Optimal simulation of two-qubit Hamiltonians using general local operations

C. H. Bennett,[1] J. I. Cirac,[2] M. S. Leifer,[3] D. W. Leung,[1,*] N. Linden,[3] S. Popescu,[4] and G. Vidal[2,†]
[1]*IBM T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598*
[2]*Institut für Theoretische Physik, Universität Innsbruck, A-6020 Innsbruck, Austria*
[3]*Department of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, United Kingdom*
[4]*H. H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United Kingdom and BRIMS, Hewlett-Packard Laboratories, Stoke Gifford, Bristol BS12 6QZ, United Kingdom*
(Received 17 October 2001; published 19 July 2002)

We consider the simulation of the dynamics of one nonlocal Hamiltonian by another, allowing arbitrary local resources but no entanglement or classical communication. We characterize notions of simulation, and proceed to focus on deterministic simulation involving one copy of the system. More specifically, two otherwise isolated systems A and B interact by a nonlocal Hamiltonian $H \neq H_A + H_B$. We consider the achievable space of Hamiltonians H' such that the evolution $e^{-iH't}$ can be simulated by the interaction H interspersed with local operations. For any dimensions of A and B, and any nonlocal Hamiltonians H and H', there exists a scale factor s such that for all times t the evolution $e^{-iH'st}$ can be simulated by H acting for time t interspersed with local operations. For two-qubit Hamiltonians H and H', we calculate the optimal s and give protocols achieving it. The optimal protocols do not require local ancillas, and can be understood geometrically in terms of a polyhedron defined by a partial order on the set of two-qubit Hamiltonians.

DOI: 10.1103/PhysRevA.66.012305 PACS number(s): 03.67.Hk, 03.67.Lx

I. INTRODUCTION

A. Motivation

A central problem of quantum information theory is to understand what kinds and quantities of nonlocal resources, such as entanglement and communication, are necessary and sufficient to accomplish a desired state transformation of a multipartite quantum system, if the parties are allowed unlimited local resources, including local unitary operations and the change of local Hilbert space dimension by measurements and/or the juxtaposition of local ancillas. It can be argued that the most fundamental nonlocal resource, from which all others are in practice derived, is interaction, represented in nonrelativistic quantum mechanics by a Hamiltonian that is not a sum of local terms. Given two nonlocal Hamiltonians H and H', one would like to know whether one can simulate the other, and if so, how efficiently.

The qualitative answer to this question is quite simple, as shown by the following parable. Let there be two parties who desire their joint state to evolve according to an arbitrarily intense and complex Hamiltonian H'. Unfortunately, like the mythical lovers Pyramus and Thisbe, they are almost completely isolated from one another, living on opposite sides of a wall pierced by a hole so small that only one atom of Pyramus can interact with one atom of Thisbe, via the two-atom Hamiltonian H (Fig. 1). Can H, together with local operations, be used to simulate H'? Yes, given enough time, because any nontrivial bipartite interaction can be used both to generate entanglement and to perform classical communication. Therefore they can use H, along with local ancillary degrees of freedom on each side of the wall, to generate enough entanglement, and perform enough classical communication to teleport Thisbe's entire original state to Pyramus' side. Now that they are (virtually) together, they can interact to their hearts' content. When it is time for Thisbe to go home, they teleport her back to her side, in whatever entangled state they have gotten themselves into, again using H to generate the needed entanglement and perform the needed classical communication.

A more practical motivation for studying the ability of nonlocal Hamiltonians to simulate one another comes from quantum control theory [1], in particular the problem of using an experimentally available interaction, together with local operations, to simulate the evolution that would have occurred under some other Hamiltonian not directly accessible to experiment. A more mathematical motivation comes from the desire to parametrize the nonlocal properties of interaction Hamiltonians, so as to characterize the efficiency with which they can be used to simulate one another, and perform other tasks such as generating entanglement [2,3] or performing quantum computation [4–7]. This parallels the

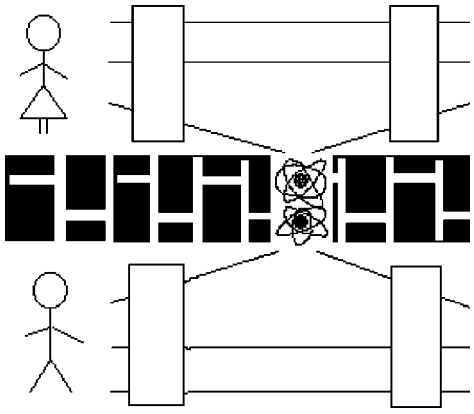

FIG. 1. Thisbe and Pyramus, separated by a wall, through which they can only interact by a two-atom Hamiltonian H.

*Email address: wcleung@watson.ibm.com
†Email address: Guifre.Vidal@uibk.ac.at

many recent efforts to parametrize the nonlocal properties of quantum states, so as to understand when, and with what efficiency, one quantum state can be converted to another by local operations, or local operations and classical communication. It is not difficult to see, by the Pyramus and Thisbe argument, that all nonlocal Hamiltonians are *qualitatively* equivalent, in the sense that for any positive t' and ϵ, there is a time t such that t' seconds of evolution under H' can be simulated, with fidelity at least $1-\epsilon$, by t seconds of evolution under H, interspersed with local operations; but much work remains to be done on the *quantitative* efficiency of such simulations.

In this paper we derive bounds on the time efficiency with which one Hamiltonian can simulate another using local resources. In the case of two interacting qubits, we show that these bounds are optimal. The structure of the paper is as follows. In Sec. II, we define the allowed resources and the type of simulation we consider. In Sec. III, we prove some general results on the type of simulation we consider along with some examples. In Sec. IV, we define our goal and summarize our main results for two-qubit Hamiltonians that are proved in Secs. V and VI. Some discussions and conclusions, and more auxiliary results can be found in Sec. VII, Sec. VIII, and Appendixes A and B. We first describe in more detail some related results.

B. Related work

The qualitative equivalence of nonlocal Hamiltonians noted above, and the use of interaction as an infinitesimal generator of entanglement, was already noted several years ago [8]. These discussions also considered the question of interconverting discrete nonlocal primitives, such as nonlocal gates, shared entanglement, and uses of a classical bit channel. More generally and quantitatively one may ask, given a nonlocal Hamiltonian $H_{AB} \neq H_A + H_B$, what is the optimal efficiency with which it can be used, in conjunction with local operations, (1) to generate entanglement between A and B, (2) to transmit classical or quantum information from A to B, or vice versa, (3) to simulate the operation of another nonlocal Hamiltonian H'. A partial answer to the first question, for two-qubit Hamiltonians, was given by Ref. [2]. The current work is a continuation of previous efforts to study the efficiency of simulating one Hamiltonian by another.

Hamiltonian simulation has been considered in the context of quantum computation [4–7,9–11,23]. In these works the system consists of n qubits, with some given *pairwise* interaction Hamiltonian. In Refs. [4–6], the given Hamiltonian was a sum of $\sigma_z \otimes \sigma_z$ interaction terms between distinct qubits (see Sec. III C for definitions) and the goal was to simulate a particular one of these terms. This was extended in Refs. [7,10,11] to arbitrary pairwise interactions, in both the simulating and the simulated Hamiltonians. In these papers the main concern was to obtain methods for simulation, and therefore upper bounds on the resources as a function of n.

Independent results on optimizing the time used of a given Hamiltonian for performing certain tasks are reported in Refs. [9,12,13]. Reference [9] gives a necessary condition for simulating one n-qubit pairwise interaction Hamiltonian by another, and gives a necessary and sufficient condition for simulation with a particular given Hamiltonian. Time resources for simulating the inverse of a Hamiltonian are discussed in Refs. [9,10,12]. Reference [13] considers simulating a unitary gate using a given Hamiltonian and a set of controllable gates in the shortest time. A general framework is set up in terms of Riemannian geometry. A time optimal protocol is obtained for the specific Hamiltonian $\sigma_z \otimes \sigma_z$ in the two-qubit case.

Finally, some more recent results have appeared since the original posting of this paper, extending it and related work in various ways [14–21].

II. SIMULATION FRAMEWORK

In this section we describe our framework of Hamiltonian simulation, i.e., the rules under which the simulation is to be performed. We also describe other possible frameworks and their relations to the one we adopt.

A. Available resources

Let H and H' each be a nonlocal Hamiltonian acting on two isolated systems A and B, possessed by Alice and Bob. We consider the problem of simulating H' by H using unlimited local resources. These include instantaneous local operations and uncorrelated local ancillas of any finite dimensions. It is also necessary to allow some initial classical correlation—Alice and Bob are assumed to have agreed beforehand on their time and spatial coordinates and the simulation protocol to be followed. Besides this, no other nonlocal resources are allowed, neither prior entanglement nor any form of communication beyond what can be achieved through the interaction H itself. Our goal is to minimize the time required of the given Hamiltonian H to simulate another Hamiltonian H'. This will be defined more formally in Sec. IV.

Note that either the simulating or the simulated system or both can be given the freedom of bringing in *local* degrees of freedom (ancillas) and allowing interaction between each ancilla with the corresponding local system. Ancillas on the simulated system can make it more powerful and therefore harder to simulate. Ancillas on the simulating system potentially make the simulation easier. We will allow ancillas on the simulating system, though they may not always help (Sec. VI).

B. One-shot and deterministic simulations

In this paper we only concern ourselves with protocols that are one-shot—i.e., operate on a single copy of each of the simulated and simulating systems—and which are required to succeed with probability 1.

More generally, a simulation can be "blockwise," in which $H^{\otimes n}$ is used for the simulation of $H'^{\otimes n}$, or in which H is time shared among many copies of the system and the amortized cost is considered. A simulation can also be stochastic and fail with finite probability, in which case the expected cost is considered.

C. Gate versus dynamics simulations

One possible notion of simulation is that, given H' and t', we simulate the final unitary evolution $e^{-iH't'}$ by composing local operations with elements in the one-parameter family $\{e^{-iHt}\}_t$.[1] The final evolution needs to be correct, but the intermediate evolution need not correspond to $e^{-iH't''}$ for $0 \leqslant t'' < t'$. The efficiency, given by the ratio t/t' can depend on t'. For example, a protocol can use H to generate entanglement and classical communication to teleport A to B, apply $e^{-iH't'}$ locally, and teleport A back. Viewing the cost t as a function of t', t does not increases indefinitely with t', rather, t can be made constant after it reaches a sufficiently large value. As another example, if the nonlocal Hamiltonian $H' = \sigma_z \otimes \sigma_z$ acts for time $t' = \pi/2$, the resulting unitary gate $i\sigma_z \otimes \sigma_z$ is local, and requires no nonlocal interaction time at all to simulate. This type of simulation, with very different primitives, is much studied in the context of universality of quantum gates [22] (composing a small set of available gates to obtain any desired unitary gate). More recently, simulation of a unitary gate using a fixed given Hamiltonian for a minimal amount of time and local manipulations was studied in Ref. [13] and some partial results were obtained. From now on, we call this type of simulation "gate simulation" or "finite time simulation."

A natural direction to strengthen the above notion of Hamiltonian simulation is to require not only the end result, but also the intervening *dynamics* of H' to be simulated. Intuitively, one might expect this to mean that the application of H, interspersed with instantaneous local operations, produces a trajectory that remains continuously close to the trajectory $e^{-iH't}$ that one wishes to simulate. However, this is impossible in general, because the needed local operations cause the simulating trajectory to be discontinuous, agreeing only intermittently with the trajectory one wishes to simulate. Accordingly we adopt the following definition of dynamics simulation. The Hamiltonian H simulates the dynamics of H' with efficiency μ if $\forall t' > 0, \forall \epsilon > 0$ the unitary operation $e^{-iH't'}$ can be simulated with fidelity $\geqslant 1-\epsilon$ by some protocol using H for a total time t'/μ and local operations. While this characterization may appear to have given up the idea of approximating the simulated system at intermediate times, in fact it has not, because it can be shown to imply the existence of a μ-efficient "stroboscopic" simulation, which approximates the simulated trajectory arbitrarily closely not only at the beginning and end, but also at an arbitrarily large set of intermediate times. We discuss this and other simulation notions in Appendix A. We also show that the existence of a protocol for dynamics simulation is equivalent to the existence of one for simulating an infinitesimal time (see Sec. III A), which in turns implies the ability to create protocols for arbitrary finite times by appropriately rescaling and repeating the infinitesimal-time protocol (see Appendix B).

