Stoquastic PCP vs. Randomness

Dorit Aharonov
School of Computer Science and Engineering
The Hebrew University of Jerusalem
Jerusalem, Israel
Email: doria@cs.huji.ac.il

Alex B. Grilo
CWI and QuSoft
Amsterdam, The Netherlands
Email: Alex.Grilo@cwi.nl

***Abstract*—The derandomization of MA, the probabilistic version of NP, is a long standing open question. In this work, we connect this problem to a variant of another major problem: the quantum PCP conjecture. Our connection goes through the surprising quantum characterization of MA by Bravyi and Terhal. They proved the MA-completeness of the problem of deciding whether the groundenergy of a uniform stoquastic local Hamiltonian is zero or inverse polynomial. We show that the gapped version of this problem, i.e. deciding if a given uniform stoquastic local Hamiltonian is frustration-free or has energy at least some constant ε, is in NP. Thus, if there exists a gap-amplification procedure for uniform stoquastic Local Hamiltonians (in analogy to the gap amplification procedure for constraint satisfaction problems in the original PCP theorem), then MA = NP (and vice versa). Furthermore, if this gap amplification procedure exhibits some additional (natural) properties, then P = RP. We feel this work opens up a rich set of new directions to explore, which might lead to progress on both quantum PCP and derandomization.**

We also provide two small side results of potential interest. First, we are able to generalize our result by showing that deciding if a uniform stoquastic Local Hamiltonian has negligible or constant frustration can be also solved in NP. Additionally, our work reveals a new MA-complete problem which we call SetCSP, stated in terms of classical constraints on strings of bits, which we define in the appendix. As far as we know this is the first (arguably) natural MA-complete problem stated in non-quantum CSP language.

I. Introduction

It is a long standing open question, whether the randomized version of NP, called MA (for Merlin-Arthur) can be derandomized, namely, whether MA = NP. MA is defined like NP, except the verifier is probabilistic: If x is a yes-instance, there exists a proof y of polynomial size, such that the verifier always accepts (x, y). If x is a no-instance, the verifier rejects (x, y) for any polynomial y, with high probability.

The derandomization of MA is implied by widely believed conjectures such as NEXP does not have polynomial size circuits [IKW02], as well as by the stronger P = BPP conjecture [GZ11], which itself is implied by the existence of one-way functions [HILL99] and also by commonly conjectured circuit lower bounds [BFNW93], [NW94], [IW97], [STV01], [KI04]. The (somewhat counter-intuitive at first) connection between lower-bounds on computation, and derandomization (which can be seen as an upper-bound result) coined the intriguing name "Hardness versus Randomness" [NW94]. Our work follows this path, and provides a result of a similar flavor: we connect the derandomization of MA (as well as, more weakly, to that of RP) with a different hardness problem in computational complexity – that of quantum PCP – hence the title of this paper.

Our starting point is a seminal and beautiful paper of Bravyi and Terhal [BT09], where they introduced the first natural MA complete problem, which surprisingly is defined in the quantum setting[1]. To explain this quantum characterization of MA, and how we use it to make the connection to quantum PCP, let us make a detour to quantum Hamiltonian complexity.

A. Hamiltonian complexity and stoquastic Hamiltonians

The power of QMA [KSV02] the quantum version of NP, had been a major area of study in the past decade, leading to enormous progress in our understanding of the complexity of quantum states and the reductions between Hamiltonians (see [Osb12], [GHLS15]). In QMA, a polynomial time *quantum* verification algorithm receives a *quantum* proof $|\psi\rangle$ for some classical input x. and applies a quantum polynomial algorithm on both x and $|\psi\rangle$. If x is a yes-instance, there is a $|\psi\rangle$ which causes the verifier to accept with high probability; otherwise, no matter what the quantum proof $|\psi\rangle$ is, x should be rejected with high probability. In addition to being a natural generalization of classical proof systems, the relevance of QMA was evidenced by Kitaev, who showed that estimating the lowest eigenvalue of a local Hamiltonian, a central problem in many body quantum physics, is complete for QMA [KSV02]. Kitaev's theorem is the quantum analog of the Cook-Levin theorem [Coo71], [Lev73], and it makes a strong connection between a major question in condensed matter physics (namely, groundstates

[1] For PromiseMA, it is folklore that one can define complete problems by extending NP-complete problems (see, e.g. [SW]): we define an exponential family of 3SAT formulas (given as input succinctly) and we have to decide if there is an assignment that satisfies all of the formulas, or for every assignment, a random formula in the family will not be satisfied with good probability.

2575-8454/19/$31.00 ©2019 IEEE
DOI 10.1109/FOCS.2019.00065

of local Hamiltonians), and a major problem in Theoretical Computer Science, (namely, optimal solutions for constraint satisfaction problems). In fact, the connection is even deeper since what is shown is that the latter is simply a special case of the former.

More concretely[2], physicists associate with an n qubit system a self-adjoint operator called a Hamiltonian, which corresponds to the *energy* of the system, and can be usually decomposed as a sum of terms corresponding to interactions between a small number of particles. Of major interest is the lowest eigenvalue of this operator, and its corresponding eigenstates - called *groundstates*. Looking at this problem through the Theoretical Computer Science lens, Kitaev defined the k-Local Hamiltonian problem [KSV02], [AN02], whose input is a Hamiltonian H on an n particle system given as a sum of m local terms, each of them acting non-trivially on at most k out of the n particles (the term local only refers to the fact that k is assumed to be small, there are no geometrical restrictions on the interaction). We are also given as input two parameters, α and β. We then ask if the smallest eigenvalue of H is smaller than α, or all states have energy larger than β. The hardness of the Local Hamiltonian problem depends on the input *promise gap* defined as $\beta-\alpha$; Kitaev showed that the problem is QMA-complete for some inverse polynomial promise gap [KSV02], [AN02].

Bravyi, DiVincenzo, Oliveira and Terhal [BDOT08] asked how the problem behaves when the terms are restricted such that their off-diagonal elements are all non-positive, a property that they named "stoquastic" (as a combination of the words stochastic and quantum). This property implies a lot of structure on groundstates (See Lemma 13), and in physics it is associated with the lack of the "sign problem", in which case one can associate with the Hamiltonian a classical Markov Chain Monte Carlo experiment and study it (See [BDOT08]); such systems are considered far easier than general Hamiltonians. [BBT06] showed that the stoquastic Local Hamiltonian problem is StoqMA-complete, where StoqMA is a complexity class that sits between MA and QMA.

Importantly for this paper, Bravyi and Terhal [BT09] then showed that the stoquastic Local Hamiltonian problem is MA-complete if we pick $\alpha = 0$ and $\beta \geq \frac{1}{poly(n)}$, or in other words, if we want to decide whether the Hamiltonian is *frustration-free*[3] or the lowest eigenvalue is at least inverse polynomial. This was, to the best of our knowledge, the first MA-complete problem which is not an extension of NP-complete problems into the randomized setting[4] (see also [Bra14]). A simple observation important for the current work is that in the MA-complete problem of [BT09], the groundspaces of the local terms are all spanned by *subset-states*, i.e., states which are the *uniform* superposition of a subset of strings. We call such Hamiltonians *uniform* stoquastic Hamiltonians.

This paper is concerned with the *gapped* version of the uniform stoquastic Local Hamiltonian problem. Gapped versions of NP-hard problems have played a crucial role in the topic of probabilistically checkable proofs (PCPs) which had revolutionized classical Theoretical Computer Science over the past three decades. Before we define the gapped version of the uniform stoquastic Hamiltonian problem, let us introduce the by-now-standard notion of PCPs in more detail.

B. The PCP theorem

The "mother" of all NP-complete problems is 3SAT. An instance to this problem is a Boolean formula ϕ in the form $\phi(x) = C_1 \wedge C_2 \wedge ... \wedge C_m$, where $C_i = (y_{i,1} \vee y_{i,2} \vee y_{i,3})$ is a clause and $y_{i,j} \in \{x_1, ..., x_n, \overline{x_1}, ...\overline{x_n}\}$. We ask if there exists an assignment $x \in \{0,1\}^n$ such that ϕ is satisfied.

The problem MAX3SAT$_\delta$, parametrized by some function $\delta(n)$, is a generalization of 3SAT. In this problem we have to distinguish between the cases where ϕ is satisfiable or for every assignment of the input variables, at least a $\delta(n)$ fraction of the clauses are not satisfied.

By picking $\delta(n) = \frac{1}{m}$, MAX3SAT$_\delta$ becomes equivalent to 3SAT, and therefore it is NP-complete. In PCPs we are interested in versions of the problem with significantly larger δ. The celebrated PCP theorem [ALM+98], [AS98] states the remarkable result that there exists some constant ε independent of n, such that the problem MAX3SAT$_\varepsilon$ is NP-complete. This problem with constant ε is called the *gapped* version of the problem, and the PCP theorem proves that the gapped version of this problem is as hard as the original one.

In her celebrated alternative proof to the PCP theorem, Dinur [Din07] used an explicit *gap amplification* procedure, or *reduction*. The reduction takes an n-bit instance ϕ of 3SAT (or equivalently of MAX3SAT$_{\frac{1}{m}}$) to an instance ϕ' of MAX3SAT$_\varepsilon$, acting on not many more than n bits, for some constant ε. It is required that ϕ is satisfiable iff ϕ' is satisfiable. Importantly, when ϕ is not satisfiable, every assignment to the variables of ϕ' leaves at least ε fraction of the clauses of ϕ' unsatisfied. This shows that even when there is a constant ε promise gap between the yes- and no-cases, the problem remains NP-hard. The PCP theorem is one of the crown jewels of Computational Complexity Theory, with far-reaching applications such as hardness of approximation, verifiable delegated computation, program obfuscation, and cryptocurrencies (e.g., [Hås01], [GKR15], [BISW17], [BSCG+14]).

Does a quantum version of the PCP theorem [AN02], [AALV09], [AAV13] hold? The gap amplification version of the quantum PCP conjecture can be stated as follows: the Local Hamiltonian problem remains QMA-complete even when the promise gap is a constant ε. This PCP conjecture

[2] See Section III for a detailed definition.

[3] A Hamiltonian is frustration-free if there is a state in the groundspace of all local terms.

[4] As commented on in Footnote 1

has implications also to our understanding of multipartite entanglement, specifically, whether multipartite entanglement can persist at “room-temperature” (see [AAV13]). Despite much work on this direction [AALV09], [Ara11], [Has13b], [BH13], [FH14], [AE15], [EH17], [NVY18], progress had so far been limited, and it remains a major open problem [5].

What about PCPs for MA? To the best of our knowledge, the only work on PCPs for randomized classes is Drucker's [Dru11], which proves a PCP theorem for AM, another randomized version of NP. The following is a complete problem: an instance is a (succinctly given) *collection* of Boolean formulas $\{\phi_r\}$, and we want to decide between the yes-case: every formula in this family is satisfiable, or the no-case: with high probability a uniformly random formula in this family is not satisfiable. Drucker proved that if one replaces ”not satisfiable” in the description of no instances, by ”ε-far from satisfiable” (for some fixed constant ε) - namely every assignment violates at least ε fraction of the constraints - the problem remains AM hard. The proof relies on the PCP theorem for NP, and does not seem to provide insight as to how to carry it over to MA.

C. Our Contribution

This work creates a surprising link between the long standing problem of derandomizing MA, to the seemingly unrelated question of quantum PCP for stoquastic Hamiltonians. By the work of [BT09] the latter question can be viewed as a variant of PCP for MA. Before stating the results, let us provide some definitions.

Definition (Informal):: The (ε, k, d)- Gapped, Uniform, Stoquastic, Frustration-Free, Local Hamiltonian problem.) An input is a set of m positive semi-definite uniform stoquastic terms $H_1, \ldots, H_m$, where each H_i acts on k out of the n qudit system and $\|H_i\| \leq 1$; moreover every qubit participates in at most d terms. For $H = \frac{1}{m}\sum_{j=1}^{m} H_j$, decide which of the following holds, given the promise that one is true:
Yes. There exists a quantum state $|\psi\rangle$ such that $\langle\psi|H|\psi\rangle = 0$. (the groundenergy of H is 0)
No. For all quantum states $|\psi\rangle$ it holds that $\langle\psi|H|\psi\rangle \geq \varepsilon$. (all eigenvalues are at least ε)

Thus, we are given a bounded degree uniform stoquastic k-local Hamiltonian which is either frustration-free or constantly frustrated. Our main result is that distinguishing these two is in NP.

Theorem 1 (Main: uniform stoquastic frustration-free Local Hamiltonian is in NP). *For any constants $\varepsilon > 0, k, d$, the problem of deciding whether a uniform d-bounded degree stoquastic k-local Hamiltonian H is frustration-free or ε-frustrated, is in* NP.

[5] We mention that a the *game* version of quantum PCP, whose connection to the above quantum PCP is not well understood, was recently resolved [NV18].

We note that the same problem, except with inverse polynomial gap, is MA-complete.

The restriction in Theorem 1 that the yes-instances have perfect completeness (frustration free) seems too strong. Indeed, further work enables us to strengthen Theorem 1 and relax the requirement, such that yes-instances are just negligibly frustrated:

Theorem 2 (uniform stoquastic Local Hamiltonian with Imperfect completeness is in NP). *The problem of deciding whether a uniform stoquastic Hamiltonian H on n qubits has negligible frustration[6] or is at least ε-frustrated is in* NP*, for any constant ε.*

Hence, first of all, this provides a new tighter upper-bound on the hardness of ground energy and groundstate of stoquastic Hamiltonians, in case the promise gap is constant. This is of interest first of all in the context of quantum Hamiltonian complexity [Osb12], [GHLS15]. We next discuss the implications of these results from the complexity-theoretical point of view. To this end we propose the following conjecture:

Conjecture 3 (Stoquastic PCP conjecture)**.** *(Informal) There exist constants $\varepsilon > 0$, $k', d' > 0$ and an efficient gap amplification procedure that reduces the problem deciding if a uniform stoquastic d-degree k Local-Hamiltonian is frustration-free or at least inverse polynomially frustrated, to the problem of deciding if a uniform stoquastic d'-degree k' Local-Hamiltonian is frustration-free or at least ε frustrated.*

An immediate consequence of Theorem 1 is that Conjecture 3 implies MA = NP.

Corollary 4 (uniform stoquastic PCP conjecture implies derandomization of MA). *If the stoquastic* PCP *conjecture is true, then* MA = NP.

From an optimistic perspective, our result thus opens a way towards proving that MA = NP (as well as circuit lower bounds by [IKW02]) via quantum arguments, in particular by proving specific types of quantum PCPs. This path could of course be very hard, but under the belief that MA = NP, such a gap amplification procedure is in fact *known* to exist. Taking the opposite point of view, and assuming the less commonly believed assumption that MA is strictly larger than NP, our work proves that no PCP exists for stoquastic local Hamiltonians, or, loosely speaking, there is no PCP for MA.

It is natural to ask whether these results have implications to the derandomization of RP and BPP, and in particular, would stoquastic quantum PCPs imply anything in this scaled down context. We provide some (weaker, namely with more assumptions) results in this direction, stated in Appendix A. We describe these briefly here. First, one can

[6] A function f is called *negligible* if $f = o\left(\frac{1}{n^c}\right)$ for every constant c.

slightly modify the uniform stoquastic Local Hamiltonian problem by requiring that in the yes-case, the groundstate has the all 0 string in its support (see Definition 47); then no witness is needed. We call this the *pinned* version of the problem. The same proof of Theorem 1 would imply that the gapped version of this pinned uniform stoquastic problem is in P. A somewhat strengthened version of Conjecture 3, with the additional natural requirement that the gap amplification reduction is also *witness preserving* [7] would then imply that P = co-RP, and since P is closed under complement, P = RP. One might expect that Theorem 2 would imply similar implications for BPP; however we do not know how to overcome a technical obstacle in this argument and thus clarifying a reasonable set of assumptions that would imply BPP=P remains open. For details see Appendix A.

Finally, as a small side result, we present in Appendix A an alternative proof for the result of [GLSW15] that shows that the *commuting* version of stoquastic Hamiltonian problem is in NP (for any promise gap.)

Theorem 5 (commmuting stoquastic Local Hamiltonian problem is in NP)**.** *The problem of deciding if a commuting stoquastic Hamiltonian H is frustration-free is in* NP.

D. Proof overview and main ideas

We prove Theorem 1 by derandomizing the verification algorithm used in [BT09] in their proof of the containment in MA of the inverse-polynomial version of the stoquastic Hamiltonian problem. The derandomization becomes possible when the gap is constant, namely, when we know that the Hamiltonian is either frustration free or there is a large amount of frustration.

We briefly explain now the main ideas behind the randomized verification procedure of [BT09], using a random walk; then we overview our approach to derandomize it. They start by defining an (exponential-size) graph whose vertices are all possible n-bit strings. The edges are defined based on the stoquastic Hamiltonian: two strings x, y are adjacent in the graph, iff they are connected by some H_i, one of the local terms of the Hamiltonian, namely, if x and y appear together in some groundstate of H_i. The paper considers the following random-walk on the graph: starting from a given n-bit string, pick one of the terms uniformly at random, and go to any of the (constantly many) strings connected to the current string by that term, uniformly at random. This is called a *step*.[8] Bravyi and Terhal also define the notion of a *bad string*, which is a string that does not appear in the support of any of the groundstates of some local term. All other strings are *good*.

