Abstract

Analog quantum simulation – simulation of one Hamiltonian by another – is one of the major goals in the noisy intermediate-scale quantum computation (NISQ) era, and has many applications in quantum complexity. We initiate the rigorous study of the physical resources required for such simulations, where we focus on the task of Hamiltonian sparsification. The goal is to find a simulating Hamiltonian \( \tilde{H} \) whose underlying interaction graph has bounded degree (this is called degree-reduction) or much fewer edges than that of the original Hamiltonian \( H \) (this is called dilution). We set this study in a relaxed framework for analog simulations that we call gap-simulation, where \( \tilde{H} \) is only required to simulate the groundstate(s) and spectral gap of \( H \) instead of its full spectrum, and we believe it is of independent interest.

Our main result is a proof that in stark contrast to the classical setting, general degree-reduction is impossible in the quantum world, even under our relaxed notion of gap-simulation. The impossibility proof relies on devising counterexample Hamiltonians and applying a strengthened variant of Hastings-Koma decay of correlations theorem. We also show a complementary result where degree-reduction is possible when the strength of interactions is allowed to grow polynomially. Furthermore, we prove the impossibility of the related sparsification task of generic Hamiltonian dilution, under a computational hardness assumption. We also clarify the (currently weak) implications of our results to the question of quantum PCP. Our work provides basic answers to many of the “first questions” one would ask about Hamiltonian sparsification and gap-simulation; we hope this serves as a good starting point for future research of these topics.

2012 ACM Subject Classification Theory of computation → Quantum computation theory

Keywords and phrases quantum simulation, quantum Hamiltonian complexity, sparsification, quantum PCP

Digital Object Identifier 10.4230/LIPIcs.ITCS.2019.2


Acknowledgements We are grateful to Oded Kenneth for suggesting the construction of \( H_B \), and for fruitful discussions; to Itai Arad for insightful remarks about the connection to quantum PCP; to Eli Ben-Sasson for discussions about PCP; to Ashley Montanaro and Toby Cubitt for clarifications about Hamiltonian simulation.

1 Supported by ERC grant 280157, ISF grant No. 039-9494.
2 Supported by ERC grant 280157, MIT MISTI.
1 Introduction

A major theme in quantum computation is the idea of analog quantum simulation. This is the task of simulating one Hamiltonian $H$ by another Hamiltonian $\tilde{H}$, which might be more readily or easily implemented. In fact, this goal was identified as a main motivation for realizing quantum computers as early as 1981 by Feynman [37], with the idea that such analog quantum simulations can shed light on properties of physical quantum systems that are hard to simulate efficiently on classical computers. Cirac and Zoller [26] further developed this idea, and explained that such simulators are likely to be achievable well before fully fault-tolerant quantum computation [4, 48, 46] becomes practical, which might take a long time. While fault-tolerant quantum computers when realized can be used to apply digital quantum simulations [52] (where a quantum circuit simulates the time-evolution $e^{-iHt}$ under a local Hamiltonian $H$), analog quantum simulations are more accessible for near-term experiments because they do not require full-fledged quantum computer. Many groups are designing implementations in a variety of experimental platforms [59, 19, 18, 12, 42, 38], and we have recently witnessed some experiments in intermediate-sized quantum systems in regimes where classical simulations are difficult [17, 64, 44]. It has been argued that analog quantum simulation constitutes one of the more interesting challenges in the noisy intermediate-scale quantum computing (NISQ) era [55].

Beyond their natural physical applications, analog simulations of Hamiltonians are also very important for quantum complexity theory. For example, in the theory of quantum NP, one is often interested in reducing problems defined by one class of Hamiltonians to another (e.g. [47, 8, 22, 27, 45]). These reductions are often derived using perturbative gadgets (e.g. [45, 54, 21, 43, 24, 25]). Moreover, analog Hamiltonian simulators might also be useful for the design of Hamiltonian-based quantum algorithms, such as adiabatic algorithm [35] and QAOA [36]. In those settings, it is often desirable to tailor the Hamiltonians being used, while maintaining the properties essential for the algorithm.

In this paper, we initiate the rigorous study of the minimal resources required to simulate a given target Hamiltonian, and ask: When can we simulate a Hamiltonian $H$ by another $\tilde{H}$ that is simpler, easier, or more economic to implement? Of course, this vaguely stated question can take several forms if made more rigorous; here we focus on a natural goal which we loosely call Hamiltonian sparsification, which aims to simplify the interaction graph of the Hamiltonian. For a 2-local $n$-qubit Hamiltonian $H$, the interaction graph has $n$ vertices, with edges connecting any pairs of qubits that participate in a local term in $H$. For a $k$-local Hamiltonian, we consider an interaction hypergraph, where each term acting on $k$ qubits is represented by a hyperedge. A generic $k$-local Hamiltonian has $\Theta(n^k)$ edges, and $\Theta(n^{k-1})$ degree per vertex. Roughly speaking, Hamiltonian sparsification aims to simulate a Hamiltonian using another whose interaction graph is more “economic”, e.g., it has less edges (we refer to this as dilution) or its degree is bounded (we refer to this as degree-reduction).

Hamiltonian sparsification has several important motivations. First, it can help physicists tackle the immense hurdles they face when trying to realize Hamiltonians in the lab. In addition, in many settings in quantum complexity, such as in the study of quantum PCP [3] and recent approaches to the area law question [7], simulating a Hamiltonian by one with constant degree or fewer edges is a potentially important primitive. Indeed, sparsification is used ubiquitously in classical computer science, in a variety of applications; we mention two important ones. The first, graph sparsification (and more generally, matrix sparsification) is a central tool in matrix algorithms [61, 62, 60, 15]. Famously, Ref. [14] proved that any graph can be replaced by another which is sparse (namely, has small degree on average),
such that their Laplacian matrices are spectrally similar. Another common application of sparsification in classical computer science is degree-reduction (DR), used in the study of local Constraint Satisfaction Problems (CSPs) and PCPs [32]. We believe that this natural and basic notion deserves to be studied in the quantum context as well, and might have interesting applications beyond those we can foresee today.

1.1 Gap-Simulations: Simulating only the low-lying part of the spectrum

Before embarking on the study of Hamiltonian sparsification, we first need an appropriate definition of analog simulation. The study of analog Hamiltonian simulation was set on rigorous footing in a recent work by Cubitt, Montanaro, and Piddock [27]; their definition refines that of Bravyi and Hastings [22], and it roughly goes as follows: A given Hamiltonian $H$ is simulated by “encoding” its full spectrum into the low-lying part of the spectrum of $\tilde{H}$ acting on a larger Hilbert space. When $\tilde{H}$ is implemented, then the low-lying part of its spectrum can be used to derive properties and information about the original Hamiltonian $H$. For obvious reasons, we will refer to this definition as full-spectrum simulation. In Ref. [27], the notion of universal Hamiltonians was defined and studied: these are families of Hamiltonians which are capable of performing full-spectrum simulations of any given Hamiltonian, albeit generally with exponential overhead in energy.

While this strong notion of full-spectrum simulation is necessary for simulating all dynamical properties of a system, it is common in physics that one is only interested in the properties of the low-energy states and, particularly, the groundstates. In addition, the spectral gap separating the groundstates from the rest of the spectrum is an intimately related quantity that is usually physically important. For example, the groundstates encode exotic quantum behaviors such as topological order, and the spectral gap protects them [63, 50]. Also, they are used together to define quantum phases of matter and characterize phase transitions [56, 28]. Moreover, both are the main objects of interest in quantum computational complexity: In quantum adiabatic algorithms [35], the goal is to prepare a groundstate of a problem Hamiltonian, and the spectral gap governs the efficiency of the process. In quantum NP theory [8], only the groundstate(s) of the Hamiltonian matters as it is the witness for the problem. The spectral gap also determines the temperature of a thermal equilibrium (Gibbs) state that can be used to approximate the groundstate. Hence, we believe that a natural and minimal notion of analog Hamiltonian simulation, which is still meaningful for many physical contexts, should require that both the space of groundstates and the spectral gap above it be preserved.

Therefore, we suggest to consider sparsification, or more generally Hamiltonian simulation, using this minimal notion, which we formally define as gap-simulation. To the best of our knowledge, despite its naturalness, this relaxed notion of Hamiltonian simulation was not formally defined and rigorously studied previously in the quantum complexity literature.