[1]The evolution due to a Hamiltonian H is given by e^{-iHt}. Note the $-$ sign in the exponent.

III. GENERAL RESULTS AND EXAMPLES

Having defined the simulation framework, we derive some important general results and provide some examples of dynamics simulation, which motivate our main results and simplify some of the later discussions.

A. Infinitesimal and time independent simulation

First of all we show that dynamics simulation is equivalent to "infinitesimal simulation," the problem of simulating the evolution of H' for an infinitesimal amount of time t'. On one hand, any protocol for dynamics simulation simulates the initial evolution, therefore is a protocol for infinitesimal simulation. On the other hand, iterating an infinitesimal simulation results in dynamics simulation. We restrict our attention to infinitesimal simulation from now on, and focus on the lowest order effects in t'. Note that this property may not hold for other types of simulation described in Appendix A.

Infinitesimal simulation has a very special structure—the optimal simulation protocol is independent of the infinitesimal value of t'. The proof is included in Appendix B.

B. Local Hamiltonians are irrelevant

A general bipartite Hamiltonian K can be written as

$$K = K_A \otimes I + I \otimes K_B + \sum_{ij} M_{ij} \eta_i \otimes \eta_j , \tag{1}$$

where I denotes the identity throughout the paper, K_A, K_B are local Hamiltonians acting on A, B, respectively, and $\{\eta_i\}$ is a basis for traceless Hermitian operators acting on each of A and B. We can "dispose" of the local Hamiltonians K_A and K_B by undoing them with local unitaries on A and B,

$$(e^{iK_A t} \otimes e^{iK_B t}) e^{-itK} = e^{-i(K - K_A \otimes I - I \otimes K_B)t + O(t^2)}. \tag{2}$$

In other words, K can be made to simulate its own nonlocal component.

Likewise, any Hamiltonian can simulate itself with additional local terms. Therefore, given unlimited local resources, the problem of simulating an arbitrary Hamiltonian H' by another arbitrary one H reduces to the case when both are purely nonlocal.

C. Possible inefficiencies in simulation

Consider the simplest case of two-qubit systems. We introduce the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3}$$

and the useful identity

$$U e^M U^\dagger = e^{UMU^\dagger}, \tag{4}$$

where M is any bounded square matrix and U is any unitary matrix of the same dimension.

As an example, let $H=\sigma_x\otimes\sigma_x$ and $H'=\frac{1}{3}(\sigma_x\otimes\sigma_x+\sigma_y\otimes\sigma_y+\sigma_z\otimes\sigma_z)$. To simulate H' by H, let $U_1=\frac{1}{\sqrt{2}}(\sigma_x+\sigma_y)$ and $U_2=\frac{1}{\sqrt{2}}(\sigma_x+\sigma_z)$, so that $\sigma_y=U_1\sigma_xU_1^\dagger$ and $\sigma_z=U_2\sigma_xU_2^\dagger$. Using Eq. (4), it is easily verified that

$$e^{-iH't'}=(e^{-iHt'/3})(U_1\otimes U_1\,e^{-iHt'/3}U_1^\dagger\otimes U_1^\dagger)\times(U_2\otimes U_2\,e^{-iHt'/3}U_2^\dagger\otimes U_2^\dagger). \tag{5}$$

Conversely, we can simulate H with H',

$$e^{-iHt}=e^{-iH'3t/2}(\sigma_z\otimes I\,e^{-iH'3t/2}\sigma_z\otimes I). \tag{6}$$

Note that simulating H' for a duration of t' requires applying H for a duration of t' whereas simulating H for a duration t requires applying H' for a duration of $3t$. As the time required of the given Hamiltonian is a resource to be minimized, we see that some simulations are less efficient than the others. In this paper, we are concerned with the inefficiencies of simulation intrinsic to the Hamiltonians H and H' that are not caused by a bad protocol. For example, we will show later that the inefficiency in the above example is intrinsic.

D. Simulating the zero Hamiltonian—stopping the evolution

In some applications, the given Hamiltonian H cannot be switched on and off. Simulating the zero Hamiltonian $\mathbf{0}$ can be viewed as a means for switching off the Hamiltonian H [4–6]. This can always be done for any dimensions of A and B.

First, let A and B be 2^n-dimensional, and

$$H=\sum_{\mathbf{ij}} c_{\mathbf{ij}}P_{\mathbf{i}}\otimes P_{\mathbf{j}}, \tag{7}$$

where $\mathbf{i}$ is a binary vector $(i_1,i_2,\ldots,i_{2n})$ that labels the n-qubit Pauli matrix $P_{\mathbf{i}}=\sigma_x^{i_1}\sigma_z^{i_2}\otimes\cdots\otimes\sigma_x^{i_{2n-1}}\sigma_z^{i_{2n}}$. It is easily verified that

$$\frac{1}{2^{2n}}\sum_{\mathbf{i}} P_{\mathbf{i}}MP_{\mathbf{i}}=\mathrm{tr}\,M\,\frac{I}{2^n}. \tag{8}$$

A protocol for simulating $\mathbf{0}$ by H is given by

$$\Pi_{\mathbf{ij}}(P_{\mathbf{i}}\otimes P_{\mathbf{j}})e^{-iHt/2^{4n}}(P_{\mathbf{i}}^\dagger\otimes P_{\mathbf{j}}^\dagger)=\exp\left(-\frac{it}{2^{4n}}\sum_{\mathbf{ij}}(P_{\mathbf{i}}\otimes P_{\mathbf{j}})H(P_{\mathbf{i}}^\dagger\otimes P_{\mathbf{j}}^\dagger)+O(t^2)\right)\approx e^{-it\,\mathrm{tr}\,H/2^{2n}}, \tag{9}$$

in which the net evolution is just an overall phase to the lowest order in t.

When A and B are d-dimensional, one can embed each of A and B in a larger, 2^n-dimensional system for $n=\lceil\log_2 d\rceil$ to perform the simulation. Physically, this can be done on each of A and B, by attaching a qubit ancilla, extending the Hilbert space to $2d$ dimensions, and applying the simulation to a 2^n-dimensional subspace, such as one spanned by $|i\rangle\otimes|0\rangle$ for $i=1,\ldots,d$ and $|i\rangle\otimes|1\rangle$ for $i=1,\ldots,2^n-d$. Such simulation can also be done without ancillary degrees of freedom, and an alternative method based on Ref. [23] is given in Appendix C.

E. Arbitrary but inefficient simulations

We now show that any nonlocal bipartite Hamiltonian can be used to simulate any other, albeit with inefficiencies. In other words, for any H and H', operating H for time t can simulate the evolution of H' for time t' with $t'/t>0$. This holds for any dimensions. We keep all definitions from the previous example in the following protocol.

First, let A and B be 2^n-dimensional, $H=\Sigma_{\mathbf{ij}}c_{\mathbf{ij}}P_{\mathbf{i}}\otimes P_{\mathbf{j}}$ and $H'=\Sigma_{\mathbf{ij}}c'_{\mathbf{ij}}P_{\mathbf{i}}\otimes P_{\mathbf{j}}$. Without loss of generality the coefficient for $P_{\mathbf{k}}\otimes P_{\mathbf{k}}$ is positive, i.e., $c_{\mathbf{kk}}>0$, where $\mathbf{k}=(0,1,0,\ldots,0)$ and $P_{\mathbf{k}}=\sigma_z\otimes I\otimes\cdots\otimes I$. It is known that for any $P_{\mathbf{i}}$ and $P_{\mathbf{j}}$, there exist unitary operations $U_{\mathbf{ij}\pm}$ in the *Clifford group* [24], such that

$$U_{\mathbf{ij}\pm}P_{\mathbf{i}}U_{\mathbf{ij}\pm}^\dagger=\pm P_{\mathbf{j}}. \tag{10}$$

In other words, one can always transform any $P_{\mathbf{i}}$ to any other or to its negation. In our protocol, H simulates H' in two steps. First, H simulates $P_{\mathbf{k}}\otimes P_{\mathbf{k}}$ by

$$\Pi_{\mathbf{i},\mathbf{i}'|i_1,i_1'=0}P_{\mathbf{i}}\otimes P_{\mathbf{i}'}\,e^{-iHt/2^{4n-2}}P_{\mathbf{i}}\otimes P_{\mathbf{i}'}\approx\exp\left(-\frac{it}{2^{4n-2}}\sum_{\mathbf{i},\mathbf{i}'|i_1,i_1'=0}P_{\mathbf{i}}\otimes P_{\mathbf{i}'}HP_{\mathbf{i}}\otimes P_{\mathbf{i}'}\right)=e^{-itc_{\mathbf{kk}}P_{\mathbf{k}}\otimes P_{\mathbf{k}}+\text{local terms}}. \tag{11}$$

Alice and Bob independently apply an averaging over all Pauli operators commuting with $P_{\mathbf{k}}$, removing all operators except for $I=P_{\mathbf{0}}$ and $P_{\mathbf{k}}$ in each of their systems. The local terms can be ignored, following Sec. III B. Second, $P_{\mathbf{k}}\otimes P_{\mathbf{k}}$ simulates H' by

$$\Pi_{\mathbf{ij}}(U_{\mathbf{ki}\,\mathrm{sgn}(c'_{ij})}\otimes U_{\mathbf{kj}+})e^{-iP_{\mathbf{k}}\otimes P_{\mathbf{k}}|c'_{ij}|t'}(U_{\mathbf{ki}\,\mathrm{sgn}(c'_{ij})}\otimes U_{\mathbf{kj}+})^\dagger\approx\exp\left(-it'\sum_{\mathbf{ij}}P_{\mathbf{i}}\otimes P_{\mathbf{j}}c'_{ij}\right)=e^{-iH't'}, \tag{12}$$

where $\mathrm{sgn}(x)=x/|x|$ if $x\neq0$ and we omit terms with $c'_{ij}=0$.

When A and B are d-dimensional, the simulation of sH' by H can again be performed in a larger $2^n\times2^n$ system. This method implies a lower bound on the maximum possible value of s, $s\geqslant[1/(2^{2\lceil\log_2 d\rceil})](\max_{\mathbf{ij}}|c_{\mathbf{ij}}|)/(\Sigma_{\mathbf{ij}}|c'_{\mathbf{ij}}|)$. It is also possible to perform the simulation without ancillas. The proof is given in Appendix D. Other methods for such simulation were independently reported in Refs. [18–20].

F. Equivalent classes of local manipulations

Under our simulation framework, Alice and Bob are given unlimited local resources. In this subsection, we show that they only need a relatively small class of manipulations. To facilitate the discussion, we introduce classes of operations C, that can be LU, LO, LU+anc, and LO+anc, to be defined as follows. LU is the class of all local unitaries that act on $A\otimes B$. LU+anc is similar, but acts on $(A\otimes A')\otimes(B\otimes B')$ where A' and B' are uncorrelated ancillary systems of any finite dimension. LO and LO+anc are similarly defined, with the unitaries replaced by general trace-preserving quantum operations. Note that the largest class LO+anc corresponds to what is most generally allowed under our simulation framework.

We now show that LU+anc, LO, and LO+anc are equivalent under our framework. First, we show that LU+anc is at least as powerful as LO+anc. Any trace preserving quantum operation can be implemented by performing a unitary operation on a larger Hilbert space, followed by discarding the extra degrees of freedom (see, for example, Ref. [25]). The exact difference between LO+anc and LU +anc is that measurements and tracing are disallowed in the latter. However, these are not needed when simulating Hamiltonian in LU+anc, due to the following facts. (1) Measurements can be delayed until the end of the protocol, as operations conditioned on intermediate measurement results can be implemented unitarily. (2) In Hamiltonian simulation, the ancillary systems $A'B'$ have to be disentangled from AB at the end of the simulation.

Thus no actual measurement or discard is needed. These facts allow any LO+anc protocol to be reexpressed as an LU+anc protocol with pure product state ancillas, meaning that LO and LO+anc are no more powerful than LU+anc. Conversely, due to fact (2) above, any LU+anc protocol can be viewed as an LO protocol. Thus, we establish the equivalence between LO, LU+anc, and LO+anc. From now on, we focus on LU+anc protocols for full generality, and on LU protocols as a possible restriction.