Bravyi and Terhal then showed that if the stoquastic Hamiltonian is frustration-free, then the connected component of any string in the support of some groundstate of the Hamiltonian, does not contain bad strings. In particular, any walk on the above defined graph, starting from some string in a groundstate, does not reach bad strings. On the other hand, if the Hamiltonian is at least $\frac{1}{p(n)}$ frustrated, for some polynomial p, then there exists some polynomial q such that a $q(n)$-step random walk starting from any initial string reaches a bad string with high probability. The MA verification algorithm then proceeds by the Prover sending some x, which is supposed to lie in the support of some groundstate of the Hamiltonian and the verifier performs a $q(n)$-step random walk starting from x, as above. The algorithm rejects if a bad string is encountered in the random-walk.

Our main technical result is showing that if the Hamiltonian is ε frustrated, for some constant ε independent of n, then from any initial string it is possible to reach a bad string in r steps, where r is a constant that only depends on ε, k and d. Therefore, we can define an NP verification algorithm which, given some initial string x, tries all possible r-size paths, and this can be performed in polynomial time since r is constant. We describe now the main ideas on how to prove that for highly frustrated Hamiltonians, such a short path always exists.

Our proof is based on the following two key ideas. First, we notice that if we start with any initial quantum state which is a uniform superposition of good strings, then in case the Hamiltonian is highly frustrated, there must be a term H_i which has large energy with respect to that string (in fact, there must be many, but for now we focus on one). When we apply on the state the projection P_i onto the groundspace of that frustrated term H_i, then it is not very difficult to see that the number of strings in the support of the new state, after this projection, will be larger by a constant factor. Moreover, the value of this *expansion factor* is directly related to the amount of frustration of H_i with respect to the state we started with. In other words, the more frustrated the term is, the larger the expansion of the set of strings would be, due to projection with respect to that term. We call this phenomenon "one term expansion"; it is proven in Lemma 31.

Now, the idea is to start with one good string provided by the prover, and expand it to an increasingly larger set of good strings by such projections. Our goal is to perform such expansions by a "circuit of parallel non-overlapping[9] projections", as in Figure 1a. We would like to argue that if the frustration of the Hamiltonian is high for any state, as we assume now, then there is a constant fraction of the m terms, given by one layer in the circuit, which

[7] A reduction is witness preserving if there is a polynomial time algorithm that maps a witness of the original problem into a witness of the target problem.

[8] In the non-uniform case, there are weights involved and the random-walk becomes more complicated, but here we focus only on the uniform case.

[9] Two terms are non-overlapping if the sets of qudits on which they act are disjoint

are all at least constantly frustrated. By the single term expansion argument, each such term would contribute a constant multiplicative factor to the number of good strings in our set, and thus the size of the set of good strings accumulates an exponential factor due to each *layer* in the circuit. If this is true, then it must be the case that after at most constantly many layers, the argument breaks down (namely, a bad string is found) since otherwise the number of strings would exceed the number of all possible strings. The implication is that after constantly many layers, a bad string is reached.

Unfortunately, there is a problem in applying the above line of thought directly: the amount of expansion of two different terms might be strongly correlated. Let us see an example of that.

Example 6. *Let* $S = \{0000, 0011, 1100, 1111\}$ *and let* $P_{1,4} = P_{2,3} = |\Phi^+\rangle\langle\Phi^+| + |\Psi^+\rangle\langle\Psi^+|$, *where* $P_{i,j}$ *acts on qubits* i *and* j *(and are implicitly tensored with identity on the other qubits), and* $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ $|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ *are two of the Bell states. We notice that* $\langle S|P_{1,4}|S\rangle = \frac{1}{2}$ *and the same holds for* $P_{2,3}$.

However, if we take the support of $P_{1,4}|S\rangle$, $S' = \{0000, 0110, 0011, 0101, 1100, 1010, 1111, 1001\}$, *it follows that* $\langle S'|P_{2,3}|S'\rangle = 0$, *so* $P_{2,3}$ *has no frustration after we correct the frustration of* $P_{1,4}$.

This example means that we cannot use the above argument as stated, for many non-overlapping terms applied in parallel: even though there are indeed a linear number of non-overlapping terms which are all frustrated, we cannot simply multiply the expansions due to each of them.

We overcome this difficulty by resorting to some "online" version of the claim: it turns out that by using an adaptive argument, a constant fraction of terms can be found which will all contribute independent multiplicative factors to the increase in size of the set of good strings. This means that each layer in the circuit does contribute an exponential increase in the number of strings.

We can now provide a sketch of the first part of the proof: in the case of ε-frustration, assume we start with a subset-state of good strings $|S\rangle$, and let $L|S\rangle$ be the state which we arrive at, after applying all non-overlapping projections in the sequence L which we have found above, and taking the subset-state of all strings we have reached. We can show that $L|S\rangle$ contains $\left(1 + \frac{\varepsilon}{4}\right)^{\frac{\varepsilon n}{2kd}}$ more strings than S. Then, we repeat this constantly many times. More concretely, set $S_0 = \{x\}$, for some initial string x. The above argument shows that either $|S_0\rangle$ contains a bad string (i.e. x is a bad string), or there is a set of terms L_1 such that the set S_1 with the strings in the support of $L_1|S_0\rangle$ has exponentially more strings than S_0. We now repeat this process starting with the state $|S_1\rangle$ instead of $|S_0\rangle$, and so on, until we reach a bad string. Since the number of strings in the set increases

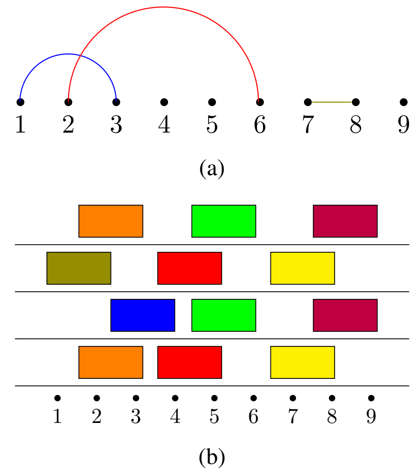


Figure 1: In Figure 1a, we show an example of non-overlapping 2-local terms, where each term corresponds to an edge of the graph. In Figure 1b, we depict a constant-depth circuit where each layer contains non-overlapping 2-local terms.

exponentially at every step, there exists some constant ℓ, that depends only on ε, k and d, such that S_ℓ (which we prove to be the strings in the support of $L_\ell ... L_1|x\rangle$) contains a bad string. This shows that a constant depth circuit of non-overlapping projections, applied to an input string, leads to a bad string. We depict this constant depth circuit in Figure 1b. The proof of the claim that within constantly many layers a bad string is reached, is given in Lemma 37.

We notice that such a constant depth circuit implies that a bad string can be found within a constant number of rounds, where each round consists of a set of local steps, each changing a different local part of the string. However the number of steps in each round might be linear, leading to exponentially many possibilities; thus, the brute-force search for such a path is intractable.

The next part of the proof is where we show how to find a bad string efficiently, given that one is reached in the above constant depth projection circuit. We call this "the light cone argument". To retrieve a constant size path from $L_\ell ... L_1|x\rangle$, the key point is noticing that badness of a string is a *local* property, namely, if a string is bad, we can point at at least one local term which it is bad for; let us refer to this term as the frustrated term (this is a meaningful name since if a state contains that bad string, then that term will indeed be frustrated). The crux of the matter is that the fact that badness of a string is local, implies also that projections on one set of qubits does not affect the badness of terms restricted to the complementary set of qudits (see Claim 16). This implies, by a simple argument, that even if we remove all of the terms in $L_1, ..., L_\ell$ that are not in the *lightcone* of the frustrated term, we will still achieve a bad string. This is because if $L_\ell ... L_1|x\rangle$ contains a bad string, which is bad for some term, then any projection in $L_\ell ... L_1$ which is not in the light cone of that term, cannot influence its badness. We depict such argument in Figure 2; It is proven in Claim 16.

Using the light-cone argument, we deduce that instead of applying the layers $L_\ell ... L_1$, we can apply just the terms in these layers which are contained in the lightcone, denoted by $L_1^{\triangle}$, ..., $L_\ell^{\triangle}$, to conclude that also the state $L_\ell^{\triangle} ... L_1^{\triangle}|x\rangle$ contains a bad string. Since the lightcone operators are of constant size, every string in $L_i^{\triangle} ... L_1^{\triangle}|x\rangle$ can be reached from a string in $L_{i-1}^{\triangle} ... L_1^{\triangle}|x\rangle$ in a constant number of steps, and by induction we deduce that there is a short path from x to a bad string.

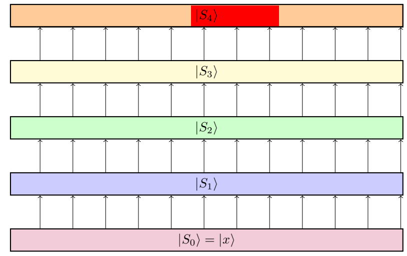
E. Related work, implications and open problems

PCP for AM vs. PCP for MA.: The PCP theorem for AM proved by Drucker [Dru11] and mentioned above, relies strongly on the classical PCP theorem: since the randomness in AM is public, both Prover and Verifier agree on the same Boolean formula, and then they can apply the original PCP theorem with this formula. In the case of MA, such an argument does not hold since the Prover does not know the formula that will be tested by the Verifier.

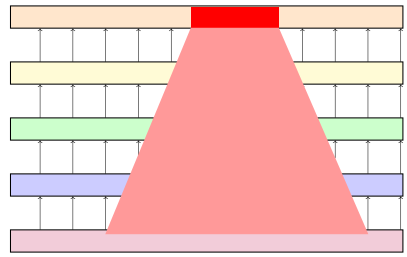
PCP for QMA vs. PCP for MA.: We notice the difference between the quantum and the uniform stoquastic PCP conjecture. It is widely believed that NP = MA, and therefore the gapped uniform stoquastic Local Hamiltonian problem is believed to be both in NP, which we prove in this work, and MA-hard, which is the uniform stoquastic PCP conjecture (Conjecture 3). In this sense, the result that we prove is somehow “expected”, and it does not change our belief that the stoquastic PCP conjecture does hold (as it is implied by BPP = P). However, in the fully quantum setting, we believe that NP $\neq$ QMA, and if the constant promise gap Local Hamiltonian problem is in NP, then this is a strong indication that the quantum PCP conjecture is false[10]. Our result implies also that if one expects to prove the quantum PCP conjecture without causing the extra “side-effect” of proving NP = MA, the gap-amplification process should not preserve uniform-stoquasticity.

Detectability lemma.: Our setting resembles that of the Detectability lemma (DL) [AALV09], a useful tool in quantum Hamiltonian complexity [AALV11], [ALV12], [GH16], [ALVV17]. Like in our setting, in the DL setting (see the formulation of [AAV16]) one considers a given local Hamiltonian, where each term is associated with a local projector on its groundspace. Starting with some state, one considers applying the local projectors one by one (in an arbitrary order). Under the assumption that every qudit participates in at most constantly many local terms, this can be viewed as applying all local projections, organized in a

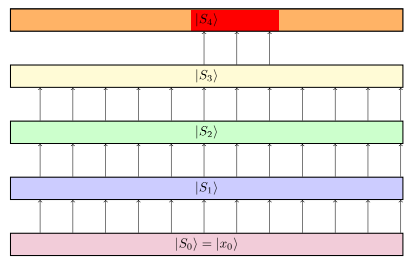
[10]Indeed, proving that constant promise gap Local Hamiltonian problem is in NP is often considered as disproving the quantum PCP conjecture.



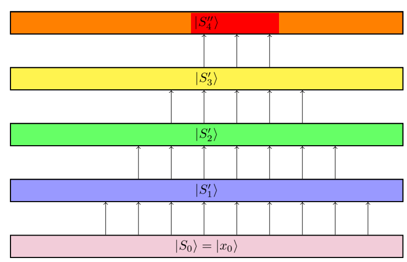
(a)



(b)



(c)



(d)

Figure 2: Example of the light-cone argument used to find a constant-step path from the initial string to a bad string. The red rectangle marks the qudits of the term frustrated by a bad string. In Figure 2a, we have a constant number of layers reaching a state with a bad string. In Figure 2b, we show the light-cone from the frustrated term. In Figure 2c, we remove the projections outside of the light-cone on the last layer. Notice that the new state on the last layer still contains a bad string. In Figure 2d, we show the projections that are left after removing all the terms outside of the light-cone. Again, the last layer still contains a bad string.

constant depth circuit made of local projections - similar to our setting. If the state we start with is the groundstate, and there is no-frustration in the Hamiltonian, then the norm of the state after all these projections of course remains one; this is the easy part. The DL says that if the Hamiltonian is frustrated, then the norm of the state after all these projections will have shrunk by at least some factor; the key in the lemma is to upper bound that factor. When the Hamiltonian is highly frustrated, the DL says that the factor will be a constant strictly less than 1. This is a strong statement; could it possibly be used to deduce our result?

A closer look reveals important differences between the two questions. While our goal is to argue containment in NP, and thus we need an efficient classical witness (which we take as an n-dit string), the DL requires full knowledge of the quantum state on which the projections are applied, in order to deduce the behavior of the norm. We thus do not know how to make any usage of the DL in our setting, though interestingly, there seem to be some conceptual connections; In particular, like in the current paper, the proof of the DL relies on arguing that the correlations between the different projections do not matter. It would be interesting if more can be said in this direction.

Uniform vs. non-uniform case.: Deriving our results without the uniformity restriction on our stoquastic Hamiltonians seems to be an interesting challenge. We note that since the uniform case is already MA hard, this will not affect our complexity implications, but seems conceptually important, as well as interesting from the complexity of groundstates point of view. Unfortunately we do not know how to do this, and it is left for future work. The main difference between the uniform and the general case is that in the uniform case, the only source of frustration is the existence of bad strings. When we move to the non-uniform case, the frustration might also appear due to amplitude inconsistency. Let us see a simple example of this.

Example. Let ε be some small value. Let us consider a one-qubit system, and the Hamiltonian consists in the sum of two terms whose groundstate projectors are

$$P_1 = \frac{1}{2}\left(|0\rangle + |1\rangle\right)\left(\langle 0| + \langle 1|\right)$$

and

$$P_2 = \left(\sqrt{1-\varepsilon}|0\rangle + \sqrt{\varepsilon}|1\rangle\right)\left(\sqrt{1-\varepsilon}\langle 0| + \sqrt{\varepsilon}\langle 1|\right).$$

We notice that the Hamiltonian is frustrated but there are no bad strings! The source of the frustration is the fact that the first terms requires the amplitude of the groundstate to be the same, but on the other hand the second term pushes them far apart.

BT [BT09] deal with this problem in their random walk by assigning weights to the edges, where the weights depend on the Hamiltonian term connecting the two strings. Then, frustration implies that the weight of different paths between two pairs of strings in the support of the groundstate, have different weights (a weight of a path is the product of the weights of the edges). BT [BT09] then add extra tests to find these inconsistent paths. At every step of the random walk, the verifier rejects if the weight of the path is larger than one (for this the verifier should start the random walk from the string with maximal amplitude in the groundstate). They prove that if the verifier is provided a string whose amplitude was not maximal or that inconsistent paths exist, then with high probability a random walk finds a path whose weight is larger than one, leading to rejection.

Interestingly, our proof goes through almost all the way, also for the non-uniform case. The only problem preventing us from extending the proof to the non-uniform case is the light-cone lemma, which does not seem to hold when attempting to detect inconsistencies rather than reachability of bad strings. A different way to say it is that as far as we can tell, the random-walk proposed by Bravyi and Terhal deals with such cases in a non-local way. The situation seems to reminiscent of the hardness of sampling from constant depth quantum circuits [TD04], [BGK18], [CSV18], [Gal18].

PCP for StoqMA.: As mentioned previously, the stoquastic Local Hamiltonian problem where we have to decide if the groundstate energy is below some threshold α or above another threshold β is StoqMA-complete for inverse polynomial $\beta - \alpha$. In fact, the StoqMA-completeness holds even for a restricted class of stoquastic Hamiltonians: the Transverse-field Ising mode [BH17]. We show in our result that if α is negligible, then the problem is still in NP, and we leave as an open question if the problem is still in NP when $\beta - \alpha$ is constant for an arbitrary α.

Adiabatic evolution of Hamiltonians.: Bravyi and Terhal used their random walk for stoquastic Hamiltonians to prove that the adiabatic evolution of frustration-free stoquastic Hamiltonians with inverse polynomial spectral-gap can be performed in randomized polynomial time. A major open problem remains to extend their result to the general case, in which the frustration-free assumption is relaxed. This would lead to a classical simulation of adiabatic optimization, e.g., of D-Wave type algorithms [FGGS00], [BT09], [Has13a], [CCD15]. We leave as an open question whether our techniques can have any implications in that context.

