QMA, **QMA**-Complete LOCAL-HAMILTONIAN, & Beyond

Andrew Yang, Mark Chen

12 April 2024

An Introduction: QMA and Local-Hamiltonian

1 Quantum Merlin-Arthur

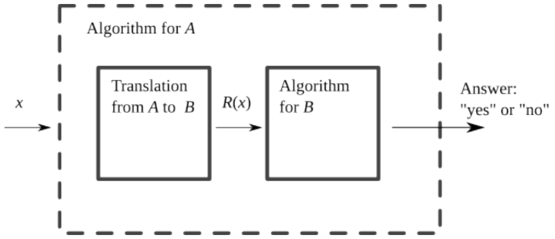
1.1 Classical Motivation

In classical complexity, non-deterministic polynomial time (**NP**) is a class of decision problems that can be verified in polynomial time. Decision problems are problems where each language L is split into yes (1) and no (0) answers $L = (L_1, L_0)$.

Definition 1. *A decision problem $L = (L_1, L_0) \in$ **NP** iff there exists a deterministic verifier algorithm $V(x, y)$ such that*

1. *(Efficiently Verifiable) A runs in poly-time with respect to $n = |x|$.*
2. *(Completeness) If $x \in L_1$, there exists a string $y \in \{0,1\}^{p(n)}$ such that $V(x, y) = 1$.*
3. *(Soundness) If $x \in L_0$, $\forall y \in \{0,1\}^{p(n)}$, $A(x, y) = 0$.*

Definition 2 (**NP**-Completeness)**.** *Any **NP** problem can be poly-time reduced to any **NP**-complete problem ($\exists$ a poly-time deterministic algorithm that maps all YES/NO instances of the original problem to the **NP**-complete problem so that solving the complete problem solves the original problem, up to the difference of a poly-time run-time factor).*



Example 1 (**NP**-Complete Problems)**.** *The most famous such problem is* 3-SAT *(proved by Cook-Levin Theorem). Other examples include* TRAVELING-SALES-PERSON *(*TSP*) and* 0/1-INTEGER-PROGRAMMING*. At this point there are tens of thousands of **NP**-complete problems that imply various applications.*

1.2 Quantum Analogy of NP

QMA is defined by slackening the conditions for **NP** to allow probabilistic in the process.

Definition 3 (**QMA**)**.** *A promise problem $L = (L_1, L_0, L_*)$ is in **QMA**(b, a) iff there exists a uniform family $\{C_n\}$ of poly-size quantum circuits that take in two input registers, $x \in \{0,1\}^n$, and output a single qubit:*

1. *(Completeness) If $x \in L_1 \cap \{0,1\}^n$, then $\exists p(n)$-qubit state $|\psi\rangle$ s.t.*

$$\Pr[C_n(x, |\psi\rangle) = 1] \geq b.$$

2. *(Soundness) If $x \in L_1 \cap \{0,1\}^n$, then $\forall p(n)$-qubit state $|\psi\rangle$ s.t.*

$$\Pr[C_n(x, |\psi\rangle) = 1] \leq a.$$

Proposition 1. *Generally, the class* **QMA** *is defined as* **QMA**$(2/3, 1/3)$*, but we can prove that for $b - a \geq 1/poly(n)$,* **QMA**$(b, a) =$ **QMA**$(2/3, 1/3)$*.*

Proof. The proof is the same as the classical **BPP** amplification proof, which states that as long as there is an inverse polynomial gap between the accepting and rejecting probabilities, one can always amplify the probability to be equivalent to $(2/3, 1/3)$.

One thing that was thought to also be different is that such a naive approach would actually increase the witness-size by a factor of $O(\log(1/\delta))$ as well, which is not ideal. [MW05] came up with a surprising yet beautiful approach that, instead, increases the run-time by a factor of $O(\log(1/\delta))$ and leaves the length of the witness being $p(n)$ qubits. ■

Remark 1. *What is special about* **QMA** *compared to* **NP** *and* **BPP***? The quantum power comes at the properties that:*

1. *(Quantum Witness): The witness is a quantum state, $|\psi\rangle$.*
2. *(Quantum Verifier): The verifier algorithm is a circuit that takes in two quantum registers and and outputs a qubit.*

We call the special case where we take out the first property and make the witness classical bit string the **QCMA** *class ("C" for classical). Then, we call the special case where we make the verifier algorithm classical* **MA** *(this step cannot be done before the first special case, because as long as the witness is still quantum it doesn't make sense to run it by a classical verifier).*

Proposition 2. *By the chain of special cases in remark 1, it is easy to see that*

$$\mathbf{NP} \subseteq \mathbf{MA} \subseteq \mathbf{QCMA} \subseteq \mathbf{QMA}$$

Conjecture 1. *It is conjectured that* **MA** $=$ **NP** *and* **NP** $\subsetneq$ **QMA***. In other words, it is believed that "quantum proofs" can efficiently prove more than "classical proofs" can.*

1.3 Open Problem

The following problem is in **QMA**, but it is unknown if it is **QMA**-complete and unknown if it is in **NP**.