IV. FORMAL STATEMENT OF THE PROBLEM AND SUMMARY OF RESULTS

Let H, H', A, B, A', B' be defined as before.

Definition. H' can be *efficiently simulated* by H,

$$H' \leqslant_C H, \tag{13}$$

if the evolution according to $e^{-iH't'}$ for any time t' can be simulated by using the Hamiltonian H for the same time t' and using manipulations in the class C.

Definition. H' and H are *equivalent* under the class C,

$$H' \equiv_C H, \tag{14}$$

if $H' \leqslant_C H$ and $H \leqslant_C H'$.

Throughout the paper, we only consider LU+anc protocols following Sec. III F. We also restrict our attention to H and H' that are purely nonlocal, following Sec. III B.

An LU+anc protocol simulates H' with H by interspersing the evolution of H with local unitaries on AA' and BB'. More specifically, the most general protocol for simulating H' using H for a total time t is to attach the ancillas $A'B'$ in the state $|0_{A'}\rangle\otimes|0_{B'}\rangle$, apply some $U_1\otimes V_1$, evolve AB according to H for some time t_1, apply $U_2\otimes V_2$, further evolve AB according to H for time t_2, and iterate "apply $U_i\otimes V_i$ and evolve with H for time t_i" some n times. At the end, it applies a final $U_f\otimes V_f$. The $t_i>0$ are constrained[2] by $\Sigma_{i=1}^n t_i = t$. Suppose the protocol indeed simulates an evolution for time t' according to H'. Then we can write

$$\begin{aligned}&(U_f\otimes V_f U_n\otimes V_n\, e^{-iHt_n}U_n^\dagger\otimes V_n^\dagger\times\cdots\\ &\quad\times U_1\otimes V_1 e^{-iHt_1}U_1^\dagger\otimes V_1^\dagger)|\psi\rangle\otimes|0_{A'}\rangle\otimes|0_{B'}\rangle\\ &=[e^{-iH't'}|\psi\rangle]\otimes[W_{A'B'}(t_1,\ldots,t_n)|0_{A'}\rangle\otimes|0_{B'}\rangle],\end{aligned} \tag{15}$$

where we have redefined $U_{i=1,2,\ldots,n}$ and $V_{i=1,2,\ldots,n}$, and $|\psi\rangle$ denotes the initial state in AB. In Eq. (15), e^{-iHt_i} acts on AB and implicitly means $e^{-iHt_i}\otimes I_{A'B'}$. The operator $W_{A'B'}(t_1,\ldots,t_n)$ describes the residual transformation of $A'B'$, and can be chosen to be unitary since the operation on the left-hand side of Eq. (15) is unitary. The problem we are concerned with can be stated in two equivalent ways.

Optimal and efficient simulation. Let H be arbitrary. The *optimal simulation problem* is to, for each H', find a solution $\{U_i\},\{V_i\},\{t_i\}$ of Eq. (15) such that t'/t is maximal. The *efficient simulation problem* is to characterize every H' that admits a solution for Eq. (15) with $t'=t$, i.e., $H' \leqslant_{\mathrm{LU+anc}} H$.

Definition. The optimal simulation factor $s_{H'|H}$ under class C of operations is the maximal $s>0$ such that $sH' \leqslant_C H$.

The optimal and efficient simulation problems are equivalent because inefficient simulation is always possible (see Sec. III). The efficient simulation problem can be solved by finding the optimal solution for each H' and characterizing those with $t'/t\geqslant 1$. The optimal simulation problem can be solved by finding the maximum s for which sH' is efficiently simulated. With this in mind, we may talk of solving either problem throughout the paper.

We now summarize our results. We show in Appendix B that, in the infinitesimal regime, the most general simulation protocol Eq. (15) using LU+anc is equivalent to

$$\begin{aligned}&sH'\\ &=\langle 0_{A'}|\otimes\langle 0_{B'}|\sum_i p_i U_i\otimes V_i(H\otimes I_{A'B'})U_i^\dagger\otimes V_i^\dagger|0_{A'}\rangle\otimes|0_{B'}\rangle.\end{aligned} \tag{16}$$

In the LU case (without ancillas), Eq. (16) reads

[2]Without loss of generality, a protocol with $\Sigma_{i=1}^n t_i<t$ can be turned to one with $\Sigma_{i=1}^n t_i=t$ by simulating the zero Hamiltonian as described in Sec. III.

$$sH' = \sum_i p_i U_i \otimes V_i H U_i^\dagger \otimes V_i^\dagger , \tag{17}$$

where $t = t_1 + \cdots + t_n$, $p_k = t_k/t$, and $s = t'/t$. Thus, the set $\{H' \leqslant_{\mathrm{LU}} H\}$ is precisely the convex hull of the set $\{U \otimes V H U^\dagger \otimes V^\dagger\}$ when U and V range over all unitary matrices on A and B, respectively. The linear dependence of $(t'/t)H'$ on H is manifest in both Eq. (16) and Eq. (17).

Our main results apply to the simulation of two-qubit Hamiltonians, and are summarized as follows.

Result 1. *Any* simulation protocol using LU+anc can be replaced by one using LU with the same simulation factor. This will be proved in Sec. VI. Thus, the four partial orders $\leqslant_{\mathrm{LU}}$, $\leqslant_{\mathrm{LU+anc}}$, $\leqslant_{\mathrm{LO}}$, $\leqslant_{\mathrm{LO+anc}}$ are equivalent for two-qubit Hamiltonians.

Result 2. We present the necessary and sufficient conditions for $H' \leqslant_{\mathrm{LU}} H$, for arbitrary two-qubit Hamiltonians H and H', and find the optimal simulation factor $s_{H'|H}$ and the optimal simulation strategy in terms of $\{U_i\},\{V_i\},\{t_i\}$. This will be discussed in Sec. V.

These results naturally endow the set of two-qubit Hamiltonians with a partial order $\leqslant_C$. This induces for each H, a set $\{H' : H' \leqslant_C H\}$ that is convex: if $H' \leqslant_C H$ and $H'' \leqslant_C H$, $pH' + (1-p)H'' \leqslant_C H$ for any $0 \leqslant p \leqslant 1$. Our method relies on the convexity of the set $\{H' : H' \leqslant_C H\}$, which has a simple geometric description, and in turns allows the partial order $\leqslant_C$ to be succinctly characterized by a majorization-like relation. The geometric and majorization interpretations offer two different methods to obtain, in practice, the optimal protocol and the simulation factor.

V. OPTIMAL LU SIMULATION OF TWO-QUBIT HAMILTONIANS

We will prove that $\leqslant_{\mathrm{LU}}$ is equivalent to $\leqslant_{\mathrm{LU+anc}}$ in the following section. In this section, we focus on LU simulations. We first adapt a result from Ref. [2] to reduce the problem to a smaller set of two-qubit Hamiltonians H and H'. Then, for any H, we identify the set $\{H' : H' \leqslant_C H\}$ with a simple polyhedron and obtain simple geometric and algebraic characterizations of it. The optimal solution for each pair of H and H' is derived. Finally, the problem is rephrased in the language of majorization.

A. Normal form for two-qubit Hamiltonians

The most general purely nonlocal two-qubit Hamiltonian K can be written as

$$K = \sum_{ij} M_{ij} \sigma_i \otimes \sigma_j , \tag{18}$$

where the summation is over Pauli matrices $i,j = x,y,z$ or 1,2,3 throughout the discussion for two-qubit Hamiltonians. Let

$$H = \sum_i h_i \sigma_i \otimes \sigma_i , \tag{19}$$

where $h_1 \geqslant h_2 \geqslant |h_3|$ are the singular values of the 3×3 matrix M with entries M_{ij}, and $h_3 = \mathrm{sgn}(\det M)|h_3|$. We say H is the *normal* form of K.

Theorem. Let H be the normal form of K. Then $H \equiv_{\mathrm{LU}} K$.

Proof. If the local unitaries $U^\dagger \otimes V^\dagger$ and $U \otimes V$ are applied before and after e^{-iKt}, the resulting evolution is given by

$$\begin{aligned} e^{-iK't} &= (U \otimes V)\, e^{-iKt}\, (U^\dagger \otimes V^\dagger) \\ &= e^{-i(U \otimes V)K(U^\dagger \otimes V^\dagger)t}, \end{aligned} \tag{20}$$

with

$$\begin{aligned} K' &= (U \otimes V)K(U^\dagger \otimes V^\dagger) \\ &= \sum_{ij} M_{ij} (U \sigma_i U^\dagger) \otimes (V \sigma_j V^\dagger) \\ &= \sum_{ij} M_{ij} \Big(\sum_l R_{il} \sigma_l \Big) \otimes \Big(\sum_k S_{jk} \sigma_k \Big) \qquad (21) \\ &= \sum_{lk} (R^T M S)_{lk} \sigma_l \otimes \sigma_k \equiv \sum_{lk} M'_{lk} \sigma_l \otimes \sigma_k . \qquad (22) \end{aligned}$$

In Eq. (21), $R,S \in \mathrm{SO}(3)$ since conjugating $\vec{r} \cdot \vec{\sigma}$ by SU(2) matrices corresponds to rotating $\vec{r}$ by a matrix in SO(3) (and vice versa). Equation (22) implies $K' = (U \otimes V)K(U^\dagger \otimes V^\dagger)$ for some unitary U,V if and only if $M' = R^T M S$. In particular, there is a choice of R and S that makes $K' = H$,

$$\begin{aligned} R^T &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det O_1 \end{pmatrix} \times O_1^T , \\ S &= O_2^T \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det O_2 \end{pmatrix} , \end{aligned} \tag{23}$$

where $M = O_1 D O_2$ is the singular value decomposition of M, with $O_1, O_2 \in O(3)$ and $D = \mathrm{diag}(h_1, h_2, |h_3|)$. Thus K and H are related by a conjugation by local unitaries, which implies $K \equiv_{\mathrm{LU}} H$. ■

As suggested by the above proof, we define a few useful notations.

Definitions. We call the 3×3 real matrix M_{ij} the "Pauli representation" of K, when M and K are related by Eq. (18). We use D_K to denote a diagonal Pauli representation of K.

Since any two-qubit Hamiltonian is equivalent to its normal form, we assume H', H are in normal forms from now on. We now turn to LU simulation of H' by H.

B. General LU simulation of normal form two-qubit Hamiltonians

Recall from Eq. (17) in Sec. IV that the most general simulation using LU is given by

$$sH' = p_1(U_1\otimes V_1)H(U_1^\dagger\otimes V_1^\dagger)+\cdots + p_n(U_n\otimes V_n)H(U_n^\dagger\otimes V_n^\dagger), \quad (24)$$

where $s=t'/t$. Following the discussion in Sec. V A, we only need to consider $H=\Sigma_i h_i\sigma_i\otimes\sigma_i$ and $H'=\Sigma_i h_i'\sigma_i\otimes\sigma_i$ that are in their normal forms. The Pauli representation of $(U\otimes V)H(U^\dagger\otimes V^\dagger)$ is given by RD_HS for some $R,S\in \mathrm{SO}(3)$. We can reexpress Eq. (24) as

$$sD_{H'} = p_1R_1D_HS_1+\cdots+p_nR_nD_HS_n, \quad (25)$$

where $R_i,S_i\in\mathrm{SO}(3)$. Since H and H' are in their normal form, $h_1\geqslant h_2\geqslant|h_3|$ and $h_1'\geqslant h_2'\geqslant|h_3'|$. Without loss of generality, we can make two assumptions. First, we can assume $h_3\geqslant 0$. If $h_3<0$, we can multiply Eq. (25) on the right side by $S=\mathrm{diag}(1,1,-1)$,

$$sD_{H'}S = p_1R_1(D_HS)(SS_1S)+\cdots+p_nR_n(D_HS)(SS_nS), \quad (26)$$

in which $SS_iS\in\mathrm{SO}(3)$, and $D_HS=\mathrm{diag}(h_1,h_2,|h_3|)$ is of the desired form. Thus, we can assume $h_3\geqslant 0$. Second, note that $s_{H'|H}=a\ s_{H'|aH}=(1/a)\ s_{aH'|H}$. The protocol is unchanged when Eq. (25) is divided by $\mathrm{tr}\,D_H=h_1+h_2+h_3$. Therefore, without loss of generality, the normalization $h_1+h_2+h_3=1$ can be assumed.