A Hamiltonian $\tilde{H}$ is said to gap-simulate $H$ if it mimics the groundstate(s) and the spectral gap of $H$; no constraints are imposed on the excited part of the spectrum. To provide a sensible definition requires some care, since in the quantum world we can allow inaccuracies and entanglement to an ancilla. We provide two versions of the definition: In the weaker one (Definition 3), the groundspace is mimicked faithfully, i.e. the support of any groundstate of $\tilde{H}$, when reduced to the original Hilbert space, is close to the groundspace of $H$. However, this definition does not require quantum coherence within the groundspace be maintained. Such coherence is guaranteed by our stronger definition (Definition 2), in which all superpositions within the groundspace are simulated. The extent to which the gap-simulation is incoherent (or unfaithful) is quantified via a small constant $\epsilon$ (or $\delta$). It
seems that the coherent notion is the “correct” one for most quantum applications, though the weaker one might also be useful in certain contexts (see Sec. 5). We mention that here, like in Ref. [22, 27], we allow encoding of the qubits. Typically, we consider “localized” encodings, though this is not explicitly required in the definition.

To set the stage, some basic results about the framework are provided: We show in Lemma 4 that for Hamiltonians with unique groundstates, our two definitions of gap-simulations coincide. Moreover, both coherent and incoherent gap-simulation definitions are shown to be stable under compositions (Lemma 6).

How does the gap-simulation framework compare with the stricter definitions of full-spectrum simulations developed in Ref. [22, 27]? In Sec. 2.1.1, this connection is discussed formally; roughly, our definition is indeed a relaxed version of full-spectrum simulations whose spectral error is smaller than the spectral gap, up to varying restrictions about encoding (Lemma 8). We choose to work here with the more relaxed definition of gap-simulation, since impossibility results for a weaker definition are of course stronger. More generally, it seems that this framework is an important and relevant one to consider in physics and quantum complexity contexts. Being less demanding, gap-simulation is likely achievable in certain cases where full-spectrum simulation is difficult or even impossible.

1.2 Main Results

Equipped with this framework of Hamiltonian sparsification via gap-simulations, we ask: When are sparsifications possible in the quantum world? It is conceivable that, like in the analogous classical settings mentioned above [32, 15], they ought to be always possible. The main result of in this paper (Theorem 11) shows that in stark contrast to the classical setting, both coherent and incoherent degree-reductions are not generally possible in the quantum world, even if one uses the relaxed notion of gap-simulation. This impossibility phenomenon is due to the existence of many-body entanglement in some quantum groundstates; we show, using a strengthened version of Hastings-Koma decay of correlation theorem [41], that there exist local Hamiltonians whose groundstates cannot be coherently mapped into the groundspace of a gapped Hamiltonian with constant degree and bounded interaction strength. Though one might suspect this is a consequence of degeneracy in the groundspace, we show that it holds even in the case of a unique groundstate. We believe this is a surprising and curious phenomenon, which demonstrates the richness in the subject, and highlights the difference in the resources required for classical versus quantum Hamiltonian simulation.

This impossibility result on degree-reduction is essentially tight, as we provide a complementary result (Theorem 13) based on a somewhat sophisticated application of the circuit-to-Hamiltonian construction, stating that degree-reduction becomes possible for any local Hamiltonian with non-negligible spectral gap, when polynomially large overhead in interaction strength is allowed.

We also study a related important sparsification task: dilution. While our main result of Theorem 11 is an information-theoretic result that rules out existence of degree-reducers regardless of computational power, we are unable to provide such a strong result in the case of dilution. Information-theoretically, we can only rule out dilution with perfect (or inverse-polynomially close to perfect) coherence (Theorem 15). Nevertheless, we are able to prove impossibility of any efficient classical algorithm to find diluters with constant unfaithfulness, for generic (even classical) Hamiltonians (Theorem 16). The proof of this theorem (relying on Ref. [29]) works under the assumption that $\text{coNP} \not\subseteq \text{NP/poly}$ (alternatively, the polynomial hierarchy does not collapse to its third level). Although generic constructive dilution is ruled out by our Theorem 16, the question of existence of diluters for general Hamiltonian, with bounded or large interaction strengths, remains open.
The paper provides quite a few further results complementing the above-mentioned main contributions. These build on ideas in classical PCP reductions and perturbative gadgets. In addition, the ideas studied here are strongly reminiscent of questions arising in the context of the major open problem of quantum PCP [3]. We clarify this connection and provide some preliminary results along those lines.

We believe that the study of the resources required for Hamiltonian simulations in various contexts, as well as the framework of gap-simulation, are of potential deep interest for physics as well as quantum complexity. The questions raised touch upon a variety of important challenges, from quantum simulations, to algorithm design, to quantum PCP and NP reductions, to physical implementations on near-term quantum processors, and more. Their study might also shed light on questions in many-body physics, by developing tools to construct “equivalent” Hamiltonians, from the point of view of the study of groundstate physics. The discussion in Sec. 5 includes a more detailed list of some of the open questions and implications.

1.3 Overview

In Sec. 2, we set the stage by providing definitions of gap-simulation and sparsification, and proving basic facts about this new framework. In Sec. 3, we state our results formally. Subsequently, Sec. 4 provides elaborated and intuitive proof sketches, and Sec. 5 provides further discussion. All technical proofs are deferred to the extended version [1].

2 Definition of the Framework: Setting the Stage

2.1 Gap-Simulations of Hamiltonians

We restrict our attention to k-local Hamiltonians \( H = \sum_{i=1}^{M} H_i \) acting on \( n \) qudits (with internal states \( \{|0\rangle, \ldots, |d-1\rangle\} \)), where each term \( H_i \) acts nontrivially on a (distinct) subset of at most \( k \) qudits. We denote \( \lambda_j(X) \) as the \( j \)-th lowest eigenvalue of \( X \), and \( \|X\| \) as the spectral norm of \( X \). In addition, for any Hermitian projector \( P \), we denote \( P^\perp \equiv 1 - P \), and \( |\psi\rangle \in P \iff P|\psi\rangle = |\psi\rangle \).

Definition 1 (groundspace, energy spread and gap). Consider a family of \( n \)-qudit Hamiltonians \( \{H(n)\}_{n=1}^{\infty} \). Let \( E_n^\lambda = \lambda_1(H(n)) \), and suppose \( P(n) \) is a Hermitian projector onto the subspace of eigenstates of \( H(n) \) with energy \( \leq E_n^\gamma + w_n \gamma_n \), for some \( \gamma_n > 0 \), \( 0 \leq w_n < 1 \), such that

\[
[H(n), P(n)] = 0, \quad \|P(n)(H(n) - E_n^\gamma)P(n)\| \leq w_n \gamma_n, \\
\text{and} \quad \lambda_j(P^\perp(n)(H(n) - E_n^\gamma)P^\perp(n) + \gamma_n P(n)) \geq \gamma_n \quad \forall j.
\]

We call the subspace onto which \( P(n) \) projects a quasi-groundspace, \( w_n \) its energy spread, and \( \gamma_n \) its quasi-spectral gap. When we choose \( w_n = 0 \) and \( \gamma_n = \min_j \{\lambda_j(H(n)) - E_n^\gamma : \lambda_j(H(n)) \neq E_n^\gamma\} \), we call the quasi-groundspace that \( P(n) \) projects onto simply the groundspace of \( H(n) \), and \( \gamma_n \) the spectral gap of \( H(n) \). Let \( w_\infty = \sup_n w_n \) and \( \gamma_\infty = \inf_n \gamma_n \). If \( \gamma_\infty > 0 \) and \( w_\infty < 1 \), we say \( \{H(n)\}_{n=1}^{\infty} \) is spectrally gapped.