A classical version of the problem.: Finally, we note that our work suggests a completely classical way to phrase the MA-complete problem of Bravyi and Terhal [BT09]. It turns out, as we present more formally in Appendix A, that *uniform* stoquastic local Hamiltonians, can be rephrased as a problem which is a generalization of constraint satisfaction problems (CSP), and which we call here Set-Constraints Satisfaction problem, or SetCSP in short. In SetCSP instead of having constraints that should be satisfied by an assignment $x \in \{0,1\}^n$, we have "set-constraints" that should be satisfied by some *set* of strings $S \subseteq \{0,1\}^n$. A k-local single set-constraint acting on the set B of k bits, is a collection of non-intersecting sets $T_1, ...T_\ell$ of k bit strings. Roughly, a set S of n-bit strings satisfies such a k-local set-constraint if first, the k-bit restriction of any string in S to B must be contained in one of the sets. Secondly, there is a notion of *uniformity*: if an n-bit string $x \in S$ restricts to some string in T_j in that set-constraint, then by replacing in x the k-bits to any other string in T_j, the string is still in S. Namely, all elements of the subset T_j appear together (with the same extension to the remaining bits) or none of them appears. An instance of the SetCSP consists of m such k-local Set Constraints, and we ask if there is a set of n-bit strings that satisfies each of the set constraints, or if all sets of n-bit strings are far from satisfying them, namely, frustrated (of course, the notion of being "far" needs to be defined). We show that deciding whether a SetCSP instance is satisfiable or inverse polynomially frustrated is MA-complete.

Our result can be presented in SetCSP language, e.g. Theorem 1 implies that for constant promise gap, this problem is actually in NP. However we prefer to present our results here using quantum language since the equivalent stoquastic Hamiltonian problem seems to be a more natural problem. Never the less, this first complete problem for MA defined in CSP language might be useful in future works.

Organization of the remainder of the paper:: We start with some preliminaries in Section II. We discuss stoquastic Hamiltonians and the proof of MA-completeness in Section III. Our main result is proven in Section IV. In Section V, we show how to replace the frustration-free condition by allowing the frustration to be negligible. We finish by proving that commuting stoquastic Hamiltonians are in NP in Appendix A. We leave to Appendix A the definition of the variants of the stoquastic Hamiltonian problems which are hard for co-RP and BPP. In Appendix A we provide the description of the uniform stoquastic Hamiltonian problem in CSP language, and the proof it is MA complete.

Acknowledgments

The authors thank Ayal Green for the helpful discussion at early stages of this work. DA is also grateful for a very short discussion with Noam Nisan which was as insightful as much as it was short, as well as for very useful remarks from Avi Wigderson, and a question from Umesh Vazirani. We also thank Henry Yuen for the help on improving the clarity of the paper. AG is supported by ERC Consolidator Grant 615307-QPROGRESS. DA's research on this project was supported by ERC grant 280157 and ISF grant 1721/17. Part of this work was done when AG was a member of IRIF, Université Paris Diderot, Paris, France, where he was supported by ERC QCC. The authors thank also the French-Israeli Laboratory on Foundations of Computer Science (FILOFOCS) and Simons collaboration grant number 385590 that allowed AG to visit Hebrew University of Jerusalem.

II. Preliminaries and Notations

A. Complexity classes, NP and MA

A (promise) problem $A = (A_{yes}, A_{no})$ consists of two non-intersecting sets $A_{yes}, A_{no} \subseteq \{0,1\}^*$. We define now the main complexity classes that are considered in this work. We start by formally defining the well-known class NP.

Definition 7 (NP)**.** *A problem $A = (A_{\text{yes}}, A_{\text{no}})$ is in* NP *if and only if there exist a polynomial p and a deterministic algorithm D, where D takes as input a string $x \in \Sigma^*$ and a $p(|x|)$-bit witness y and decides on acceptance or rejection of x such that:*
Completeness. *If $x \in A_{\text{yes}}$, then there exists a witness y such that D accepts (x, y).*
Soundness. *If $x \in A_{\text{no}}$, then for any witness y, D rejects (x, y).*

We can then generalize this notion, by giving the verification algorithm the power of flip random coins, leading to the complexity class MA.

Definition 8 (MA)**.** *A problem $A = (A_{\text{yes}}, A_{\text{no}})$ is in* MA *if and only if there exist a polynomial p and a probabilistic algorithm R, where R takes as input a string $x \in \Sigma^*$ and a $p(|x|)$-bit witness y and decides on acceptance or rejection of x such that:*
Completeness. *If $x \in A_{\text{yes}}$, then there exists a witness y such that R accepts (x, y) with probability* 1*.*
Soundness. *If $x \in A_{\text{no}}$, then for any witness y, R accepts (x, y) with probability at most $\frac{1}{3}$.*

The usual definition of MA requires yes-instances to be accepted with probability at least $\frac{2}{3}$, but it has been shown that there is no change in the computational power if we require the verification algorithm to always accept yes-instances [ZF87], [GZ11].

B. Quantum states

We review now the concepts and notation of Quantum Computation that are used in this work. We refer to Ref. [NC00] for a detailed introduction of these topics.

Let $\Sigma = \{0, ..., q-1\}$ be some alphabet. A qudit of dimension q is associated with the Hilbert space $\mathbb{C}^\Sigma$, whose canonical (also called computational) basis is $\{|i\rangle\}_{i\in\Sigma}$. A pure quantum state of n qudits of dimension q is a unit vector in the Hilbert space $\left(\mathbb{C}^\Sigma\right)^{\otimes n}$, where $\otimes$ is the Kroeneker (or tensor) product. The basis for such Hilbert space is $\{|i\rangle\}_{i\in\Sigma^n}$. For some quantum state $|\psi\rangle$, we denote $\langle\psi|$ as its conjugate transpose. The inner product between two vectors $|\psi\rangle$ and $|\phi\rangle$ is denoted by $\langle\psi|\phi\rangle$ and their outer product as $|\psi\rangle\langle\phi|$. For a vector $|\psi\rangle \in \mathbb{C}^{|\Sigma|^n}$, its 2-norm is defined as $\||\psi\rangle\| := \left(\sum_{i\in\Sigma^n} |\langle\psi|i\rangle|^2\right)^{\frac{1}{2}}$.

We now introduce some notation which is somewhat less commonly used and more specific for this paper: the support of $|\psi\rangle$, $supp(|\psi\rangle) = \{i \in \Sigma^n : \langle\psi|i\rangle \neq 0\}$, is the set strings with non-zero amplitude. We call quantum state $|\psi\rangle$ *non-negative* if $\langle i|\psi\rangle \geq 0$ for all $i \in \Sigma^n$. For any $S \subseteq \Sigma^n$, we define the state $|S\rangle := \frac{1}{\sqrt{S}}\sum_{i\in S}|i\rangle$ as the subset-state corresponding to the set S [Wat00]. For a non-negative state $|\psi\rangle$, we define $|\widehat{\psi}\rangle := |supp(|\psi\rangle)\rangle$ as the subset-state induced by the strings in the support of $|\psi\rangle$. We say that this is the subset state corresponding to the state $|\psi\rangle$. Analogously, for some linear operator P the state $\widehat{P|\psi\rangle}$ means the subset-state corresponding to the state $P|\psi\rangle$.

C. Hamiltonians, Groundstates, Energies, Frustration

Definition 9 (Hamiltonian). *A* Hamiltonian *on n qudits is a Hermitian operator on $\mathbb{C}^{|\Sigma|^n}$, namely, a complex Hermitian matrix of dimension $|\Sigma|^n \times |\Sigma|^n$. A Hamiltonian on n qudits is called k-*Local *if it can be written as $H = \sum_{i=1}^m \widetilde{H}_i$, where each $\widetilde{H}_i$ can be written in the form $\widetilde{H}_i = H_i \otimes I$, where H_i acts on at most k out of the n qudits.*

Hamiltonians describe the evolution of physical systems, using Schrodinger's equation. Their eigenvalues correspond to the energy of the system; more generally, the energy of a state $|\psi\rangle$ with respect to a Hamiltonian $H = \frac{1}{m}\sum_{i=1}^m H_i$ is given by $\langle\psi|H|\psi\rangle$. Notice that we use the term energy even though we average by the number of terms, so this is the average energy *per term*; this is different from the usual usage of the term energy, or energy density, in the physics literature, where one usually considers the average energy *per particle*. This normalization is more convenient in the context of PCPs [Din07]. We also consider the energy of the state with respect to a specific term H_i, which is $\langle\psi|H_i|\psi\rangle$. The minimal energy is the smallest eigenvalue of the Hamiltonian, and an eigenstate which has this energy is called a *groundstate*.

Definition 10 (Groundstate, groundspace, frustration and frustration-free). *A* groundstate *of a Hamiltonian $H = \frac{1}{m}\sum_{i=1}^m H_i$ is an eigenvector associated with its minimum eigenvalue, which is called the* groundstate energy. *The* groundspace *of a Hamiltonian is the subspace spanned by its groundstates. H is called ε-frustrated if for every state $|\psi\rangle$, $\frac{1}{m}\sum_i\langle\psi|H_i|\psi\rangle \geq \varepsilon$. Finally, H is called* frustration-free *if there exists some $|\psi\rangle$ such that for every i, the local term H_i is positive definite and $\langle\psi|H_i|\psi\rangle = 0$.*

Throughout this paper we use the following notation for a local Hamiltonian $H = \frac{1}{m}\sum_{i=1}^m \widetilde{H}_i$. We set P_i to be the *local* projection on the groundspace of H_i; while $\widetilde{P}_i = P_i \otimes I$ corresponds to the projection on the groundspace of $\widetilde{H}_i$.

We prove now a useful lower-bound on the number of frustrated terms of a highly frustrated Hamiltonian.

Claim 11 (Lower bound on frustrated terms). *Let $H = \frac{1}{m}\sum_{i=1}^m \widetilde{H}_i$ be a Local Hamiltonian that is ε-frustrated. Then for every state $|\psi\rangle$, there exist at least $\frac{\varepsilon m}{2}$ terms that are at least $\frac{\varepsilon}{2}$ frustrated.*

Proof: We prove this by contradiction. Let $F = \{i : \langle\psi|\widetilde{H}_i|\psi\rangle \geq \frac{\varepsilon}{2}\}$. We assume then that $|F| < \frac{\varepsilon m}{2}$. Then the energy of the state is

$$\begin{aligned}
&\langle\psi|H|\psi\rangle \\
&= \frac{1}{m}\left(\sum_{i\in F}\langle\psi|\widetilde{H}_j|\psi\rangle + \sum_{i:\notin F}\langle\psi|\widetilde{H}_j|\psi\rangle\right) \\
&\leq \frac{|F|}{m} + (m - |F|)\frac{\varepsilon}{2m} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon,
\end{aligned}$$

where in the first inequality we used the fact that the norm of all terms is at most 1, and that the terms outside of F contribute at most $\frac{\varepsilon}{2m}$ to the above sum by definition of the set F. In the second inequality we used our assumption that $|F| < \frac{\varepsilon m}{2}$. We then have that H is not ε-frustrated, which is a contradiction. ■

III. Background: the Stoquastic Hamiltonian problem

In this section, we define stoquastic Hamiltonians, prove certain basic properties, as well as state their relation to the complexity class MA.

A. Stoquastic Hamiltonians

In this work, we deal with a special type of Hamiltonians, which are called stoquastic.

Definition 12 (Stoquastic Hamiltonian [BDOT08]). *A k-Local Hamiltonian $H = \sum_{i=1}^m \widetilde{H}_i$ is called* stoquastic *in the computational basis if for all i, the off-diagonal elements of H_i (the local terms) in this basis are non-positive*[11].

As remarked in [MLH18], every Hamiltonian is stoquastic in the basis that diagonalizes it. However, the length of such description might be exponential in the number of qubits, since it may be impossible to write it as a sum of local terms. Some recent works [MLH18], [KT18], provide evidence that deciding if a given local Hamiltonian can be made stoquastic by local basis change is computationally hard. Therefore, in our definition we assume that the stoquastic Hamiltonian is *given* in the basis where each of the *local* terms is stoquastic, i.e., has non-positive off-diagonal elements.

We remark that WLOG we assume that each term H_i is positive semi-definite since we could just add a constant cI to the term which only causes a constant shift in the

[11]Klassen and Terhal [KT18] have a different nomenclature. They call a matrix Z-symmetric if the off-diagonal elements of the local terms are non-positive and they call a Hamiltonian stoquastic if all local terms can be made Z-symmetric by local rotations.

eigenstates of the total Hamiltonian, hence does not change the nature of the problems we discuss here but can be easily seen to make the terms PSD.

A property of a stoquastic local Hamiltonian is that the groundspace of the local terms can be decomposed in a sum of orthogonal non-negative rank-1 projectors.

Lemma 13 (Groundspace of stoquastic Hamiltonians, Proposition 4.1 of [BT09])**.** *Let H be a stoquastic Hamiltonian and let P be the projector onto its groundspace. It follows that*

$$P = \sum_j |\phi_j\rangle\langle\phi_j|, \tag{1}$$

where for all j, $|\phi_j\rangle$ is non-negative and for $j \neq j'$, $\langle\phi_{j'}|\phi_j\rangle = 0$.

Proof: We start by showing that if all of the entries of P are non-negative, then the statement holds. Let x, y, z be some strings such that $\langle x|P|y\rangle > 0$ and $\langle y|P|z\rangle > 0$. Then

$$\langle x|P|z\rangle = \langle x|P^2|z\rangle = \sum_w \langle x|P|w\rangle\langle w|P|z\rangle > \langle x|P|y\rangle\langle y|P|z\rangle > 0, \tag{2}$$

where in the first inequality we use the fact that P has only non-negative entries. Therefore, we can partition the string in equivalent classes $T_1, ...T_t$ regarding the property $\langle x|P|y\rangle > 0$.

It follows that the subspace spanned by the strings in T_i is P-invariant and therefore P is block-diagonal with respect to the direct sum of such subspaces. Using the Perron-Frobenius theorem for each of the blocks, we have that its largest eigenvalue is non-degenerate, and in this case the block is rank-one, since P is a projector with eigenvalues 1 and 0. Since all the entries of P are non-negative, then each one of these rank-one blocks correspond to a non-negative state.

This finishes the proof of the case where P has non-negative entries. We show now that this property holds for stoquastic Hamiltonians.

We have that the groundspace projector P of the Hamiltonian consists of the Gibbs state for temperature tending to 0, i.e., $P = \lim_{\beta\to\infty} q\frac{e^{-\beta H}}{Tr(e^{-\beta H})}$, where $q > 0$ is the dimension of the groundspace (This is a well known easy fact, see for example Proposition 4.1 in [BT09]). Thus, it suffices to prove that $e^{-\beta H}$ is a matrix of non-negative entries (we already know that the trace is non-negative by the fact that the eigenvalues of $e^{-\beta H}$ are positive).

Let s be some value such that $-\beta H + sI$ has only non-negative entries. Write

$$e^{-\beta H} = e^{-\beta H + sI - sI} = e^{-\beta H + sI} e^{-sI},$$

where the last equality holds because $-sI$ and $(-\beta H + sI)$ commute.

Note that the Taylor expansion of e^A is

$$e^A = \sum_{i=0}^{\infty} \frac{A^k}{k!}.$$

Thus we have that the entries of $e^{-\beta H + sI}$ are non-negative. Write $e^{-s} = p > 0$, and we have that all entries of $e^{-\beta H} = pe^{-\beta H + sI}$ are non-negative as well. ■

Definition 14. *(Stoquastic Projector) Given a projection matrix P acting on k qudits, if there exists a set of orthogonal k-qudit non-negative states $\{|\Phi_j\rangle\}_j$, such that*

$$P = \sum_j |\Phi_j\rangle\langle\Phi_j|,$$

we say P is a stoquastic projector, and we refer to this unique decomposition as a sum of projections to non-negative states as the non-negative decomposition of P.

Remark 15. *Notice that Lemma 13 implies that the projection on the groundspace of a stoquastic Hamiltonian is a stoquastic projector.*

A crucial point in the paper is the fact that when applying a local stoquastic projector P on some set of qudits Q, we do not introduce new strings in the reduced density matrix of the set of qudits outside of Q. Notice that since we are considering non-unitary operators, namely projections, then even though they are local, such projections *can* in fact have the effect of *removing* strings away from the density matrices of qubits which they do not touch; the point of this claim is that they cannot *add* new strings away from where they act.

Claim 16 (Local action of projectors)**.** *Let P be a stoquastic projector on a subset Q of $k \leq n$ qudits. Consider the projection $\widetilde{P}$ on n qudits derived from P by $\widetilde{P} = P_Q \otimes I_{\overline{Q}}$. Then $\widetilde{P}$ is also a stoquastic projector, it can be written as the following non-negative decomposition:*

$$\widetilde{P} = \sum_{z\in\Sigma^{n-k}, j} |\phi_j\rangle\langle\phi_j|_Q \otimes |z\rangle\langle z|_{\overline{Q}}, \tag{3}$$

where $\{|\phi_j\rangle\}_j$ is the non-negative decomposition of P, and moreover, for any non-negative state $|\psi\rangle$, we have:

$$\{x_{\overline{Q}} : x \in supp(\widetilde{P}|\psi\rangle)\} \subseteq \{x_{\overline{Q}} : x \in supp(|\psi\rangle)\}.$$

Proof: For the first part of the claim, the fact that $\widetilde{P}_i = P_i \otimes I$ implies that it can be written in the desired form, and this implies that it is a stoquastic projector.