Equations (24) and (25) have a simple physical interpretation: the protocol partitions the allowed usage of H $[D_H]$ into different $U_k\otimes V_kHU_k^\dagger\otimes V_k^\dagger$ $[R_kDS_k]$, resulting in an "average Hamiltonian" H' $(D_{H'})$, which is a convex combination of the $U_k\otimes V_kHU_k^\dagger\otimes V_k^\dagger$ $[R_kDS_k]$.

The Hamiltonians, represented by $D_{H'}$, that can be efficiently simulated $(s=1)$ correspond to the diagonal elements of the convex hull of $\{RD_HS:R,S\in\mathrm{SO}(3)\}$. We call this diagonal subset, which is also convex, $\mathcal{C}_H$. Note that the zero Hamiltonian is in the *interior* of $\mathcal{C}_H$, because H can simulate any sH' for small s without ancillas (see Sec. III). Thus $\forall D_{H'}\neq 0$, the optimal solution is a *boundary* point of $\mathcal{C}_H$. The problem of efficient or optimal simulation can be rephrased.

Given H, let $\mathcal{C}_H$ be the diagonal subset of the convex hull of $\{RD_HS:R,S\in\mathrm{SO}(3)\}$. Then H' can be efficiently simulated by H if and only if $D_{H'}\in\mathcal{C}_H$. For any H', $s_{H'|H}D_{H'}$, which represents the optimal simulation, is the unique intersection of the semiline $\lambda D_{H'}$ $(\lambda\geqslant 0)$ with the boundary of $\mathcal{C}_H$. The optimal protocol can be obtained by decomposing $s_{H'|H}D_{H'}$ in terms of the extreme points of $\mathcal{C}_H$.

Since each point in $\mathcal{C}_H$ can be decomposed as a convex combination of the extreme points of $\mathcal{C}_H$, each efficiently simulated Hamiltonian can be identified with a simulation protocol and vice versa. We will refer to elements in $\mathcal{C}_H$ as Hamiltonians or simulation protocols interconvertibly.

Central to our problem is the structure of $\mathcal{C}_H$. We investigate its structure by first defining another object $\mathcal{P}_H$. $\mathcal{P}_H$ is a simple polyhedron defined by its set of 24 vertices, P_{24}, which is a subset of $\mathcal{C}_H$ (thus $\mathcal{P}_H\subseteq\mathcal{C}_H$). They are obtained from D_H by permuting the diagonal elements and putting an even number of $-$ signs. More explicitly, the vertices of $\mathcal{P}_H$ are $\pi_iD_H\pi_is_j$, where

$$\pi_0=I,\pi_1=\begin{bmatrix}-1&0&0\\0&0&1\\0&1&0\end{bmatrix},\ \pi_2=\begin{bmatrix}0&0&1\\0&-1&0\\1&0&0\end{bmatrix},$$

$$\pi_3=\begin{bmatrix}0&1&0\\1&0&0\\0&0&-1\end{bmatrix},\ \pi_4=\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix},$$

$$\pi_5=\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix},$$

$$s_0=I,s_1=\begin{bmatrix}1&0&0\\0&-1&0\\0&0&-1\end{bmatrix},\ s_2=\begin{bmatrix}-1&0&0\\0&1&0\\0&0&-1\end{bmatrix},$$

$$s_3=\begin{bmatrix}-1&0&0\\0&-1&0\\0&0&1\end{bmatrix}.$$

The transformation $D_H\rightarrow\pi_iD_H\pi_i^\dagger s_j$ is physically achieved by $H\rightarrow(U_{\pi_i}^\dagger\otimes U_{s_j}U_{\pi_i})H(U_{\pi_i}\otimes U_{s_j}^\dagger U_{\pi_i}^\dagger)$, where $U_{\pi_i}=(1/\sqrt{2})(\sigma_j+\sigma_k)$ for $i=1,2,3$ and i,j,k distinct, $U_{\pi_i}=\cos(\pi/3)I\pm i\sin(2\pi/3)(\sigma_x+\sigma_y+\sigma_z)/\sqrt{3}$ for $i=4,5$, and $U_{s_i}=\sigma_i$ for $i=1,2,3$. These can be verified using Eq. (21).

We will study the geometry of $\mathcal{P}_H$ in Sec. V C. We are interested in $\mathcal{P}_H$ because we will show in Sec. V D that $\mathcal{C}_H=\mathcal{P}_H$. Then we can find the optimal solution for any H' using our knowledge of $\mathcal{P}_H$. Moreover, $\mathcal{C}_H=\mathcal{P}_H$ means that P_{24} is the set of extreme points of $\mathcal{C}_H$ so that any optimal simulation protocol only involves the transformations $D_H\rightarrow\pi_iD_H\pi_is_j$. We restate the solution in terms of a majorizationlike relation in Sec. V F.

C. The polyhedron $\mathcal{P}_H$

Since P_{24} and $\mathcal{P}_H$ consist of diagonal matrices only, their elements can be represented by real three-dimensional vectors. The defining characterization of $\mathcal{P}_H$ is the polyhedron with 24 (not necessarily distinct) vertices that are elements of P_{24}. We now turn to a useful characterization of $\mathcal{P}_H$ as the region enclosed by its faces,

$$(x,y,z)\in\mathcal{P}_H \ \text{iff}\begin{cases}|x|\leqslant h_1,\ |y|\leqslant h_1,\ |z|\leqslant h_1,\\ -(1-2h_3)\leqslant +x+y+z\leqslant 1,\\ -(1-2h_3)\leqslant -x-y+z\leqslant 1,\\ -(1-2h_3)\leqslant +x-y-z\leqslant 1,\\ -(1-2h_3)\leqslant -x+y-z\leqslant 1,\end{cases} \quad (27)$$

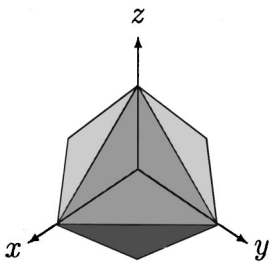

FIG. 2. $\mathcal{P}_H$ for $(h_1,h_2,h_3)=(1,0,0)$.

where the facts that H is in normal form, $h_3 \geqslant 0$, and that $h_1+h_2+h_3=1$ are used to replace the bounds $\Sigma_i h_i$ and $-(\Sigma_i h_i - 2\min_i h_i)$ by 1 and $-(1-2h_3)$ in Eq. (27). Equation (27) can be used to determine whether a point, as specified by its coordinates, is in $\mathcal{P}_H$ or not. The validity of Eq. (27) can be proved by plotting P_{24} (and therefore $\mathcal{P}_H$) and verifying that the faces are as given in Eq. (27). We first plot $\mathcal{P}_H$ for the simple case $(h_1,h_2,h_3)=(1,0,0)$, for which P_{24} has six distinct points: $(\pm 1,0,0)$, $(0,\pm 1,0)$, $(0,0,\pm 1)$ and Eq. (27) holds trivially (Fig. 2). Now, we plot $\mathcal{P}_H$ for the most complicated case, $h_1>h_2>h_3>0$ in Fig. 3.

Just like Fig. 2, Fig. 3 is viewed from the direction $(1,1,1)$. Three faces are removed to show the structure in the back. There are three types of faces. There are six identical rectangular dark grey faces on the planes $x=\pm h_1, y=\pm h_1, z=\pm h_1$. There are two groups of four identical hexagonal faces. The first group of four consists of the three light grey faces in the back, and the light grey face in the front. These are the truncated faces of the original octahedron, lying on the planes $x+y+z=1, -x+y-z=1, -x-y+z=1, x-y-z=1$. The second group consists of the three empty faces in the front, and the white face in the back. They are *inside* the original octahedron and are parallel to the original faces. They lie on the planes $-x-y-z=1-2h_3$, $-x+y+z=1-2h_3, x-y+z=1-2h_3, x+y-z=1-2h_3$. Note that each hexagon in one group has a parallel counterpart in the other group. Altogether, there are seven pairs of parallel faces, each pair bounds one expression in Eq. (27). It is straightforward to verify Fig. 3 and Eq. (27).

The plots for other cases, such as when $h_3=0$ or $h_1=h_2$, can be likewise obtained and Eq. (27) be verified. These are generally simpler than Fig. 3, and may admit simpler solutions in Sec. V E. However, we leave the details to the interested readers and move on to prove that $\mathcal{C}_H=\mathcal{P}_H$.

D. Proof of $\mathcal{C}_H=\mathcal{P}_H$

We now show that $\mathcal{C}_H=\mathcal{P}_H$. By definition $\mathcal{P}_H \subseteq \mathcal{C}_H$, thus we only need to show $\mathcal{C}_H \subseteq \mathcal{P}_H$. Recall that $\mathcal{C}_H$ consists of Hamiltonians that can be expressed as $D_{H'}=\Sigma_i p_i R_i D_H S_i^T$ [by putting $s=1$ in Eq. (25) and using S_i^T in place of S_i]. The fact that $D_{H'}$ is diagonal implies that only the diagonal elements in each $R_i D_H S_i^T$ contribute to $D_{H'}$; it is possible for an individual $R_i D_H S_i^T$ to be off-diagonal, but the off-diagonal elements have to cancel out in the sum. To show that $\mathcal{C}_H=\mathcal{P}_H$, it suffices to show that the diagonal part of each $R_i D_H S_i^T$ is in $\mathcal{P}_H$, because any $D_{H'} \in \mathcal{C}_H$ will then be in $\mathcal{P}_H$.

We represent the diagonal part of any $RD_H S^T$ as a three-dimensional vector (g_1,g_2,g_3). We need to show that

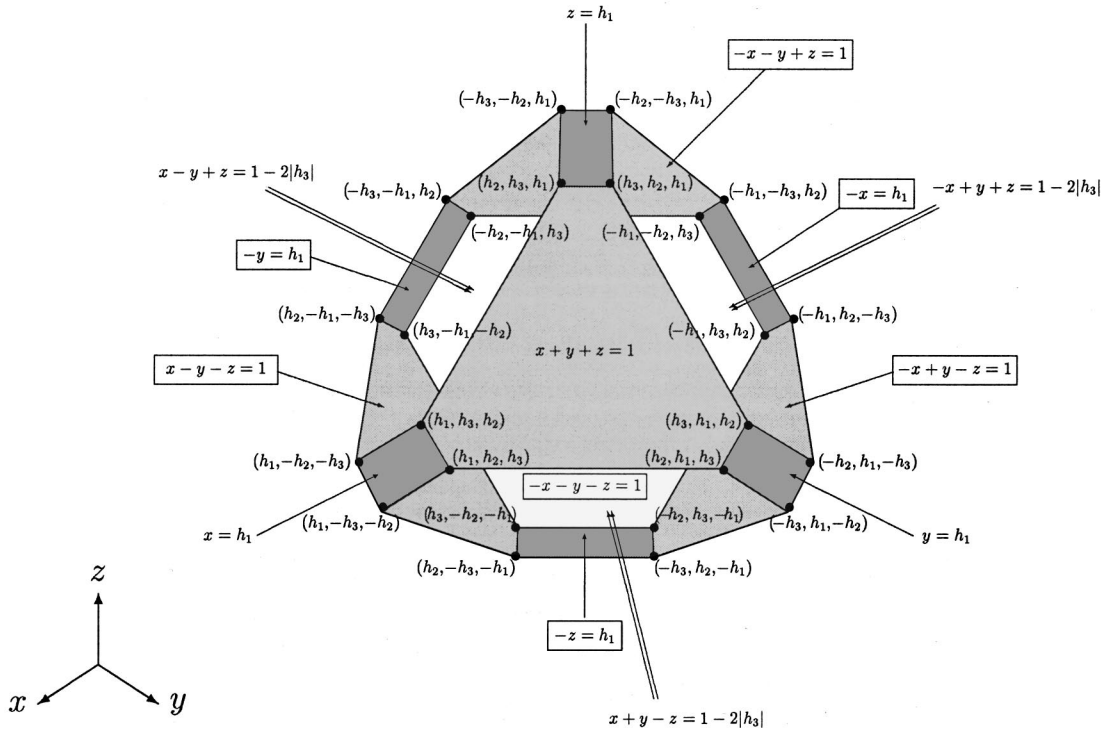

FIG. 3. $\mathcal{P}_H$ for $h_1>h_2>h_3>0$. The equations for the faces in the background are given in boxes. The empty faces are given by double arrows.