Below, we omit the subscript \( n \) in \( H(n) \), referring to a single \( H \), with the usual implicit understanding that we consider families of Hamiltonians, where \( n \rightarrow \infty \). All explicit Hamiltonians we gap-simulate here have \( w_n = 0 \), but Definition 1 is more general and allows \( w_n > 0 \), so that it can capture situations with slightly perturbed groundstates (or when simulating a larger low-energy part of the spectrum). We now define Hamiltonian gap-simulation, visualized in Fig. 1:
**Definition 2** (gap-simulation of Hamiltonian). Let $H$ and $\tilde{H}$ be two Hamiltonians, defined on Hilbert spaces $\mathcal{H}$ and $\tilde{\mathcal{H}}$ respectively, where $\tilde{\mathcal{H}}$. Let $V : \mathcal{H} \otimes \mathcal{H}_{\text{anc}} \to \tilde{\mathcal{H}}$ be an isometry ($V^\dagger V = I$), where $\mathcal{H}_{\text{anc}}$ is some ancilla Hilbert space. Denote $E^g \equiv \lambda_1(\tilde{H})$. Per Definition 1, let $P$ be a quasi-groundspace projector of $H$, $\gamma$ its quasi-spectral gap. We say that $\tilde{H}$ gap-simulates $(H, P)$ with encoding $V$, incoherence $\epsilon \geq 0$ and energy spread $0 \leq \tilde{\epsilon} < 1$ if the following conditions are both satisfied:

1. There exists a Hermitian projector $\tilde{P}$ onto a subspace of eigenstates of $\tilde{H}$ such that

$$[\tilde{H}, \tilde{P}] = 0, \quad \|\tilde{P}(\tilde{H} - E^g)\tilde{P}\| \leq \tilde{\epsilon}, \quad \text{and} \quad \lambda_j(\tilde{P}^\perp(\tilde{H} - E^g)\tilde{P}^\perp + \gamma \tilde{P}) \geq \gamma \quad \forall j. \quad (2)$$

I.e., $\tilde{P}$ projects onto a quasi-groundspace of $\tilde{H}$ with quasi-spectral gap not smaller than that of $P$ in $H$, and energy spread $\tilde{\epsilon}$.

2. There exists a Hermitian projector $P_{\text{anc}}$ acting on $\mathcal{H}_{\text{anc}}$, so that

$$[\text{bounded incoherence}] \quad \|\tilde{P} - V(P \otimes P_{\text{anc}})V^\dagger\| \leq \epsilon. \quad (3)$$

When $P$ projects onto the groundspace of $H$, rather than a quasi-groundspace, we usually do not mention $P$ explicitly, and simply say that $\tilde{H}$ gap-simulates $H$.

Requiring $\epsilon$ from Eq. (3) to be small ensures that coherence in the groundspace is maintained by the gap-simulation. This is illustrated by considering a Hamiltonian $H$ with two orthogonal groundstates $|g_1\rangle$ and $|g_2\rangle$. The condition of Eq. (3) essentially says that for any coherent superposition $|g\rangle = c_1|g_1\rangle + c_2|g_2\rangle$, and a state $|a\rangle \in P_{\text{anc}}$ on the ancilla, there exists a groundstate of $\tilde{H}$ that looks like $|\tilde{g}\rangle = V(|g\rangle \otimes |a\rangle) + O(\epsilon)$. Moreover, any groundstate of $\tilde{H}$ could be written in this form. This would preserve the expectation value of any observable in the groundspace, i.e., $\langle g|\hat{\sigma}|g\rangle \approx \langle \tilde{g}|V\hat{\sigma}V^\dagger|\tilde{g}\rangle + O(\epsilon)$. In contrast, one can consider an alternative situation where the groundspace of a simulator $\tilde{H}$ is spanned by states of the form $|\tilde{g}_i\rangle \approx V(|g_i\rangle \otimes |a_i\rangle)$, where $\langle a_i|a_i\rangle \ll 1$. This situation remains interesting, as finding a groundstate $|\tilde{g}_i\rangle$ of $\tilde{H}$ reveals information about a groundstate of $H$ by decoding:

$$|g_i\rangle\langle g_i| \approx \text{Tr}_{\text{anc}}(V^\dagger|\tilde{g}_i\rangle\langle \tilde{g}_i|V).$$

However, the coherence among groundstates is destroyed, since $|g\rangle = |g_1\rangle + |g_2\rangle$ is mapped to $|\tilde{g}\rangle \approx V(|g_1\rangle \otimes |a_1\rangle + |g_2\rangle \otimes |a_2\rangle)$, and observables such as $\hat{\sigma} = |g_1\rangle\langle g_2|$ are not preserved: $\langle g|\hat{\sigma}|g\rangle \not\approx \langle \tilde{g}|V\hat{\sigma}V^\dagger|\tilde{g}\rangle$.

Although coherence seems important to maintain in most quantum settings, we also define incoherent gap-simulation, which may be relevant in some situations (see discussion in Sec. 5).
Definition 3 (incoherent gap-simulation). Consider two Hamiltonians $H$ and $\tilde{H}$, $P$ a quasi-groundspace projector of $H$, and $V$ some isometry in the same setting as in Definition 2. We say that $\tilde{H}$ incoherently gap-simulates $(H, P)$ with encoding $V$, unfaithfulness $\delta \geq 0$ and energy spread $0 \leq \tilde{w} < 1$ if it satisfies the first condition of Definition 2 and, instead of the second condition of Eq. (3),

\[
\|\tilde{P} - V(P \otimes I_{\text{anc}})V^\dagger\tilde{P}\| \leq \delta.
\]  

(4)

Again, when $P$ projects onto the groundspace of $H$, we simply say $\tilde{H}$ incoherently gap-simulates $H$.

Small unfaithfulness essentially means that the support of the vectors in the groundspace of $\tilde{H}$ is roughly contained in a subspace spanned by encoding the groundspace of $H$ with some ancilla.

It is easy to see that small incoherence implies small unfaithfulness, namely $\delta \leq 2\epsilon$ [1]. However, small unfaithfulness is a strictly weaker condition than small incoherence; we will see an example in Prop. 17. Importantly, when $H$ has a unique groundstate, the two notions are equivalent up to a constant (the proof of this fact is surprisingly not trivial; see Sec. 4.2):

Lemma 4 (incoherent gap-simulation is coherent when groundstate is unique). Suppose $H$ has a unique groundstate, with groundspace projector $P = |g\rangle\langle g|$. If $\tilde{H}$ incoherently gap-simulates $H$ with unfaithfulness $\delta < 1$, then it also gap-simulates $H$ with incoherence $\epsilon \leq \sqrt{2\delta/\sqrt{1 - \delta^2}}$.

While we do not explicitly restrict the form of encoding $V$ in the above definitions, we need to specify them for the impossibility proofs, where we will consider localized encoding:

Definition 5 (localized encoding). Consider a (possibly incoherent) gap-simulation of $H$ by $\tilde{H}$ encoded by an isometry $V : H \otimes H_{\text{anc}} \rightarrow \tilde{H}$. Let $H \otimes H_{\text{anc}} = [\otimes_{i=1}^n (H_i \otimes A_i)] \otimes A_E$, where $H_i$ is the $i$-th qudit in $H$, $A_i$ is its associated ancilla subsystem, and $A_E$ contains the remaining ancillas; also let $\tilde{H} = \otimes_{i=1}^m \tilde{H}_i$, $m \geq n$. We say $V$ is a localized encoding if either of the following is true:

1. $V = [\otimes_{i=1}^n V_i] \otimes V_E$, where $V_i : H_i \otimes A_i \rightarrow \tilde{H}_i$ is an isometry, and $\tilde{H}_i$ consists of $O(1)$ qudits in $H$ for $i = 1, \ldots, n$. Also, $V_E : A_E \rightarrow \otimes_{i=n+1}^m \tilde{H}_i$ is an isometry.

2. $V$ is a constant-depth quantum circuit: $V = \prod_{a=1}^D U_a$, where $D = O(1)$, $U_a = \otimes_{\mu} U_{a,\mu}$, and $U_{a,\mu}$ is a unitary operator acting on $O(1)$ number of qudits.

We say $V$ is an $\eta$-localized encoding if there is a localized encoding $V_L$ such that $\|V - V_L\| \leq \eta$.

In addition to constant-depth quantum circuits, any quantum error-correcting code where each logical qudit is encoded as $O(1)$ qudits is also a localized encoding. Note it is easy to see that if a gap-simulation has $\eta$-localized encoding $V$ and incoherence $\epsilon$ (or unfaithfulness $\delta$), it is also a gap-simulation with localized encoding $V_L$ and incoherence $\epsilon' \leq \epsilon + 2\eta$ (or unfaithfulness $\delta' \leq \delta + 2\eta$). Hence, we usually restrict our attention to fully localized encoding in the remainder of the paper.