For the moreover part, write $|\psi\rangle = \sum_{x\in\Sigma^n} \alpha_x|x\rangle$. We have:

$$\widetilde{P}|\psi\rangle = \sum_{x\in supp(\psi)} \sum_j \sum_{z\in\Sigma^{n-k}} \alpha_x |\phi_j\rangle\langle\phi_j|x_Q\rangle|z\rangle\langle z|_{\overline{Q}}|x_{\overline{Q}}\rangle_{\overline{Q}}$$
$$= \sum_{x\in supp(\psi)} \sum_j \alpha_x \langle\phi_j|x_Q\rangle|\phi_j\rangle_Q|x_{\overline{Q}}\rangle_{\overline{Q}}.$$

Let $y \in supp(\widetilde{P}|\psi\rangle)$; so $\langle y|\widetilde{P}|\psi\rangle \neq 0$. By the above expression there must be $x \in supp(|\psi\rangle)$ such that $x_{\overline{Q}} = y_{\overline{Q}}$. ∎

Remark 17 (Notation of local and global projectors, and their non-negative decompositions)**.** *As can be seen in Claim 16, we use the tilde to denote the global projector, i.e., $\widetilde{P} = P \otimes I_{\overline{Q}}$. We also extend (in a slightly different way) this notation to the rank-1 projectors and we denote $|\widetilde{\phi}_{j,z}\rangle := |\phi_j\rangle_Q|z\rangle_{\overline{Q}}$, for $z \in \Sigma^{n-k}$.*

Remark 18 (Uniqueness of groundstate containing a string)**.** *Consider a stoquastic projector P and its global version, $\widetilde{P} = \sum_{j,z}|\widetilde{\phi}_{j,z}\rangle\langle\widetilde{\phi}_{j,z}|$(this can be done by Claim 16 and we use the notation of Remark 17). Then for every n-dit string $x \in \Sigma^n$, there exists at most one pair of values j^*, z^* such that $\langle\widetilde{\phi}_{j^*,z^*}|x\rangle > 0$. Clearly, a similar uniqueness statement holds for the local stoquastic projector P and its non-negative decomposition, with respect to x being a k-dit string.*

We now define a particular type of strings, called *bad* strings, which play a crucial role in our result. Consider a string x such that $\langle x|\widetilde{P}_i|x\rangle = 0$; we notice that in this case x cannot belong to any groundstate of $\widetilde{P}_i$. This leads to the following definition:

Definition 19 (Bad strings [BT09])**.** *Given a stoquastic projector P on a set Q of k out of n qudits, we say that a string $x \in \Sigma^n$ is bad for P (or P-bad) if $\langle x|\widetilde{P}|x\rangle = 0$ for $\widetilde{P} = P \otimes I_{\overline{Q}}$. Equivalently, x is bad for P, if $\langle x_Q|P|x_Q\rangle = 0$. This means that x_Q is not in the support of any of the states in the non-negative decomposition, as in Lemma 13, of P. If a string is not bad for P, we say it is P-good. Given a local Hamiltonian $H = \frac{1}{m}\sum_i \widetilde{H}_i$, and the corresponding stoquastic projectors P_i, We say that x is bad for H, or H-bad, if there exists some $i \in [m]$ such that x is bad for P_i.*

One other property that we use is that starting with a non-negative state $|\psi\rangle$, and applying $\widetilde{P}_i$, the projector onto the groundspace of some local term H_i, maintains all the H_i-good strings in $|\psi\rangle$.

Lemma 20 (Strings added by Local Stoquastic Projectors)**.** *Let $|\psi\rangle$ be a non-negative n-qudit state. Consider $P = \sum_j |\phi_j\rangle\langle\phi_j|$ a stoquastic projector and its non-negative decomposition, acting on the subset Q of k qudits out of these n qudits. Then all P-good strings of $|\psi\rangle$ are also in the support of $\widetilde{P}|\psi\rangle$, where $\widetilde{P} = P \otimes I_{\overline{Q}}$. Moreover, it follows that*

$$supp(\widetilde{P}|\psi\rangle) = \bigcup_{j,z:\langle\widetilde{\phi}_{j,z}|\psi\rangle>0} supp(|\widetilde{\phi}_{j,z}\rangle).$$

Proof: Let S be the support of $|\psi\rangle$ and let also $|\psi\rangle = \sum_{x\in S}\alpha_x|x\rangle$. Let also G and B be the sets of P-good and P-bad strings of n dits, respectively. We have that

$$\begin{aligned}\widetilde{P}|\psi\rangle &= \sum_{x\in S\cap G}\alpha_x\widetilde{P}|x\rangle + \sum_{x\in S\cap B}\alpha_x\widetilde{P}|x\rangle \\ &= \sum_{x\in S\cap G}\alpha_x\widetilde{P}|x\rangle. \end{aligned} \tag{4}$$

We now use Equation (3) from Claim 16, to apply the projector $\widetilde{P}$. We have

$$\begin{aligned}&\widetilde{P}|\psi\rangle \\ &= \sum_{x\in S\cap G, j, z\in\Sigma^{n-k}} \alpha_x \left(|\phi_j\rangle\langle\phi_j|_Q \otimes |z\rangle\langle z|_{\overline{Q}}\right)|x_Q\rangle_Q|x_{\overline{Q}}\rangle_{\overline{Q}} \\ &= \sum_{x\in S\cap G, j}\alpha_x\langle x_Q|\phi_j\rangle|\phi_j\rangle_Q|x_{\overline{Q}}\rangle_{\overline{Q}}\end{aligned} \tag{5}$$

If x is a P-good string, then there exists a (unique) $|\phi_j\rangle$ such that $\langle\phi_j|x_Q\rangle > 0$. Using also that $x \in S$, we see that the amplitude of x in the above expression $\widetilde{P}|\psi\rangle$ is non-zero.

For the moreover part, notice that Equation (5) can be written as

$$\widetilde{P}|\psi\rangle = \sum_{x\in S\cap G, j, z}\alpha_x\langle x|\widetilde{\phi}_{j,z}\rangle|\widetilde{\phi}_{j,z}\rangle \tag{6}$$

and the statement holds directly. ∎

Corollary 21 (Composition of stoquastic projectors)**.** *Let $\widetilde{P}_1$ and $\widetilde{P}_2$ be two n-qudit stoquastic projectors (which may or may not be global versions of local projections) and let $|\psi\rangle$ be a non-negative state. Then $\widehat{\widetilde{P}_1\widetilde{P}_2|\psi\rangle} = \widetilde{P}_1\widehat{\widetilde{P}_2|\psi\rangle} = |supp(\widetilde{P}_1\widetilde{P}_2|\psi\rangle)\rangle$. Moreover, if $\widetilde{P}_1$ and $\widetilde{P}_2$ commute, then the above states are also equal to $\widehat{\widetilde{P}_2\widetilde{P}_1|\psi\rangle} = \widetilde{P}_2\widehat{\widetilde{P}_1|\psi\rangle}$.*

Proof:
First, we claim that for two non-negative states $|\psi\rangle$ and $|\psi'\rangle$, if $supp(|\psi\rangle) = supp(|\psi'\rangle)$, then $supp(\widetilde{P}|\psi\rangle) = supp(\widetilde{P}|\psi'\rangle)$, for any stoquastic projector $\widetilde{P}$. This can be argued as follows. Let y be a string in the support of $\widetilde{P}|\psi\rangle$, i.e. $\langle y|\widetilde{P}|\psi\rangle \neq 0$, and $|\widetilde{\phi}_{j,z}\rangle$ be the unique state in the non-negative decomposition of $\widetilde{P}$ that contains y, as stated in Remark 18. Since $|\psi\rangle$ and $|\widetilde{\phi}_{j,z}\rangle$ are both non-negative states, we have that $\langle y|\widetilde{P}|\psi\rangle = \langle y|\widetilde{\phi}_{j,z}\rangle\langle\widetilde{\phi}_{j,z}|\psi\rangle > 0$ and in particular $\langle\widetilde{\phi}_{j,z}|\psi\rangle > 0$. Notice that since $|\psi'\rangle$ is also a non-negative state and its support is equal to the support of $|\psi\rangle$, we also have that $\langle\widetilde{\phi}_{j,z}|\psi'\rangle > 0$ and thus $\langle y|\widetilde{P}|\psi'\rangle = \langle y|\widetilde{\phi}_{j,z}\rangle\langle\widetilde{\phi}_{j,z}|\psi'\rangle > 0$, i.e. y is in the support of $\widetilde{P}|\psi'\rangle$. The converse follows from a similar argument.

In particular, by definition, $supp(\widetilde{P}_2|\psi\rangle) = supp(\widehat{\widetilde{P}_2|\psi\rangle})$; and thus applying $\widetilde{P}_1$ on both states, we get $supp(\widetilde{P}_1\widetilde{P}_2|\psi\rangle) = supp(\widetilde{P}_1\widehat{\widetilde{P}_2|\psi\rangle})$. By definition, we also have $supp(\widetilde{P}_1\widetilde{P}_2|\psi\rangle) = supp(\widehat{\widetilde{P}_1\widetilde{P}_2|\psi\rangle})$, which proves the first part of the Corollary.

For the moreover part, we can apply the previous argument to show $\widehat{\widetilde{P}_2\widetilde{P}_1|\psi\rangle} = \widetilde{P}_2\widehat{\widetilde{P}_1|\psi\rangle} = supp(\widehat{\widetilde{P}_2\widetilde{P}_1}|\psi\rangle)$. If $\widetilde{P}_1$ and $\widetilde{P}_2$ commute, we have $supp(\widehat{\widetilde{P}_1\widetilde{P}_2}|\psi\rangle) = supp(\widehat{\widetilde{P}_2\widetilde{P}_1}|\psi\rangle)$, and the result follows. ■

B. Uniform Stoquastic Hamiltonians

In this work, we focus on a restricted class of stoquastic Hamiltonian which we call uniform stoquastic Hamiltonian.

Definition 22 (uniform stoquastic Local Hamiltonian)**.** *A stoquastic Local Hamiltonian $H = \frac{1}{m}\sum_{i=1}^{m}\widetilde{H}_i$ is called uniform if the states of the unique non-negative decompositions of each local stoquastic projector (P_i) are subset-states.*

Following Claim 16 and remark 17, for uniform stoquastic Local Hamiltonians, the groundspace projector of $\widetilde{H}_i$ is $\widetilde{P}_i = (P_i)_Q \otimes I_{\overline{Q}} = \sum_{j,x}|T_{i,j}\rangle\langle T_{i,j}| \otimes |x\rangle\langle x|$, with $T_{i,j} \subseteq \Sigma^k$ and $T_{i,j} \cap T_{i,j'} = \emptyset$ for $j \neq j'$. We also denote $|\widetilde{T}_{i,j,x}\rangle := |T_{i,j}\rangle|x\rangle$ (and $\widetilde{T}_{i,j,x}$ as the corresponding set of n-bit strings).

We provide now a lemma which we do not strictly use in the proof, regarding stoquastic frustration-free Local Hamiltonians; that we can always assume that the groundstate of the entire Hamiltonian is a subset-state. Though the claim itself is not used, it is helpful to conceptually hold it in mind, when reading the proof.

Lemma 23 (The structure of groundstates of uniform stoquastic Hamiltonian)**.** *Let H be a uniform stoquastic frustration-free Local Hamiltonian. Let H_i be a local term of H. Then if $|\psi\rangle$ is a groundstate of H, it can be written in the form $|\psi\rangle = \sum_{j,z}\alpha_{i,j,z}|\widetilde{T}_{i,j,z}\rangle$, for some choice of coefficients $\alpha_{i,j,z} \in \mathbb{C}$. Moreover, the subset-state $|S\rangle$, for $S = \bigcup_{j,z:\alpha_{j,z}\neq 0}\widetilde{T}_{i,j,z}$, is also a groundstate of H.*

Proof: The first claim just follows from the fact that H is frustration free so any groundstate must be spanned by groundstates of a fixed term H_i, namely

$$|\psi\rangle = \sum_{j,z}\alpha_{i,j,z}|\widetilde{T}_{i,j,z}\rangle. \quad (7)$$

We show now that $|S\rangle$ is also a ground-state of H. Let us consider some term $H_{i'}$ and the decomposition of $|\psi\rangle$ from Equation (7) in respect to its non-negative decomposition. It follows that $\bigcup_{j,z:\alpha_{i',j,z}\neq 0}\widetilde{T}_{i',j,z} = S$. This implies that

$$|S\rangle = \sum_{j,z:\alpha_{i',j,z}\neq 0}\frac{\sqrt{\widetilde{T}_{i',j,z}}}{\sqrt{S}}|\widetilde{T}_{i',j,z}\rangle,$$

and therefore $|S\rangle$ is in the groundspace of $H_{i'}$, for any $i' \in [m]$. ■

We formally define now the frustration-free Uniform Stoquastic k-Local Hamiltonian problem.

Definition 24 (uniform stoquastic frustration-free k-Local Hamiltonian problem)**.** *The* uniform stoquastic frustration-free k-Local Hamiltonian *problem, where $k \in \mathbb{N}^*$ is called the locality and $\varepsilon : \mathbb{N} \to [0,1]$ is a non-decreasing function, is the following promise problem. Let n be the number of qudits of a quantum system. The input is a set of $m(n)$ uniform stoquastic Hamiltonians $H_1, \ldots, H_{m(n)}$ where m is a polynomial, $\forall i \in m(n) : 0 \leq H_i \leq I$ and each H_i acts on k qudits out of the n qudit system. We also assume that there are at most d terms that act non-trivially on each qudit, for some constant d, and that $m \geq n$. For $H = \frac{1}{m(n)}\sum_{i=1}^{m(n)}H_i$, one of the following two conditions hold.*
***Yes.** There exists a n-qudit quantum state $|\psi\rangle$ such that $\langle\psi|H|\psi\rangle = 0$*
***No.** For all n-qudit quantum states $|\psi\rangle$ it holds that $\langle\psi|H|\psi\rangle \geq \varepsilon(n)$.*

Remark 25. *At some points in this work, we assume that H_i is a projector. This holds without loss of generality, since we could replace an arbitrary Hamiltonian H_i by $H'_i = I - P_i$, where P_i is the projector onto the space of H_i, and this could just change frustration by a constant factor.*

C. MA-completeness

Bravyi and Terhal [BT09] showed that there exists some polynomial $p(n) = O(n^2)$ such that the frustration-free uniform stoquastic 6-Local Hamiltonian problem with $\varepsilon(n) = \frac{1}{p(n)}$ is MA-hard. They also proved that for every polynomial p' and every constant k, the frustration-free uniform stoquastic k-Local Hamiltonian problem with $\varepsilon(n) = \frac{1}{p'(n)}$ is in MA.

Let us start with the direction which is of less technical interest to us, and thus we will not need to go into details. The MA-hardness is proved by analyzing the quantum Cook-Levin theorem [KSV02], [AN02] when considering an MA verifier. A verification circuit for MA can be described as a quantum circuit consisting only of the (classical) gates from the universal (classical) gateset Toffoli and NOT, operating on input qubits in the state $|0\rangle$ (the NOT gates can then fix them to the right input) and ancillas which are either in the state $|0\rangle$ used as workspace, or in the state $|+\rangle$, used as random bits. At the end of the circuit, the first qubit is measured in the computational basis and the input is accepted iff the output is 1. It is not difficult to check that for such gateset and ancillas, the stoquastic Local Hamiltonian resulting from the circuit-to-Hamiltonian construction of the quantum Cook-Levin theorem (which forces both the correct propagation as well as the correct input state, as well as the output qubit accepting), is a uniform stoquastic Hamiltonian; in particular all the entries are in $\{0, \pm 1, -\frac{1}{2}\}$. We can also assume that each qubit is used in at most d gates, for some $d \geq 3$. This is true because all the computation done by the verifier is classical and therefore the information can be copied to fresh ancilla bits (initialized on $|0\rangle$) with a

CNOT operation. Notice then that each qubit takes place on at most 3 steps: as the target of the CNOT, in some actual computation, and as the source of the next CNOT.

We now explain the other direction, namely Bravyi and Terhal's approach for showing that the stoquastic Hamiltonian problem is in MA. We actually show a simplified version of their result, since we are only interested in *uniform* stoquastic Hamiltonian. Notice that by our above description, this problem is sufficient to achieve MA-hardness.

Following [BT09] We now define the graph on which the random walk will take place; this graph is based on a given uniform stoquastic Hamiltonian.

Definition 26 (Graph from uniform stoquastic Hamiltonian)**.** *Let $H = \frac{1}{m}\sum_i H_i$ be a uniform stoquastic Hamiltonian on n qudits of dimension $|\Sigma|$. We define the undirected graph $G(H) = (|\Sigma|^n, E)$ where $(x, y) \in E$ iff there exists a local term H_i with corresponding groundstate projector P_i such that*

$$\langle x|P_i|y\rangle > 0. \tag{8}$$

From Remark 18 and claim 16, we have that for a fixed i, the neighbor strings form an equivalence class and in each class the strings differ only in the positions where H_i acts non-trivially. We also remark that given some string x, one can compute in polynomial time if x is bad for H, by just inspecting the groundspace of each local term.

