Subset operator)**.** For any k-local operator $R = \sum R_i$ and a subset $S \subseteq [n]$, we define the subset operator R_S as

$$R_S = \sum_{\substack{i \text{ s.t. } R_i \text{ acts non-trivially} \\ \text{on at least one qubit in } S}} R_i \,.$$

The decomposition of a k-local operator $R = \sum_i R_i$ we have considered so far is not unique. It will occasionally be useful to define a canonical such decomposition. This can easily be done as follows: noting that the n-qubit Pauli matrices (with identity) form a basis of $L(\mathcal{H})$, for any operator $R \in L(\mathcal{H})$ there is a unique decomposition in terms of the n-qubit Pauli matrices. Furthermore, if R is k-local, only basis elements that act non-trivially on at most k qubits will appear in this decomposition. Starting from this unique decomposition, we can group basis elements that act non-trivially on the same set of qubits. This way, we obtain a unique decomposition of a k-local operator R as

$$R = \sum_{T \subseteq [n] \text{ s.t. } |T| \leq k} O_{(T)} \,, \tag{2.1}$$

where $O_{(T)}$ is a (weighted) sum of all Pauli operators that act non-trivially exactly on qubits in the set T. The following lemma bounds the operator norm of an individual term in this decomposition.

Lemma 2.7. *Given an operator R, consider the unique decomposition $R = \sum_T O_{(T)}$ from Equation (2.1). Then, for all $T \subseteq [n]$:*

$$\|O_{(T)}\| \leq 2^{|T|}\|R\| \,.$$

Proof. Defining $O'_{(S)} := \sum_{T \subseteq S} O_{(T)}$, we can write $O'_{(S)}$ as a Pauli twirl applied to the operator R:

$$O'_{(S)} = \mathbb{E}_i \sigma^i_{[n]\setminus S}(R)\sigma^i_{[n]\setminus S} \,,$$

where $\{\sigma^i_{[n]\setminus S}\}_{i=1}^{4^{n-|S|}}$ is the set of all multi-qubit Pauli operators that act as identity on S. By the inclusion-exclusion principle (see e.g. [GGL95, Thm. 12.1]),

$$O_{(T)} = \sum_{U \subseteq T} (-1)^{|T \setminus U|} O'_{(U)} \,.$$

Thus,

$$\|O_{(T)}\| \leq 2^{|T|} \max_{U \subseteq T} \|O'_U\| \leq 2^{|T|} \max_{U \subseteq T} \|\mathbb{E}_i \sigma^i_{[n]\setminus U}(R)\sigma^i_{[n]\setminus U}\| = 2^{|T|}\|R\| \,,$$

where the last equality holds because the Pauli twirl is a unital channel. □

2.3 Polynomial approximations

Definition 2.8 (Chebyshev polynomials)**.** The Chebyshev polynomials are defined as

$$T_k(x) = \frac{k}{2} \sum_{r=0}^{\lfloor \frac{k}{2} \rfloor} \frac{(-1)^r}{k-r} \binom{k-r}{r} (2x)^{k-2r} := \sum_{r=0}^{\lfloor \frac{k}{2} \rfloor} \alpha_r x^{k-2r} \,.$$

We have already seen the utility of the powering function $x \mapsto x^s$ in creating local operator approximations in Example 2.5. However, the powering function does not result in approximations with an optimal tradeoff between approximation error and locality. For this purpose, we require an approximation to the function $x \mapsto x^s$ as a linear combination of Chebyshev polynomials.

Lemma 2.9. *Let T_k be the degree-k Chebyshev polynomial and $p_{s,k}$ the probability that an s-step symmetric random walk on integers (starting from 0) is at k or $-k$. For any $a \leq s$, we define the degree-a polynomial*

$$P_{s,a}(x) = \sum_{k=0}^{a} p_{s,k} T_k(x) \,.$$

Then, $P_{s,a}$ provides a good approximation to x^s in the sense that for any operator O with $\|O\| \leq 1$,

$$\|O^s - P_{s,a}(O)\| \leq 2e^{-\frac{(a+1)^2}{2s}} \,.$$

Proof. As shown in [SV14, Theorem 3.1], the monomial x^s can be viewed as a random walk over the Chebyshev polynomials. More precisely, for any $s > 0$ we have

$$x^s = \sum_{k=0}^{s} p_{s,k} T_k(x) = P_{s,s}(x) \,. \tag{2.2}$$

By the Chernoff bound, $\sum_{k \geq a} p_{s,k} \leq 2e^{-\frac{a^2}{2s}}$. This means that the contributions from the high-degree Chebyshev polynomials in Equation (2.2) are suppressed, and we can obtain a good approximation to x^s by truncating Equation (2.2) at a degree $a < s$. Specifically, for any operator O with $\|O\| \leq 1$ we can bound

$$\left\| O^s - \sum_{k=0}^{a} p_{s,k} T_k(O) \right\| = \left\| \sum_{k=a+1}^{s} p_{s,k} T_k(O) \right\| \leq \sum_{k=a+1}^{s} |p_{s,k}| \leq 2e^{-\frac{(a+1)^2}{2s}} \,.$$

Here, the first inequality holds by the triangle inequality and because $|T_k(x)| \leq 1$ for $|x| \leq 1$. □

2.4 Concentration bounds from local operator approximations

Our general strategy for proving concentration bounds on quantum states will be to show that these states can be approximated by local operators. From this we can obtain concentration bounds from the following lemma, adapted from [KAAV17]. We state this lemma for the Hamming weight distribution W_ρ of a state ρ, but it can easily be generalised to the distribution of any function of bitstrings that varies slowly as the Hamming weight is changed (see Corollary 4.8 for an example).

Lemma 2.10. *Let $\psi = |\psi\rangle\langle\psi|$ be a (k,ϵ)-local pure quantum state and let m be the median of W_ψ. Then*

$$\Pr[W_\psi > m+k] \leq 4\epsilon^2 \qquad \text{and} \qquad \Pr[W_\psi < m-k] \leq 4\epsilon^2 \,.$$

Proof. We only prove the first bound, the second is analogous. We define

$$\Pi_{>m+k} = \sum_{x\in\{0,1\}:|x|>m+k} |x\rangle\langle x|$$

as the projector onto all computational basis states with Hamming weight greater than $m+k$. We define $\Pi_{\leq m}$ analogously. Because strings in the support of $\Pi_{>m+k}$ and $\Pi_{\leq m}$ differ on more than k positions, for any operator O that only acts non-trivially on k qubits we have that $\Pi_{>m+k} O \Pi_{\leq m} = 0$. By linearity we also get $\Pi_{>m+k} R \Pi_{\leq m} = 0$ for any k-local operator R.

Because ψ is (k,ϵ)-local, there exists a k-local operator R that is ϵ-close to ψ. We therefore get that

$$\epsilon \geq \|\psi - R\| \geq \|\Pi_{>m+k}(\psi - R)\Pi_{\leq m}\| = \|\Pi_{>m+k}\psi\Pi_{\leq m}\| \,.$$

By the variational definition of the operator norm and inserting $\psi = |\psi\rangle\langle\psi|$:

$$\begin{aligned}
\|\Pi_{>m+k}\psi\Pi_{\leq m}\| &\geq \|\Pi_{>m+k}|\psi\rangle\langle\psi|\Pi_{\leq m}|\psi\rangle\|_2 \\
&= \langle\psi|\Pi_{\leq m}|\psi\rangle \, \|\Pi_{>m+k}|\psi\rangle\|_2 \\
&= \mathrm{Tr}[\Pi_{\leq m}\psi] \cdot (\mathrm{Tr}[\Pi_{>m+k}\psi])^{1/2} \\
&= \Pr[W_\psi \leq m] \cdot (\Pr[W_\psi > m+k])^{1/2} \,.
\end{aligned}$$

The lemma now follows by noting that since m is the median of W_ψ, $\Pr[W_\psi \leq m] \geq 1/2$. □

More generally, we observe that approximately local quantum states have the following clustering property (see also [AN22, AB22, ABN22]).

Lemma 2.11. *Let $\psi = |\psi\rangle\langle\psi|$ be a (k,ϵ)-local operator and consider two sets $S, S' \subset \{0,1\}^n$ with Hamming distance more than k from each other. Then, either $\mathrm{Tr}[\Pi_S\rho] \leq \sqrt{\epsilon}$ or $\mathrm{Tr}[\Pi_{S'}\rho] \leq \sqrt{\epsilon}$ (or both).*

Proof. Letting R be the k-local operator ϵ-close to ψ, we can bound

$$\mathrm{Tr}[\Pi_S\psi]\,\mathrm{Tr}[\Pi_{S'}\psi] = \langle\psi|\Pi_S\psi\Pi_{S'}|\psi\rangle \leq \|\Pi_S\psi\Pi_{S'}\| \leq \|\Pi_S R\Pi_{S'}\| + \epsilon = \epsilon \,.$$

The last equality follows by the same reasoning as in the proof of Lemma 2.10. This implies the lemma. □

3 Concentration bounds for shallow circuits and matrix product states

[AAG22] showed that the ground states of various local Hamiltonians can be approximated by local operators using polynomials. Combining this with Lemma 2.10, we can easily obtain concentration bounds for such states. As concrete examples, we prove concentration bounds for the output states of shallow quantum circuits and injective matrix product states (MPSs).

We begin by considering the special case of ground states (or, for convenience, maximum-energy eigenstates) of commuting local Hamiltonians $H = \sum H_i$, where each H_i is a projector.

Lemma 3.1. *Let $H = \sum_{i=1}^{r} H_i$ be a sum of r commuting projectors each acting on ℓ qubits, and $|\psi\rangle$ the maximum-energy eigenstate of H. Then, for any $d \in (\sqrt{r}, r)$, the state $|\psi\rangle\langle\psi|$ is $\left(\ell d, 2^{-\frac{d^2}{2^8 r}}\right)$-local and*

$$\Pr[W_\psi > m + \ell d] \leq 4 \cdot 2^{-\frac{d^2}{2^7 r}} \qquad \text{and} \qquad \Pr[W_\psi < m - \ell d] \leq 4 \cdot 2^{-\frac{d^2}{2^7 r}} ,$$

where m is the median of W_ψ.