It is also fairly straightforward to show that compositions of gap-simulation behave intuitively:

Lemma 6 (Composition). Suppose $H_1$ (incoherently) gap-simulates $(H_0, P_0)$ with encoding $V_1$, incoherence $\epsilon_1$ (or unfaithfulness $\delta_1$), energy spread $\tilde{w}_1$, and a corresponding quasi-groundspace projector $P_1$. Also suppose $H_2$ (incoherently) gap-simulates $(H_1, P_1)$ with encoding $V_2$, incoherence $\epsilon_2$ (or unfaithfulness $\delta_2$), and energy spread $\tilde{w}_2$. Then $H_2$ (incoherently) gap-simulates $(H_0, P_0)$ with encoding $V_2(V_1 \otimes I_{\text{anc},1})$, incoherence $\leq \epsilon_2 + \epsilon_1$ (or unfaithfulness $\leq 2\delta_2 + \delta_1$), and energy spread $\tilde{w}_2$. 
2.1.1 Comparison of gap-simulation to full-spectrum simulation

To formally compare gap-simulations to full-spectrum simulations of Hamiltonians, we use the following well-motivated definition of full-spectrum simulation developed by Ref. [27]:

Definition 7 (Full-spectrum simulation, adapted from Definition 1 of [27]). A Hamiltonian \( \tilde{H} \) full-spectrum-simulates a Hamiltonian \( H \) to precision \((\eta, \xi)\) below an energy cut-off \( \Delta \) if there exists an encoding \( \mathcal{E}(H) = V(H \otimes P + \tilde{H} \otimes Q)V^\dagger \), where \( V \) is an isometry, \( P \) and \( Q \) are mutually orthogonal projectors, such that

1. There exists an encoding \( \tilde{\mathcal{E}}(H) = \tilde{V}(H \otimes P + \tilde{H} \otimes Q)\tilde{V}^\dagger \) such that
   \[ \|\tilde{V} - V\| \leq \eta \] and
   \[ \tilde{\mathcal{E}}(1) = P_{\leq \Delta} \eta \]
2. \( \|H_{\leq \Delta} - \tilde{\mathcal{E}}(H)\| \leq \xi \).

Here, \( P_{\leq \Delta} \eta \) projects onto eigenstates of \( \tilde{H} \) with eigenvalues \( \leq \Delta \), and \( H_{\leq \Delta} = \tilde{H} P_{\leq \Delta} \eta \).

The appearance of \( \tilde{H} \), the complex-conjugate of \( H \), is necessary to allow for encoding of complex Hamiltonians into real ones. Note that for any real-valued Hamiltonian \( H \), we can simply write \( \mathcal{E}(H) = V(H \otimes P_{\text{anc}})V^\dagger \), where \( P_{\text{anc}} = P + Q \) is a projector since \( P \) and \( Q \) are orthogonal.

In the main definition used by Ref. [27], they have also motivated a natural restriction that the encoding should be local. Specifically, their restriction to “local encoding” requires that \( V = \bigotimes_i V_i \) for some isometries \( V_i \) each acting on at most one qudit from \( H \); moreover, \( P \) and \( Q \) are locally orthogonal projectors, in the sense that there are projectors \( P_i \perp Q_i \) acting on the same qudits as \( V_i \) where \( P_i P = P \) and \( Q_i Q = Q \). We note that our localized encodings per Definition 5 is somewhat different than this notion of “local encoding”. For example, constant-depth circuit qualifies as a localized encoding but not a “local encoding”, due to the possibility of overlaps between supports of encoded qubits (and hence \( V \) cannot be written in tensor-product form). On the other hand, Ref. [27] does not place any explicit restriction on the size of the support of each encoded qubit in their definition. Nevertheless, the specific encodings utilized for simulating general spin-Hamiltonians by universal Hamiltonians in [27] all have \( O(1) \)-sized supports, and thus also qualify as localized encodings.

It is not difficult to show that full-spectrum simulation by Definition 7 with sufficiently small precision \( (\xi \ll (1 - w)\gamma) \) implies a coherent gap-simulation by our Definition 2, if we restrict to encodings of the form \( \mathcal{E}(H) = V(H \otimes P_{\text{anc}})V^\dagger \). This restriction of the encoding format simplifies the comparison, and has no loss of generality when considering real-valued Hamiltonians.

Lemma 8 (Full-spectrum simulation implies coherent gap-simulation). Let \( H \) be a Hamiltonian that has a quasi-groundsparse projector \( P \) with quasi-spectral gap \( \gamma \) and energy spread \( w \leq 1/2 \). Suppose \( \tilde{H} \) full-spectrum-simulates \( H \) to precision \((\eta, \xi)\) according to Definition 7 with encoding \( \mathcal{E}(H) = V(H \otimes P_{\text{anc}})V^\dagger \), such that \( \xi \leq (1 - w)\gamma/8 \). Then \( \tilde{H}' = \frac{1}{\gamma} \tilde{H} \) gap-simulates \((H, P)\) with encoding \( V \), incoherence \( \epsilon \leq 32\xi/\gamma + 2\eta \), and energy spread \( \tilde{w} \leq (w + 2\xi/\gamma)/(1 - 2\xi/\gamma) \), per Definition 2.

2.2 Hamiltonian Sparsification: Degree-Reduction and Dilution

We define here the set of parameters of interest when considering minimizing resources in gap-simulations:

1. \( k \) – locality of individual Hamiltonian term; typically \( O(1) \) in physical systems, but we parametrize it to allow minimization, as well as to allow \( O(\log^a n) \)-local Hamiltonians, for some constant \( a \).
2. $r$ — maximum degree of Hamiltonian, the main objective in degree-reduction.
3. $M$ — number of terms in the Hamiltonian, the main objective in dilution.
4. $J$ — the interaction strength of individual Hamiltonian terms. This is typically restricted to $O(1)$ in physical systems, but allowing it to grow with $n$ leads to more possibilities of gap-simulators. Equivalently, a gap-simulator with $J$ growing with $n$ can be converted to one that simulates the original Hamiltonian but has a vanishing gap if we restrict to bounded-strength Hamiltonian terms.
5. $\epsilon$ and $\delta$ — incoherence $\epsilon$ and unfaithfulness $\delta$ that capture how well the Hamiltonian gap-simulates the original Hamiltonian in terms of groundspace projectors.
6. $\tilde{w}$ — energy spread in the gap-simulator Hamiltonian; allowing it to be different from the gap-simulates the original Hamiltonian in terms of groundspace projectors.

We will use the notation of $[r, M, J]$-gap-simulator to indicate that the maximum degree is $r$, the number of local terms is $M$, and for each term $H_i$ we have $\|H_i\| \leq J$. We define:

**Definition 9** (Degree-reduction (DR) and dilution). Let $\tilde{H}$ be a $k$-local $[r, M, J]$-gap-simulator of $H$ with $\epsilon$-incoherence (or $\delta$-unfaithfulness) and energy spread $\tilde{w}$. Additionally suppose $H = \sum_{i=1}^{M_0} H_i$ is a sum of $M_0 = M_0(n)$ terms, each of which is $O(1)$-local. Then
- We call $\tilde{H}$ an $[r, M, J]$-degree-reducer of $H$ if $r = O(1)$.
- We call $\tilde{H}$ an $[r, M, J]$-diluter of $H$ if $M = o(M_0(n))$.

We also call any degree-reducer or diluter of $H$ a sparsifier of $H$.

### 3 Results

Our impossibility results are based on two families of 2-local $n$ qubits Hamiltonians, which can both be expressed in terms of the collective angular momentum operator $J_{\alpha} = \sum_{i=1}^{n} \sigma_{\alpha}^{(i)}/2$ for $\alpha \in \{x, y, z\}$, where $\sigma_{\alpha}^{(i)}$ are the standard Pauli matrices.

**Example A** (degenerate groundstates).

$$H_A = \left( J_z + \frac{n}{2} \right) \left( J_z + \frac{n}{2} - 1 \right) = \frac{1}{4} \sum_{i < j} (1 - \sigma_z^{(i)}) \otimes (1 - \sigma_z^{(j)}) = \sum_{i < j} |1\rangle \langle 1|^{(i)} \otimes |1\rangle \langle 1|^{(j)}. \quad (5)$$

There are $M_0(n) = \frac{1}{2} n(n-1)$ terms in $H_A$, and each qubit has degree $n - 1$. The terms in $H_A$ mutually commute, and its groundspace is spanned by the following $n + 1$ zero-energy orthonormal states that have $J_j = -n/2$ or $J_z = -n/2 + 1$:

$$GS(H_A) = \text{span}\{|00\cdots00\rangle, |00\cdots01\rangle, |00\cdots10\rangle, \ldots, |10\cdots00\rangle\}. \quad (6)$$

If we consider a qubit in $|1\rangle$ to be an “excitation,” the groundstates are states with one or zero “excitations.” Observe that $w_n = 0$ and $\gamma_n = 1$, independent of $n$; the system is thus spectrally gapped.