(g_1,g_2,g_3) satisfies Eq. (27) and belongs to $\mathcal{P}_H$. Since $D_H = \mathrm{diag}(h_1,h_2,h_3)$,

$$g_i=(RD_HS^T)_{ii}=\sum_k R_{ik}h_kS^T_{ki}=\sum_k R_{ik}S_{ik}h_k. \quad (28)$$

The vectors (h_1,h_2,h_3) and (g_1,g_2,g_3) are linearly related by

$$\begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} = R*S \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}, \quad (29)$$

where * denotes the entry-wise multiplication of two matrices, also known as the Schur product or the Hadamard product. It is useful to expand g_i in Eq. (28) explicitly,

$$g_i=R_{i1}S_{i1}h_1+R_{i2}S_{i2}h_2+R_{i3}S_{i3}h_3. \quad (30)$$

Then, we can prove the first group of inequalities in Eq. (27),

$$\begin{aligned} |g_i| &\leqslant |R_{i1}S_{i1}|h_1+|R_{i2}S_{i2}|h_2+|R_{i3}S_{i3}|h_3 \\ &\leqslant \max_i h_i = h_1. \end{aligned} \quad (31)$$

We have used the fact that $R,S\in \mathrm{SO}(3)$ to prove the second inequality in Eq. (31): R,S consist of orthonormal rows and columns. Hence, $(|R_{i1}|,|R_{i2}|,|R_{i3}|)$ and $(|S_{i1}|,|S_{i2}|,|S_{i3}|)$ are unit vectors, and their inner product $|R_{i1}S_{i1}|+|R_{i2}S_{i2}|+|R_{i3}S_{i3}|\leqslant 1$. We refer to this argument, which we use frequently, as the ''inner product argument.'' The second group of inequalities can be proved by

$$\begin{aligned} \sum_i |g_i| &= \sum_i \left| \sum_k R_{ik}S_{ik}h_k \right| \\ &\leqslant \sum_k \left(\sum_i |R_{ik}||S_{ik}| \right) |h_k| \leqslant \sum_k h_k = 1. \end{aligned} \quad (32)$$

The second inequality in Eq. (32) is due to $\Sigma_i|R_{ik}||S_{ik}|\leqslant 1$, obtained again by the inner product argument. This proves all of

$$\begin{aligned} &g_1+g_2+g_3\leqslant 1, \quad g_1-g_2-g_3\leqslant 1, \\ &-g_1+g_2-g_3\leqslant 1, \quad -g_1-g_2+g_3\leqslant 1. \end{aligned} \quad (33)$$

Finally,

$$\begin{aligned} &g_1+g_2+g_3 \\ &= \begin{pmatrix} R_{11}S_{11} \\ +R_{21}S_{21} \\ +R_{31}S_{31} \end{pmatrix} h_1 + \begin{pmatrix} R_{12}S_{12} \\ +R_{22}S_{22} \\ +R_{32}S_{32} \end{pmatrix} h_2 + \begin{pmatrix} R_{13}S_{13} \\ +R_{23}S_{23} \\ +R_{33}S_{33} \end{pmatrix} h_3 \\ &= \lambda_1h_1+\lambda_2h_2+\lambda_3h_3, \end{aligned} \quad (34)$$

where each λ_i is the coefficient of h_i in the parenthesis. The inner product argument implies $|\lambda_i|\leqslant 1$. Moreover, we will prove $\Sigma_i\lambda_i\geqslant -1$ shortly, which implies

$$\begin{aligned} g_1+g_2+g_3 &\geqslant \lambda_1h_1+\lambda_2h_2+(-1-\lambda_1-\lambda_2)h_3 \\ &= \lambda_1(h_1-h_3)+\lambda_2(h_2-h_3)-h_3 \\ &\geqslant -h_1-h_2+h_3 \\ &= -(1-2h_3), \end{aligned} \quad (35)$$

where Eq. (35) is the minimum of the preceding line, attained at $\lambda_1=\lambda_2=-1$ and $\lambda_3=1$. We now prove $\Sigma_i\lambda_i\geqslant -1$. First,

$$\begin{aligned} \sum_i \lambda_i &= R_{11}S_{11}+R_{21}S_{21}+R_{31}S_{31}+R_{12}S_{12}+R_{22}S_{22}+R_{32}S_{32} \\ &\quad +R_{13}S_{13}+R_{23}S_{23}+R_{33}S_{33} \\ &= \mathrm{tr}(R^TS). \end{aligned} \quad (36)$$

As $R,S\in \mathrm{SO}(3)$, $R^TS\in \mathrm{SO}(3)$. Each SO(3) matrix is a spatial rotation, therefore having the eigenvalue $+1$ that corresponds to the vector defining the rotation axis. Moreover, any SO(3) matrix has determinant 1. Therefore, the eigenvalues are generally given by 1, $e^{\pm i\phi}$ and the trace is $1+2\cos\phi \geqslant -1$. This completes the proof of Eq. (35). The last three of the four inequalities

$$\begin{aligned} &+g_1+g_2+g_3\geqslant -(1-2h_3), \\ &+g_1-g_2-g_3\geqslant -(1-2h_3), \\ &-g_1+g_2-g_3\geqslant -(1-2h_3), \\ &-g_1-g_2+g_3\geqslant -(1-2h_3), \end{aligned} \quad (37)$$

can be proved similarly. For example, consider

$$\begin{aligned} &g_1-g_2-g_3 \\ &= \begin{pmatrix} R_{11}\ S_{11} \\ -R_{21}\ S_{21} \\ -R_{31}\ S_{31} \end{pmatrix} h_1 + \begin{pmatrix} R_{12}\ S_{12} \\ -R_{22}\ S_{22} \\ -R_{32}\ S_{32} \end{pmatrix} h_2 + \begin{pmatrix} R_{13}\ S_{13} \\ -R_{23}\ S_{23} \\ -R_{33}\ S_{33} \end{pmatrix} h_3. \end{aligned} \quad (38)$$

The previous argument for $g_1+g_2+g_3$ applies by redefining R to be

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \times R.$$

Altogether, the inequalities in Eqs. (31), (32), and (37) satisfied by (g_1,g_2,g_3) are precisely the defining inequalities for $\mathcal{P}_H$ in Eq. (27). Therefore, the diagonal part of any RD_HS^T is in $\mathcal{P}_H$, and $\mathcal{C}_H=\mathcal{P}_H$.

E. Optimization over $\mathcal{P}_H$

Having proved $\mathcal{C}_H=\mathcal{P}_H$, we can solve the optimal simulation problem given D_H and $D_{H'}$ by finding the unique intersection of the semiline $\lambda D_{H'}$ with the boundary of $\mathcal{P}_H$

(see Sec. V B). We now explicitly work out $s_{H|H'}$, i.e., the value of λ in the intersection, as a function of H and H'.

Let all the symbols be as previously defined. The intersection is given by $\vec{v}=s_{H'|H}\,(h_1',h_2',h_3')$, so that

$$s_{H'|H}=\frac{||\vec{v}||_1}{||(h_1',h_2',h_3')||_1}=\frac{||\vec{v}||_1}{h_1'+h_2'+|h_3'|}\ , \qquad (39)$$

where $||\vec{v}||_1$ for a vector $\vec{v}$ is the sum of the absolute values of the entries. The set $\mathcal{P}_H$ has only three types of boundary faces. Therefore, there are only three possibilities where the intersection can occur.

(1) On the group of faces given by $x+y+z=1$, $-x+y-z=1$, $-x-y+z=1$, $x-y-z=1$. In this case, $||\vec{v}||_1=1$, and $s_{H'|H}=1/(h_1'+h_2'+|h_3'|)$.

(2) On the group of faces $x+y-z=1-2h_3$, $x-y+z=1-2h_3$, $-x+y+z=1-2h_3$, $-x-y-z=1-2h_3$. In this case, $||\vec{v}||_1=1-2h_3$, and $s_{H'|H}=(1-2h_3)/(h_1'+h_2'+|h_3'|)$.

(3) On the group of faces $x=\pm h_1, y=\pm h_1, z=\pm h_1$. In this case, $\vec{v}=(h_1/h_1')(h_1',h_2',h_3')$ (note $h_1'/h_1\geqslant 0$), $||\vec{v}||_1=(h_1/h_1')(h_1'+h_2'+|h_3'|)$ (not constant on the face), and $s_{H'|H}=(h_1/h_1')$.

Note that when H' is in normal form, $\vec{v}$ can only fall on $x+y+z=1$, $x+y-z=1-2h_3$, and $x=h_1$ in each of cases 1, 2, and 3. We now characterize (h_1',h_2',h_3') belonging to each case.

Case 1. Note that the face of $\mathcal{P}_H$ on $x+y+z=1$ is the convex hull of (h_1,h_2,h_3) and all permutations of the entries. The hexagon contains exactly all vectors $\vec{v}$ *majorized* by (h_1,h_2,h_3), $\vec{v}\prec(h_1,h_2,h_3)$ (see the following section for definition of majorization). Hence, (h_1',h_2',h_3') is in case 1 if and only if it is proportional to some $\vec{v}\prec(h_1,h_2,h_3)$.[3]

Case 3. In this case, $\vec{v}=(h_1,h_1h_2'/h_1',h_1h_3'/h_1')$. Thus (h_1',h_2',h_3') is in case 3 iff $(h_1h_2'/h_1',h_1h_3'/h_1')$ is within the rectangle with vertices $(h_2,h_3),(h_3,h_2),(-h_2,-h_3),(-h_3,-h_2)$ (Fig. 4). Hence, (h_1',h_2',h_3') is of case 3 iff

$$\left|\frac{h_1h_2'}{h_1'}+\frac{h_1h_3'}{h_1'}\right|\leqslant h_2+h_3 \quad \text{and} \quad \left|\frac{h_1h_2'}{h_1'}-\frac{h_1h_3'}{h_1'}\right|\leqslant h_2-h_3,$$

iff

$$\frac{h_1}{h_2+h_3}\leqslant\frac{h_1'}{h_2'+h_3'} \quad \text{and} \quad \frac{h_1}{h_2-h_3}\leqslant\frac{h_1'}{h_2'-h_3'}.$$

Case 2. This contains all (h_1',h_2',h_3') not in case 1 or 3.

The intersection on a boundary face can be easily decomposed as a convex combination of at most three vertices in P_{24}. The decomposition directly translates to an optimal protocol (using the discussion at the end of Sec. V B) with at most three types of conjugation.

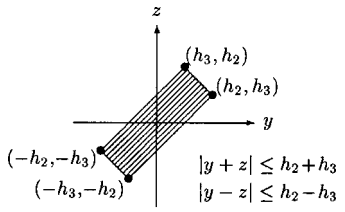

FIG. 4. ■■■■.

F. Optimal simulation, polyhedron $\mathcal{P}_H$, and *s*-majorization

The problem of Hamiltonian simulation also motivates a majorizationlike relation, which in turns provides a compact language to present the main results of this paper.

Let us recall the standard notions of majorization and weak majorization as defined in the space of n-dimensional real vectors. Let u be an n-dimensional vector with real components u_i, $i=1,\ldots,n$. We denote by $u^{\downarrow}$ the vector with components $u_1^{\downarrow}\geqslant u_2^{\downarrow}\geqslant\cdots\geqslant u_n^{\downarrow}$, corresponding to $|u_i|$ decreasingly ordered. Then, for two vectors u and v, u is submajorized or weakly majorized by v, written $u\prec_w v$, if

$$\begin{aligned} u_1^{\downarrow}&\leqslant v_1^{\downarrow},\\ u_1^{\downarrow}+u_2^{\downarrow}&\leqslant v_1^{\downarrow}+v_2^{\downarrow},\\ &\vdots\\ u_1^{\downarrow}+u_2^{\downarrow}+\cdots u_n^{\downarrow}&\leqslant v_1^{\downarrow}+v_2^{\downarrow}+\cdots v_n^{\downarrow}. \end{aligned} \qquad (40)$$

In case of equality in the last equation, we say that u is majorized by v, and write $u\prec v$.