Proof. Because H is a sum of commuting projectors, $|\psi\rangle\langle\psi| = \prod H_i$. [AAG22, Theorem 3.1] (see also [KLS96, BCDZ99]) construct multi-variate degree-d polynomials P_d (with $d \in (\sqrt{r}, r)$) such that for any $x_1, \ldots, x_r \in \{0, 1\}$:

$$|P_d(x_1, \ldots, x_r) - \prod_i x_i| \leq 2^{-\frac{d^2}{2^8 r}} .$$

If we insert the ℓ-local projectors $x_i = H_i$, whose spectrum is $\{0, 1\}$, we get that

$$\|P_d(H_1, \cdots, H_r) - |\psi\rangle\langle\psi|\| \leq 2^{-\frac{d^2}{2^8 r}} ,$$

and $P_d(H_1, \cdots, H_r)$ is a (ℓd)-local operator by Example 2.4. This shows that ψ is $\left(\ell d, 2^{-\frac{d^2}{2^8 r}}\right)$-local, and the concentration bound follows directly from Lemma 2.10. □

As an application of this, we consider quantum circuits with arbitrary 2-local gates between any two qubits, arranged into t layers. The number of layers t is called the circuit depth. The key property of shallow circuits is that they spread locality in a controllable way: by a standard lightcone argument, it is easy to see that if O is a k-local operator and U is a unitary implemented by a depth-t circuit, then $UOU^\dagger$ is a $(2^t \cdot k)$-local operator. Therefore, the output state $U|0\rangle$ of the circuit is the maximum-energy eigenstate of the 2^t-local Hamiltonian $H = \sum_i U(|0\rangle\langle 0|_i \otimes \mathbb{1}_{\setminus i})U^\dagger$, and the local terms of the Hamiltonian are clearly commuting projectors. Hence, we obtain the following corollary, answering an open question from [DPMRF23].

Corollary 3.2. *Let U be a unitary implemented by a depth-t circuit. Then, the output state $|\psi\rangle\langle\psi| = U|0\rangle\langle 0|^{\otimes n}U^\dagger$ is $\left(k, 2^{-\frac{k^2}{2^{2t+8} \cdot n}}\right)$-local operator for $k \in (2^t\sqrt{n}, 2^t n)$. Furthermore, denoting the median of W_ψ by m, the following Gaussian concentration bounds hold:*

$$\Pr[W_\psi > m + k] \leq 4 \cdot 2^{-\frac{k^2}{2^{2t+7} \cdot n}} \qquad \text{and} \qquad \Pr[W_\psi < m - k] \leq 4 \cdot 2^{-\frac{k^2}{2^{2t+7} \cdot n}} .$$

A more general case considered in [AAG22] is that of a 1D local Hamiltonian $H = \sum H_i$, where $0 \leq H_i \leq \mathbb{1}$ and H needs to satisfy the local gap property (see [AAG22] for a definition). For such Hamiltonians, [AAG22, Theorem 3] gives local approximations to the ground state, and we can use this and our Lemma 2.10 to obtain concentration bounds. We do not spell out the full statement and instead consider a useful example, injective matrix product states (MPSs) with constant bond dimension. We refer to [CPGSV21] for an introduction to MPSs. For our purposes, the main property of injective MPSs is that they are the unique ground state of a "parent Hamiltonian" with constant locality spectral gap (see e.g. [CPGSV21, Section IV.C] for details). In fact, the proof of the spectral gap lower bound also implies a constant local gap lower bound [PGVWC07]. From this, the following result is immediate.

Lemma 3.3. *Let $|\psi\rangle$ be an injective MPS on a chain with n qubits with a constant bond dimension. Then, for every $\delta \in (0, \frac{1}{4})$, $|\psi\rangle\langle\psi|$ is a $\left(k, e^{-c_1(\delta)\frac{k^2}{n}}\right)$-local operator for $k \leq c_2(\delta)n^{1-\delta}$ for some $c_{1,2}(\delta)$ independent of n. Furthermore, denoting by m the median of $W_{|\psi\rangle\langle\psi|}$,*

$$\Pr[W_\rho > m + k] \leq e^{-2c_1(\delta)\frac{k^2}{n}} \quad \text{and} \quad \Pr[W_\rho < m - k] \leq e^{-2c_1(\delta)\frac{k^2}{n}} .$$

An injective MPS can be used to encode conditionally independent distributions on a line. For this, consider a distribution $P(x_1, \dots x_n)$ such that $P(x_i|x_1, \dots x_{i-1}) = P(x_i|x_{i-1})$ and assume that all these conditional probabilities are positive. Note that $P(x_1, \dots x_n) = P(x_1)P(x_2|x_1) \dots P(x_n|x_{n-1})$. Then, it is easy to verify that the state $\sum_{x_1,\dots x_n} \sqrt{P(x_1, \dots, x_n)}|x_1, x_2, \dots x_n\rangle$ can be written as an injective MPS with constant bond dimension. The output distribution after a computational basis measurement is precisely P, which allows us to show Gaussian concentration using Lemma 3.3, reproducing a (version of) Azuma's inequality. In this sense, Lemma 3.3 can be understood as a quantum version of Azuma's inequality.

4 Concentration bounds for dense Hamiltonian evolution

In the previous section we considered the output states of shallow quantum circuits and showed that they can be approximated by local operators. A depth-t circuit can be written as a product of t unitaries $U_1 U_2 \dots U_t$, where each U_i is a tensor product of one- or two-qubit gates acting on disjoint sets of qubits. Each U_i can also be written as a Hamiltonian evolution $e^{\iota H^{(i)}}$, where $H^{(i)}$ is a 2-local Hamiltonian and each qubit participates in at most one local term. Therefore, we can rephrase the result from the previous section as saying that for any sequence of t 2-local Hamiltonians $H^{(i)}$ where each qubit participates in at most one local term, the state $e^{\iota H^{(t)}} \cdots e^{\iota H^{(1)}}|0\rangle^{\otimes n}$ can be approximated by a local operator. Since each local term in $H^{(i)}$ acts on different qubits, $H^{(i)}$ is obviously a commuting Hamiltonian.

In this section, we generalise this result to more general families of commuting Hamiltonians $H^{(i)}$. In particular, we drop the requirement that each qubit can only participate in at most one local term of $H^{(i)}$ and instead allow local commuting Hamiltonians with dense interaction graphs, i.e. any qubit is allowed to participate in an arbitrary number of terms. Evolution under such dense Hamiltonians cannot be implemented by a shallow circuit. Nonetheless, we show that as long as such Hamiltonians satisfy a norm constraint explained in Equation (4.1), the output state $e^{\iota H^{(t)}} \cdots e^{\iota H^{(1)}}|0\rangle^{\otimes n}$ can still be approximated by a local operator just like for shallow circuits, and as a consequence also obeys concentration bounds. This is of particular relevance as quantum optimisation algorithms such as the QAOA apply an evolution of this form when applied to dense constrained optimisation problems (COPs). Therefore, our concentration bounds apply to the output of the QAOA for dense COPs, which previously required a highly technical analysis that only yielded an asymptotic statement without explicit bounds [BGMZ22]. We can also use these concentration bounds to prove limitations on the performance of the QAOA (and dense evolutions more generally) at solving COPs; see Section 5 for details.

More formally, let $H^{(1)}, H^{(2)}, \dots H^{(p)}$ be a collection of ℓ-local commuting Hamiltonians (i.e. the local terms in each Hamiltonian commute, but the different $H^{(i)}$ need not commute), where each qubit is allowed to participate in arbitrarily many local terms of each Hamiltonian $H^{(i)}$. Define $U_i = e^{-\iota H^{(i)}}$. Because each Hamiltonian $H^{(i)}$ may have a dense interaction graph, we call $U_1 \cdots U_p$ a *dense Hamiltonian evolution* with *level* p. To prove concentration bounds, we will require that

there exist constants $\alpha \in [0,1)$ and $\tilde{C} > 1$ such that for every i and every subset $S \subseteq [n]$ of qubits, the subset Hamiltonian $H_S^{(i)}$ (see Definition 2.6) satisfies

$$\left\| H_S^{(i)} \right\| \leq \tilde{C} n^{\alpha} |S|^{1-\alpha} . \tag{4.1}$$

In the special case of sparse Hamiltonians, we get that Equation (4.1) is satisfied with $\alpha = 0$. However, crucially, for random *dense* classical COPs, Equation (4.1) can still be satisfied with high probability. For example, in Section 4.3 we show that a class of random spin models satisfies Equation (4.1) with high probability even though it allows for constraints between any variables, i.e. its constraint (hyper-)graph is the complete (hyper-)graph. The well-known Sherrington-Kirkpatrick model is an example of such a spin model.

We consider any pure product state ρ_0 and denote by

$$\rho_p = \Big(U_p \cdots U_1 \Big) \rho_0 \Big(U_1^\dagger \cdots U_p^\dagger \Big) \tag{4.2}$$

the output of the dense Hamiltonian evolution. The purpose of this section is to show that the output state ρ_p satisfies certain concentration properties even if p grows (slowly) with n. Our strategy for proving such concentration bounds will be to approximate the state ρ_p by a local operator (Theorem 4.1). Once we have established such a local approximation, Lemma 2.10 immediately implies a concentration bound for the Hamming weight distribution W_{ρ_p} (Corollary 4.7). In Corollary 4.8, we extend this to a concentration bound for the energy density of ρ_p with respect to any classical Hamiltonian satisfying Equation (A.2).