**Example B** (unique groundstate). In this example we require that $n$ is even, $n = 2s$:

$$H_B = J_z^2 - \frac{1}{2} J^2 + b_n = \frac{1}{2} (J_z^2 - J_x^2 - J_y^2) + b_n$$

$$= \frac{1}{4} \sum_{i < j} (\sigma_z^{(i)} \sigma_z^{(j)} - \sigma_x^{(i)} \sigma_x^{(j)} - \sigma_y^{(i)} \sigma_y^{(j)}) - \frac{n}{8} + b_n \quad (7)$$
where \( b_n \equiv \frac{1}{2} n(n + 2) \) is a constant chosen so that \( \lambda_1(H_B) = 0 \). Similarly to \( H_A \), this Hamiltonian has \( M_{0}(n) = \frac{1}{2} n(n - 1) \) 2-local terms, and each qubit has degree \( n - 1 \). Since \( \{\hat{J}^2, \hat{J}_z\} = 0 \), the eigenstates of \( H_B \) can be written in eigenbasis of both \( \hat{J}^2 \) and \( \hat{J}_z \). As \( \hat{J}_z \) are spin-\( s \) angular momentum operators, \( \hat{J}^2 \) has eigenvalues \( \{0, 1, 2^2, \ldots, s^2\} \) and \( \hat{J}^2 \) has eigenvalues \( \{s(s + 1), (s - 1)s, \ldots, 6, 2, 0\} \). The groundstate of \( H_B \) is thus a state that has minimal \( \hat{J}^2 = 0 \) and maximal total angular momentum \( J = s = \frac{n}{2} \). This is a unique state, which is known as a Dicke state [30]:

\[
|g_B\rangle = |J = \frac{n}{2}; J_z = 0\rangle = \left( \frac{n}{n/2} \right)^{-1/2} \sum_{|\{i : x_i = 1\}| = n/2} |x_1 \cdots x_n\rangle. \tag{8}
\]

where the state can be explicitly written as a symmetric superposition of all strings \( x \) with Hamming weight \( h(x) = |\{i : x_i = 1\}| = n/2 \). This groundstate \( |g_B\rangle \) has energy 0. Meanwhile, all other eigenstates must have energy at least 1. Thus, the system is spectrally gapped with energy spread \( \gamma_n = 0 \) and \( \gamma_n = 1 \).

It turns out that these deceptively simple examples form a challenge for Hamiltonian sparsification.

### 3.1 Limitations on Degree-Reduction

For didactic reasons, we start by ruling out generic perfectly coherent DR. This is done by showing that such DR is impossible for \( H_A \).

**Lemma 10 (Impossibility of generic 0-incoherence DR).** There does not exist any \( k \)-local Hamiltonian \( \tilde{H}_A \) that is an \( \{o(n/k), M, J\} \)-degree-reducer of the \( n \)-qubit Hamiltonian \( H_A \) with localized encoding, 0-incoherence, and energy spread \( \tilde{\omega} < 1/2 \), regardless of number of terms \( M \) or interaction strength \( J \).

A closer inspection of the proof implies a trade-off between \( \epsilon \) and \( J \), from which it follows that if \( J = O(1) \) then generic DR is impossible even if we allow \( \epsilon \) which is inverse polynomially small (see exact statement in extended version [1]). We note that this result in fact rules out any improvement of the degree for \( H_A \), to some sub-linear degree.

However, perfect (or even inverse-polynomially close to perfect) coherence is a rather strong requirement. Indeed, by improving our proof techniques, we manage to improve our results for \( H_A \) to show impossibility even for constant coherence. Moreover, by devising another Hamiltonian with a unique groundstate, \( H_B \), and proving such an impossibility result also for this Hamiltonian, we arrive at the following theorem. Our main result is a strong impossibility result, ruling out generic DR with constant unfaithfulness and energy spread \( \tilde{\omega} \), for any number of Hamiltonian terms \( M \).

**Theorem 11 (Main: Impossibility of constant coherence (faithfulness) DR for \( H_A \) (\( H_B \))).** For sufficiently small constants \( \epsilon \geq 0 \) (\( \delta \geq 0 \)) and \( \tilde{\omega} \geq 0 \), there exists system size \( n_0 \) where for any \( n \geq n_0 \), there is no \( O(1) \)-local \( O(1), M, O(1)\)-degree-reducer of the \( n \)-qubit Hamiltonian \( H_A \) (\( H_B \)) with localized encoding, \( \epsilon \)-incoherence (\( \delta \)-unfaithfulness), and energy spread \( \tilde{\omega} \), for any number of Hamiltonian terms \( M \).

We deduce that generic quantum DR, with even constant unfaithfulness, is impossible. This stands in striking contrast to the classical setting. It is well known that classical DR is possible for all CSPs in the context of PCP reductions[32]. This construction easily translates to a 0-unfaithfulness degree-reducer for any classical local Hamiltonian:
Proposition 12 (Incoherent DR of classical Hamiltonians). Consider an $n$-qudit $k$-local classical Hamiltonian $H = \sum_{S \subseteq \{1, \ldots, n\} \subseteq [k]} C_S$, where each $C_S : \{z_i : i \in S\} \rightarrow [0, 1]$ is a function of $d$-ary strings of length $|S| \leq k$ representing states of qudits in $S$. Let the number of terms in $H$ be $M_0 = |\{S\}| = O(n^k)$. Then there is a $k$-local $[O(kM_0), O(1)]$-degree-reducer of $H$ with 0-unfaithfulness, energy spread $\bar{w} = 0$, and trivial encoding $V = 1$.

This demonstrates a large difference between the quantum and classical settings in the context of Hamiltonian sparsification. Characterizing which quantum Hamiltonians can be degree-reduced (with bounded interaction strength), either coherently or just faithfully, remains open.

The impossibility of DR by Theorem 11, which heavily relies on the interaction strength $J$ being a constant, is essentially tight. We prove this in a complementary result showing that degree-reduction is possible when $J$ is allowed to grow polynomially for any local Hamiltonian whose spectral gap closes slower than some polynomial (which is the case of interest for gap-simulation):

Theorem 13 (Coherent DR with polynomial interaction strength). Suppose $H$ is an $O(1)$-local Hamiltonian with a quasi-groundspace projector $P$, which has quasi-spectral gap $\gamma = \Omega(1/\text{poly}(n))$ and energy spread $w$. Also assume $\|H\| = O(\text{poly}(n))$. Then for any $\epsilon > 0$, one can construct an $O(1)$-local $[O(1), O(\text{poly}(n)/\epsilon^2), O(\text{poly}(n, \epsilon^{-1}))]$-degree-reducer of $H$ with incoherence $\epsilon$, energy spread $w + O(1/\text{poly}(n))$, and trivial encoding.

The proof is constructive: we map any given Hamiltonian to the quantum phase-estimation circuit, make the circuit sparse with new ancilla qudits and swap gates, and transform it back to a Hamiltonian using Kitaev’s circuit-to-Hamiltonian construction [47]. Some innovations are required to ensure coherence within the groundspace isn’t destroyed. For the most general local Hamiltonian whose spectral gap may close exponentially (or possibly even faster, see e.g. [2]), we can show that coherent DR is possible with interaction strength that scales exponentially with inverse gap and incoherence:

Theorem 14 (Coherent DR with exponential interaction strength). Let $H$ be an $n$-qubit $O(1)$-local Hamiltonian with $M_0$ terms, each with bounded norm. Suppose $H$ has quasi-spectral gap $\gamma$ and energy spread $w$. For any $\epsilon > 0$, one can construct a 2-local $[O(1), O(M_0), O((\gamma \epsilon)^{-\text{poly}(n)})]$-degree-reducer of $H$ with incoherence $\epsilon$, energy spread $w + O(\epsilon)$, and trivial encoding.

The proof uses a construction from perturbative gadgets, and is similar to other results in the Hamiltonian simulation literature [54, 27]. Due to significantly more resource required compared to Theorem 13, this construction is only useful in situations where we want to preserve some extremely small spectral gap.

3.2 Limitations on Dilution

For perfect or near-perfect dilution, we can prove a similar impossibility result to Lemma 10:

Theorem 15 (Impossibility of generic 0-incoherence dilution). There does not exist any $k$-local Hamiltonian $\tilde{H}_A$ that is an $[\epsilon, O(n^2/k^2), J]$-diluter of the $n$-qubit Hamiltonian $H_A$ with localized encoding, 0-incoherence, and energy spread $\bar{w} < 1/2$, regardless of degree $r$ or interaction strength $J$.