Weak majorization of vectors induces a similar partial order in real matrices. More precisely, suppose M and N are two $n\times n$ real matrices, with respective singular values $\mathrm{sing}(M)$ and $\mathrm{sing}(N)$. Then, the weak majorization of real matrices can be defined as

$$M\prec_w N \qquad \text{iff} \qquad \mathrm{sing}(M)\prec_w \mathrm{sing}(N). \qquad (41)$$

Since the transformation $M\rightarrow O_1MO_2$ preserves the singular values when O_i are orthogonal, weak majorization also defines an equivalence relation,

$$M\sim O_1MO_2 \qquad \forall O_1,O_2\in \mathrm{O}(n). \qquad (42)$$

A useful result [26] in weak majorization is that the following ''convex sum'' characterization is equivalent to Eq. (42):

$$M\prec_w N \Leftrightarrow M=\sum_i p_iO_{i1}NO_{i2}. \qquad (43)$$

Our results in Secs. V B–V D show that the partial order $H'\leqslant_{LU}H$ strongly resembles weak majorization. $H'\leqslant_{LU}H$ when the convex sum $D_{H'}=\Sigma_i p_iR_iD_HS_i$ holds or equivalently when h_1',h_2',h_3' satisfy the inequalities in Eq. (27). This motivates the definition of an *s*-majorization for all

[3]The fact that $(h_1',h_2',h_3')\prec(h_1,h_2,h_3)$ is a necessary condition for efficient simulation is independently proved in Ref. [9].

$n\times n$ real matrices without restricting to the special form of D_H. Let M and N be real $n\times n$ matrices. We say that M is s-majorized by N, denoted by $M\prec_s N$, when M is a (left and right) special orthogonal mixing of N,

$$M\prec_s N \Leftrightarrow M=\sum_i p_i R_i N S_i . \tag{44}$$

This defines an equivalence relation,

$$M\sim_s RMS \qquad \forall\ R,S\in \mathrm{SO}(n), \tag{45}$$

and associates each matrix M with a vector $M^{\downarrow s}=(u_1,u_2,\ldots,u_{n-1},\det(M)u_n)$, where u_i are its singular values in decreasing order. This also defines an "s-ordered" vector $v^{\downarrow s}$ for any real vector v, viewed as a diagonal real matrix: the absolute values of the entries of v are arranged in decreasing order, and the product of all the original signs is added to the last element. Note that $u\prec_s v \Rightarrow u\prec_w v$. Moreover, when $\mathrm{sgn}(\Pi_i u_i)=\mathrm{sgn}(\Pi_i v_i)$ and $\Sigma_i u_i=\Sigma_i v_i$, $\prec_w$, $\prec_s$, and $\prec$ are all equivalent.

For $n=3$, our results in Secs. V C–V D can be extended to obtain the following characterization of s-majorization in terms of inequalities.

Let u and v be three-dimensional vectors, $u^{\downarrow s}=(u_1,u_2,u_3)$ and $v^{\downarrow s}=(v_1,v_2,v_3)$. Then $u\prec_s v$ if and only if

$$\begin{aligned} u_1 &\leqslant v_1,\\ u_1+u_2-u_3 &\leqslant v_1+v_2-v_3,\\ u_1+u_2+u_3 &\leqslant v_1+v_2+v_3. \end{aligned} \tag{46}$$

Let M and N be 3×3 real matrices. Then $M\prec_s N$ [defined by Eq. (44)] if and only if $M^{\downarrow s}\prec_s N^{\downarrow s}$.

Equation (46) follows from Eq. (27) and $\mathcal{P}_H=\mathcal{C}_H$, and from the fact that Eq. (46) is unchanged when the signs of u_3 and v_3 are flipped simultaneously and when u_i and v_i are rescaled by $v_1+v_2+v_3$.

Finally, we restate our result in Hamiltonian simulation in the language of s majorization.

Theorem. Let $H=\Sigma_i h_i\sigma_i\otimes\sigma_i$ and $H'=\Sigma_i h_i'\sigma_i\otimes\sigma_i$, $h=(h_1,h_2,h_3)$, and $h'=(h_1',h_2',h_3')$. Then

$$H'\leqslant_{\mathrm{LU}} H \quad\Leftrightarrow\quad h'\prec_s h. \tag{47}$$

The optimal simulation factor is given by $s_{H'|H}=\max_{\{sh'\prec_s h\}} s$.

VI. HAMILTONIAN SIMULATION WITH LU+anc

In this section we will show that the use of uncorrelated ancillas does not help when simulating one two-qubit Hamiltonian with another, so that all results on efficient and optimal simulation under LU hold under LU+anc. We prove this by describing the most general LU+anc protocol and reducing it to an LU protocol.

In this scenario, qubits A and B are, respectively, appended with ancillas A' and B', which have finite but arbitrary dimensions. The initial state of $A'B'$ can be chosen to be a pure product state $|0_{A'}\rangle\otimes|0_{B'}\rangle$. At the final stage of the simulation, the ancillas A' and B' may be correlated, but $A'B'$ is uncorrelated with AB if the latter is to evolve unitarily according to H'. The local unitary transformations U_i and V_i can act on AA' and BB', respectively. This feature distinguishes LU+anc from LU.

The most general LU+anc protocol to simulate H' with H can be described as

$$\begin{aligned} &(U_f\otimes V_f\times U_n\otimes V_n\, e^{-iHt_n}U_n^\dagger\otimes V_n^\dagger\times\cdots\times U_1\otimes V_1\, e^{-iHt_1}U_1^\dagger\otimes V_1^\dagger)|\psi\rangle\otimes|0_{A'}\rangle\otimes|0_{B'}\rangle\\ &\quad=(e^{-iH't'}|\psi\rangle)\otimes[W_{A'B'}(t_1,\ldots,t_n)|0_{A'}\rangle\otimes|0_{B'}\rangle]. \end{aligned} \tag{48}$$

In Appendix B we have shown that for infinitesimal times Eq. (48) leads to

$$sH'_{AB}=\langle 0_{A'}|\otimes\langle 0_{B'}|\left[\sum_k p_k U_k\otimes V_k(H\otimes I_{A'B'})U_k^\dagger\otimes V_k^\dagger\right]|0_{A'}\rangle\otimes|0_{B'}\rangle, \tag{49}$$

where $p_k\equiv t_k/t$ and $s\equiv t'/t$. Let $M_k\equiv\langle 0_{A'}|U_k$ and $N_k\equiv\langle 0_{B'}|V_k$. We can write Eq. (51) as

$$sH'=\sum_k p_k M_k\otimes N_k(H\otimes I_{A'B'})M_k^\dagger\otimes N_k^\dagger. \tag{50}$$

Note that this is the LU+anc analog of Eq. (24) for LU. In this case, H is replaced by $H\otimes I_{A'B'}$ and the local unitaries are replaced with more general transformations.

We focus on just one term in the convex combination of Eq. (50), $M\otimes N(H_{AB}\otimes I_{A'B'})M^\dagger\otimes N^\dagger$, with $M\equiv\langle 0_{A'}|U_k$ and $N\equiv\langle 0_{B'}|V_k$. We will show how to obtain the same contribution to H' using only local unitaries on A and B to establish the equivalence of LU and LU+anc. First, note that

$$M\otimes N(H_{AB}\otimes I_{A'B'})M^\dagger\otimes N^\dagger=\mathcal{E}_A\circ\mathcal{E}_B(H), \tag{51}$$

where $\mathcal{E}_A(\tau)\equiv M(\tau\otimes I_{A'})M^\dagger$ and similarly for $\mathcal{E}_B$. We emphasize that $\mathcal{E}_{A,B}$ are linear operators on matrices that are *not*

necessarily quantum operations [25], despite various resemblances to the latter. One can check that $\mathcal{E}_A$ is *unital*, i.e., $\mathcal{E}_A(I)=I$, by using $M=\langle 0_{A'}|U$. Furthermore, $\mathcal{E}_A$ is *completely positive* [25], because an operator-sum representation $\mathcal{E}_A(\tau)\equiv\Sigma_i F_i \tau F_i^\dagger$ can be obtained by expanding $I_{A'}$ in terms of some basis $\{|i_{A'}\rangle\}$, and by writing $F_i=M|i_{A'}\rangle = \langle 0_{A'}|U|i_{A'}\rangle$. However, in general, $\mathcal{E}_A$ is neither trace nonincreasing nor trace nondecreasing, though $\mathrm{tr}_A\Sigma_i F_i^\dagger F_i = \mathrm{tr}_A\Sigma_i\langle i_{A'}|U^\dagger|0_{A'}\rangle\langle 0_{A'}|U|i_{A'}\rangle=2$. For each F_i, we can obtain the singular value decomposition $F_i=W_{2i}Q_iW_{1i}$, where W_{1i} and W_{2i} are unitary, and

$$Q_i=\begin{bmatrix} q_{i1} & 0 \\ 0 & q_{i2}\end{bmatrix} \tag{52}$$

is diagonal and positive semidefinite. Altogether,

$$\mathcal{E}_A(\tau)=\sum_i W_{2i}Q_iW_{1i}\tau W_{1i}^\dagger Q_i W_{2i}^\dagger \tag{53}$$

$$=\sum_i \frac{1}{2}(q_{i1}^2+q_{i2}^2)W_{2i}Q_i'W_{1i}\tau W_{1i}^\dagger Q_i' W_{2i}^\dagger, \tag{54}$$

where

$$Q_i'=\sqrt{2}\begin{bmatrix}\cos\theta_i & 0\\ 0 & \sin\theta_i\end{bmatrix} \text{ and } \frac{\cos\theta_i=q_{i1}}{\sqrt{q_{i1}^2+q_{i2}^2}}.$$

We now show that, without affecting the Hamiltonian simulation, the conjugation by Q_i' in $\mathcal{E}_A(\tau)$ [Eq. (54)] can be replaced by the operation $\mathcal{Q}_i(\tau)=(1-\cos\theta\sin\theta)I\tau I + \cos\theta\sin\theta\sigma_z\tau\sigma_z$, i.e., replacing $\mathcal{E}_A$ by the following:

$$\tilde{\mathcal{E}}_A(\tau)=\sum_i \frac{1}{2}(q_{i1}^2+q_{i2}^2)W_{2i}\mathcal{Q}_i(W_{1i}\tau W_{1i}^\dagger)W_{2i}^\dagger. \tag{55}$$

It is straightforward to verify that

$$Q_i'IQ_i'=I+\cos(2\theta)\sigma_z, \quad Q_i'\sigma_xQ_i'=\sin(2\theta)\sigma_x,$$
$$Q_i'\sigma_yQ_i'=\sin(2\theta)\sigma_y, \quad Q_i'\sigma_zQ_i'=\cos(2\theta)I+\sigma_z, \tag{56}$$

$$\mathcal{Q}_i(I)=I, \quad \mathcal{Q}_i(\sigma_x)=\sin(2\theta)\sigma_x,$$
$$\mathcal{Q}_i(\sigma_y)=\sin(2\theta)\sigma_y, \quad \mathcal{Q}_i(\sigma_z)=\sigma_z. \tag{57}$$

Conjugation by Q_i' differs from the operation $\mathcal{Q}_i$ only when the input has an I or σ_z component. Their differences do not affect Hamiltonian simulation for the following reasons. As H is purely nonlocal, the input to $\mathcal{E}_A$ in Eq. (51) is traceless and has no I component. For the σ_z component in the input, $\mathcal{Q}_i(\sigma_z)$ and $Q_i'\sigma_zQ_i'$ differ only by an I component in the output, which contributes as a local term in Eq. (51). Hence, $\tilde{\mathcal{E}}_A$ can be used in place of $\mathcal{E}(\tau)$. Finally, we note that $\Sigma_i\frac{1}{2}(q_{i1}^2+q_{i2}^2)=\frac{1}{2}\Sigma_i\,\mathrm{tr}\,F_i^\dagger F_i=1$, so that $\tilde{\mathcal{E}}_A$ is indeed a convex combination of the individual terms, each in turn a mixture of unitary operations on A. Applying the same argument to $\mathcal{E}_B$, Alice and Bob only need to perform local unitaries in the simulation step of Eq. (51).