4.1 Local approximations to output states of dense Hamiltonian evolution

We start by giving the main result of this section, a local approximation to the output state ρ_p. The proof of this result proceeds inductively: we first show how to approximate the starting state ρ_0 by a local operator; then, we can analyse how the locality and approximation error evolves under application of the unitaries U_i. We emphasise that the bounds in Theorem 4.1 are optimised for ease of use, not tightness; one can easily obtain a tighter final result by keeping around more parameters.

Theorem 4.1. *Consider the output state ρ_p from Equation (4.2) for a family of ℓ-local commuting Hamiltonians $H^{(1)}, \dots H^{(p)}$ satisfying Equation (4.1) for some $\alpha \in [0,1)$ and $\tilde{C} > 0$ (independent of n). Then, for sufficiently large n and $p = o(\log(n))$, there exists a constant $c_1 > 0$ such that ρ_p is (k_p, ϵ_p)-local with*

$$k_p \leq c_1^p n^{1-(1-\alpha)^p/4} , \quad \epsilon_p \leq 4 e^{-n^{1/8}/\sqrt{2}} .$$

Remark 4.2. In general, the above bound is useful when $k = o(n)$. For sparse COPs, where $\alpha = 0$, the bound on k_p simplifies to $k_p \leq c_1^p n^{3/4}$. Therefore, we get $k_p = o(n)$ for $p = o(\log n)$. (Note that for sparse Hamiltonians, the circuit is in fact low depth, so one can alternatively use Corollary 3.2.) For dense COPs, where $\alpha > 0$, we need $p = o(\log \log n)$ for $k_p = o(n)$. For the remainder of this section, we will focus on dense Hamiltonians and always impose the requirement that $p = o(\log \log n)$.

Remark 4.3. As an example of why such a local operator approximation is useful, we observe that Theorem 4.1 combined with Lemma 2.11 gives a clustering property for the output of dense evolutions. Such a clustering property is used in the proof of the NLTS theorem [ABN22], which shows that the low energy states of recently discovered [PK22, LZ22] linear-rank and linear-distance quantum LDPC code Hamiltonians require $\Omega(\log n)$ quantum circuit depth to prepare. Replacing [ABN22, Fact 4] with the clustering property for the output states of dense evolutions, we can show that if one were to try to prepare the low-energy states of LDPC codes using dense Hamiltonian evolution instead of shallow circuits, one would need at least $\Omega(\log\log n)$ levels of dense evolution.

Proof of Theorem 4.1. We invoke Lemma 4.4 shown below, which gives general (albeit complicated) expressions for k_p and ϵ_p. Setting $c_1 = 40 \cdot \ell\tilde{C}$, which is a constant by assumption, we immediately obtain the bound on k_p. To bound ϵ_p, we use the expression from Lemma 4.4:

$$\epsilon_p \leq 3\, e^{-n^{1/8}/\sqrt{2}} + 6 \sum_{j=1}^{p} e^{-4\cdot(20\cdot l)^{j-1}\tilde{C}^j n^{1-(1-\alpha)^j/4} + \sqrt{n} + 2(j-1)\log(2n)k_{j-1}}\,.$$

We first bound the exponent of the second term for $j \in \{1, \ldots, p\}$:

$$\begin{aligned} &-4\cdot(20\cdot l)^{j-1}\tilde{C}^j n^{1-(1-\alpha)^j/4} + \sqrt{n} + 2(j-1)\log(2n)k_{j-1} \\ &\leq -\tilde{C}n^{1-(1-\alpha)^j/4} + \sqrt{n} + 2(p-1)\log(2n)c_1^p n^{1-(1-\alpha)^{j-1}/4} \end{aligned}$$

For $p = o(\log n)$ and sufficiently large n and p, we see that the first term $-\tilde{C}n^{1-(1-\alpha)^j/4}$ is dominant for any j. Therefore, there exist constants c_3, c_4 such that

$$\epsilon_p \leq 3e^{-n^{1/8}/\sqrt{2}} + 6pc_3 e^{-c_4 n^{3/4}} \leq 4e^{-n^{1/8}/\sqrt{2}}$$

for sufficiently large n and $p = o(\log(n))$. □

As mentioned above, Theorem 4.1 is a simplification of the following technical lemma.

Lemma 4.4. *Consider the output state ρ_p from Equation (4.2) for a family of ℓ-local commuting Hamiltonians $H^{(1)}, \ldots H^{(p)}$ satisfying Equation (4.1) for some $\alpha \in [0,1)$ and $\tilde{C}$ such that $n \geq \frac{k}{\tilde{C}^{1/\alpha}}$. Then, ρ_p is (k_p, ϵ_p)-local for*

$$k_p \leq 2\cdot(20\cdot\ell\tilde{C})^p n^{1-(1-\alpha)^p/4}\,, \quad \epsilon_p \leq 3\, e^{-n^{1/8}/\sqrt{2}} + 6\sum_{j=1}^{p} e^{-4\cdot(20\cdot l)^{j-1}\tilde{C}^j n^{1-(1-\alpha)^j/4} + \sqrt{n} + 2(j-1)\log(2n)k_{j-1}}\,.$$

Proof. For $i \leq p$ we define the intermediate states ρ_i in the obvious way, i.e.

$$\rho_i = \Big(U_i \cdots U_1\Big)\rho_0\Big(U_1^\dagger \cdots U_i^\dagger\Big)\,.$$

We now claim that each ρ_i is (k_i, ϵ_i)-local with

$$\begin{aligned} k_i &\leq 2\cdot(20\cdot\ell\tilde{C})^i n^{1-(1-\alpha)^i/4}\,, \\ \epsilon_i &\leq 3\, e^{-n^{1/8}/\sqrt{2}} + 6\sum_{j=1}^{i} e^{-4\cdot(20\cdot l)^{j-1}\tilde{C}^j n^{1-(1-\alpha)^j/4} + \sqrt{n} + 2(j-1)\log(2n)k_{j-1}}\,, \\ \mathrm{TLN}_{\epsilon_i}(\rho_i) &\leq 2\, e^{\sqrt{n} + 2i\log(2n)k_i}\,. \end{aligned} \tag{4.3}$$

We prove this by induction. The base case $i = 0$, i.e. the approximation of the starting state ρ_0, follows from Lemma 4.5. For the inductive step, we will make use of a simplification of Lemma A.2, stated as Lemma 4.6 below. Concretely, suppose that Equation (4.3) holds for ρ_i. Since

$$\rho_{i+1} = U_{i+1}\rho_i U_{i+1}^\dagger \,,$$

we can apply Lemma 4.6 with

$$\delta = (1-\alpha)^i/4 \qquad \text{and} \qquad C = 2\cdot(20\cdot l\tilde{C})^i \,.$$

This implies the bounds in Equation (4.3) after minor simplifications. □

We now show the two missing statements in the proof of Lemma 4.4: Lemma 4.5 for the base case of the induction and Lemma 4.6 for the inductive step. We will show both of these statements in slightly more generality than is required for Lemma 4.4, with additional parameters that could be optimised to obtain tighter bounds in Lemma 4.4 for specific applications.

For the base case of the induction, we need to approximate the starting state ρ_0. By assumption, this is a pure product state. Since all pure product states are related to each other by 1-local unitaries, it suffices to show the lemma for any one particular product state. For simplicity, we show it for $\rho_0 = |+\rangle\langle+|^{\otimes n}$, which we have already considered in Example 2.5. There, we gave the following simple approximation: we defined the local Hamiltonian $H_0 = \frac{1}{n}\sum_{i=1}^n |+\rangle\langle+|_i$ and noted that $|+\rangle\langle+|^{\otimes n}$ can be approximated by powers of this Hamiltonian. Here, we will require a better tradeoff between locality and approximation error. This will be achieved using the approximation to the powering function established in Lemma 2.9. Because we additionally need to bound the total local norm of the approximation, which becomes large when using the function from Lemma 2.9, we will combine the simple powering from Example 2.5 with the function from Lemma 2.9 to achieve an approximation that has both low locality and low total local norm. The requirement that our approximation should have small total local norm is also the reason why we cannot use Lemma 3.1 to approximate ρ_0 in this setting.

Lemma 4.5. *The state $\rho_0 \coloneqq |+\rangle\langle+|^{\otimes n}$ is (k_0, ϵ_0)-local with*

$$k_0 = 2n^{3/4}\,, \quad \epsilon_0 = 3e^{-(n^{1/8})/\sqrt{2}}\,, \quad \mathit{TLN}_{\epsilon_0}(\rho_0) \le 2e^{\sqrt{n}}.$$

Proof. Let $H_0 = \frac{1}{n}\sum_{i=1}^n |+\rangle\langle+|_i$ and consider the polynomial $P_{s,a}(x) = \sum_{j=0}^a p_{s,j}T_j(x)$ from Lemma 2.9. We define the $(a\cdot m)$-local operator

$$R = P_{s,a}(H_0^m)\,,$$

where a, m, and s will be chosen later. We can then bound

$$\|\rho_0 - R\| \le \|\rho_0 - (H_0^m)^s\| + \|(H_0^m)^s - P_{s,a}(H_0^m)\| \le e^{-\frac{ms}{n}} + 2e^{-\frac{(a+1)^2}{2s}}\,,$$

where the second inequality uses Example 2.5 and Lemma 2.9. If we now choose $m = \lceil n^{1/4}\rceil$, $a = \lceil n^{1/2}\rceil$, and $s = \lceil \frac{1}{\sqrt{2}}n^{7/8}\rceil$, then we get that

$$\frac{ms}{n} \ge \frac{1}{\sqrt{2}}n^{1/8} \qquad \text{and} \qquad \frac{(a+1)^2}{2s} = \frac{(\lceil n^{1/2}\rceil+1)^2}{2\lceil\frac{1}{\sqrt{2}}n^{7/8}\rceil} \ge n^{1/2}\frac{n^{1/2}+1}{\sqrt{2}n^{7/8}+1} \ge \frac{1}{\sqrt{2}}n^{1/8}\,,$$

where the last inequality holds because $\sqrt{2}n^{7/8} \geq n^{1/2}$. Consequently, R has locality $k_0 = a \cdot m \leq 2n^{3/4}$ and approximates ρ_0 up to error $\epsilon_0 \leq e^{-\frac{ms}{n}} + 2e^{-\frac{(a+1)^2}{2s}} \leq 3e^{-(n^{1/8})/\sqrt{2}}$ as claimed.