Similar to Lemma 10, this in fact holds even if we allow inverse polynomial incoherence [1]; and like above, this seems to be a rather weak impossibility result since requiring inverse polynomial incoherence may be too strong in many situations. Can we strengthen this to rule
out dilution with constant incoherence? The proof technique in Theorem 11 does not apply for dilution, since it relies on the decay of correlation between distant nodes in the interaction graph of $H$ (see Sec. 4.1). On the other hand, a diluter $\tilde{H}$ can have unbounded degree, and hence constant diameter, e.g., the star graph. Nevertheless, under a computational hardness assumption, no efficient classical algorithm for generic constant-unfaithfulness dilution exists, even for all $k$-local classical Hamiltonians:

**Theorem 16** (Impossibility of dilution algorithm for classical Hamiltonians). If $\text{coNP} \not\subseteq \text{NP/poly}$, then for any $\xi > 0$, $\delta < 1/\sqrt{2}$, $\tilde{w} \leq 1/2$, there is no classical algorithm that given a $k$-local $n$-qubit classical Hamiltonian $H$, runs in $O(\text{poly}(n))$ time to find an $[r, O(n^{k-\xi}), J]$-diluter of $H$ with $\delta$-unfaithfulness, energy spread $\tilde{w}$, and any encoding $V$ that has an $O(n^{k-\xi})$-bit description. This holds for any $r$ and $J$.

The above result rules out general (constructive) dilution even when the Hamiltonians are classical. For specific cases, however, dilution is possible. Our $H_A$ (which is also a classical Hamiltonian) provides such an example, for which we can achieve dilution even with 0-unfaithfulness, in the incoherent setting:

**Proposition 17** (0-unfaithfulness incoherent dilution and DR for $H_A$). There is a 3-local incoherent $[2, n - 1, 1]$-diluter of $H_A$ with 0-unfaithfulness, energy spread $\tilde{w} = 0$, and trivial encoding. This is also an incoherent $[2, n - 1, 1]$-degree-reducer of $H_A$.

Furthermore, combining ideas from the construction in Proposition 17 and Theorem 13, we can show that coherent dilution of $H_A$ with polynomial interaction strength is also possible:

**Proposition 18** (Coherent dilution and DR for $H_A$ with polynomial interaction strength). There is a 6-local $[6, O(n/\epsilon^2), O(\text{poly}(n, \epsilon^{-1}))]$-diluter of $H_A$ with $\epsilon$-incoherence, energy spread $\tilde{w} = 0$, and trivial encoding. This is also a $[6, O(n/\epsilon^2), O(\text{poly}(n, \epsilon^{-1}))]$-degree-reducer of $H_A$.

Note since Theorem 16 rules out constructive dilution regardless of interaction strength $J$, we cannot hope to prove an analogue of Theorem 13 or 14 to build coherent diluters for generic Hamiltonians, even allowing arbitrarily large interaction strength. Nevertheless, it remains an interesting open question to characterize Hamiltonians for which diluters exist, whether coherent or incoherent, with constant or large interaction strengths.

### 3.3 Connection to Quantum PCP

It might appear that our results rule out quantum degree-reduction (DR) in the context of quantum PCP (which would add to existing results [23, 5, 11, 20, 40, 6] ruling out quantum generalizations of other parts of Dinur’s PCP proof [32]). However, our results in this context (detailed in extended version [1]) currently have rather weak implications towards such a statement. The catch is that despite the apparent similarity, our gap-simulating DR is a very different notion from DR transformations used in the context of classical and quantum PCP. Gap-simulation seeks the existence of a Hamiltonian $\tilde{H}$ that reproduces the properties of the groundstate(s) and spectral gap of an input Hamiltonian $H$. On the other hand, a qPCP reduction is an algorithm that given $H$, it is merely required to output some $\tilde{H}$, such that if the groundstate energy of $H$ is small (or large), then so is the groundstate energy of $\tilde{H}$; in other words, qPCP preserves the promise gap. Notice that such a $H$ always exists, and the difficulty in qPCP reductions is to generate $\tilde{H}$ efficiently, without knowing the groundstate energy of $H$. Thus, we cannot hope for an information-theoretical impossibility result (as in Theorem 11 and 15) in the qPCP setting without further restriction on the output.
To circumvent this issue, we generalize to the quantum world a natural requirement, which seems to hold in the classical world for all known PCP reductions, namely that the reduction is constructive: roughly, it implies a mapping on not only the CSPs (Hamiltonians) but also individual assignments (states) [16, 33]. Specifically, we require that the reduction from \( \mathcal{H} \) to \( \tilde{\mathcal{H}} \) preserves groundstate properties in the similar sense as gap-simulation does, and maps excited states of \( \mathcal{H} \) (above the promise gap) to high-energy states \( \tilde{\mathcal{H}} \). Under this restriction, we prove that degree-reduction and dilution for quantum PCP with near-perfect coherence is impossible:

\[ \textbf{Theorem 19} \text{ (Limitation on qPCP-DR and qPCP-dilution (rough))}. \] There is no “constructive” qPCP reduction that works to degree-reduce or dilute the Hamiltonian \( \mathcal{H}_A \) with localized encoding, if we require small incoherence \( \epsilon \) relative to the energy of the output Hamiltonian \( \tilde{\mathcal{H}}_A \), namely, \( \epsilon \leq o(\|\tilde{\mathcal{H}}_A\|^{-1/2}) \).

The proof of Theorem 19 approximately follows that of impossibility results of Lemma 10 and Theorem 15 for sparsification with close-to-perfect coherence. Unfortunately, as we explain in Sec. 4.1, strengthening these results to prove impossibility for constant error (the regime of interest for qPCP), as is done in Theorem 11, seems to require another new idea.

### 4 Proofs Overview

#### 4.1 Proof Sketch for Main Theorem 11 (and related results: Theorem 15, 19 and Lemma 10)

We start with the idea underlying the impossibility of degree-reduction and dilution with (close to) perfect coherence (Lemma 10 and Theorem 15), which we refer to as “contradiction-by-energy”. For simplicity, let’s first examine the case of gap-simulation without encoding. Consider all pairs of original qubits \((i, j)\). The groundstates of \( \mathcal{H}_A \) include basis states with zero or one excitations (namely, 1’s), but not 2-excitation states. Importantly, the groundstates can be obtained from the 2-excitation state by local operations \( \sigma_x(i) \) and \( \sigma_x(j) \). Assuming the gap-simulator \( \tilde{\mathcal{H}}_A \) of \( \mathcal{H}_A \) does not interact the qubits \((i, j)\), we can express the energy of the 2-excitation state as a linear combination of the energy of 0- and 1-excitation states, up to an error of \( O(\tilde{w}) \) and \( O(\epsilon\|\tilde{\mathcal{H}}_A\|) \), using the fact that we can commute \( \sigma_x(i) \) and \( \sigma_x(j) \) through independent parts of \( \tilde{\mathcal{H}}_A \). If we assume \( \tilde{w} \) is small and \( \epsilon = 0 \), the energy of the 2-excitation state cannot be distinguished from these groundstates. Thus any gap-simulator \( \tilde{\mathcal{H}}_A \) must directly interact all pairs of qubits, which easily proves the impossibility without encoding. We can also see that if \( \epsilon > 0 \), then DR and dilution remain impossible if \( \|\tilde{\mathcal{H}}_A\| \leq O(\epsilon^{-1}) \), e.g. when \( \epsilon \) is polynomially small. This impossibility easily extends to localized encoding, where each original qubit is encoded into \( O(1) \) qudits in the gap-simulator Hamiltonian either independently or via some constant-depth circuit. In both cases, the required \( \Omega(n) \) degree and \( \Omega(n^2) \) interaction terms implied for the non-encoded version translate to the same requirements for the encoded version up to a constant factor, proving Lemma 10 and Theorem 15.