VII. DISCUSSION

First, we point out that the normal form for Hamiltonians acting on two qubits (Sec. V A) is *symmetric* with respect to exchanging the systems A and B. More formally, define $\mathcal{S}(M_1\otimes M_2)=M_2\otimes M_1$ as the (nonlocal) SWAP operation. Then $H\equiv_{\mathrm{LU}}\mathcal{S}(H)$. This has important consequence—any task generated by the Hamiltonian can be done equally well with the role of Alice and Bob interchanged.

In higher dimensions, the property $H\equiv_{\mathrm{LU}}\mathcal{S}(H)$ no longer holds. For example, $H\not\equiv_{\mathrm{LU}}\mathcal{S}(H)$ and $\mathcal{S}(H)\not\equiv_{\mathrm{LU}}H$ for the Hamiltonian (see Ref. [29] for a proof)

$$H=\begin{bmatrix}1&0&0\\0&-1&0\\0&0&0\end{bmatrix}\otimes\begin{bmatrix}1&0&0\\0&1&0\\0&0&-2\end{bmatrix}. \tag{58}$$

In fact, if $H=H_1\otimes H_2$ where H_1 and H_2 are members of a traceless orthogonal basis with different eigenvalues, $\mathcal{S}(H)\not\equiv_{\mathrm{LU}}H$ and $H\not\equiv_{\mathrm{LU}}\mathcal{S}(H)$. This also has important consequences—in higher dimensions, the nonlocal degrees of freedom of a Hamiltonian cannot be characterized by quantities that are symmetric with respect to A and B, such as eigenvalues of H (independently reported in Ref. [21]). Any normal form necessarily contains terms of the form $c_{ij}\eta_i\otimes\eta_j$ for some nonzero c_{ij} and the matrix with entries c_{ij} *cannot* be symmetric.

Second, we revisit the notion of efficiency in Hamiltonian simulation. Our definition of $H'\leqslant H$ depends on the normalization of both H and H'. One method to remove the normalization dependence is to require $h_1'+h_2'+|h_3'|=1$. Alternatively, we can consider the product $s_{H|H'}s_{H'|H}$ that measures the inefficiency of interconverting H and H' independent of the normalization of the Hamiltonians. We found that (proof omitted) when $h_3'\geqslant 0$, $s_{H|H'}s_{H'|H}\geqslant\frac{1}{3}$. Otherwise, $s_{H|H'}s_{H'|H}\geqslant\frac{1}{9}$, with equality when $h=(1/3,1/3,1/3)$ and $h'=(1/3,1/3,-1/3)$.

Third, we have considered the optimal simulation of one two-qubit Hamiltonian using another, both arbitrary but known. We can apply the characterization of $\mathcal{P}_H$ to analyze other interesting problems. For example, inverting a known Hamiltonian is equivalent to setting $H'=-H$. Without loss, assume $h_3\geqslant 0$ and $h_1+h_2+h_3=1$. Using the analysis in Sec. V E, the intersection is of case 2. Therefore, $s_{-H|H}=-(1-2h_3)$. The worst case is inverting $\frac{1}{3}(\sigma_x\otimes\sigma_x+\sigma_y\otimes\sigma_y+\sigma_z\otimes\sigma_z)$ in which case $s_{-H|H}=1/3$. In contrast, any protocol for inverting an unknown Hamiltonian can invert the worst known Hamiltonian, thus $s_{-H|H}\leqslant 1/3$. This is achievable using the following protocol:

$$(\sigma_x\otimes I)e^{-iHt'}(\sigma_x\sigma_y\otimes I)e^{-iHt'}(\sigma_y\sigma_z\otimes I)e^{-iHt'}(\sigma_z\otimes I)$$
$$=e^{-i(-H)t'/3}. \tag{59}$$

We can also improve on the time requirement for simulation protocols for *n*-qubit pairwise coupling Hamiltonians [7] with our construction. Instead of selecting a term by term simulation using a single nonlocal Pauli operator acting on a pair of qubit, one can directly simulate the desired coupling between the pair with any given one in a time optimal manner.

VIII. CONCLUSION

We have discussed various notions of Hamiltonian simulation. Focusing on dynamics simulation, we show its equivalence to infinitesimal simulation, and the intrinsic time independence of the protocols. We also show the possibility of simulating one nonlocal Hamiltonian with another without ancillas in any two *d*-dimensional systems. Our main results are on two-qubit Hamiltonians, in which case, for any Hamiltonian H, we characterize all H' that can be simulated efficiently, and obtain the optimal simulation factor and protocol. We obtain our results by considering a simple polyhedron that is related to some majorizationlike relations. Our results show that the two-qubit Hamiltonians are endowed with a partial order, in close analogy to the partial ordering of bipartite pure states under local operations and classical communication [27].

We have restricted our attention to simulation protocols that are infinitesimal, one-shot, deterministic, and without the use of entangled ancillas and classical communication. We also restricted our attention to bipartite systems. Extensions to the unexplored regime, and alternative directions such as other nonlocal tasks will prove useful, and are being actively pursued.

ACKNOWLEDGMENTS

We thank Wolfgang Dür, John Smolin, and Barbara Terhal for suggestions and discussions that were central to this work, as well as Herbert Bernstein, Ben Recht, and Aram Harrow for additional helpful discussions and comments. Part of this work was finished when G.V. was visiting the IBM T. J. Watson Research Center, and D.W.L. was visiting the IQI at Caltech. C.H.B. and D.W.L. were supported in part by the NSA and ARDA under the U.S. Army Research Office, Grant No. DAAG55-98-C-0041. G.V. was supported by the European Community project EQUIP (Contract No. IST-1999-11053), and Contract No. HPMF-CT-1999-00200.

APPENDIX A: NOTIONS OF SIMULATION

We consider various notions of using a Hamiltonian H to simulate the evolution due to H' for time t'.

In dynamics simulation, the evolution of the system is close to $e^{-iH't''}$ after an operation time of $\mu t''$ for constant μ and $\forall t'' \in [0,t']$. It is possible to relax this requirement, so that, $\mu(t'')$ is a function of t'', and without loss of generality, $\mu(t'')$ is nondecreasing. We call this ''variable rate dynamics simulation.'' Finally, in gate simulation, the only requirement is that, the final evolution is given by $e^{-iH't'}$.

As an analogy, let H' be driving along a particular highway from house A to house B at 100 km/hr. Dynamics simulation is like driving, biking, or walking along the same highway at any *constant* speed. Variable rate dynamics simulation is like driving along the highway at variable speed, for example, when there is stop-and-go traffic. The vehicle is always on the trajectory defined by H'. Finally, gate simulation is like going from house A to house B by any means, for example, using local roads, or flying a helicopter.

It is important to note the difference between dynamics simulation (or infinitesimal simulation) and variable rate dynamics simulation. For example, iterating infinitesimal simulations to perform dynamics simulation, the ancillas are implicitly discarded after each iteration, and new ones be used next. However, it is possible in variable dynamics simulation that used ancillas can subsequently be used to accelerate the simulation. Such phenomena are known in entanglement generation [2]. The more complicated analysis for variable dynamics simulation will be addressed in future work.

APPENDIX B: INFINITESIMAL SIMULATION AND TIME INDEPENDENCE

In this appendix, we show that the optimal protocol for infinitesimal simulation is independent of t', the time of evolution to be simulated.

The most general simulation protocol of H' with H using LU+anc can be described by

$$(U_f\otimes V_f\times U_n\otimes V_n\, e^{-iHt_n}U_n^\dagger\otimes V_n^\dagger\cdots U_1\otimes V_1\, e^{-iHt_1}U_1^\dagger\otimes V_1^\dagger)(|\psi\rangle\otimes|0_{A'}\rangle\otimes|0_{B'}\rangle)=(e^{-iH't'}|\psi\rangle)\otimes(W_{A'B'}|0_{A'}\rangle\otimes|0_{B'}\rangle), \tag{B1}$$

where the equality must hold for all possible states $|\psi\rangle$ of system AB. Here the unitaries U_i and V_i, acting on AA' and BB', respectively, and the partition $\{t_i\}$ of the time interval $t=\Sigma_i t_i$, correspond to all the degrees of freedom available for the simulation of H' for time t'. The initial state of the ancillas A' and B' is $|0_{A'}\rangle\otimes|0_{B'}\rangle$, and $W_{A'B'}$ is their residual, unitary evolution, which is determined by the other degrees of freedom and may create entanglement between A' and B'.

We have argued earlier that optimal dynamics simulation can always be achieved by a protocol for simulating infinitesimal evolution times t'. This also implies t being infinitesimal. Recall that $p_i\equiv t_i/t$ and $s\equiv t'/t$. We can expand Eq. (B1) to first order in t to obtain

$$U_f\otimes V_f\times\left[I-it\sum_i p_i U_i\otimes V_i(H\otimes I_{A'B'})U_i^\dagger\otimes V_i^\dagger\right]|0_{A'}\rangle\otimes|0_{B'}\rangle=(I-itsH')\otimes(W_{A'B'}|0_{A'}\rangle\otimes|0_{B'}\rangle). \tag{B2}$$

The validity of Eq. (B1) for all $|\psi\rangle$ is used to obtain Eq. (B2), each term of which is taken to be an operator on *AB*. It follows from Eq. (B2) that

$$(U_f|0_{A'}\rangle)\otimes(V_f|0_{B'}\rangle)=I_{AB}\otimes(W_{A'B'}|0_{A'}\rangle\otimes|0_{B'}\rangle)+O(t), \tag{B3}$$

which implies that

$$U_f|0_{A'}\rangle=I_A\otimes(W_{A'}|0_{A'}\rangle)+O(t), \tag{B4}$$

$$V_f|0_{B'}\rangle=I_B\otimes(W_{B'}|0_{B'}\rangle)+O(t), \tag{B5}$$

$$W_{A'B'}|0_{A'}\rangle\otimes|0_{B'}\rangle=(W_{A'}|0_{A'}\rangle)\otimes(W_{B'}|0_{B'}\rangle)+O(t). \tag{B6}$$

Equation (B6) implies $W_{A'B'}$ is a product operator to zeroth order in t. Redefining U_f and V_f is necessary, we can assume $W_{A'}=I_{A'}$ and $W_{B'}=I_{B'}$. Explicitly writing down the most general $O(t)$ terms in Eqs. (B4)–(B6), we obtain

$$U_f|0_{A'}\rangle=(I_{AA'}-itK_{AA'})|0_{A'}\rangle+O(t^2), \tag{B7}$$

$$V_f|0_{B'}\rangle=(I_{BB'}-itK_{BB'})|0_{B'}\rangle+O(t^2), \tag{B8}$$

$$\begin{aligned}W_{A'B'}|0_{A'}\rangle\otimes|0_{B'}\rangle&=[(I_{A'}-itK_{A'})|0_{A'}\rangle)]\otimes[(I_{B'}-itK_{B'})|0_{B'}\rangle]\\&\quad-itK_{A'B'}|0_{A'}\rangle\otimes|0_{B'}\rangle+O(t^2),\end{aligned} \tag{B9}$$

where the unitarity of the operators on the left-hand side implies the hermiticity of $K_{AA'}$, $K_{BB'}$, $K_{A'}$, $K_{B'}$, and $K_{A'B'}$. Substituting Eqs. (B7)–(B9) in Eq. (B2) implies

$$\begin{aligned}&sH'\otimes(I_{A'B'}|0_{A'}\rangle\otimes|0_{B'}\rangle)\\&\quad=O(t)+\Bigg(\sum_i p_iU_i\otimes V_i(H\otimes I_{A'B'})U_i^\dagger\otimes V_i^\dagger\\&\qquad+K_{AA'}+K_{BB'}-K_{A'}-K_{B'}-K_{A'B'}\Bigg)|0_{A'}\rangle\otimes|0_{B'}\rangle.\end{aligned} \tag{B10}$$

Projecting this equation on the left onto $\langle 0_{A'}|\otimes\langle 0_{B'}|$, the terms $K_{AA'}+K_{BB'}-K_{A'}-K_{B'}-K_{A'B'}$ become local or identity terms. Taking into account that H' has zero trace and no local terms (recall Sec. II B), their contributions vanish, and we obtain

$$sH'=\langle 0_{A'}|\otimes\langle 0_{B'}|\Bigg[\sum_i p_iU_i\otimes V_i(H\otimes I_{A'B'})U_i^\dagger\otimes V_i^\dagger\Bigg]|0_{A'}\rangle\otimes|0_{B'}\rangle. \tag{B11}$$

In the case we do not have ancillary systems, U_i and V_i only act on *A* and *B*, and Eq. (B11) reads

$$sH'=\sum_i p_iU_i\otimes V_iHU_i^\dagger\otimes V_i^\dagger. \tag{B12}$$

In Eqs. (B11) and (B12), the dependence of the equation on the original infinitesimal times t and t' is only through $s=t'/t$. This implies any protocol for t and t' applies to αt and $\alpha t'$ within the infinitesimal regime. Thus the protocol, namely, the set $\{U_i,V_i,p_i\}$ can be considered being independent of t in the infinitesimal regime.