To bound $\text{TLN}_{\epsilon_0}(\rho_0) = \text{TLN}(R)$, we can first bound the norms of the coefficients in the Chebyshev polynomial (Definition 2.8) by

$$\sum_{r=0}^{\lfloor \frac{j}{2} \rfloor} |\alpha_r| = \sum_{r=0}^{\lfloor \frac{j}{2} \rfloor} \frac{j}{2} \cdot \frac{1}{j-r} \binom{j-r}{r} 2^{j-2r} \leq \sum_{r=0}^{\lfloor \frac{j}{2} \rfloor} \binom{j-r}{r} 2^{j-2r} \leq 2^j \sum_{r=0}^{\lfloor \frac{j}{2} \rfloor} \binom{j}{r} \frac{1}{4^r} \leq 2^j \cdot \frac{5^j}{4^j} \leq e^j.$$

The second-to-last inequality holds because $\sum_{r=0}^{\lfloor \frac{j}{2} \rfloor} \binom{j}{r} \frac{1}{4^r} \leq \sum_{r=0}^{j} \binom{j}{r} \frac{1}{4^r} = (1 + \frac{1}{4})^j$ by the binomial formula. Therefore,

$$\text{TLN}(R) = \text{TLN}\left(\sum_{j=0}^{a} p_{s,j} T_j(H_0^m)\right) \leq \sum_{j=0}^{a} \sum_{r=0}^{\lfloor \frac{j}{2} \rfloor} p_{s,j} \cdot |\alpha_r| \cdot \text{TLN}\left(H_0^{m(j-2r)}\right) \leq \sum_{j=0}^{a} p_{s,j} \cdot e^j \leq e^a.$$

The second inequality holds because $\text{TLN}\left(H_0^{m(j-2r)}\right) = 1$ as noted in Example 2.5, and the last inequality holds because $\sum_{j=0}^{a} p_{s,j} \leq 1$ as $p_{s,j}$ forms a probability distribution. Inserting $a = \lceil n^{1/2} \rceil \leq n^{1/2} + \frac{1}{2}$, we get $e^a \leq 2e^{n^{1/2}}$. □

For the inductive step in Lemma 4.4, we need to analyse how the locality and approximation error change under application of one unitary U_i. This is done in the following lemma, which is a simplification of Lemma A.2. If one wishes to obtain the tightest possible bounds in Lemma 4.4, one can use Lemma A.2 directly instead of this simplified statement.

Lemma 4.6. *Let ρ be a (k, ϵ)-local quantum state with $k \leq Cn^{1-\delta}$ for some $C > 1, \delta > 0$, and H an ℓ-local commuting Hamiltonian satisfying Equation (4.1) for some $\alpha \in [0, 1)$ and $n \geq \frac{k}{\tilde{C}^{1/\alpha}}$. Then, the state*

$$\rho' \coloneqq e^{-\iota H} \rho e^{\iota H}$$

is (k', ϵ')-local for

$$k' \leq 20\ell\tilde{C}Cn^{1-(1-\alpha)\delta}, \quad \epsilon' = 3e^{-4\tilde{C}Cn^{1-(1-\alpha)\delta}} \text{TLN}_\epsilon(\rho) + \epsilon, \quad \text{TLN}_{\epsilon'}(\rho') \leq e^{2\log(2n)k'} \text{TLN}_\epsilon(\rho).$$

Proof. Let R be a k-local operator within ϵ-distance from ρ. We apply Lemma A.2 to R for the Hamiltonian H, which satisfies $\|H_S\| \leq \tilde{C}n^\alpha |S|^{1-\alpha}$. Setting $\mu = 1 + 4/e$, we get that $e^{-\iota H} R e^{\iota H}$ is a $(k', \tilde{\epsilon})$-local operator with

$$k' \leq 2\ell\lceil 7\tilde{C}n^\alpha \kappa^{1-\alpha} \rceil + k, \quad \tilde{\epsilon} = 3e^{-4\tilde{C}n^\alpha \kappa^{1-\alpha}} \text{TLN}(R), \quad \text{TLN}_{\tilde{\epsilon}}\left(e^{-\iota H} R e^{\iota H}\right) \leq (2n)^{k'} \left(\text{TLN}(R) + \tilde{\epsilon}\right),$$

for $\kappa \geq k$ to be chosen later. Since $e^{-\iota H} R e^{\iota H}$ is $(k', \tilde{\epsilon})$-local, it immediately follows from the triangle inequality that $e^{-\iota H} \rho e^{\iota H}$ is (k', ϵ')-local for $\epsilon' = \tilde{\epsilon} + \epsilon$. We can now simplify the resulting bounds as follows choosing $\kappa = C^{\frac{1}{1-\alpha}} n^{1-\delta} \geq k$:

(i) Because $l \geq 1$, $n \geq \frac{k}{\tilde{C}^{1/\alpha}}$, and $\kappa \geq k$, it is clear that $2\ell\lceil 7\tilde{C}n^\alpha \kappa^{1-\alpha} \rceil + k \leq 20\ell\tilde{C}n^\alpha \kappa^{1-\alpha}$. Inserting $\kappa = C^{\frac{1}{1-\alpha}} n^{1-\delta}$, we get the claimed bound on k'.

(ii) The bound on ϵ' follows immediately from $\epsilon' = \tilde{\epsilon} + \epsilon$, $\textsc{tln}(R) = \textsc{tln}_\epsilon(\rho)$, and $\kappa = C^{\frac{1}{1-\alpha}} n^{1-\delta}$.

(iii) To bound $\textsc{tln}_{\epsilon'}(\rho')$, observe that since ρ is a quantum state, we are only interested in approximations where $\tilde{\epsilon} \leq 1 \leq (2n)^{k'} \textsc{tln}(R)$, as otherwise the claim becomes trivial. Combining this and inserting $\textsc{tln}(R) = \textsc{tln}_\epsilon(\rho)$ yields the claimed bound. □

4.2 Concentration bounds for output states of dense Hamiltonian evolution

Combining Lemma 2.10 and Theorem 4.1, we immediately obtain the following concentration bound. One can of course also derive an analogous but tighter and more explicit statement from Lemma 4.4 instead of Theorem 4.1.

Corollary 4.7. *Let ρ_p and c_1 be as in Theorem 4.1, $p = o(\log\log n)$, and m be the median of W_{ρ_p}. Then for sufficiently large n and p,*

$$\Pr\Big[|W_{\rho_p} - m| > c_1^p n^{1-(1-\alpha)^p/4}\Big] \leq 128\, e^{-\sqrt{2}n^{1/8}}.$$

While the above statement is about concentration with respect to Hamming weight, we can also prove concentration with respect to other observables. Let $G = \sum_i G_i$ be a classical local Hamiltonian (i.e. a local Hamiltonian that is diagonal in the computational basis) that satisfies the following condition analogous to Equation (4.1): for all $S \subset [n]$,

$$\|G_S\| \leq D n^{\alpha'} |S|^{1-\alpha'}.$$

In this case, we can also prove a concentration bound on the expectation of ρ_p with respect to G. More specifically, we can define a random variable E_{ρ_p} indicating the "energy" of ρ_p according to G, i.e. if we take the spectral decomposition $G = \sum g_i \Pi_i$ for orthogonal projectors Π_i, then E_{ρ_p} takes value g_i with probability $\mathrm{Tr}[\Pi_i \rho_p]$.

Corollary 4.8. *Let ρ_p and c_1 be as in Theorem 4.1 with $p = o(\log\log n)$, and G and E_{ρ_p} as above. Let E be the median energy, i.e. the median of E_{ρ_p}. Then*

$$\Pr\Big[|E_{\rho_p} - E| > 2D c_1^{p(1-\alpha')} n^{1-(1-\alpha)^p(1-\alpha')/4}\Big] \leq 128\, e^{-\sqrt{2}n^{1/8}}.$$

Proof. Consider strings $x, y \in \{0,1\}^n$ and let $S := \{i : x_i \neq y_i\}$. Suppose $|S| \leq c_1^p n^{1-(1-\alpha)^p/4}$. Expanding $G = \sum_i G_i$, consider the energy difference

$$\begin{aligned}
\mathrm{Tr}[(|x\rangle\langle x| - |y\rangle\langle y|)G] &= \sum_i \mathrm{Tr}[(|x\rangle\langle x| - |y\rangle\langle y|)G_i] \\
&= \sum_{\substack{i \text{ s.t. } G_i \text{ acts non-trivially} \\ \text{on at least one qubit in } S}} \mathrm{Tr}[(|x\rangle\langle x| - |y\rangle\langle y|)G_i] \\
&= \mathrm{Tr}[(|x\rangle\langle x| - |y\rangle\langle y|)G_S] \\
&\leq 2\|G_S\| \leq 2D n^{\alpha'} |S|^{1-\alpha'} \leq 2D c_1^{p(1-\alpha')} n^{1-(1-\alpha)^p(1-\alpha')/4}.
\end{aligned}$$

Let

$$\Pi_{>E+f} = \sum_{x \in \{0,1\}^n : \langle x|H|x\rangle > E+f} |x\rangle\langle x|, \quad \Pi_{\leq E} = \sum_{x \in \{0,1\}^n : \langle x|H|x\rangle \leq E} |x\rangle\langle x|.$$

If $f > 2D c_1^{p(1-\alpha')} n^{1-(1-\alpha)^p(1-\alpha')/4}$, then by the above argument strings in the support of $\Pi_{>E+f}$ and $\Pi_{\leq E}$ differ on more than $c_1^p n^{1-(1-\alpha)^p/4}$ positions. Thus, a $(c_1^p n^{1-(1-\alpha)^p/4})$-local operator O satisfies $\Pi_{>E+f} O \Pi_{\leq E} = 0$. The corollary now follows along the same lines as Lemma 2.10. □