We now explain the proof of Theorem 11 that rules out degree-reduction even with constant incoherence. Let us first consider the statement for \( \mathcal{H}_A \) with constant \( \epsilon \) incoherence. The challenge is that the contradiction-by-energy trick used in the proof of Lemma 10 and Theorem 15 does not work for \( \epsilon = \Theta(1) \) incoherence. The problem is that the error in energy is of the order of \( O(\epsilon\|\tilde{\mathcal{H}}_A\|) \); this is too large for constant \( \epsilon \), and does not allow one to distinguish the energy of ground and excited states. Instead of contradiction-by-energy, we derive a contradiction using the groundspace correlations between qubits \((i, j)\), where \( \epsilon \)-incoherence
only induces an error of $O(\epsilon)$. Since $H_A$ is gapped, then any degree-reducer Hamiltonian $\bar{H}_A$ of $H_A$ must be gapped (while allowing some small energy spread $\tilde{w}$) by Definition 2. We can therefore apply a result (strengthened to accommodate non-vanishing energy spread[1]) of Hastings-Koma [41] stating that groundspace correlation decays exponentially with the distance on the graph where $\bar{H}_A$ is embedded. Since we assume bounded degree, we can find a pair $(i,j)$ among the original $n$ qubits such that their supports $(S_i,S_j)$ after a localized encoding are $\Omega(\log n)$ distance apart, with respect to the graph metric. Hence, their correlation $\langle V\sigma_x^{(i)}\sigma_x^{(j)}V^\dagger \rangle$ in the groundspace of $\bar{H}_A$ must decay as $e^{-\Omega(\log n)} = O(1/\text{poly}(n))$. Contradiction is achieved by the fact that for any pair of original qubits $(i,j)$, the groundspace of $\bar{H}_A$ contains a state of the form $V(|0_1,1_j\rangle + |1_i,0_j\rangle)|0^{n-2},\text{rest}\rangle + O(\epsilon)$, which has correlation at least $\langle V\sigma_x^{(i)}\sigma_x^{(j)}V^\dagger \rangle = 1 - O(\epsilon)$. For sufficiently small $\epsilon$ and $\tilde{w}$, this constant correlation from the latter lower bound contradicts the $O(1/\text{poly}(n))$ upper bound from the Hastings-Koma result.

The second part of Theorem 11 proves impossibility of incoherent DR for $H_B$ with $\delta$-unfaithfulness. Since $H_B$ has a unique groundstate that can be shown to have constant correlation between any pair of original qubits $(i,j)$, we can apply the same argument above for $H_A$ and show a contradiction with the Hastings-Koma’s vanishing upper bound of $O(1/\text{poly}(n))$ for small $\delta$ and $\tilde{w}$.

We now remark how these impossibility proofs can be extended to the context of quantum PCP. The contradiction-by-energy idea in Lemma 10 and Theorem 15 can indeed be generalized in this context. Under a reasonable restriction on the reduction – namely that the energy of non-satisfying assignments (frustrated or excited states) after the mapping is lower bounded by the promise gap – degree-reduction or dilution for quantum PCP is not generally possible with close-to-perfect (namely inverse polynomial) coherence (Theorem 19). However, this impossibility proof would not work when constant incoherence is allowed. To move to contradiction-by-correlation as in Theorem 11, we need to use some form of Hastings-Koma, which requires a spectral gap in $\hat{H}$. Thus, more innovation is needed as it may be an unnecessarily strong requirement for quantum PCP to preserve the spectral gap.

4.2 Overview of Remaining Proofs

Proof sketch: Equivalence between coherent and incoherent gap-simulations for unique groundstates (Lemma 4). We want to show that incoherent gap-simulation implies coherent gap-simulation, in the case of unique groundstate of the original Hamiltonian $H$. A naïve approach using the small error per groundstate of the gap-simulator will not work due to possible degeneracy in the groundspace of the simulator $\hat{H}$; this (possibly exponential) degeneracy could add an unwanted exponential factor. Hence, we explicitly construct the subspace on which the ancilla qubits should be projected by $P_{\text{anc}}$. The main observation is that since faithful gap-simulation implies that any state in the groundspace of $\hat{H}$ must be close to the space spanned by this $P_{\text{anc}}$, the dimensions of $P_{\text{anc}}$ and the groundspace of $\hat{H}$ must be the same. A sequence of simple arguments then allows us to derive a bound on the incoherence of any state (i.e., its norm after the incoherence operator in Eq. (3) is applied).

Proof sketch: DR of any classical Hamiltonian (Proposition 12). Here we follow the standard classical DR (as in [32]) in which each variable (of degree $d$) is replaced by $d$ variables, and a ring of equality constraints on these variables is added to ensure they are the same. The proof that this satisfies our gap-simulation definition is straightforward.
Proof sketch: Coherent DR of any Hamiltonian with $\Omega(1/\text{poly}(n))$ spectral gap using polynomial interaction strength (Theorem 13). The construction is based on mapping the quantum phase estimation (PE) circuit\[53\] to a Hamiltonian, using a modified version of Kitaev’s circuit-to-Hamiltonian construction\[47\]. The PE circuit can write the energy of any eigenstate of a given $H$ in an ancilla register, up to polynomial precision using polynomial overhead. The degree of the Hamiltonian is reduced by “sparsifying” the circuit before converting to the Hamiltonian. To repair the incoherence due to different histories, we run the circuit backwards, removing entanglement between the ancilla and the original register. To achieve $\epsilon$-incoherence, we add $O(\text{poly}(n)/\epsilon^2)$ identity gates to the end of the circuit. The eigenvalue structure of the original Hamiltonian $H$ is restored by imposing energy penalties on the energy bit-string written on the ancilla by the PE circuit. This yields a full-spectrum simulation of $H$, which also implies a gap-simulation of $H$.

Proof sketch: Impossibility of generic dilution algorithm (Theorem 16). Ref. \[29\] shows that under the assumption \text{coNP} \not\subseteq \text{NP/poly}, there is no poly-time algorithm to “compress” vertex-cover problems on $n$-vertex $k$-uniform hypergraphs and decide the problem by communicating $O(n^{k-\xi})$ bits for any $\xi > 0$ to a computationally unbounded oracle. Suppose towards a contradiction that $A$ is a poly-time algorithm to dilute any $k$-local classical Hamiltonian; we use it to derive a compression algorithm for vertex cover. To this end, $A$ is given a classical $k$-local Hamiltonian $H$ encoding a vertex cover problem; $A$ produces the diluter $\tilde{H}$ with $O(n^{k-2\xi})$ terms and some encoding $V$ described by $O(n^{k-2\xi})$ bits. Using Green’s function perturbation theory, we show that $\tilde{H}$ can be written using only $\log(n)$-bit precision as $\tilde{H}'$ with $O(1)$ error in the quasi-groundspace (even accounting for degeneracy). We then communicate $(\tilde{H}', V)$ to the oracle by sending $O(n^{k-2\xi} \log n) = O(n^{k-\xi})$ bits. The oracle then uses any groundstate of $\tilde{H}'$, which has large overlap with groundstates of $H$ for small $\delta$ and high precision, to decide the vertex cover problem and transmit back the answer.

Proof sketch: Incoherent dilution and DR of $H_A$ (Proposition 17). We use here the usual translation of a classical circuit to a CSP: $n-1$ qubits in a tree structure are used to simulate counting of the number of 1s among the original qubits, and the CSP checks the correctness of this (local) computation. The “history” of the computation is written on the ancilla qubits. Since different strings have different such histories, the construction is incoherent.

Proof sketch: Coherent dilution and DR of $H_A$ with polynomial interaction strength (Proposition 18). We improve upon the construction in Prop. 17 and Theorem 13 to obtain a coherent diluter of $H_A$ with polynomial interaction strength. The key is an $O(n)$-length circuit similar to that of Prop. 17 with a circuit that counts the number of 1s in the same tree geometry. Using the same tricks in Theorem 13 to uncompute computational histories and add identity gates at the end, we show that this leads to a coherent gap-simulator of $H_A$ with $\epsilon$-incoherence and $O(n/\epsilon^2)$ terms.

Proof sketch: Coherent DR for any Hamiltonian using exponential or larger interaction strength (Theorem 14). In order to provide generic coherent degree-reduction for any local Hamiltonian without restriction on the spectral gap $\gamma$, we first show that perturbative gadgets\[45, 54\] can be used for gap-simulation. The proofs make use of Green’s function machinery to bound incoherence $\epsilon$, which can be made small after every application of the gadgets using interactions of strength $O(\text{poly}(n)/(\gamma\epsilon)^O(1))$. This allows us to construct a
degree-reducer for any $k$-local Hamiltonian by a sequence of perturbative gadget applications. In the first part of the sequence, we reduce the locality of individual Hamiltonian terms to 3-local via $O(\log k)$ serial applications of subdivision gadgets [54], and each 3-local term is further reduced to 2-local via “3-to-2-local” gadgets [54]. Then, each original qubit is isolated from each other by subdivision gadgets so that they only interact with $O(n^{k-1})$ ancilla qubits that mediate interactions. Finally, applying fork gadgets [54] in $O(\log n)$ iterations allows us to reduce maximum degree of these original qubits to 6, generating our desired degree-reducer. It is this last part that causes the exponential blow-up in the interaction strength relative to $\gamma\epsilon$, so as to maintain the gap-simulation.