The random walk starts from some initial string x_0 sent by the prover. If x_0 is bad for H, then the algorithm rejects. Otherwise, a term H_i is picked uniformly at random and a string x_1 is picked uniformly at random from $|\widetilde{T}_{i,j,z}\rangle$, which is the unique rank-one subset-state from $\widetilde{P}_i$ such that $x_0 \in \widetilde{T}_{i,j,z}$ (see Remark 18). The random walk proceeds by repeating this process with x_1. We describe the random walk proposed by BT (simplified for the uniform case) in Figure 3.

1) Let x_0 be the initial string.
2) Repeat for T steps
   a) If x_t is bad for H, reject
   b) Pick $i \in [m]$ uniformly at random
   c) Pick x_{t+1} uniformly at random from the strings in the unique $\widetilde{T}_{i,j,z}$ that contains x_t
3) Accept

Figure 3: BT Random Walk

We state now the lemmas proved in [BT09].

Lemma 27 (Completeness, adapted from Section 6.1 of [BT09])**.** *If H is frustration-free, then there exists some string x such that there are no bad-strings in the connected component of x.*

The proof goes by showing that if H is frustration-free, then for any string x in some groundstate of H, the uniform superposition of the connected component of x is a groundstate of H. In this case, since all strings in the connected component of x are good (this is by definition of the connected component), the verifier will accept.

Lemma 28 (Soundness, adapted from Section 6.2 of [BT09])**.** *For every polynomial p, there exists some polynomial q such that if H is at least $\frac{1}{p(n)}$-frustrated, then for every string x, for $T = q(n)$, the random walk from Figure 3 rejects with constant probability.*

The intuition of the proof is that since the Hamiltonian is frustrated, one can upper bound the expansion on any set of good-strings by $1 - \frac{1}{p(n)}$, otherwise the Hamiltonian would not be $\frac{1}{p(n)}$-frustrated. In this case, there exists some polynomial q such that a random walk with $q(n)$ steps escape of any set of good strings with high probability.

IV. Uniform Gapped stoquastic Hamiltonians are in NP

Our main technical result in this work is showing that if a stoquastic uniform Hamiltonian is ε-frustrated for some constant ε, then every string x is constantly-close to a bad string.

Lemma 29 (Short path to a bad string)**.** *If the stoquastic uniform k-Local Hamiltonian H is ε-frustrated, then for every string x, there is a bad string y such that the distance between x and y in $G(H)$ is at most $k^{\frac{2kd}{\varepsilon}\log\left(1+\frac{\varepsilon}{4}\right)|\Sigma|}$.*

Using this lemma, we can prove our main result:

Theorem 1 (restated). *For any constants $\varepsilon > 0, k, d$, the problem of deciding whether a uniform d-bounded degree stoquastic k-local Hamiltonian H is frustration-free or ε-frustrated, is in* NP.

Proof: The NP witness for the problem consists in some initial string x that is promised to be in the support of the groundstate of H. The verification proceeds by running over all possible $k^{\frac{2kd}{\varepsilon}\log\left(1+\frac{\varepsilon}{4}\right)|\Sigma|}$-step paths from x. Since for each one of the m terms there are constantly many possible steps, the number of possibilities for one step is polynomial, and so the number constantly-long paths is also polynomial. Therefore, such enumeration can be performed efficiently. For each path, we check if one of the strings it reaches is bad - again this can be done in polynomial time since badness is with respect to the local terms (see Remark 30 for the precision issues). The verifier rejects if any of the paths reached a bad string, otherwise it accepts.

Let x be the string sent by the Prover. If H is frustration-free, then by Lemma 27 all strings in the connected component of x are good. On the other hand, if H is ε-frustrated, then by Lemma 29, there exists a $k^{\frac{2kd}{\varepsilon}\log\left(1+\frac{\varepsilon}{4}\right)|\Sigma|}$-step path from x to some string y that is bad for H, and such path will be found by the brute-force search. ■

Remark 30 (Deciding on badness of a string with respect to a local uniform stoquastic term)**.** *We note that while in the non-uniform case the question of whether a string is bad for a local term or not, may depend on precision issues, this is not a problem when considering uniform stoquastic Hamiltonians. In the uniform case, the set of strings comprising the subset states in the non-negative decomposition of every projector, as in Equation* (1)*, can be calculated exactly given the matrix description of the local Hamiltonian term (even if we need to apply approximations when computing the groundstates). This is because the locality of the Hamiltonian, together with uniformity, imply that if a string is in the support of one of the groundstates, its weight must be $\frac{1}{\sqrt{q}}$ for some positive integer q smaller than some constant.*

The remainder of this section is dedicated to proving Lemma 29. Section IV-A gives the one term expansion argument, Section IV-B provides the proof that a constant number of layers consisting of parallel non-overlapping projections suffices to reach a bad string; and Section IV-C provides the light-cone argument to show that if a bad string is reached within constantly many layers, then in fact we there is a bad string within *constantly* many steps from the initial string. This then allows searching for such a string by brute-force. Finally, Section IV-D just puts all the pieces together to finish the proof of Lemma 29.

A. Expansion

We start by showing that if subset-state $|S\rangle$ does not contain any bad string but a term P is highly frustrated by $|S\rangle$, then the support of $\widetilde{P}|S\rangle$ is larger than that of $|S\rangle$ by a constant factor.

Lemma 31 (One term expansion)**.** *Let $\widetilde{P} = \sum_{j,z} |\widetilde{T}_{j,z}\rangle\langle\widetilde{T}_{j,z}|$ be a uniform stoquastic projector on the set Q of k out of n qudits. Let $S \subseteq \Sigma^n$ be such that $|S\rangle$ does not contain P-bad-strings, and $\|\widetilde{P}|S\rangle\|^2 \leq 1-\delta$. It follows that the size of the support of $\widetilde{P}|S\rangle$ is at least $(1+\frac{\delta}{2})|S|$.*

Proof: Since S does not contain bad strings, we start by noticing that from Lemma 20, S is contained in the support of $\widetilde{P}|S\rangle$, and $\widetilde{P}$ only adds the neighbors of strings in $|S\rangle$.

We have that

$$1-\delta \geq \|\widetilde{P}|S\rangle\|^2 = \langle S|\widetilde{P}|S\rangle = \sum_{j,z} \langle S|\widetilde{T}_{j,z}\rangle\langle\widetilde{T}_{j,z}|S\rangle$$
$$= \sum_{j,z} |\langle S|\widetilde{T}_{j,z}\rangle|^2. \tag{9}$$

Let $\mathbf{T} = \bigcup_{j,z:S\cap\widetilde{T}_{j,z}\neq\emptyset} \widetilde{T}_{j,z}$. It follows that

$$\begin{aligned}
\sum_{j,z} |\langle S|\widetilde{T}_{j,z}\rangle|^2 &= \sum_{j,z} \frac{|\widetilde{T}_{j,z}\cap S|^2}{|\widetilde{T}_{j,z}||S|} \\
&= \sum_{j,z} \frac{(|\widetilde{T}_{j,z}| - |\widetilde{T}_{j,z}\setminus S|)^2}{|\widetilde{T}_{j,z}||S|} \\
&= \sum_{j,z:\widetilde{T}_{j,z}\cap S\neq\emptyset} \frac{|\widetilde{T}_{j,z}|^2 - 2|\widetilde{T}_{j,z}||\widetilde{T}_{j,z}\setminus S| + |\widetilde{T}_{j,z}\setminus S|^2}{|\widetilde{T}_{j,z}||S|} \\
&\geq \frac{|\mathbf{T}| - 2|\mathbf{T}\setminus S|}{|S|} \\
&\geq 1 - \frac{2|\mathbf{T}\setminus S|}{|S|}.
\end{aligned} \tag{10}$$

where in first inequality we remove some non-negative terms and use the fact that $\widetilde{T}_{j,z}$ and $\widetilde{T}_{j',z}$ are disjoint for $j \neq j'$ (Remark 18) and in the second inequality we use the fact that $S \subseteq \mathbf{T}$ since there are no bad strings in S.

By putting together Equations (9) and (10), and noticing that $\mathbf{T} = supp(\widetilde{P}|S\rangle)$ from Lemma 20, we have that

$$|supp(\widetilde{P}|S\rangle)| = |\mathbf{T}| = |S| + |\mathbf{T}\setminus S| \geq |S| + \frac{\delta}{2}|S|$$
$$= \left(1+\frac{\delta}{2}\right)|S|.$$

■

B. Bad string in a constant number of layers

We prove in this section that with a constant number of "layers", it is possible to reach a state with a bad string.

We first want find a linear number of non-overlapping terms that are (roughly) simultaneously frustrated by some subset-state $|S\rangle$. Let us first define what we mean by non-overlapping terms.

Definition 32 (Non-overlapping projectors)**.** *A sequence of local projectors $L = (Q_1, ..., Q_\ell)$ is non-overlapping if for any $i \neq j$, Q_i and Q_j act on disjoint sets of qudits.*

Now, we need to be more careful in order to define the notion of "simultaneous". Recall that if H is ε-frustrated, by Claim 11, there must exist at least $\frac{\varepsilon m}{2}$ terms that are at least $\frac{\varepsilon}{2}$-frustrated. However, as we explained in Example 6, their frustration may be correlated due to entanglement, and when we "correct" the frustration of one term, we could also be correcting the frustration of other terms,

Because of this, we need to choose the sequence of projectors more carefully. We are looking for projectors which are frustrated, even after applying the previous projectors in the sequence.

Definition 33 (Sequentially frustrated terms)**.** *A sequence of projectors $L = (Q_1, ..., Q_\ell)$ is* sequentially *δ-frustrated by some subset-state $|S_0\rangle$, if for all $i = 1, ..., \ell$, $\|\widetilde{Q}_i\widetilde{Q}_{i-1}...\widetilde{Q}_1|S_0\rangle\|^2 \leq 1-\delta$.*

We show now that we can find a linear-size sequence that is non-overlapping and sequentially highly frustrated, in a greedy way. At iteration i, we fix a projector Q_i such that $\widetilde{Q}_i$ is highly frustrated by $\widetilde{\widetilde{Q}_{i-1}...\widetilde{Q}_1|S_0\rangle}$ and Q_i does not overlap with any Q_j for $j < i$. More concretely, we choose Q_i arbitrarily from the intersection of the following sets:

- F_i, the set of terms that are at least $\frac{\varepsilon}{2}$-**F**rustrated by $\widetilde{\widetilde{Q}_{i-1}...\widetilde{Q}_1|S_0\rangle}$
- A_i, the set of **A**vailable terms, i.e. the terms that do not overlap with Q_j, for $j < i$.

We describe such an algorithm in Figure 4 and analyze its correctness in Lemma 34.

Let $H = \frac{1}{m}\sum_{j=1}^{m} H_j$ be a stoquastic k-Local Hamiltonian with frustration at least ε and some subset-state $|S_0\rangle$. For each term H_j, we denote by P_j the projector onto its groundspace.

1) Let $i = 0$, $A_0 = \{P_1, ..., P_m\}$ and $F_0 = \{P_j : \|\widetilde{P}_j|S_0\rangle\|^2 \leq 1 - \frac{\varepsilon}{2}\}$
2) While $A_i \cap F_i \neq \emptyset$
   a) Pick any $P_j \in A_i \cap F_i$ and set $Q_i = P_j$
   b) Let $A_{i+1} = A_i \setminus \{P_j : P_j \text{ overlaps with } Q_i\}$ and $F_{i+1} = \{P_j : \|\widetilde{P}_j\widetilde{\widetilde{Q}_i...\widetilde{Q}_0|S_0\rangle}\|^2 \leq 1 - \frac{\varepsilon}{2}\}$
   c) Let $i = i + 1$
3) Output $L = (Q_0, ..., Q_{i-1})$

Figure 4: Algorithm for finding non-overlapping frustrated terms

Lemma 34 (Linear number of sequential non-overlapping frustrated terms)**.** *Let $H = \frac{1}{m}\sum_{i=1}^{m} H_i$ be a ε-frustrated uniform stoquastic k-Local Hamiltonian, $S_0 \subseteq \Sigma^n$, and $L = (Q_0, ..., Q_{i^*-1})$ be the output of Figure 4. Then $i)$ L is non-overlapping, $ii)$ L is sequentially $\frac{\varepsilon}{2}$-frustrated by $|S_0\rangle$, and $iii)$ $i^* \geq \frac{\varepsilon n}{2kd}$.*

Proof: Properties $i)$ and $ii)$ follow by construction: $Q_i \in A_i$, and thus it does not overlap with Q_j for $j < i$; and $Q_i \in F_i$, therefore $\|\widetilde{Q}_i\widetilde{\widetilde{Q}_{i-1}..\widetilde{Q}_0|S_0\rangle}\|^2 \leq 1 - \frac{\varepsilon}{2}$.

We prove now property $iii)$. Notice that $|A_{i+1}| - |A_i| \leq kd$, since the only difference between these two sets are the overlapping terms of Q_i, Q_i acts on at most k qudits, and there are at most d other terms that overlap with Q_i due to a specific qudit. Therefore, we have that

$$|A_{i^*}| \geq m - i^*kd. \tag{11}$$

Notice that for every i, we have that $|F_i| \geq \frac{\varepsilon m}{2}$ by Claim 11. We also have that if $|A_i| + |F_i| > m$, then $A_i \cap F_i \neq \emptyset$ by the pigeonhole principle. Therefore, $A_{i^*} \cap F_{i^*} = \emptyset$ implies that

$$\left(1 - \frac{\varepsilon}{2}\right) m \geq |A_{i^*}| \tag{12}$$

Putting Equations (11) and (12) together we have

$$m - i^*kd \leq |A_{i^*}| \leq \left(1 - \frac{\varepsilon}{2}\right) m$$

and therefore it follows that

$$i^* \geq \frac{\varepsilon m}{2kd} \geq \frac{\varepsilon n}{2kd},$$

where we use the fact that $m \geq n$. ■

Definition 35 (A layer acting on $|S_0\rangle$)**.** *We denote $L|S_0\rangle$ to be $\widetilde{Q}_{i^*-1}...\widetilde{Q}_0|S_0\rangle$. We notice that this state is equal to $\widetilde{Q}_{\sigma(i^*-1)}...\widetilde{Q}_{\sigma(0)}|S_0\rangle$ for every permutation σ on the indices $0, ..., i^*-1$, since by property i) of Lemma 34 the terms in L are non overlapping, and thus commuting. Thus the resulting state does not depend on the order of application of these projections, hence the notion of applying a layer of non-overlapping terms on a state is well defined.*

By combining Lemmas 31 and 34 we can now prove that if we apply all the projections output by Figure 4, namely all projections in L, the resulting state has exponentially more strings than the original one.

Corollary 36 (Multiple terms expansion)**.** *Let $H = \frac{1}{m}\sum_{i=1}^{m} H_i$ be an ε-frustrated uniform stoquastic k-Local Hamiltonian and $|S\rangle$ be a subset-state such that $|S\rangle$ does not contain any bad string for H. Then there is a sequence L of non-overlapping terms of H such that the number of strings in the support of $L|S\rangle$ is at least $(1 + \frac{\varepsilon}{4})^{\frac{\varepsilon n}{2kd}}$ times the number of strings in $|S\rangle$.*

Proof: Let $L = (Q_1, ..., Q_{i^*})$ be the output of Figure 4. Let us argue now that we can use Lemma 31 for each one of the $i^* \geq \frac{\varepsilon n}{2kd}$ terms in L sequentially, which would imply the statement.

First, let us claim that for all j, $\widetilde{Q}_j...\widetilde{Q}_1|S\rangle$ does not contain $Q_{j'}$-bad strings, for $j' > j$. This follows by induction; for $j = 0$ this is true by assumption. Also from Lemma 34, the terms in L are non-overlapping. Hence, by Claim 16 when we apply $\widetilde{Q}_j$ no $Q_{j'}$-bad strings appear, for all $j' \neq j$. Secondly, directly from Lemma 34 we have that $\|\widetilde{Q}_{j+1}\widetilde{\widetilde{Q}_j...\widetilde{Q}_1|S\rangle}\| \leq 1 - \frac{\varepsilon}{2}$.

This means that the conditions of Lemma 31 are satisfied, and it follows that the number of strings in the support of the state $\widetilde{Q}_{j+1}\widetilde{\widetilde{Q}_j...\widetilde{Q}_1|S\rangle}$ is larger by a factor of $(1 + \frac{\varepsilon}{4})$ than that of the state $\widetilde{\widetilde{Q}_j...\widetilde{Q}_1|S\rangle}$. By Corollary 21 we deduce that the support of $\widetilde{Q}_{j+1}...\widetilde{Q}_1|S\rangle$ is bigger by the same factor, than that of $\widetilde{Q}_j...\widetilde{Q}_1|S\rangle$. ■

We now observe that this expansion is too large to be applied for many layers, since the number of strings reached will just exceed the number of n-bit strings.

Lemma 37 (Bad string in constant number of layers)**.** *Let $\ell^* = \lceil \frac{2kd}{\varepsilon} \log_{(1+\frac{\varepsilon}{4})} |\Sigma| \rceil$. Consider a uniform*[12] *stoquastic k-*

[12]In fact, uniformity is not needed for this claim, but this requires more work.

Local Hamiltonian $H = \frac{1}{m}\sum_{i=1}^{m} H_i$ which is ε-frustrated. Then for every good string x, there exists some $\ell < \ell^*$ and a sequence $L_1, ...L_\ell$, where each L_i consists of a set of non-overlapping local projectors (corresponding to the projections on the local terms of H), such that $L_\ell ... L_1|x\rangle$ contains a bad string.