4.3 Example: random spin models

As an example of a family of dense Hamiltonians that satisfies Equation (4.1), we consider the pure q-spin model, which was also considered in [BGMZ22]. The pure q-spin model is a random COP with cost function

$$C_n^q(z;J) = \frac{1}{n^{(q+1)/2}} \sum_{i_1,\dots i_q=1}^{n} J_{i_1,\dots i_q} z_{i_1} \dots z_{i_q} \,, \tag{4.4}$$

where the coefficients $J = (J_{i_1,\dots i_q})_{i_1,\dots i_q \in [n]}$ are i.i.d. standard Gaussian random variables $J_{i_1,\dots i_q} \sim \mathcal{N}(0,1)$. Here, $z_i \in \{\pm 1\}$ and the objective is to maximise $C^q(z_1,\dots,z_n)$. This can be identified with a q-local Hamiltonian

$$H_n^q(J) = \frac{1}{n^{(q-1)/2}} \sum_{i_1,\dots i_q=1}^{n} J_{i_1,\dots i_q} \sigma_{i_1}^Z \dots \sigma_{i_q}^Z \,. \tag{4.5}$$

We note the different normalisation factors: $C_n^q(z)$ is normalised such that on average over J, $\max_z C_n^q(z;J) = \Theta(1)$. In contrast, H_n has an additional factor of n, so that on average over J, $\|H_n^q(J)\| = \Theta(n)$. We use these different normalisations because the former is common in the classical literature (see e.g. [GJW20]), whereas the latter is common in the quantum literature.

The following lemma shows that this model satisfies Equation (4.1) with overwhelming probability, and as a result we can apply Corollary 4.7 and Corollary 4.8 to obtain concentration bounds.

Lemma 4.9. *With probability at least $1-e^{-n}$ over the choice of $J_{i_1,\dots i_q} \sim \mathcal{N}(0,1)$, the Hamiltonian $H_n^q(J) = \frac{1}{n^{(q-1)/2}} \sum_{i_1,\dots i_q=1}^{n} J_{i_1,\dots i_q} \sigma_{i_1}^Z \dots \sigma_{i_q}^Z$ satisfies Equation (4.1) with $\alpha = \frac{1}{2}$ and $\tilde{C} = \sqrt{6}$ for every subset $S \subseteq [n]$.*

Proof. Fix any subset $S \subseteq [n]$. Recall the definition of the subset Hamiltonian from Definition 2.6, and define analogously

$$C_{n,S}^q(z;J) = \frac{1}{n^{(q+1)/2}} \sum_{\{i_1,\dots i_q\} \cap S \neq \emptyset}^{n} J_{i_1,\dots i_q} z_{i_1} \dots z_{i_q} \,.$$

Because all terms of $H_{n,S}^q$ are proportional to tensor products of Pauli-Z operators, it is easy to see that $\left\| H_{n,S}^q \right\| = n \cdot \max_{z_1,\dots,z_n \in \{\pm 1\}} C_{n,S}(z;J)$. For any fixed choice of $z_1,\dots,z_n \in \{\pm 1\}$, the random variable $C_{n,S}(z;J)$ is a sum of $\ell \leq |S| \binom{n}{q-1} \leq |S| n^{q-1}$ standard Gaussians with a normalisation factor $\frac{1}{n^{(q+1)/2}}$, and is therefore distributed as $\mathcal{N}(0, \ell/n^{q+1})$. By the standard upper deviation inequality for Gaussians, we have that

$$\Pr_{J_{i_1,\dots i_q} \sim \mathcal{N}(0,1)} \left[C_{n,S}^q(z;J) \geq \sqrt{6|S|/n} \right] \leq e^{-3n} \,.$$

Since we are interested in upper-bounding the probability that $C_S^q(z;J) \geq \sqrt{6|S|/n}$ *for any* choices of $z_1,\dots,z_n$ and $|S|$, we can apply the union bound over the possible $2^n \cdot 2^n \leq e^{2n}$ choices of $z_1,\dots,z_n$ and $|S|$. We therefore see that the probability that Equation (4.1) is violated (for any $S \subseteq [n]$) is at most $e^{-3n} \cdot e^{2n} = e^{-n}$ as claimed. □

[BGMZ22] also consider a mixed q-spin model, which is a sum over pure j-spin models for $j \leq q$. Specifically, the cost function can be written as

$$C_n^{q,\text{mixed}}(z;J) = \sum_{j=1}^{q} c_j \, C_n^j(z;J) \,,$$

where c_j are arbitrary real coefficients and C_n^j is as defined in Equation (4.4). We can again associate a Hamiltonian $H_n^{q,\text{mixed}}$ with this cost function. The following corollary follows immediately from Lemma 4.9 by the triangle inequality.

Corollary 4.10. *With probability at least* $1-e^{-n}$*, the Hamiltonian* $H_n^{q,\text{mixed}}$ *satisfies Equation* (4.1) *for* $\alpha = \frac{1}{2}$ *and* $\tilde{C} = \sqrt{6}\sum |c_j|$.

We can use this property of the (mixed) random spin model to obtain concentration bounds for the output states of the QAOA applied to the COPs $C_n^q(z;J)$ and $C_n^{q,\text{mixed}}(z;J)$. This in turn can also be used to prove limitations on the success probability of the QAOA on these COPs. We spell this out in detail in Corollary 5.6.

5 Limitations on dense evolutions for constraint optimisation problems

Using our local operator approximations and concentration bounds for the output states of dense evolutions, we can show that such states have limitations as optimisers for COPs. We begin by introducing a structural property of random COPs, called the overlap gap property (OGP), that roughly says that good solutions to a COP must cluster, i.e. different good solutions must either be close to each other or far from each other. We then combine the OGP with our concentration results and the symmetry of the QAOA output to show that for most instances of random spin models, the QAOA can only produce a good solution with negligible probability.

5.1 Overlap gap property and existence of high-weight sets for local quantum optimisers

We consider an objective function $C_n(z)$ for $z = (z_1, \ldots, z_n) \in \{\pm 1\}^n$ that we want to maximise. We begin by recalling a few general definitions from [GJW20], adapted to the case where algorithms for COPs output probability distributions or quantum states rather than a single element of $\{\pm 1\}^n$.

Definition 5.1. For parameters $\mu \in \mathbb{R}$ and $\delta \in [0,1]$, we say that a probability distribution P over $\{\pm 1\}^n$ (μ, δ)-optimises the objective $C_n(z)$ if

$$\Pr_{z\sim P}[C_n(z) \geq \mu] \geq 1 - \delta \,.$$

We will use the same notation for quantum states ρ, which we identify with a probability distribution over $\{\pm 1\}^n$ in the natural way, i.e. $\Pr_\rho[(-1)^x] = \langle x|\rho|x\rangle$ for $x \in \{0,1\}^n$.

Definition 5.2. An objective function $C_n(z)$ satisfies the (μ, ν_1, ν_2)-*overlap gap property (OGP)* with parameters $\mu \in \mathbb{R}$ and $0 \leq \nu_1 < \nu_2 \leq 1$ if for every $z_1, z_2 \in \{\pm 1\}^n$ satisfying $C_n(z_1) \geq \mu$ and $C_n(z_2) \geq \mu$, we have that $\frac{1}{n}|\langle z_1, z_2\rangle| \in [0, \nu_1] \cup [\nu_2, 1]$. Here, $\langle z_1, z_2\rangle$ denotes the usual inner product of vectors.

Suppose that the objective function $C_n(z)$ has the (μ, ν_1, ν_2)-OGP. Then, we can define the set of "good outputs"

$$G_n = \{x \in \{0,1\}^n \mid C_n((-1)^x) \geq \mu\}.$$

Because $C_n(z)$ has the (μ, ν_1, ν_2)-OGP, any two $x, x' \in G_n$ must satisfy $\frac{1}{n}|\langle (-1)^x, (-1)^{x'}\rangle| \in [0, \nu_1] \cup [\nu_2, 1]$. Since the Hamming distance between x and x' is given by $|x - x'| = \frac{n - \langle (-1)^x, (-1)^{x'}\rangle}{2}$, this implies that

$$|x - x'| \in [0, \tilde{\nu}_1 \cdot n] \cup [\tilde{\nu}_2 \cdot n, n]$$

for $\tilde{\nu}_1 := \frac{1-\nu_2}{2} < \tilde{\nu}_2 := \frac{1-\nu_1}{2}$. Assuming that $2\tilde{\nu}_1 < \tilde{\nu}_2$, we can partition $G_n = \cup_i S_n^i$ into sets (or clusters) S_n^i such that for all $i \neq j$:

$$x, x' \in S_n^i \implies |x - x'| \leq \tilde{\nu}_1 \cdot n \quad \text{and} \quad x \in S_n^i, x' \in S_n^j \implies |x - x'| \geq \tilde{\nu}_2 \cdot n. \tag{5.1}$$

We note that the condition $2\tilde{\nu}_1 < \tilde{\nu}_2$ is necessary for this clustering property to hold. To see this intuitively, consider three points $x_1, x_2, x_3 \in G_n$. The (μ, ν_1, ν_2)-OGP then requires that $|x_i - x_j| \in [0, \tilde{\nu}_1 n] \cup [\tilde{\nu}_2 n, n]$ for all pairs (i, j). Without any condition on $\tilde{\nu}_1$ and $\tilde{\nu}_2$, this would allow the following situation: the three points could be arranged "on a line" in the sense that $|x_1 - x_2| \leq \tilde{\nu}_1 n$ and $|x_2 - x_3| \leq \tilde{\nu}_1 n$, but $|x_1 - x_3| \geq \tilde{\nu}_2 n$. This means that the points are not clustered. However, if $2\tilde{\nu}_1 < \tilde{\nu}_2$, then by the triangle inequality $|x_1 - x_2| \leq \tilde{\nu}_1 n$ and $|x_2 - x_3| \leq \tilde{\nu}_1 n$ together imply $|x_1 - x_3| \leq \tilde{2}\nu_1 n < \tilde{\nu}_2 n$. By the OGP this means that we must in fact have $|x_1 - x_3| \leq \tilde{\nu}_1 n$, so we get the clustering behaviour described above.