**Proof sketch: Generalized Hastings-Koma.** In Ref. [41], Hastings and Koma proved the exponential decay of correlations in the quasi-groundspace of a Hamiltonian $H$ consisting of finite-range (or exponentially decaying) interactions between particles embedded on a graph. They assume that the system is spectrally gapped, and has vanishing energy spread as the system size $n \to \infty$. Their proof is based on the relationship between the correlation $\langle \sigma^{(i)} \sigma^{(j)} \rangle$ they want to upper bound, and the commutator $\langle [e^{-iHt}\sigma^{(i)}e^{iHt},\sigma^{(j)}] \rangle$. By applying the Lieb-Robinson bound[51] on the latter, and integrating out the time $t$, they show that under the above conditions, the correlations between operators acting on particles $i$ and $j$ decay exponentially with the graph-theoretic distance between the particles. For application to the gap-simulation framework, we need to generalize their result to cases where the energy spread is not assumed to vanish with the system size. This is done by a careful modification of their proofs where we optimize the bounds and integration parameters so that errors due to the non-zero energy spread are suppressed.

5 Discussion and outlook

We have initiated the rigorous research of resources required for analog simulations of Hamiltonians, and proved unexpected impossibility results for Hamiltonian sparsification. Instead of working with full-spectrum simulations [22, 27], we use a new, relaxed definition of gap-simulation that is motivated by minimal requirements in physics. We note that impossibility results proven in a relaxed framework are of course stronger.

It would be very interesting to improve our understanding of the new framework of gap-simulations presented here, and clarify its applicability. As a start, it will be illuminating to find applications of gap-simulations in cases where full-spectrum simulations as in Ref. [22, 27] are unknown or difficult to achieve. For example, our Prop. 18 achieves dilution of $H_A$ with gap-simulation, but we do not know how to do so with full-spectrum simulation. Such simulations can enable experimental studies of these physical systems, by reducing resources required for analog simulations. Moreover, in many-body quantum physics, tools to construct “equivalent” Hamiltonians that preserve groundstate properties are of great utility. In this context, the study of gap-simulations can potentially lead to better understanding of universal behaviors in quantum phases of matter, which are characterized only by groundstate physics [56]. Another possible application of gap-simulators may be in the design of Hamiltonian-based quantum algorithms. In adiabatic algorithms [35], it is well known that the higher parts of the spectrum of the final and initial Hamiltonians can significantly affect the adiabatic gap [34, 31, 13]; gap-simulating these final and initial Hamiltonians by others will not affect the final groundstate, and can sometimes dramatically improve on the gap along the adiabatic path. Gap-simulations may also be a useful tool for tailoring the Hamiltonians used in other Hamiltonian-based algorithms such as QAOA [36].
We note that incoherent but faithful gap-simulations can be very interesting despite the apparent violation of the quantum requirement for coherence. For example, in adiabatic algorithms [35], we only want to arrive at one of the solutions (groundstates) to a quantum constraint satisfaction problem. In addition, in quantum NP [8], one is interested only in whether a certain eigenvalue exists, and not in the preservation of the entire groundspace. However, in the context of quantum simulation and many-body physics, maintaining coherence seems to be crucial for transporting all the physical properties of the groundspace. One would also expect maintaining coherence to be important when gap-simulating a subsystem (perhaps in an unknown state) of a larger system.

We remark that our framework deliberately avoids requiring that the eigenvalue structure of the spectrum be maintained even in its low-lying part, so as to provide a minimal but still interesting definition. Indeed, when simulating the groundspace, or a quasi-groundspace with small energy spread, this structure is not important. Nevertheless, one can imagine an intermediate definition, in which full-spectrum simulation is too strong, but the structure of a significant portion of the lower part of the spectrum matters. It might be interesting to extend the framework of gap-simulations to allow for such intermediate cases in which, for example, Gibbs states at low (but not extremely low) temperatures are faithfully simulated.

A plethora of open questions arise in the context of sparsification. First, it will be very interesting to find more examples where degree-reduction and/or dilution are possible, or are helpful from the perspective of physical implementations. Assuming bounded interaction strength, which is generally a limitation of physical systems, can we rigorously characterize which Hamiltonians can be coherently (or incoherently) degree-reduced? Of course, similar questions can be asked about dilution. It will also be interesting to consider saving other resources such as the dimensionality of the particles, which would be a generalization of alphabet-reductions from the context of PCP to Hamiltonian sparsification.

Our results on the impossibility of dilution are weaker than those for DR. Can we strengthen these to stronger information-theoretical results, by finding a quantum Hamiltonian for whom a diluter does not exist with constant incoherence, or even constant unfaithfulness?

We mention here that the classical graph sparsification results of Ref. [14, 15] can be viewed as dilution of a graph while approximately maintaining its spectrum. These results have been generalized to the matrix setting in Ref. [58]; however, this generalization does not seem to be useful in the context of diluting the interaction graph of a local Hamiltonian. The result of Ref. [58] shows that for sums of $d \times d$ positive Hermitian matrices, $O(d)$ matrices are sufficient to reproduce the spectral properties to good approximation, improving over Chernoff-like bounds [10]. While this in principle allows one to approximate a sum of terms by a sum of fewer terms, the required number of terms grows as $d = 2^Ω(n)$ for quantum Hamiltonians on $n$ qubits, and is thus irrelevant in our context.

Improving the geometry of the simulators is another important task that is relevant for applications of Hamiltonian sparsification to physical implementations. Ref. [49] has devised a method of converting the NP-complete Ising model Hamiltonian ($H = \sum_{ij} J_{ij} \sigma_i^z \sigma_j^z + \sum_i h_i \sigma_i^z$) on $n$ qubits to a new Hamiltonian on $O(n^2)$ qubits with interactions embedded on a 2D lattice, and sharing the same low-energy spectrum. Their construction encodes each edge $\sigma_i^z \sigma_j^z$ as a new qubit, and corresponds to an incoherent degree-reducer, where the new groundstates are non-locally encoded version of the original states. Our Prop. 12 also provides incoherent DR of these Hamiltonians, and without encoding, but the geometry is not in 2D; it will be interesting to improve our Prop. 12 as well as our other positive Theorems 13 and 14 to hold using a spatially local $\tilde{H}$. We note that if we allow the overhead of polynomial interaction strength, then it should be straightforward to extend the circuit-to-Hamiltonian
construction in Theorem 13 for analog simulations of local Hamiltonians on a 2D lattice, by ordering the gates in a snake-like fashion on the lattice similar to Ref. [54, 9]. Identifying situations where DR in 2D with bounded interaction strength is possible remains an open question.

A different take on the geometry question is to seek gap-simulators which use a single (or few) ancilla qubits that strongly interact with the rest. This may be relevant for physical systems such as neutral atoms with Rydberg blockade [57], where an atom in a highly excited level may have a much larger interaction radius, while no two atoms can be excited in each other’s vicinity.

Can we improve our results about quantum PCP, and show impossibility of qPCP-DR with constant incoherence? This would make our impossibility results interesting also in the qPCP context, as they would imply impossibility of DR in the qPCP regime of constant error, under a rather natural restriction on the qPCP reduction. This would complement existing impossibility results on various avenues towards qPCP [23, 5, 11, 20, 40, 6, 3]. Nevertheless, it seems that proving such a result might require a significantly further extension of Hastings-Koma, which may be of interest on its own.

Finally, we mention a possibly interesting variant of gap-simulation, which we call weak gap-simulation (see details in extended version[1]). Here, the groundspace is simulated in an excited eigenspace of the simulating Hamiltonian, spectrally gapped from above and below, rather than in its groundspace. This can be useful in the context of Floquet Hamiltonian engineering in periodically driven quantum systems, where eigenvalues are meaningful only up to a period, and thus a spectral gap in the middle of the spectrum is analogous to a spectral gap above the groundspace [39]. We are able to show how to weakly gap-simulate \( H_A \) to provide dilution with constant incoherence and bounded interaction strength – a task which we currently do not know how to do using “standard” gap-simulation. It remains open whether one can show stronger possibility results under weak gap-simulation. If not, can the impossibility results presented here be extended to the weak-gap-simulation setting? This might require an even stronger extension of Hastings-Koma’s theorem.

Overall, we hope that the framework, tools, and results presented here will lead to progress in understanding the possibilities and limitations in simulating Hamiltonians by other Hamiltonians – an idea that brings the notion of reduction from classical computer science into the quantum realm, and constitutes one of the most important contributions of the field of quantum computational complexity to physics.

References

1 See extended version at arXiv:1804.11084.