Proof: Let L_1 be the output of Figure 4 for the initial state $|x\rangle$, and recursively, for $\ell \leq \ell^*$, let L_ℓ be the output of Figure 4 for the state $\widehat{L_\ell ... L_1|x\rangle}$. Let S_ℓ be the set of strings in the support of the state $\widehat{L_\ell ... L_1|x\rangle}$.

Now, if for some $\ell < \ell^*$, S_ℓ contains a bad string for H, we are done. Otherwise, for all $\ell < \ell^*$, the state $\widehat{L_{\ell-1} ... L_1|x\rangle}$ contains no bad string for H.

From Corollary 36, we have that for each $1 \leq \ell \leq \ell^* - 1$,

$$|supp(L_\ell(\widehat{L_{\ell-1}...L_1|x\rangle}))| \geq \left(1 + \frac{\varepsilon}{4}\right)^{\frac{\varepsilon n}{2kd}} |S_{\ell-1}|$$

Moreover, by Corollary 21, we have

$$|S_\ell| = |supp(L_\ell(\widehat{L_{\ell-1}...L_1|x\rangle}))|.$$

Thus, by a trivial induction, we arrive at

$$\begin{aligned} |S_{\ell^*}| &\geq \left(\left(1 + \frac{\varepsilon}{4}\right)^{\frac{\varepsilon n}{2kd}}\right)^{\frac{2kd}{\varepsilon}\log\left(1+\frac{\varepsilon}{4}\right)|\Sigma|} \\ &= \left(\left(1 + \frac{\varepsilon}{4}\right)^{\log\left(1+\frac{\varepsilon}{4}\right)|\Sigma|}\right)^{n} = |\Sigma|^n. \end{aligned}$$

Thus S_{ℓ^*} contains all possible strings. However if H is frustrated, bad strings exist and thus there must be a bad string in $S_{\ell^*} = supp(L_{\ell^*}...L_1|x\rangle)$. ■

C. Finding the bad string

In this section, we prove that if for some constant ℓ we have that $L_\ell...L_1|x\rangle$ contains a bad string, and each L_i consists of non-overlapping terms, then there exists a constant path (namely a sequence of constantly many local steps) from x to a bad string.

We start by showing how to retrieve (possibly polynomial-size) paths from strings in some non-negative state $|\psi\rangle$ to string in some state $L|\psi\rangle$, for a non-overlapping set of projections L.

Lemma 38 (From non-overlapping projections to paths)**.** *Let $|\psi\rangle$ be a non-negative state and L be an arbitrary set of non-overlapping stoquastic projectors. Then for every string y in $L|\psi\rangle$, there exists a string x in $|\psi\rangle$ such that there is a $|L|$-step path between x and y in $G(H)$.*

Proof: Let $L = \{Q_1, ..., Q_g\}$. Since L is a set of non-overlapping stoquastic projectors, if $y \in supp(L|\psi\rangle)$, then there must exist some x in $supp(|\psi\rangle)$ such that

$$0 < \langle x|L|y\rangle.$$

This is because writing $\langle\psi| = \sum_x \psi_x \langle x|$ we have (using $L^\dagger = L$ due to the fact that L consists of non-overlapping projections):

$$0 < \langle\psi|L|y\rangle = \sum_{x \in supp(|\psi\rangle)} \psi_x \langle x|L|y\rangle$$

and the coefficients ψ_x are all non-negative. We have

$$\begin{aligned} 0 < \langle x|L|y\rangle &= \langle x|\widetilde{Q}_1...\widetilde{Q}_g|y\rangle \qquad (13) \\ &= \sum_{w_1,...,w_{g-1}} \langle x|\widetilde{Q}_1|w_1\rangle\langle w_1|\widetilde{Q}_2...|w_{g-1}\rangle\langle w_{g-1}|\widetilde{Q}_g|y\rangle, \end{aligned}$$

where we can write the $\widetilde{Q}_i$ in any order since they are non-overlapping, from Definition 35; the w's above run over all n-dit strings.

Since every $\widetilde{Q}_i$ has only non-negative entries, for every pair of strings z and z', $\langle z|\widetilde{Q}_i|z'\rangle \geq 0$. Then Equation (13) holds iff there exists values $w_1^*, ..., w_{g-1}^*$ such that

$$0 < \langle w_0^*|\widetilde{Q}_1|w_1^*\rangle\langle w_1^*|\widetilde{Q}_2...|w_{g-1}^*\rangle\langle w_{g-1}^*|\widetilde{Q}_g|w_g^*\rangle,$$

where we have set $x := w_0^*$ and $y := w_g^*$. It follows that for every $i \in [g]$, $\langle w_{i-1}^*|\widetilde{Q}_i|w_{i+1}^*\rangle > 0$, and then from Equation (8), w_i^* and w_{i+1}^* are neighbors in $G(H)$. Therefore, there is a $|L|$-step path from x to y. ■

Finally, we show how to find a short path between the initial string a bad string. The intuition of the proof is depicted in Figure 2.

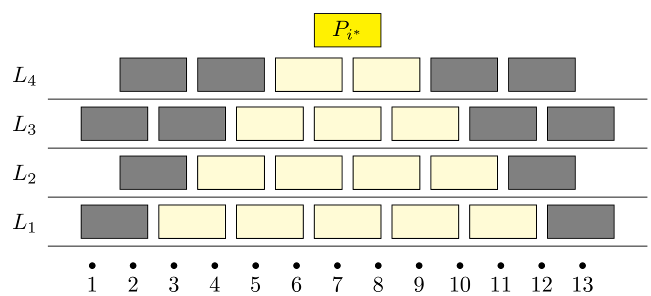


Figure 5: The layers $L_1, ..., L_4$ reach a bad string for some term P_{i^*} that does not necessarily belong to any layer, and then we consider the light-cone of P_{i*} in these layers. Notice that the light-cone is defined only within the layers, each consisting of non-overlapping terms, and *not* by considering *all* projectors which touch any qudit in P_{i^*}, then all projectors which touch those, etc.)

Lemma 39 (The light-cone argument)**.** *For some initial string x, let $L_1, ..., L_\ell$ each be a sequence of non-overlapping projectors such that $L_\ell...L_1|x\rangle$ contains in its support a string w^* which is bad for H. Then, there is a $O(k^\ell)$-steps path from x to a bad string for H (which could be different than w^*).*

Proof: Since w^* is bad for H, we have that for some $i^* \in [m]$, $\langle w^*|\widetilde{P}_{i^*}|w^*\rangle = 0$. Let $L_\ell^\triangle \subseteq L_\ell$ be the projectors of L_ℓ that touch some qudit of P_{i^*}. We define recursively $L^{\triangle}_{j-1}$ as the set of projectors in L_{j-1} that overlap with some projector in $\bigcup_{j'\geq j} L^{\triangle}_{j'}$. These are the *layers* of what we call the *light-cone* of P_{i*} and we depict it in Figure 5. Let us also define the complement of $L^{\triangle}_j$ in L_j to be L'_j.

For convenience, set $L_{\ell+1} = L^{\triangle}_{\ell+1} = \{P_{i^*}\}$. Let D_j be the set of qudits touched by the terms in $\bigcup_{j'\geq j} L^{\triangle}_{j'}$. We prove that $|D_j| \leq k^{\ell-j+2}$. We prove this by a downward induction from $j = \ell+1$ to $j = 1$. The basis step is true since P_{i^*} is k-local, and so $|D_{\ell+1}| = k$. Now, assume this is true for the j'th layer. D_{j-1} is defined by adding to D_j the qudits touched by the next layer of the lightcone, $L^{\triangle}_{j-1}$. By definition of the lightcone, these are all qudits touched by terms in L_{j-1} that overlap D_j. Since these terms are non-overlapping and k-local, this can at most multiply the number of qudits already in D_j by a factor of k. We have that the set D_1 of qudits within the entire lightcone originating from P_{i^*} contains at most $k^{\ell+1}$ qudits.

The terms in L_j commute (as they are non-overlapping), and thus $L_j = L'_j L^{\triangle}_j$. It follows that

$$L_\ell....L_1|x\rangle = L'_\ell L^{\triangle}_\ell L'_{\ell-1} L^{\triangle}_{\ell-1}...L'_1 L^{\triangle}_1|x\rangle \tag{14}$$
$$= L'_\ell L'_{\ell-1}...L'_1 L^{\triangle}_\ell...L^{\triangle}_1|x\rangle, \tag{15}$$

where in the second equality, we use in fact an iterative argument (that is common in light-cone reasoning, see, e.g., [AALV09], [BGK18]): we notice that the projectors in L'_j commute with the projectors in $L^{\triangle}_{j'}$ for all $j' \geq j$, and thus they can be commuted one by one across the lightcone operators, to the left. The fact that they commute follows by definition: level j of the lightcone, $L^{\triangle}_j$, contains *all* terms in L_j that overlap with any term $L^{\triangle}_{j'}$, for $j' > j'$; thus the remaining terms in L_j, namely L'_j, do not overlap the upper layers of the lightcone, and thus commute with them.

From Equation 14 we deduce that we can first apply on x all terms in the lightcone, and delay all terms outside of the lightcone to later. From this, we can show that $L^{\triangle}_\ell...L^{\triangle}_1|x\rangle$ also contains a bad string for $\widetilde{P}_{i*}$, and this will complete the proof. To do this, let Q be the set of positions where the term P_{i^*} acts non-trivially. We claim that the application of the terms outside of the lightcone, which do not touch Q, couldn't have added a string which is bad with respect to P_{i^*}, unless such a string was there before. This can be deduced from Claim 16, which when applied iteratively gives that

$$\{y_Q : y \in supp(L'_\ell L'_{\ell-1}...L'_1 L^{\triangle}_\ell...L^{\triangle}_1|x\rangle)\} \tag{16}$$
$$\subseteq \{y_Q : y \in supp(L^{\triangle}_\ell...L^{\triangle}_1|x\rangle)\}. \tag{17}$$

Since $L'_\ell L'_{\ell-1}...L'_1 L^{\triangle}_\ell...L^{\triangle}_1|x\rangle$ contains a bad string w^* for $\widetilde{P}_{i*}$, from Equation (16) we have that $L^{\triangle}_\ell...L^{\triangle}_1|x\rangle$ contains a string w' such that $w^*_Q = w'_Q$, thus w' is also bad for $\widetilde{P}_{i^*}$.

Finally, we can use Lemma 38 together with a (highly wasteful) bound on $|L^{\triangle}_j| \leq |D_j| \leq k^{\ell-j+2}$ recursively for $L^{\triangle}_\ell...L^{\triangle}_1|x\rangle$: for any string y in $L^{\triangle}_j...L^{\triangle}_1|x\rangle$, there exists a string y' in $L^{\triangle}_{j-1}...L^{\triangle}_1|x\rangle$, such that there is a $|L^{\triangle}_j| \leq k^{\ell-j+2}$-step path from y' to y in $G(H)$. Hence, there is a path of size $\sum_{j'=1}^{\ell} k^{\ell-j'+2} = \sum_{j'=2}^{\ell+1} k^{j'} = O(k^\ell)$ from x to w' in $G(H)$. ■

D. Proof of Lemma 29

We can finally prove Lemma 29 by composing Lemmas 37 and 39.

Lemma 29 (restated). *If the stoquastic uniform k-Local Hamiltonian H is ε-frustrated, then for every string x, there is a bad string y such that the distance between x and y in $G(H)$ is at most $k^{\frac{2kd}{\varepsilon}\log_{\left(1+\frac{\varepsilon}{4}\right)}|\Sigma|}$.*

Proof: Let $\ell^* = \frac{2kd}{\varepsilon}\log_{\left(1+\frac{\varepsilon}{4}\right)}|\Sigma|$. By Lemma 37, there exists a sequence of sets $L_1, ...L_\ell$, for some $\ell \leq \ell^*$, such that each L_i consists of non-overlapping projectors, and $L_\ell...L_1|x\rangle$ contains a bad string. We finish the proof by using $L_1, ..., L_\ell$ in Lemma 39 for some $\ell \leq \frac{2kd}{\varepsilon}\log_{\left(1+\frac{\varepsilon}{4}\right)}|\Sigma|$. ■

V. Negligible vs. high frustration

In this section we show that we can replace the frustration-free property by allowing the stoquastic Hamiltonian to be negligibly-frustrated in yes-instances.

Our proof follows by showing that in this case, there exists some string x in the groundstate of the Hamiltonian that cannot reach a bad string in a constant number of steps, and therefore the same verification algorithm of Theorem 1 would still work.

In order to prove that, we first show that if the groundenergy of the Hamiltonian is very small, then there is a also a subset state $|S\rangle$ with very low energy in respect to H. Then, in order to achieve our goal, we need bound two quantities. First, we need to show that the number of bad strings in $|S\rangle$ is small. Secondly, we consider the good strings that are connected to the strings in $|S\rangle$ but are not in this state. We notice that these strings might be harmful, since they could be connected to bad strings and then strings in the "border" of S could be close to bad strings. Finally, we can show that there must be some string in S that is far from all of these strings if the frustration of the Hamiltonian is sufficiently small. This follows directly from the upper bound on the number of such strings and from the structure of $G(H)$.

Let us start then by bounding the weight of bad strings in the groundstate of slightly frustrated Hamiltonians.

Lemma 40. *Let H be a stoquastic Hamiltonian and $|\psi\rangle = \sum_x \alpha_x|x\rangle$ be a positive state. Let B be the set of bad strings for H. Then $\sum_{x\in B} \alpha_x^2 \leq m\langle\psi|H|\psi\rangle$*

Proof: Let B_i be the set of bad strings for some term H_i, $|\psi\rangle = \alpha|\psi_{G_i}\rangle + \beta|\psi_{B_i}\rangle$ where $\alpha = \sqrt{\sum_{x \notin B_i} \alpha_x^2}$, $\beta = \sqrt{\sum_{x \in B_i} \alpha_x^2}$, $|\psi_{G_i}\rangle = \frac{1}{\alpha}\sum_{x \notin B_i} \alpha_x|x\rangle$ and $|\psi_{B_i}\rangle = \frac{1}{\beta}\sum_{x \in B_i} \alpha_x|x\rangle$.

We have that

$$\begin{aligned}\langle\psi|H_i|\psi\rangle &= (\alpha\langle\psi_{G_i}| + \beta\langle\psi_{B_i}|)(I - P_i)(\alpha|\psi_{G_i}\rangle + \beta|\psi_{B_i}\rangle)\\ &= 1 - \alpha^2\langle\psi_{G_i}|P_i|\psi_{G_i}\rangle\\ &\geq 1 - \alpha^2\langle\psi_{G_i}|\psi_{G_i}\rangle\\ &= \beta^2.\end{aligned}$$

where in the first equality we assume that H_i is a projector (see Remark 25) and in the second inequality we use the fact that that $|\psi_{G_i}\rangle$ is positive and P_i is stoquastic.

We have then that

$$\langle\psi|H|\psi\rangle \geq \frac{1}{m}\sum_i \sum_{x \in B_i} \alpha_x^2 \geq \frac{1}{m}\sum_{x \in B} \alpha_x^2$$

where we use in the second inequality that each $x \in B$ appears in at least one of the B_i. ■

Lemma 41 (Projection on boundary lower bounds energy)**.** *Let H be a stoquastic Hamiltonian and $|\psi\rangle = \sum_x \alpha_x|x\rangle$ be a positive state. Let N be the set of $\{x \in supp(|\psi\rangle : \exists i \in [m], y \notin supp(|\psi\rangle), \langle x|P_i|y\rangle > 0\}$. Then $\langle\psi|H|\psi\rangle \geq \frac{1}{|\Sigma|^k m}\sum_{x \in N} \alpha_x^2$.*

Proof: Let $N_i = \{x \in supp(|\psi\rangle : \exists y \notin supp(|\psi\rangle), \langle x|P_i|y\rangle > 0\}$, $|\psi\rangle = \alpha|\psi_{M_i}\rangle + \beta|\psi_{N_i}\rangle$ where $\alpha = \sqrt{\sum_{x \notin N_i} \alpha_x^2}$, $\beta = \sqrt{\sum_{x \in N_i} \alpha_x^2}$, $|\psi_{M_i}\rangle = \frac{1}{\alpha}\sum_{x \notin N_i} \alpha_x|x\rangle$ and $|\psi_{N_i}\rangle = \frac{1}{\beta}\sum_{x \in N_i} \alpha_x|x\rangle$.

We have that

$$\begin{aligned}&\langle\psi|H_i|\psi\rangle = (\alpha\langle\psi_{M_i}| + \beta\langle\psi_{N_i}|)(I - P_i)(\alpha|\psi_{M_i}\rangle + \beta|\psi_{N_i}\rangle)\\ &= 1 - \alpha^2\langle\psi_{M_i}|P_i|\psi_{M_i}\rangle - \beta^2\langle\psi_{N_i}|P_i|\psi_{N_i}\rangle \qquad (18)\end{aligned}$$

where in the first equality we use the fact that if $\langle x|P_i|y\rangle > 0$ and $x \in N_i$, then there exists a $z \notin supp(|\psi\rangle)$ with $\langle x|P_i|z\rangle > 0$, which implies that $\langle y|P_i|z\rangle > 0$ and therefore $y \in N_i$.