We can now show that for (approximately) local quantum states that optimise $C_n(z)$, the measurement distribution must be concentrated on one of these sets, which we will call the *high-weight set*.

Lemma 5.3. *Suppose the objective $C_n(z)$ has the (μ, ν_1, ν_2)-OGP with $2\tilde{\nu}_1 < \tilde{\nu}_2$ and there exists a (k, ϵ)-local operator $\psi_n = |\psi_n\rangle\langle\psi_n|$ for $k = o(n)$ that $(\mu, 1 - 4\sqrt{\epsilon})$-optimises $C_n(z)$. Then there exists a (unique) i such that for sufficiently large n, $\operatorname{Tr}\left[\Pi_{S_n^i}\psi_n\right] \geq \operatorname{Tr}[\Pi_{G_n}\psi_n] - \sqrt{\epsilon}$.*

Proof. Because $k = o(n)$, for sufficiently large n, we have $k < \tilde{\nu}_2 n$. Therefore, by Lemma 2.11, it suffices to show that there exists an S_n^i for which $\operatorname{Tr}\left[\Pi_{S_n^i}\psi_n\right] > \sqrt{\epsilon}$. This is because if such an S_n^i exists, we can consider the set $S' = G_n \setminus S_n^i$, which is at least $\tilde{\nu}_2 n$-far from S_n^i. By Lemma 2.11 and the assumption $\operatorname{Tr}\left[\Pi_{S_n^i}\psi_n\right] > \sqrt{\epsilon}$, this means that $\operatorname{Tr}[\Pi_{S'}\psi_n] \leq \sqrt{\epsilon}$. Therefore, $\operatorname{Tr}\left[\Pi_{S_n^i}\psi_n\right] = \operatorname{Tr}[G_n\psi_n] - \operatorname{Tr}[S'\psi_n] \geq \operatorname{Tr}[G_n\psi_n] - \sqrt{\epsilon}$.

Now suppose for the sake of contradiction that for all i, $\operatorname{Tr}\left[\Pi_{S_n^i}\psi_n\right] \leq \sqrt{\epsilon}$. Since ψ_n is assumed to be a $(\mu, 1 - 4\sqrt{\epsilon})$-optimiser, $\sum_i \operatorname{Tr}\left[\Pi_{S_n^i}\psi_n\right] = \operatorname{Tr}[G_n\psi_n] \geq 4\sqrt{\epsilon}$. Therefore, we can find two disjoint sets of indices $\mathcal{I}$ and $\mathcal{I}'$ such that $\sum_{i\in\mathcal{I}} \operatorname{Tr}\left[\Pi_{S_n^i}\psi_n\right] > \sqrt{\epsilon}$ and $\sum_{i\in\mathcal{I}'} \operatorname{Tr}\left[\Pi_{S_n^i}\psi_n\right] > \sqrt{\epsilon}$. However, this contradicts Lemma 2.11 since $\cup_{i\in\mathcal{I}} S_n^i$ and $\cup_{i\in\mathcal{I}'} S_n^i$ are separated by Hamming distance at least $\tilde{\nu}_2 n > k$.

The claim that i is unique holds because if there were at least two such sets S_n^i and $S_n^{i'}$, then $\operatorname{Tr}[\Pi_{G_n}\psi_n] \geq \operatorname{Tr}\left[\Pi_{S_n^i}\psi_n\right] + \operatorname{Tr}\left[\Pi_{S_n^{i'}}\psi_n\right] \geq 2\operatorname{Tr}[\Pi_{G_n}\psi_n] - 2\sqrt{\epsilon}$, which is a contradiction since $\operatorname{Tr}[\Pi_{G_n}\psi_n] \geq 4\sqrt{\epsilon}$. □

5.2 Limitations on symmetric optimisers for symmetric COPs

From Lemma 5.3 we immediately obtain limitations on the performance of *symmetric* local optimisers, i.e. optimisers that are invariant under the operation $\sigma_X^{\otimes n}$, on symmetric COPs $C_n(z)$, i.e. COPs that satisfy $C_n(z) = C_n(-z)$.

Corollary 5.4. *Suppose the objective $C_n(z)$ has the (μ, ν_1, ν_2)-OGP with $2\tilde{\nu}_1 < \tilde{\nu}_2$ and satisfies $C_n(z) = C_n(-z)$. Then, no (k, ϵ)-local quantum state $\psi_n = |\psi_n\rangle\langle\psi_n|$ for $k = o(n)$ that satisfies $\sigma_X^{\otimes n}\psi_n\sigma_X^{\otimes n} = \psi_n$ can $(\mu, 1 - 4\sqrt{\epsilon})$-optimise $C_n(z)$.*

Proof. Suppose for the sake contradiction that such a local symmetric ψ_n does $(\mu, 1-4\sqrt{\epsilon})$-optimise $C_n(z)$. Then it follows from Lemma 5.3 that there exists a unique S_n^i for which $\mathrm{Tr}\left[\Pi_{S_n^i}\psi_n\right] \geq \mathrm{Tr}[\Pi_{G_n}\psi_n] - \sqrt{\epsilon}$. Now consider the set $S = \{x \oplus 1^n \mid x \in S_n^i\}$, where 1^n is the all-1 string. By symmetry of $C_n(z)$ we have that for $x \in S$, $C_n((-1)^x) = C_n((-1)^{x\oplus 1^n}) \geq \mu$, where the inequality holds because $x \oplus 1^n \in S_n^i$. Furthermore, it is easy to see that $|x - x'| \leq \tilde{\nu}_1 \cdot n$ for $x, x' \in S$. Therefore, there must exist an S_n^j in the partition of G_n such that $S \subset S_n^j$. (In fact, by symmetry it is easy to see that $S = S_n^j$ for some j, but we will not need this here.) Since there exist strings $x \in S_n^i$ and $x' \in S$ with Hamming distance n, we must have $i \neq j$. Furthermore, by symmetry of ψ_n and the fact that $\Pi_S = \sigma_X^{\otimes n}\Pi_{S_n^i}\sigma_X^{\otimes n}$,

$$\mathrm{Tr}\left[\Pi_{S_n^j}\psi_n\right] \geq \mathrm{Tr}[\Pi_S\psi_n] = \mathrm{Tr}\left[\Pi_{S_n^i}\sigma_X^{\otimes n}\psi_n\sigma_X^{\otimes n}\right] = \mathrm{Tr}\left[\Pi_{S_n^i}\psi_n\right] \geq \mathrm{Tr}[\Pi_{G_n}\psi_n] - \sqrt{\epsilon}\,.$$

This contradicts the uniqueness of the high-weight set in Lemma 5.3. □

5.3 Example: QAOA on symmetric COPs

An example of practical relevance is the QAOA for approximately solving local COPs. A q-local COPs and its associated q-local Hamiltonian can be written in terms of coefficients $J = (J_{i_1,\ldots,i_q}) \in \mathbb{R}^{n^q}$ as[2]

$$C_n^q(z; J) = \frac{1}{n^{(q+1)/2}} \sum_{i_1,\ldots i_q=1}^{n} J_{i_1,\ldots i_q} z_{i_1} \ldots z_{i_q}\,, \quad H_n^q(J) = \frac{1}{n^{(q-1)/2}} \sum_{i_1,\ldots i_q=1}^{n} J_{i_1,\ldots i_q} \sigma_{i_1}^Z \ldots \sigma_{i_q}^Z\,.$$

The QAOA works by repeatedly applying the unitaries

$$V_{H_n^q(J)}(\gamma) = e^{-\iota\gamma H_n^q(J)} \qquad \text{and} \qquad V_X(\beta) = \left(e^{-\iota\beta\sigma_x}\right)^{\otimes n}$$

to the initial state $\rho_0 = |+\rangle\langle+|^{\otimes n}$ for some parameters $\gamma, \beta > 0$. The output state of the QAOA can therefore be written as

$$\rho_{p,n}(J) = \left(V_X(\beta_p)V_{H_n^q(J)}(\gamma_p)\cdots V_X(\beta_1)V_{H_n^q(J)}(\gamma_1)\right)\rho_0\left(V_{H_n^q(J)}(\gamma_1)^\dagger V_X(\beta_1)^\dagger \cdots V_{H_n^q(J)}(\gamma_p)^\dagger V_X(\beta_p)^\dagger\right). \tag{5.2}$$

The parameters γ_i and β_i may depend arbitrarily on J, but we will always assume that they are chosen from some fixed bounded range that is independent of n.

[2]Note that because here we consider arbitrary coefficients J, the normalisation factors $\frac{1}{n^{(q+1)/2}}$ and $\frac{1}{n^{(q-1)/2}}$ can be chosen arbitrarily, too; we use these ones to keep the notation consistent with Section 4.3, but emphasise that here the coefficients need not be chosen from a Gaussian distribution.

If q is even, both the COP $C_n^q(z;J)$ and the QAOA output state $\rho_{p,n}$ are symmetric in the sense of Corollary 5.4. Therefore, if $C_n^q(z;J)$ satisfies the (μ,ν_1,ν_2)-OGP (for some fixed value of J) and the associated Hamiltonian $H_n^q(J)$ satisfies the norm constraint from Equation (4.1), we can combine Theorem 4.1 and Corollary 5.4 to obtain the following result, which shows that at level $p = o(\log\log n)$, the probability that the QAOA will produce a "μ-good" string x decays with $e^{-n^{1/8}}$.