APPENDIX C: SIMULATING ZERO HAMILTONIAN IN $d\times d$ WITHOUT ANCILLAS

In Ref. [23], it is shown that for any d-dimensional square matrix M,

$$\sum_{ij} U_{ij}MU_{ij}^\dagger=(\mathrm{tr}\,M)\frac{I}{d}, \tag{C1}$$

where

$$U_{ij}=\begin{bmatrix}1&0&0&0&0\\0&\omega&0&0&0\\0&0&\omega^2&0&0\\0&0&0&\cdot&0\\0&0&0&0&\omega^{d-1}\end{bmatrix}^i\times\begin{bmatrix}0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\\1&0&0&0&0\end{bmatrix}^j \tag{C2}$$

and ω is a primitive dth root of unity. H can simulate $\mathbf{0}$ using the protocol

$$\begin{aligned}&\Pi_{ij}(U_{ij}\otimes I)e^{-iHt}(U_{ij}^\dagger\otimes I)\\&\quad\approx\exp\Bigg(-i\sum_{ij}(U_{ij}\otimes I)H(U_{ij}^\dagger\otimes I)t\Bigg)\\&\quad=e^{-iI\otimes K_Bt},\end{aligned} \tag{C3}$$

which is local and can be removed.

APPENDIX D: ARBITRARY HAMILTONIAN SIMULATION IN $D\times D$ WITHOUT ANCILLAS

Let H and H' act on two d-dimensional systems. We use the following (nonorthonormal) basis for traceless Hermitian operators acting on a d-dimensional system:

$$\eta_1=\begin{bmatrix}1&0&0&\dots\\0&-1&0&\dots\\0&0&0&\dots\\\vdots&\vdots&\vdots&\vdots\end{bmatrix},\quad \eta_2=\begin{bmatrix}1&0&0&\dots\\0&0&0&\dots\\0&0&-1&\dots\\\vdots&\vdots&\vdots&\vdots\end{bmatrix},\quad \dots,\quad \eta_{d-1}=\begin{bmatrix}1&0&0&\dots\\0&0&0&\dots\\0&0&0&\dots\\\vdots&\vdots&\vdots&-1\end{bmatrix},$$

$$\eta_d=\begin{bmatrix} 0 & 1 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}, \quad \eta_{d+1}=\begin{bmatrix} 0 & -i & 0 & \dots \\ i & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix},$$

$$\eta_{d+2}=\begin{bmatrix} 0 & 0 & 1 & \dots \\ 0 & 0 & 0 & \dots \\ -1 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}, \quad \eta_{d+3}=\begin{bmatrix} 0 & 0 & -i & \dots \\ 0 & 0 & 0 & \dots \\ i & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}, \quad \dots,$$

Let $H=\Sigma_{ij} c_{ij} \eta_i \otimes \eta_j$ and $H'=\Sigma_{ij} c'_{ij} \eta_i \otimes \eta_j$. To show that $sH' \leqslant_{\mathrm{LU}} H$ for some $s>0$, it suffices to show that $s\,\eta_1 \otimes \eta_1 \leqslant_{\mathrm{LU}} H$, since $\eta_1 \otimes \eta_1 \equiv_{\mathrm{LU}} (\pm \eta_i \otimes \eta_j)$ for all i,j. Furthermore, $\eta_1 \otimes \eta_1$ can be simulated if one can simulate $|i\rangle\langle i| \otimes |j\rangle\langle j|$ and $-|i'\rangle\langle i'| \otimes |j'\rangle\langle j'|$ for any i,j,i',j'.

Without loss of generality, $c_{11} \neq 0$. We first use H to simulate its diagonal components, $H_d=\Sigma_{i,j=1}^{d-1} c_{ij} \eta_i \otimes \eta_j$,

$$H_d=\frac{1}{d^2}\sum_{i,j=0}^{d-1} \begin{bmatrix} 1 & 0 & 0 & \dots \\ 0 & \omega & 0 & \dots \\ 0 & 0 & \omega^2 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}^i \otimes \begin{bmatrix} 1 & 0 & 0 & \dots \\ 0 & \omega & 0 & \dots \\ 0 & 0 & \omega^2 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}^j H \begin{bmatrix} 1 & 0 & 0 & \dots \\ 0 & \omega & 0 & \dots \\ 0 & 0 & \omega^2 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}^{i\dagger} \otimes \begin{bmatrix} 1 & 0 & 0 & \dots \\ 0 & \omega & 0 & \dots \\ 0 & 0 & \omega^2 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}^{j\dagger}. \tag{D1}$$

H_d can further be used to simulate $c_{11}|2\rangle\langle 2| \otimes |2\rangle\langle 2|$, using the protocol

$$\frac{1}{(d-1)^2}\sum_{i,j=0}^{d-2} \begin{bmatrix} 0 & 0 & 1 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}^i \otimes \begin{bmatrix} 0 & 0 & 1 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}^j H_d \begin{bmatrix} 0 & 0 & 1 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}^{i\dagger} \otimes \begin{bmatrix} 0 & 0 & 1 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}^{j\dagger}. \tag{D2}$$

This corresponds to Alice and Bob each applying an averaging over all the computation basis states except for $|2\rangle$. Since all $\eta_{i\neq 1}$ are traceless on the subspace spanned by $|i\neq 2\rangle$, they vanish after the averaging, leaving only a contribution by $\eta_1 \otimes \eta_1$,

$$\frac{c_{11}}{(d-1)^2} \begin{bmatrix} 1 & 0 & 0 & 0 & \dots \\ 0 & -(d-1) & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^{i\dagger} \otimes \begin{bmatrix} 1 & 0 & 0 & 0 & \dots \\ 0 & -(d-1) & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^{i\dagger}, \tag{D3}$$

which is equivalent to $c_{11}|2\rangle\langle 2| \otimes |2\rangle\langle 2|$ up to local terms. It remains to obtain a term with sign opposite to c_{11}. If some $c_{kl} \neq 0$ has a sign opposite to c_{11}, we can simply repeat the same procedure, with Alice applying an averaging over all $|i\neq k\rangle$ and Bob applying an averaging over all $|j\neq l\rangle$. If all c_{ij} have the same sign, Alice can apply an averaging over all $|i\neq 1\rangle$ and Bob can apply an averaging over all $|j\neq 2\rangle$ to obtain $-\Sigma_{i=1}^{d-1} c_{i1}|1\rangle\langle 1| \otimes |2\rangle\langle 2|$, completing the proof.

[1] For a review, see H. Rabitz, R. de Vivie-Riedle, M. Motzkus, and K. Kompa, Science **288**, 824 (2000).

[2] W. Dür, G. Vidal, J.I. Cirac, N. Linden, and S. Popescu, Phys. Rev. Lett. **87**, 137901 (2001).

[3] P. Zanardi, C. Zalka, and L. Faoro, Phys. Rev. A **63**, 040304(R) (2001).

[4] N. Linden, H. Barjat, R. Carbajo, and R. Freeman, Chem. Phys. Lett. **305**, 28 (1999).

[5] D.W. Leung, I.L. Chuang, F. Yamaguchi, and Y. Yamamoto, Phys. Rev. A **61**, 042310 (2000).
[6] J. Jones and E. Knill, J. Magn. Reson. **141**, 322 (1999).
[7] J.L. Dodd, M.A. Nielsen, M.J. Bremner, and R.T. Thew, e-print quant-ph/0106064.
[8] C.H. Bennett, S. Braunstein, I.L. Chuang, D.P. DiVincenzo, D. Gottesman, J.A. Smolin, B.M. Terhal, and W.K. Wootters [unpublished discussion during the 6th Quantum Computation Workshop, ISI, Torino, Italy (1997)].
[9] P. Wocjan, D. Janzing, and Th. Beth, e-print quant-ph/0106077.
[10] D. Leung, J. Mod. Opt. (to be published), e-print quant-ph/0107041v2.
[11] M. Stollsteimer and G. Mahler, e-print quant-ph/0107059v1, to appear in Phys. Rev. A.
[12] D. Janzing, P. Wocjan, and Th. Beth, e-print quant-ph/0106085v1.
[13] N. Khaneja, R. Brockett, and S.J. Glaser, Phys. Rev. A **63**, 032308 (2001).
[14] D. Janzing, e-print quant-ph/0108052.
[15] D. Janzing and Th. Beth, e-print quant-ph/0108053.
[16] G. Vidal and J. Cirac, e-print quant-ph/0108076.
[17] G. Vidal and J. Cirac, e-print quant-ph/0108077.
[18] P. Wocjan, M. Roetteler, D. Janzing, and Th. Beth, e-print quant-ph/0109063.
[19] M.A. Nielsen, M.J. Bremner, J.L. Dodd, A.M. Childs, and C.M. Dawson, e-print quant-ph/0109064.
[20] P. Wocjan, M. Roetteler, D. Janzing, and Th. Beth, e-print quant-ph/0109088.
[21] H. Chen, e-print quant-ph/0109115.
[22] A. Barenco, C. Bennett, R. Cleve, D. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H. Weinfurter, Phys. Rev. A **52**, 3457 (1995); for a complete list of work, see Ref. [25].
[23] L. Viola, S. Lloyd, and E. Knill, Phys. Rev. Lett. **83**, 4888 (1999).
[24] D. Gottesman, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 1997; e-print quant-ph/9705052.
[25] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).
[26] The forward implication follows from Exercise II.1.9 in Ref. [28] when applied to the vectors of singular values of M and N, whereas the converse implication follows from the triangle inequality of the Ky Fan k-norms (Exercise II.1.15 in Ref. [28]).
[27] M.A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).
[28] R. Bhatia, *Matrix Analysis* (Springer-Verlag, New York, 1997).
[29] Let $H=H_1\otimes H_2$, H_1,H_2 be members of a traceless orthogonal basis (under the trace norm) with different eigenvalues. We show that $\mathcal{S}(H)\not\leqslant_{\mathrm{LU}} H$, which also implies $H\not\leqslant_{\mathrm{LU}}\mathcal{S}(H)$. We can renormalize both H and $\mathcal{S}(H)$ so that the basis is orthonormal. The condition $\mathcal{S}(H)\leqslant_{\mathrm{LU}} H$ implies $H_2\otimes H_1=\Sigma_i p_i(U_iH_1U_i^\dagger)\otimes(V_iH_2V_i^\dagger)$ for some unitary U_i,V_i. Multiplying this with $H_2\otimes H_1$ and taking the trace, $\mathrm{tr}(H_2^2)\mathrm{tr}(H_1^2)=\Sigma_i p_i\mathrm{tr}(H_2U_iH_1U_i^\dagger)\mathrm{tr}(H_1V_iH_2V_i^\dagger)$. However, $\mathrm{tr}(H_2U_iH_1U_i^\dagger)\leqslant\mathrm{tr}(H_2^2)$ and $\mathrm{tr}(H_1V_iH_2V_i^\dagger)\leqslant\mathrm{tr}(H_1^2)$, with equalities hold only when $U_iH_1U_i^\dagger=H_2$ and $V_iH_2V_i^\dagger=H_1$, implying H_1 and H_2 need to have the same eigenvalues. It is well known that, in d dimensions, the traceless matrices $\mathrm{diag}(1,1,\ldots,1,-k,0,0,\ldots,0)$ are part of a traceless orthogonal basis for Hermitian $d\times d$ matrices (for example, the Gell-Mann matrices) completing the proof.