We will show below that

$$\langle\psi_{N_i}|P_i|\psi_{N_i}\rangle \leq \frac{|\Sigma|^k - 1}{|\Sigma|^k}. \qquad (19)$$

Using this and Equation (18) we have

$$\begin{aligned}&\langle\psi|H_i|\psi\rangle = 1 - \alpha^2\langle\psi_{M_i}|P_i|\psi_{M_i}\rangle - \beta^2\langle\psi_{N_i}|P_i|\psi_{N_i}\rangle\\ &\geq 1 - \alpha^2 - \frac{|\Sigma|^k - 1}{|\Sigma|^k}\beta^2 = \frac{1}{|\Sigma|^k}\beta^2 = \frac{1}{|\Sigma|^k}\sum_{x \in N_i} \alpha_x^2.\end{aligned}$$

This implies the desired claim since

$$\langle\psi|H|\psi\rangle \geq \frac{1}{m|\Sigma|^k}\sum_i \sum_{x \in N_i} \alpha_x^2 \geq \frac{1}{m|\Sigma|^k}\sum_{x \in N} \alpha_x^2$$

where we used the fact that $N \subseteq \cup_i N_i$.

It remains to prove Equation (19). Consider some $T \subseteq \{0,1\}^n$ and positive $|\phi\rangle = \sum_{x \in T} \beta_x|x\rangle$, it follows that

$$\begin{aligned}\langle\phi|T\rangle\langle T|\phi\rangle &= \sum_{x,y \in T} \frac{\beta_x\beta_y}{|T|} && (20)\\ &\leq \sqrt{\frac{|T \cap supp(|\phi\rangle)|^2}{|T|^2}\sum_{x,y \in T} \beta_x^2\beta_y^2} && (21)\\ &= \frac{|supp(|\phi\rangle)|}{|T|}, && (22)\end{aligned}$$

where we use Cauchy-Schwartz inequality.

Given the decomposition $P_i = \sum_{j,z} |T_{i,j,z}\rangle\langle T_{i,j,z}|$, we can define $|\psi_{N_i}\rangle = \gamma_{j,z}\sum_{j,z} |\psi_{N_i,j,z}\rangle$, with $supp(|\psi_{N_i,j,z}\rangle) \subseteq T_{i,j,z}$ and then

$$\begin{aligned}&\langle\psi_{N_i}|P_i|\psi_{N_i}\rangle = \sum_{j,z} \gamma_{j,z}^2\langle\psi_{N_i,j,z}|T_{i,j,z}\rangle\langle T_{i,j,z}|\psi_{N_i,j,z}\rangle\\ &\leq \sum_{j,z} \gamma_{j,z}^2 \frac{|supp(|\psi_{N_i,j,z}\rangle)|}{|T_{i,j,z}|} \leq \frac{|\Sigma|^k - 1}{|\Sigma|^k},\end{aligned}$$

where the first inequality follows from Equation (20) and in the second inequality we use that $|T_{i,j,z}| \leq |\Sigma|^k$ and $|supp(|\psi_{N_i,j,z}\rangle)| \leq |T_{i,j,z}| - 1$ (by definition of N_i). ■

We now prove that there exists a low-energy state with enough structure.

Lemma 42 (Existence of a "nice" low energy state)**.** *Let H be a stoquastic Hamiltonian with non-zero ground-energy at most $\frac{1}{f(n)}$. Then there exists some non-negative state $|\psi\rangle$ that does not contain bad strings for H, contains only strings with amplitude at least $\delta = \frac{1}{\sqrt{g(n)|S|}}$ and has frustration at most $\frac{1}{f(n)(1 - \frac{m}{f(n)} - \frac{1}{g(n)})}$.*

Proof: Let $|\phi\rangle = \sum_x \alpha_x|x\rangle$ be a non-negative ground-state of H, $S = supp(|\phi\rangle)$, B be the set of bad strings, L be the set of strings with amplitude smaller than δ, $T = B \cup L$ and $G = S \setminus T$. We define $|\psi\rangle = \frac{1}{\sqrt{\sum_{x \in G} \alpha_x^2}}\sum_{x \in G} \alpha_x|x\rangle$. We have that:

$$
\begin{aligned}
&\langle\phi|H|\phi\rangle \\
&= \sum_{x,y\in G}\alpha_x\alpha_y\langle x|H|y\rangle + \sum_{x\in T,y\in S}\alpha_x\alpha_y\langle x|H|y\rangle \\
&\quad + \sum_{x\in G,y\in T}\alpha_x\alpha_y\langle x|H|y\rangle \\
&\geq \sum_{x,y\in G}\alpha_x\alpha_y\langle x|H|y\rangle \\
&= \frac{\sum_{x\in G}\alpha_x^2}{\sum_{x\in G}\alpha_x^2}\sum_{x,y\in G}\alpha_x\alpha_y\langle x|H|y\rangle \\
&= \left(1-\sum_{x\in B}\alpha_x^2-\sum_{x\in T\setminus B}\alpha_x^2\right)\langle\psi|H|\psi\rangle \\
&\geq \left(1-\frac{m}{f(n)}-|S|\delta^2\right)\langle\psi|H|\psi\rangle,
\end{aligned}
$$

where the last inequality comes from Lemma 40 and the fact that $\sum_{x\in T\setminus B}\alpha_x^2 \leq \sum_{x\in T\setminus B}\delta^2 \leq |S|\delta^2$. ■

We finally show that the structured low-energy state implies that there is a string that is "far from the frustration".

Lemma 43 (Nice low energy state has small boundary)**.** *Let H be a stoquastic Hamiltonian such that there exists some non-negative state $|\psi\rangle$ such that $S := supp(|\psi\rangle)$ does not contain bad strings for H, and all amplitudes of $|\psi\rangle$ are at least $\delta = \frac{1}{\sqrt{g(n)|S|}}$. Moreover, it has frustration strictly smaller than $\frac{1}{h(n)}$. Then it follows that $|\{x\in S : \exists y\notin S, i\in[m]$ s.t. $\langle x|P_i|y\rangle > 0\}| < \frac{m|\Sigma|^k g(n)}{h(n)}|S|$.*

Proof: Let $N = \{x \in S : \exists y \notin S, i \in [m]$ s.t. $\langle x|P_i|y\rangle > 0\}$. Let us assume towards contradiction that $\frac{|N|}{|S|} > \frac{m|\Sigma|^k g(n)}{h(n)}$. We have from Lemma 41

$$
\begin{aligned}
\frac{1}{h(n)} &> \langle\psi|H|\psi\rangle \geq \frac{1}{m|\Sigma|^k}\sum_{x\in N}\alpha_x^2 \geq \frac{|N|\delta^2}{m|\Sigma|^k} \\
&> \frac{m\delta^2 g(n)|S|}{mh(n)} = \frac{1}{h(n)},
\end{aligned}
$$

where the next to last inequality follows since the set of strings in N are all in S the support of $|\psi\rangle$ and thus their amplitudes by assumption are at least δ. This is a contradiction.

■

Lemma 44 (A state of small boundary has a "protected" string)**.** *Fix some parameter $t\in\mathbb{N}^*$. Let H be a uniform stoquastic k-local Hamiltonian and $|\psi\rangle$ be a state such that its support S contains no bad strings for H, and $|\{x \in S : \exists y \notin S, i \in [m]$ s.t. $\langle x|P_i|y\rangle > 0\}| < \frac{1}{(|\Sigma|^k m)^{t+1}}|S|$. Then there exists some string $x\in S$ such that all strings within distance t from x in the graph $G(H)$ are H-good strings.*

Proof: Let $M = \{y \notin S : \exists x \in S, i \in [m]$ s.t. $\langle x|P_i|y\rangle > 0\}$ be the set of strings not in S that are connected with strings in S. Notice that $|M| \leq |\Sigma|^k m|\{x \in S : \exists y \notin S, i \in [m]$ s.t. $\langle x|P_i|y\rangle > 0\}| < \frac{1}{(|\Sigma|^k m)^t}|S|$.

Also, for any string x, we have that there exists at most $(|\Sigma|^k m)^t$ strings within distance t steps from x. Therefore the number of strings that could reach some string in M in t steps is strictly smaller than $\frac{|S|(|\Sigma|^k m)^t}{(|\Sigma|^k m)^t} = |S|$.

This means that there must exist a string in S from which a t step walk cannot lead to a string outside S. ■

Finally, we prove the main technical result of this section.

Theorem 45. *For any constant $\varepsilon > 0$, let $t = k^{\frac{2kd}{\varepsilon}\log\left(1+\frac{\varepsilon}{4}\right)^{|\Sigma|}}$ (with constants defined as in Definition Definition 24) and c be any function such that $c(n) < \frac{1}{(|\Sigma|^k m)^{2(t+2)}}$, the problem of deciding whether a uniform stoquastic Hamiltonian H has ground-energy at most $c(n)$ or at least ε is in* NP*.*

Proof: The NP witness for the problem consists of some initial string x that is promised to be in the support of some groundstate $|S\rangle$ of H and such that all strings within distance t of x are good and lie in S . The verification proceeds by running over all possible t-step paths from x. As argued in Theorem 1 such verification can be performed in polynomial time.

$$
\begin{aligned}
&|\{x\in S : \exists y\notin S, i\in[m] \text{ s.t. } \langle x|P_i|y\rangle > 0\}| \\
&< \frac{m|\Sigma|^k\sqrt{f(n)}}{(1-\frac{m}{f(n)}-\frac{1}{g(n)})f(n)}|S| \\
&\leq \frac{m|\Sigma|^k\sqrt{c(n)}}{1-2\sqrt{c(n)}m}|S| \\
&\leq \frac{(|\Sigma|^k m)^{(t+2)}}{(|\Sigma|^k m)^{t+2}\left((|\Sigma|^k m)^{t+2}-2m\right)}|S| \\
&\leq \frac{|S|}{(|\Sigma|^k m)^{t+1}}.
\end{aligned}
$$

Therefore, by Lemma 44, we have that for yes-instances, there must exist a string x that does not reach bad strings in t steps, and therefore the prover can send this string to verifier.

On the other hand, if H is ε-frustrated, then by Lemma 29, there must exist a t-step path from x to some string y that is bad for H, and such path will be found by the brute-force search. ■

This proves Theorem 2.

References

[AALV09] Dorit Aharonov, Itai Arad, Zeph Landau, and Umesh V. Vazirani. The detectability lemma and quantum gap amplification. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009*, pages 417–426, 2009.

[AALV11] Dorit Aharonov, Itai Arad, Zeph Landau, and Umesh V. Vazirani. The 1d area law and the complexity of quantum states: A combinatorial approach. In *IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011*, pages 324–333, 2011.

[AAV13] Dorit Aharonov, Itai Arad, and Thomas Vidick. Guest column: the quantum PCP conjecture. *SIGACT News*, 44(2):47–79, 2013.

[AAV16] Anurag Anshu, Itai Arad, and Thomas Vidick. Simple proof of the detectability lemma and spectral gap amplification. *Phys. Rev. B*, 93:205142, May 2016.

[AE15] Dorit Aharonov and Lior Eldar. The Commuting Local Hamiltonian Problem on Locally Expanding Graphs is Approximable in NP. *Quantum Information Processing*, 14(1):83–101, 2015.

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof Verification and the Hardness of Approximation Problems. *Journal of the ACM*, 45(3):501–555, 1998.

[ALV12] Itai Arad, Zeph Landau, and Umesh Vazirani. Improved one-dimensional area law for frustration-free systems. *Phys. Rev. B*, 85:195145, May 2012.

[ALVV17] Itai Arad, Zeph Landau, Umesh Vazirani, and Thomas Vidick. Rigorous RG Algorithms and Area Laws for Low Energy Eigenstates in 1D. *Communications in Mathematical Physics*, 356(1):65–105, Nov 2017.

[AN02] Dorit Aharonov and Tomer Naveh. Quantum NP - A Survey, 2002.

[Ara11] Itai Arad. A note about a partial no-go theorem for quantum PCP. *Quantum Information & Computation*, 11(11-12):1019–1027, 2011.

[AS98] Sanjeev Arora and S Safra. Probabilistic Checking of Proofs: A New Characterization of NP. *Journal of the ACM*, 45(1):70–122, 1998.

[BBT06] Sergey Bravyi, Arvid J. Bessen, and Barbara M. Terhal. Merlin-arthur games and stoquastic complexity. *arXiv preprint arXiv:0611021*, 2006.

[BDOT08] Sergey Bravyi, David P. Divincenzo, Roberto Oliveira, and Barbara M. Terhal. The complexity of stoquastic local hamiltonian problems. *Quantum Info. Comput.*, 8(5):361–385, May 2008.

[BFNW93] László Babai, Lance Fortnow, Noam Nisan, and Avi Wigderson. Bpp has subexponential time simulations unless exptime has publishable proofs. *Comput. Complex.*, 3(4):307–318, October 1993.

[BGK18] Sergey Bravyi, David Gosset, and Robert König. Quantum advantage with shallow circuits. *Science*, 362(6412):308–311, 2018.

[BH13] Fernando G. S. L. Brandão and Aram W. Harrow. Product-state approximations to quantum ground states. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing (STOC '13)*, pages 871–880, 2013.

[BH17] Sergey Bravyi and Matthew Hastings. On complexity of the quantum ising model. *Communications in Mathematical Physics*, 349(1):1–45, Jan 2017.

[BISW17] Dan Boneh, Yuval Ishai, Amit Sahai, and David J. Wu. Lattice-based snargs and their application to more efficient obfuscation. In *Advances in Cryptology - EUROCRYPT 2017 - 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques*, 2017.

[Bra14] Sergey Bravyi. Monte Carlo simulation of stoquastic Hamiltonians. *arXiv preprint arXiv:1402.2295*, 2014.

[BSCG+14] Eli Ben-Sasson, Alessandro Chiesa, Christina Garman, Matthew Green, Ian Miers, Eran Tromer, and Madars Virza. Zerocash: Decentralized anonymous payments from bitcoin. In *IEEE Symposium on Security and Privacy*, pages 459–474. IEEE Computer Society, 2014.

[BT09] Sergey Bravyi and Barbara M. Terhal. Complexity of stoquastic frustration-free hamiltonians. *SIAM J. Comput.*, 39(4):1462–1485, 2009.

[CCD15] Cristian S. Calude, Elena Calude, and Michael J. Dinneen. Guest column: Adiabatic quantum computing challenges. *SIGACT News*, 46(1):40–61, March 2015.

[Coo71] Stephen A Cook. The complexity of theorem proving procedures. In *Proceedings of the Third Annual ACM Symposium*, pages 151–158. ACM, 1971.

[CSV18] Matthew Coudron, Jalex Stark, and Thomas Vidick. Trading locality for time: certifiable randomness from low-depth circuits. *arXiv preprint arXiv:1810.04233*, 2018.

[Din07] Irit Dinur. The PCP Theorem by Gap Amplification. *Journal of the ACM*, 54(3), 2007.

[Dru11] Andrew Drucker. A pcp characterization of am. In *Proceedings of the 38th International Colloquim Conference on Automata, Languages and Programming - Volume Part I*, ICALP'11, 2011.

[EH17] Lior Eldar and Aram W Harrow. Local hamiltonians whose ground states are hard to approximate. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017*, pages 427–438, 2017.

[FGGS00] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Michael Sipser. Quantum computation by adiabatic evolution. *arXiv preprint arXiv:0001106*, 2000.

[FH14] Michael H. Freedman and Matthew B. Hastings. Quantum systems on non-k-hyperfinite complexes: a generalization of classical statistical mechanics on expander graphs. *Quantum Information & Computation*, 14(1-2):144–180, 2014.

[Gal18] Françis Le Gall. Average-Case Quantum Advantage with Shallow Circuits. *arXiv preprint arXiv:1810.12792*, 2018.

[GH16] David Gosset and Yichen Huang. Correlation length versus gap in frustration-free systems. *Phys. Rev. Lett.*, 116:097202, Mar 2016.

[GHLS15] Sevag Gharibian, Yichen Huang, Zeph Landau, and Seung Woo Shin. Quantum hamiltonian complexity. *Foundations and Trends in Theoretical Computer Science*, 10(3):159–282, 2015.

[GKR15] Shafi Goldwasser, Yael Tauman Kalai, and Guy N. Rothblum. Delegating computation: Interactive proofs for muggles. *J. ACM*, 62(4):27:1–27:64, 2015.

[GLSW15] Sevag Gharibian, Zeph Landau, Seung Woo Shin, and Guoming Wang. Tensor network non-zero testing. *Quantum Information & Computation*, 15(9&10):885–899, 2015.

[GZ11] Oded Goldreich and David Zuckerman. Another proof that BPP $\subseteq$ PH (and more). In *Studies in Complexity and Cryptography*, pages 40–53. Springer-Verlag, 2011.

[Hås01] Johan Håstad. Some optimal inapproximability results. *J. ACM*, 48(4), July 2001.

[Has13a] Matthew B. Hastings. Obstructions to classically simulating the quantum adiabatic algorithm. *Quantum Information & Computation*, 13(11-12):1038–1076, 2013.

[Has13b] Matthew B. Hastings. Trivial low energy states for commuting hamiltonians, and the quantum pcp conjecture. *Quantum Info. Comput.*, 13(5-6):393–429, May 2013.