Lemma 5.5. *Fix q even. Suppose that the COP $C_n^q(z;J)$ (for some choice of J) has the (μ,ν_1,ν_2)-OGP with $2\tilde{\nu}_1 < \tilde{\nu}_2$ and the Hamiltonian H_n^q satisfies Equation (4.1). Then, no QAOA output state $\rho_{p,n}(J)$ (for any choice of γ_i and β_i within an arbitrary bounded range independent of n) can $(\mu, 1-8e^{-n^{1/8}/\sqrt{8}})$-optimise $C_n^q(z;J)$. In other words, if one measures the QAOA output state $\rho_{p,n}(J)$ with level $p = o(\log\log n)$ in the computational basis, the probability of receiving a string x that satisfies $C_n^q((-1)^x;J) \geq \mu$ is at most $e^{-\Omega(n^{1/8})}$.*

Proof. It is easy to verify that the QAOA for this $H_n^q(J)$ implements an instance of dense Hamiltonian evolution with $H^{(1)} = \gamma_1 H_n^q(J), H^{(2)} = \beta_1 \sum_i \sigma_x^{(i)}, H^{(3)} = \gamma_2 H_n^q(J)$, etc. Since $H_n^q(J)$ satisfies the condition in Equation (4.1) for some $\tilde{C} = O(1)$ and $\alpha < 1$, each $H^{(i)}$ also satisfies Equation (4.1) for $\tilde{C} \to \max(\tilde{C}\gamma_i, \beta_i) = O(1)$ and the same α. Therefore, we can apply Theorem 4.1 and find that for $p = o(\log\log n)$, $\rho_{p,n}(J)$ is a (k,ϵ)-local operator for $k = o(n)$ and $\epsilon \leq 4e^{-n^{1/8}/\sqrt{2}}$. Furthermore, since $C_n^q(z;J)$ is symmetric under $z \mapsto -z$ for even q and any J, $H_n^q(J)$ commutes with $\sigma_x^{\otimes n}$, so the QAOA output state $\rho_{p,n}(J)$ satisfies $\sigma_x^{\otimes n}\rho_{p,n}(J)\sigma_x^{\otimes n} = \rho_{p,n}(J)$. Therefore, the lemma follows from Corollary 5.4. □

An example of a COP that satisfies the requirements of Lemma 5.5 is the pure q-spin model from Section 4.3, for which we showed in Lemma 4.9 that Equation (4.1) is satisfied except with probability e^{-n}. It is known that for fixed q, the limit $E_q(J) := \lim_{n\to\infty} \max_{z\in\{\pm1\}^n} C_n^q(z;J)$ exists almost surely [GT02]. Furthermore, for even $q \geq 4$, with probability at least $1 - O(e^{-n})$ over the choice of J (with i.i.d. standard Gaussian entries), the objective function $C_n^q(z;J)$ satisfies the (μ,ν_1,ν_2)-OGP for constants $0 < \nu_1 < \nu_2 < 1$ and $0 < \mu < E_q(J)$ [CGPR19, GJW20, GJ21]. We note that [CGPR19] do not explicitly show that $2\tilde{\nu}_1 < \tilde{\nu}_2$, although it appears to be implicit in the proof that $\tilde{\nu}_1$ can be made an arbitrarily small constant by choosing μ arbitrarily close to $E_q(J)$. We leave a detailed proof of this statement for future work. Therefore, we get the following implication of Lemma 5.5, giving the first provable (modulo the conjectured strengthened OGP) limitations on the QAOA on dense instances at super-constant level and improving upon the result of [BGMZ22].

Corollary 5.6. *Assuming the strengthened OGP stated above, with probability $1-O(e^{-n})$ over the choice of J (with i.i.d. standard Gaussian entries), the value of the solution to the COP $C_n^q(z;J)$ produced by the QAOA with level $p = o(\log\log n)$ is bounded away from the optimal value by at least a constant except with probability $e^{-\Omega(n^{1/8})}$.*

Using Corollary 4.10 instead of Lemma 4.9, this result can easily be extended to mixed spin models $C_n^{q,\text{mixed}}(z;J) = \sum_{j=1}^q c_j\, C_n^j(z;J)$ that only contain contributions from even pure spin models, i.e. c_j is non-zero only for even $j \geq 4$.

References

[AAG22] Anurag Anshu, Itai Arad, and David Gosset. An area law for 2d frustration-free spin systems. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2022, page 12–18, New York, NY, USA, 2022. Association for Computing Machinery. doi:10.1145/3519935.3519962.

[AB22] Anurag Anshu and Nikolas P. Breuckmann. A construction of combinatorial NLTS. *Journal of Mathematical Physics*, 63(12), 12 2022. doi:10.1063/5.0113731.

[ABN22] Anurag Anshu, Nikolas Breuckmann, and Chinmay Nirkhe. NLTS hamiltonians from good quantum codes. *arXiv preprint arXiv:2206.13228*, 2022. URL: https://arxiv.org/abs/2206.13228.

[Abr20] Nilin Abrahamsen. Short proof of a spectral chernoff bound for local hamiltonians. *arXiv preprint arXiv:2009.04993*, 2020. URL: https://arxiv.org/abs/2009.04993.

[Alh22] Álvaro M Alhambra. Quantum many-body systems in thermal equilibrium. *arXiv preprint arXiv:2204.08349*, 2022. URL: https://arxiv.org/abs/2204.08349.

[AN22] Anurag Anshu and Chinmay Nirkhe. Circuit Lower Bounds for Low-Energy States of Quantum Code Hamiltonians. In Mark Braverman, editor, *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, volume 215 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 6:1–6:22, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ITCS.2022.6.

[Ans16] Anurag Anshu. Concentration bounds for quantum states with finite correlation length on quantum spin lattice systems. *New Journal of Physics*, 18(8):083011, aug 2016. doi:10.1088/1367-2630/18/8/083011.

[BC15] Fernando GSL Brandao and Marcus Cramer. Equivalence of statistical mechanical ensembles for non-critical quantum systems. *arXiv preprint arXiv:1502.03263*, 2015. URL: https://arxiv.org/abs/1502.03263.

[BCDZ99] H. Buhrman, R. Cleve, R. De Wolf, and C. Zalka. Bounds for small-error and zero-error quantum algorithms. In *Proceedings of the 40th Annual Symposium on Foundations of Computer Science*, pages 358–368, 1999. doi:10.1109/SFFCS.1999.814607.

[BCG15] FGSL Brandao, Marcus Cramer, and Madalin Guta. A berry–esseen theorem for quantum lattice systems and the equivalence of statistical mechanical ensembles. *QIP2015 Talk*, 2015. URL: http://www.quantum-lab.org/qip2015/talks/125-Brandao.pdf.

[BGMZ22] Joao Basso, David Gamarnik, Song Mei, and Leo Zhou. Performance and limitations of the qaoa at constant levels on large sparse hypergraphs and spin glass models. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, 2022. doi:10.1109/FOCS54457.2022.00039.

[BKKT] Sergey Bravyi, Alexander Kliesch, Robert Koenig, and Eugene Tang. Obstacles to variational quantum optimization from symmetry protection. *Phys. Rev. Lett.*, 125:260505, Dec. doi:10.1103/PhysRevLett.125.260505.

[CGPR19] Wei-Kuo Chen, David Gamarnik, Dmitry Panchenko, and Mustazee Rahman. Suboptimality of local algorithms for a class of max-cut problems. *The Annals of Probability*, 47(3):1587 – 1618, 2019. doi:10.1214/18-AOP1291.

[CLSS22] Chi-Ning Chou, Peter J. Love, Juspreet Singh Sandhu, and Jonathan Shi. Limitations of Local Quantum Algorithms on Random MAX-k-XOR and Beyond. In *49th International Colloquium on Automata, Languages, and Programming (ICALP 2022)*, volume 229 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 41:1–41:20, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum fur Informatik. doi:10.4230/LIPIcs.ICALP.2022.41.

[CPGSV21] J Ignacio Cirac, David Perez-Garcia, Norbert Schuch, and Frank Verstraete. Matrix product states and projected entangled pair states: Concepts, symmetries, theorems. *Reviews of Modern Physics*, 93(4):045003, 2021. doi:10.1103/RevModPhys.93.045003.

[DPMRF23] Giacomo De Palma, Milad Marvian, Cambyse Rouzé, and Daniel Stilck Franca. Limitations of variational quantum algorithms: A quantum optimal transport approach. *PRX Quantum*, 4:010309, Jan 2023. doi:10.1103/PRXQuantum.4.010309.

[DPR22] Giacomo De Palma and Cambyse Rouzé. Quantum concentration inequalities. *Annales Henri Poincaré*, Apr 2022. doi:10.1007/s00023-022-01181-1.

[EH17] L. Eldar and A. W. Harrow. Local hamiltonians whose ground states are hard to approximate. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 427–438, 2017. doi:10.1109/FOCS.2017.46.

[FGG14] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014. URL: https://arxiv.org/abs/1411.4028.

[FGG20] Edward Farhi, David Gamarnik, and Sam Gutmann. The quantum approximate optimization algorithm needs to see the whole graph: A typical case. *arXiv preprint arXiv:2004.09002*, 2020. URL: https://arxiv.org/abs/2004.09002.

[Gam21] David Gamarnik. The overlap gap property: A topological barrier to optimizing over random structures. *Proceedings of the National Academy of Sciences*, 118(41):e2108492118, 2021. doi:10.1073/pnas.2108492118.

[GGL95] Ronald L Graham, Martin Grötschel, and László Lovász. *Handbook of Combinatorics*, volume 1. Elsevier, 1995.

[GJ21] David Gamarnik and Aukosh Jagannath. The overlap gap property and approximate message passing algorithms for p-spin models. *The Annals of Probability*, 49(1):180–205, 2021. URL: https://hdl.handle.net/1721.1/145311.

[GJW20] David Gamarnik, Aukosh Jagannath, and Alexander S Wein. Low-degree hardness of random optimization problems. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 131–140. IEEE, 2020. doi:10.1109/FOCS46700.2020.00021.

[GL18] David Gamarnik and Quan Li. Finding a large submatrix of a gaussian random matrix. *The Annals of Statistics*, 46(6A):2511–2561, 2018. URL: http://hdl.handle.net/1721.1/120593.

[GT02] Francesco Guerra and Fabio Lucio Toninelli. The thermodynamic limit in mean field spin glass models. *Communications in Mathematical Physics*, 230(1):71–79, 2002. doi:10.1007/s00220-002-0699-y.

[GV89] D. Goderis and P. Vets. Central limit theorem for mixing quantum systems and the CCR-algebra of fluctuations. *Communications in Mathematical Physics*, 122:249–265, 1989. doi:10.1007/BF01257415.

[Has04] M. B. Hastings. Lieb-schultz-mattis in higher dimensions. *Phys. Rev. B*, 69:104431, Mar 2004. doi:10.1103/PhysRevB.69.104431.

[HMH04] Michael Hartmann, Günter Mahler, and Ortwin Hess. Existence of Temperature on the Nanoscale. *Phys. Rev. Lett.*, 93:080402, Aug 2004. doi:10.1103/PhysRevLett.93.080402.

[KAAV17] Tomotaka Kuwahara, Itai Arad, Luigi Amico, and Vlatko Vedral. Local reversibility and entanglement structure of many-body ground states. *Quantum Science and Technology*, 2(1):015005, 2017. doi:10.1088/2058-9565/aa523d.

[KLS96] Jeff Kahn, Nathan Linial, and Alex Samorodnitsky. Inclusion-exclusion: Exact and approximate. *Combinatorica*, 16(4):465–477, Dec 1996. doi:10.1007/BF01271266.

[KS20a] Tomotaka Kuwahara and Keiji Saito. Eigenstate thermalization from the clustering property of correlation. *Phys. Rev. Lett.*, 124:200604, May 2020. doi:10.1103/PhysRevLett.124.200604.

[KS20b] Tomotaka Kuwahara and Keiji Saito. Gaussian concentration bound and ensemble equivalence in generic quantum many-body systems including long-range interactions. *Annals of Physics*, 421:168278, 2020. doi:10.1016/j.aop.2020.168278.

[Kuw16] Tomotaka Kuwahara. Connecting the probability distributions of different operators and generalization of the chernoff-hoeffding inequality. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(11):113103, nov 2016. doi:10.1088/1742-5468/2016/11/113103.

[LSM61] Elliott Lieb, Theodore Schultz, and Daniel Mattis. Two soluble models of an antiferromagnetic chain. *Annals of Physics*, 16(3):407–466, 1961. doi:10.1016/0003-4916(61)90115-4.

[LZ22] Anthony Leverrier and Gilles Zémor. Quantum tanner codes. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, 2022. doi:10.1109/FOCS54457.2022.00117.

[PGVWC07] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac. Matrix product state representations. *Quantum Info. Comput.*, 7(5):401–430, jul 2007. URL: https://dl.acm.org/doi/10.5555/2011832.2011833.

[PK22] Pavel Panteleev and Gleb Kalachev. Asymptotically good quantum and locally testable classical ldpc codes. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 375–388, 2022. doi:10.1145/3519935.3520017.

[SV14] Sushant Sachdeva and Nisheeth K. Vishnoi. Faster algorithms via approximation theory. *Foundations and Trends® in Theoretical Computer Science*, 9(2):125–210, 2014. doi:10.1561/0400000065.

[Tas18] Hal Tasaki. On the local equivalence between the canonical and the microcanonical ensembles for quantum spin systems. *Journal of Statistical Physics*, 172(4):905–926, Aug 2018. doi:10.1007/s10955-018-2077-y.

A Locality spread under Hamiltonian evolution

Lemma A.1. *Let O be an operator that acts non-trivially on at most k qubits and H an ℓ-local Hamiltonian. Then $e^{-\iota H} O e^{\iota H}$ is a (k', ϵ')-local operator for*

$$k' = 2\ell d + k\,, \quad \epsilon' = 3e^{-(d - e\|H\|)}\,\|O\|\,,$$

for any integer $d \geq e\,\|H\|$.

Proof. Let

$$Q = \mathbb{1} + \sum_{m=1}^{d} \frac{(-\iota H)^m}{m!}$$

be the Taylor series of $e^{-\iota H}$ truncated at degree d. By Taylor's theorem, $\left\| e^{-\iota H} - Q \right\| \leq e^{-(d - e\|H\|)}$, so it follows from the triangle inequality and submultiplicativity of the norm that for $d \geq e\,\|H\|$,

$$\left\| e^{-\iota H} O e^{\iota H} - QOQ^\dagger \right\| \leq 3e^{-(d - e\|H\|)}\,\|O\|\,.$$

Because H^d can be expanded as a sum of terms that each contain at most d of the ℓ-local terms H_i, we see that Q is (ℓd)-local. Therefore, $QOQ^\dagger$ is $(2\ell d + k)$-local, concluding the proof. □

Lemma A.2. *Let $R = \sum_i R_i$ be a k-local operator and $H = \sum H_i$ an ℓ-local* commuting *Hamiltonian. Suppose that there exist constants $\alpha \in [0, 1)$ and $\tilde{C} > 0$ such that for every subset $S \subseteq [n]$ of qubits, the subset Hamiltonian H_S (see Definition 2.6) satisfies $\|H_S\| \leq \tilde{C} n^\alpha |S|^{1-\alpha}$. Then, $e^{-\iota H} R e^{\iota H}$ is a (k', ϵ')-local operator for*

$$k' = 2\ell \lceil \mu \tilde{C} e n^\alpha \kappa^{1-\alpha} \rceil + k\,, \quad \epsilon' = 3e^{-(\mu-1)\tilde{C} e n^\alpha \kappa^{1-\alpha}}\,\textsc{tln}(R)\,, \tag{A.1}$$

for any $\mu > 1$ and $\kappa \geq k$. Furthermore,

$$\textsc{tln}_{\epsilon'}\left(e^{-\iota H} R e^{\iota H}\right) \leq (2n)^{k'}\left(\textsc{tln}(R) + \epsilon'\right)\,. \tag{A.2}$$

Proof. For each $i \in [t]$, we define S_i as the subset of qubits on which R_i acts non-trivially. Since R is k-local, $|S_i| \leq k$ for all i. Because H is commuting,

$$e^{-\iota H} R_i e^{\iota H} = e^{-\iota H_{S_i}} R_i e^{\iota H_{S_i}}\,.$$

Recall that $\|H_{S_i}\| \leq \tilde{C} n^\alpha k^{1-\alpha} \leq \tilde{C} n^\alpha \kappa^{1-\alpha}$. We can therefore apply Lemma A.1 with $d = \lceil \mu \tilde{C} e n^\alpha \kappa^{1-\alpha} \rceil > e\,\|H_{S_i}\|$ to find that $e^{-\iota H_{S_i}} R_i e^{\iota H_{S_i}}$ is a (k', ϵ_i')-local operator with

$$k' = 2\ell \lceil \mu \tilde{C} e n^\alpha \kappa^{1-\alpha} \rceil + k\,, \quad \epsilon_i' = 3e^{-(\mu-1)\tilde{C} e n^\alpha \kappa^{1-\alpha}}\,\|R_i\|\,.$$

We define $\tilde{R}^i$ as the local operator approximations to $e^{-\iota H_{S_i}} R_i e^{\iota H_{S_i}}$ (as given by Lemma A.1) and $\tilde{R} = \sum_i \tilde{R}^i$. (We use superscripts for $\tilde{R}^i$ because each $\tilde{R}^i$ is itself a k'-local operator, not an operator that acts non-trivially on only k' qubits, i.e. $\tilde{R} = \sum_i \tilde{R}^i$ is not our usual local decomposition.) Since the sum of k'-local operators is still k'-local, we see that $\tilde{R}$ is k'-local operator and, by the triangle inequality, approximates $e^{-\iota H} R e^{\iota H} = \sum_i e^{-\iota H_{S_i}} R_i e^{\iota H_{S_i}}$ to within error

$$\epsilon' \leq \sum_i \epsilon'_i = 3e^{-(\mu-1)\tilde{C}en^{\alpha}\kappa^{1-\alpha}} \sum_i \|R_i\| = 3e^{-(\mu-1)\tilde{C}en^{\alpha}\kappa^{1-\alpha}} \textsc{tln}(R)\,.$$

This completes the proof of Equation (A.1).

To show Equation (A.2), we first bound $\textsc{tln}(\tilde{R}^i)$ in terms of $\left\|\tilde{R}^i\right\|$. Since $\tilde{R}^i$ is k'-local, we can expand

$$\tilde{R}^i = \sum_{T \subseteq [n] \text{ s.t. } |T| \leq k'} \tilde{R}^i_{(T)}$$

using the unique decomposition from Equation (2.1). The number of terms $\tilde{R}^i_{(T)}$ is at most $\sum_{j=0}^{k'} \binom{n}{j} \leq n^{k'}$. Therefore,

$$\textsc{tln}(\tilde{R}^i) \leq \sum_T \left\|\tilde{R}^i_{(T)}\right\| \leq n^{k'} \max_T \left\|\tilde{R}^i_{(T)}\right\| \leq n^{k'} \cdot 2^{k'} \cdot \left\|\tilde{R}^i\right\|\,,$$

where the last inequality follows from Lemma 2.7. Due to unitary invariance of the norm:

$$\left\|\tilde{R}^i\right\| \leq \left\|e^{-\iota H_{S_i}} R_i e^{\iota H_{S_i}}\right\| + \epsilon'_i = \|R_i\| + \epsilon'_i$$

Finally, we combine the above bounds to obtain Equation (A.2):

$$\textsc{tln}_{\epsilon'}\left(e^{-\iota H} R e^{\iota H}\right) \leq \sum_i \textsc{tln}(\tilde{R}^i) \leq (2n)^{k'} \left(\textsc{tln}(R) + \epsilon'\right)\,.$$

□