[HILL99] Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM J. Comput.*, 28(4), March 1999.

[IKW02] Russell Impagliazzo, Valentine Kabanets, and Avi Wigderson. In search of an easy witness: exponential time vs. probabilistic polynomial time. *J. Comput. Syst. Sci.*, 65(4):672–694, 2002.

[IW97] Russell Impagliazzo and Avi Wigderson. P = bpp if e requires exponential circuits: Derandomizing the xor lemma. In *Proceedings of the Twenty-ninth Annual ACM Symposium on Theory of Computing*, STOC '97, pages 220–229. ACM, 1997.

[KI04] Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. *Computational Complexity*, 13(1):1–46, Dec 2004.

[KSV02] Alexei Kitaev, A Shen, and M N Vyalyi. *Classical and quantum computation*. Graduate studies in mathematics. American mathematical society, Providence (R.I.), 2002.

[KT18] Joel Klassen and Barbara M. Terhal. Two-local qubit Hamiltonians: when are they stoquastic? *arXiv preprint arXiv:1806.05405*, 2018.

[Lev73] Leonid A Levin. Universal sequential search problems. *Problems of Information Transmission*, 9(3):265–266, 1973.

[MLH18] Milad Marvian, Daniel A Lidar, and Itay Hen. On the Computational Complexity of Curing the Sign Problem. *arXiv preprint arXiv:1802.03408*, 2018.

[NC00] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000.

[NV18] Anand Natarajan and Thomas Vidick. Low-degree testing for quantum states, and a quantum entangled games PCP for QMA. In *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018*, pages 731–742, 2018.

[NVY18] Chinmay Nirkhe, Umesh Vazirani, and Henry Yuen. Approximate low-weight check codes and circuit lower bounds for noisy ground states. In *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018*, pages 91:1–91:11, 2018.

[NW94] Noam Nisan and Avi Wigderson. Hardness vs randomness. *J. Comput. Syst. Sci.*, 49(2):149–167, 1994.

[Osb12] Tobias J Osborne. Hamiltonian complexity. *Reports on Progress in Physics*, 75(2):022001, 2012.

[STV01] Madhu Sudan, Luca Trevisan, and Salil Vadhan. Pseudorandom generators without the xor lemma. *Journal of Computer and System Sciences*, 62(2):236 – 266, 2001.

[SW] Peter Shor and Ryan Williams. Mathoverflow: Complete problems for randomized complexity classes. https://mathoverflow.net/questions/34469/complete-problems-for-randomized-complexity-classes. Accessed: 2019-01-14.

[TD04] Barbara M. Terhal and David P. DiVincenzo. Adaptive quantum computation, constant depth quantum circuits and arthur-merlin games. *Quantum Information & Computation*, 4:134–145, 2004.

[Wat00] John Watrous. Succinct quantum proofs for properties of finite groups. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*, pages 537–546, 2000.

[ZF87] Stathis Zachos and Martin Furer. Probabilistic quantifiers vs. distrustful adversaries. In *Seventh Conference on Foundations of Software Technology and Theoretical Computer Science*, pages 443–455, 1987.

Appendix

We start with the class co-RP, which is the class of problems that can be solved by randomized algorithms with perfect completeness.

Definition 46 (co-RP)**.** *A problem $A = (A_{\text{yes}}, A_{\text{no}})$ is in co-RP if and only if there exists a probabilistic polynomial-time algorithm R, where R takes as input a string $x \in \Sigma^*$ and decides on acceptance or rejection of x such that:*
Completeness. *If $x \in A_{\text{yes}}$, then R accepts x with probability* 1*.*
Soundness. *If $x \in A_{\text{no}}$, then R accepts x with probability at most $\frac{1}{2}$.*

The only difference of Definitions 8 and 46 is that in MA there is some witness y, unknown by the verification algorithm. We can see it as if the witness in co-RP is trivially the empty string. In this case, it is not surprising that we can define a version of Definition 24, where we fix some string in the groundstate.

Definition 47 (pinned uniform stoquastic frustration-free k-Local Hamiltonian problem)**.** *The* pinned uniform stoquastic frustration-free k-Local Hamiltonian *problem, where $k \in \mathbb{N}^*$ is called the locality and $\varepsilon : \mathbb{N} \to [0,1]$ is a non-decreasing function, is the following promise problem. Let n be the number of qudits of a quantum system. The input is a set of $m(n)$ uniform stoquastic Hamiltonians $H_1, \ldots, H_{m(n)}$ where m is a polynomial, $\forall i \in m(n) : 0 \leq H_i \leq I$ and each H_i acts on k qudits out of the n qudit system. We also assume that there are at most d terms that act non-trivially on each qudit, for some constant d. For $H = \frac{1}{m(n)} \sum_{j=1}^{m(n)} H_j$, one of the following two conditions holds, and the problem is to decide which one:*
Yes. *There exists a n-qudit quantum state state $|\psi\rangle$ such that $\langle\psi|0\rangle > 0$ and $\langle\psi|H|\psi\rangle = 0$.*
No. *For all n-qudit quantum states $|\psi\rangle$, it holds that $\langle\psi|H|\psi\rangle \geq \varepsilon(n)$.*

Theorem 48. *The pinned uniform stoquastic frustration-free k-Local Hamiltonian problem is in co-RP for every constant k and $\varepsilon(n) = \frac{1}{p(n)}$ where p is some polynomial. Also, for some $p'(n) = O(n^2)$, the pinned uniform stoquastic frustration-free* 6*-Local Hamiltonian problem is co-RP complete.*

Proof: The inclusion in co-RP comes directly from the random-walk proposed by Bravyi and Terhal [BT09], starting from the fixed all-zeros string.

For the hardness part, as in [BT09], we can analyze the reduction of the quantum Cook-Levin theorem for a co-RP circuit. In the yes-instance, we have that the all-zeros string must be in the groundstate, since it is a valid initial configuration, whereas for no-instances, all states have inverse-polynomial frustration. ∎

Theorem 49. *The ε-gapped pinned k-local d-bounded-degree uniform stoquastic frustration-free Hamiltonian problem is in P for every constants k, d and $\varepsilon > 0$.*

Proof: The proof follows exactly the lines of the proof of Theorem 1 except instead of starting the walk from the string given by the prover, the algorithm treats the all 0 string as the witness, and hence no prover is needed. We omit the details. ∎

Conjecture 50 (Witness preserving Stoquastic PCP conjecture)**.** *(Informal) There exist constants $\varepsilon > 0$, $k', d' > 0$ and an efficient gap amplification procedure that reduces the problem deciding if a uniform stoquastic d-degree k Local-Hamiltonian is frustration-free or at least inverse polynomially frustrated, to the problem of deciding if a uniform stoquastic d'-degree k' Local-Hamiltonian is frustrarion-free or at least ε frustrated. In the yes-cases, it also provides an efficient mapping from any witness of the original problem to a witness of the target gapped problem.*

Corollary 51. *Conjecture 50 implies* $\mathsf{RP} = \mathcal{P}$.

Proof: (sketch) We use the stoquastic gap amplification procedure to generate an instance of the gapped version of the Hamiltonian, and consider x, the image of the all-0 string under the witness preserving reduction. Then we apply the same algorithm which the verifier of Theorem 1 does, but instead of applying it on the witness, apply it on the string y. ∎

We do not know how to extend these results to BPP without making highly artificial assumptions, since for the proof of Theorem 2 to go through, we need the pinned string to not only be in the support of the groundstate but also be far from bad strings. Hence making interesting implications for the derandomization of BPP remains for future work.

In this section, we give a simple proof that deciding if a commuting Stoquastic Hamiltonian is frustration-free is in NP. This simplifies the proof of this statement originally given in [GLSW15].

Theorem 5 (restated). *The problem of deciding if a commuting stoquastic Hamiltonian H is frustration-free is in* NP.

We notice that given that we are not assuming anything on the gap, one needs to be somewhat careful with the assumptions on how the input is given; we assume here that the local terms of the commuting stoquastic Hamiltonian H are provided by giving the matrix elements, each with $poly(m)$ bits, and the terms mutually commute exactly.

Proof: Let $H_1, ..., H_m$ be the terms in the Hamiltonian, and let $\widetilde{P}_1, ..., \widetilde{P}_m$ be the corresponding projectors onto their groundspaces. We show that H is frustration-free iff there exists a string x that is good for H.

The first direction ($\Rightarrow$) is trivial: any string in the 0-energy groundstate is good for H.

We prove now the converse ($\Leftarrow$). Let x be a string that is good for H. Let $|\phi\rangle = \widetilde{P}_1\widetilde{P}_2...\widetilde{P}_m|x\rangle$. Since $\widetilde{P}_1, ..., \widetilde{P}_m$ commute, we have that

$$|\phi\rangle = \widetilde{P}_i|\phi\rangle,$$

which means that either $|\phi\rangle = 0$ or it is a $+1$-eigenstate of every $\widetilde{P}_i$, and therefore it has energy 0 with respect to H. We show now that $|\phi\rangle \neq 0$. The string x is in the support of some groundstate of every term H_i, therefore for every state $|\alpha\rangle = \sum_y \alpha_y|y\rangle$ with $\alpha_y \in \mathbb{R}^+$ and $\alpha_x > 0$, $\widetilde{P}_i|\alpha\rangle = \sum_y \alpha'_y|y\rangle$, with $\alpha'_y \in \mathbb{R}^+$ and $\alpha'_x > 0$. Therefore we have that $\widetilde{P}_1\widetilde{P}_2...\widetilde{P}_m|x\rangle \neq 0$.

Finally, to show that the problem is in NP: the proof is supposed to be a string with largest amplitude in some groundstate of the Hamiltonian. The verification algorithm checks if this string is indeed good for all local terms, as follows. First, for each k-local term H_i, the verifier computes P_i to within constant approximation. More precisely, it computes each matrix elements of P_i to within $\frac{1}{4|\Sigma|^k}$. This can be done efficiently since we are working with local terms where each matrix element of H_i is specified by polynomially many bits. Let the approximated projector be P'_i. The verifier then checks if x is bad for P_i: to do this, it first restricts x to Q, the set of k qudits on which H_i acts, and then checks if $\langle x_Q|P'_i|x_Q\rangle \leq \frac{1}{2|\Sigma|^k}$. If yes, then it rejects. If the verifier does not reject for any of the i's, then it accepts.

To finish the proof, we need to show that the above procedure cannot lead to an error in the verifier's decision of whether x is bad or good for H. We start by arguing that in the frustration-free case, if x is the string with largest amplitude in some groundstate $|\psi\rangle$, then it passes the test. This follows since we can write $\widetilde{P}_i$ using its non-negative decomposition (see Remark 17): $\widetilde{P}_i = \sum_j |\widetilde{\phi}_{i,j,z}\rangle\langle\widetilde{\phi}_{i,j,z}|$. We then write $|\psi\rangle = \sum_{j,z} \alpha_{i,j,z}|\widetilde{\phi}_{i,j,z}\rangle$; there is a unique $|\widetilde{\phi}_{i,j^*,z^*}\rangle$ that contains x. Then x must be the string with largest amplitude in $|\widetilde{\phi}_{i,j^*,z^*}\rangle$ (since a string y in $|\widetilde{\phi}_{i,j^*,z^*}\rangle$ has amplitude $\alpha_{i,j,z}\langle y|\widetilde{\phi}_{i,j,z}\rangle$ in $|\psi\rangle$ and x maximizes this value). Since $|\widetilde{\phi}_{i,j^*,z^*}\rangle$ contains at most $|\Sigma|^k$ strings, the amplitude of x in it is at least $|\Sigma|^{-\frac{k}{2}}$. Hence $\langle x|\widetilde{P}_i|x\rangle = \langle x|\widetilde{\phi}_{i,j^*,z^*}\rangle\langle\widetilde{\phi}_{i,j^*,z^*}|x\rangle \geq \frac{1}{|\Sigma|^k}$. By our described procedure, in this case the verifier accepts. For soundness, if some string x is bad for H_i, then for some i $\langle x|P_i|x\rangle = 0$ and thus by our bound on the error due to approximation of P_i by P'_i, we know $\langle x|P'_i|x\rangle \leq \frac{1}{4|\Sigma|^k}$, and the verifier will reject. ■

As we discussed in Section I-E, we can also describe the uniform stoquastic frustration-free local Hamiltonian problem in a more classical language. In this section, we present such approach.

We start by defining what a constraint is in our setting.

Definition 52 (k-SetConstraint)**.** *A k-SetConstraint is a tuple (T, B), where $B \subseteq [n]$ with $|B| = k$ and $T = \{T_1, ..., T_l\}$ where $T_j \subseteq \Sigma^k$ and $T_j \cap T_{j'} = \emptyset$ for $j \neq j'$.*

We define now what it means for a constraint to be satisfied by a set of strings. In order to do it, we need to present some notation. For $x \in \Sigma^n$ $S \subseteq \Sigma^n$, $B \subset [n]$, and $y \in \Sigma^{n-|B|}$ we denote $x|_B$ as the substring of x on the positions contained in B, $S_B = \{x|_B : x \in S\}$ and $S_{B,y} = \{x_B : x \in S \text{ and } x_{\overline{B}} = y\}$.

Definition 53 (Satisfiability of set-constraints)**.** *We define the UNSAT-value of a k-SetConstraint (T, B) with respect to a subset of strings $S \subseteq \Sigma^n$, to be*

$$UNSAT((T,B),S) = 1 - \sum_{\substack{T_j \in T \\ y \in \Sigma^{n-k}}} \frac{|T_j \cap S_{B,y}|^2}{|T_j||S|}. \quad (23)$$

A sequence of m k-SetConstraints $I = ((T_1, B_1), ..., (T_m, B_m))$ has ε-UNSAT value with respect to $S \subseteq \Sigma^n$, if

$$UNSAT(I, S) \stackrel{\text{def}}{=} \frac{1}{m}\sum_{i=1}^{m} UNSAT((T_i, B_i), S) \geq \varepsilon.$$

We say that I is satisfiable if there exists an $S \subseteq \Sigma^n$ such that $UNSAT(I, S) = 0$. We say that I is ε-frustrated if for all $S \subseteq \Sigma^n$ $UNSAT(I, S) \geq \varepsilon$.

It is not difficult to see that the UNSAT value by Equation 23 is a value between 0 and 1. As in standard CSP, an instance to the problem consists of a sequence of constraints and we ask if they can all be satisfied simultaneously.

Definition 54 (k-local Set Constraint Satisfaction problem (SetCSP))**.** *Fix $k > 0, d > 0$. For a function $\varepsilon(n)$, an instance to the ε, d k-*local Set Constraint Satisfaction problem *is a sequence of $m(n)$ k-SetConstraints $I = \{(T_1, B_1), ..., (T_{m(n)}, B_{m(n)})\}$ on Σ^n, where m is some polynomial in n. We also assume that every $j \in [n]$ appears on at most d B_i's. Under the promise that one of the following two holds, decide whether:*
***Yes.** There exists some $S \subseteq \Sigma^n$ such that $UNSAT(I, S) = 0$*
***No.** For all $S \subseteq \Sigma^n$, $UNSAT(I, S) \geq epsilon$.*

We show now the equivalence between the uniform stoquastic Local Hamiltonians problem and the SetCSP problem.

Lemma 55. *There is a polynomial time computable transformation which given an instance to the (ε, k, d) uniform gapped stoquastic local Hamiltonian problem, outputs an instance to the*

ε, k, d-SetCSP, such that the frustration is preserved. Similarly, there is also a frustration preserving reduction between the latter problem to the former.

Proof: Let $H = \frac{1}{m}\sum_i H_i$ be an instance of the uniform stoquastic Hamiltonian problem.

Let $P_i = \sum_{ij} |T_{i,j}\rangle\langle T_{i,j}|$ be the projector onto the groundspace of H_i for disjoint $T_{i,j} \subseteq \Sigma^k$. For each local term H_i, the SetCSP instance has a constraint (T_i, B_i) where $T_i = \{T_{i,j}\}$ and B_i is the set of qubits on which H_i acts non-trivially. Denote by $\sigma(H)$ the SetCSP instance $I = ((T_1, B_1), ..., (T_m, B_m))$.

It can easily be checked that for $S \subseteq \Sigma^n$, we have that

$$\begin{aligned}
\langle S|H|S\rangle &= 1 - \frac{1}{m}\langle S|\widetilde{P}_i|S\rangle \\
&= 1 - \frac{1}{m}\sum_i \sum_{j,z} \langle S|\widetilde{T}_{i,j,z}\rangle\langle \widetilde{T}_{i,j,z}|S\rangle \\
&= UNSAT(\sigma(H), S),
\end{aligned}$$

where we used the definition of $\widetilde{P}_i$ and $\widetilde{T}_{i,j,z}$ as in Remark 17.

To move in the other direction, simply define $H_i = I - P_i$, which can easily be seen to be uniform stoquastic. ■

The previous lemma shows that SetCSP and uniform stoquastic Hamiltonians problem are equivalent. Thus, we have by Bravyi and Terhal [BT09]:

Corollary 56. *There is an inverse polynomial function ε such that the ε, k, d SetCSP is MA-complete.*

And by Theorem 1, we have the following..

Corollary 57. *For any constant $\varepsilon > 0$, the k-local Set Constraint Satisfaction problem is in* NP.