Definition 4. *The group non-membership problem. Consider a finite group G defined by its generator, subgroup $H \leq G$, and element $g \in G$. (L_1) $g \notin H$. (L_0) $g \in H$. First of all, notice that it should be easy to show that group membership problem is easily in* **NP***. Now, consider this problem, the verifier is the superposition of all states in H, and all you need to show that one of them is indeed g.*

1.4 Beyond MA & QMA

The reason why **QMA** is a quantum analogue of **NP** is really due to the fact that it is a quantum analogue of **MA**. Compare the following two definitions:

- ($L \in$ **NP**) "$x \in L \iff \exists y$, such that $V(x, y) = 1$ where V runs in poly-time."

- ($L \in \mathbf{MA}$) "$x \in L \implies \exists y$, such that $\Pr_z [V(x,y,z)=1] \geq \frac{2}{3}$ where V runs in poly-time (clearly we also need the $x \notin L$ case)."

Remark 2. *Notice that* **MA** *is only different from* **NP** *in that it allows probabilistic power. This should explain why the conjecture 1 formulates that* $\mathbf{MA} = \mathbf{NP}$ *just like how it is conjectured that* $\mathbf{P} = \mathbf{BPP}$*.*

Remark 3. *Again,* **QMA** *and* **MA** *are only different, literally, in that* **QMA** *takes a witness as a* $p(n)$*-qubit state and its verifier can be a quantum circuit.*

1.4.1 Polynomial Hierarchy

Notice that **NP** language can be characterized by a single existential statement (described as the existence of a poly-size witness). It turns out this can be a lot more generalized to be the following two kinds of classes:

- ($L \in \Pi_i^p$) $x \in L \iff \forall y_1, \exists y_2, \ldots \exists/\forall y_i, V(x, y_1, y_2, \ldots, y_i) = 1$.
- ($L \in \Sigma_i^p$) $x \in L \iff \exists y_1, \forall y_2, \ldots \forall/\exists y_i, V(x, y_1, y_2, \ldots, y_i) = 1$.

This generalization will fill up the entire polynomial hierarchy, denoted **PH**, which looks like:

$$\begin{tikzcd}
 & \vdots & \\
\Sigma_3^{\mathrm{P}} & & \Pi_3^{\mathrm{P}} \\
 & \Delta_3^{\mathrm{P}} \arrow[ul] \arrow[ur] & \\
\Sigma_2^{\mathrm{P}} \arrow[uu] \arrow[ur] & & \Pi_2^{\mathrm{P}} \arrow[uu] \arrow[ul] \\
 & \mathrm{P}^{\mathrm{NP}} = \Delta_2^{\mathrm{P}} \arrow[ul] \arrow[ur] & \\
\mathrm{NP} = \Sigma_1^{\mathrm{P}} \arrow[uu] \arrow[ur] & & \Pi_1^{\mathrm{P}} = \mathrm{coNP} \arrow[uu] \arrow[ul] \\
 & \Delta_0^{\mathrm{P}} = \Sigma_0^{\mathrm{P}} = \mathrm{P} = \Pi_0^{\mathrm{P}} = \Delta_1^{\mathrm{P}} \arrow[ul] \arrow[ur] &
\end{tikzcd}$$

Definition 5 (Totally Quantifiable Boolean Function (TQBF))**.** *This problem should capture the form of each specific instance of* **PH***, as it has the following form:*

$$Q_1 x_1, Q_2 x_2, \ldots, Q_n x_n, \phi(x_1, \ldots, x_n),$$

where each Q_i *and* Q_{i+1} *are alternating between* $\exists$ *and* $\forall$*, starting from either* $\exists$ *or* $\forall$*.* ϕ *is a Boolean formula with* $x_1, \ldots, x_n$ *as variables. The goal is to decide if* $\phi(x_1, \ldots, x_n)$ *is satisfiable.*

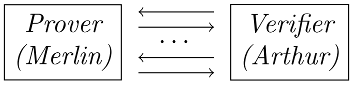
Proposition 3. *Any* Π_i^p *and* Σ_i^p *can be reduced to* TQBF*. This implies that* $\mathbf{PH} \subseteq \mathbf{PSPACE}$ *(it is not presently known if the equality holds).*

Proposition 4. TQBF *is* **PSPACE***-complete.*

1.4.2 Interactive Proof System (IP)

Definition 6 (**MA**(i))**.** *Notice that what* **MA** *really did is a protocol that captures the probabilistic version of a* Π_1^p *problem. So, we can easily generalize that to a bigger* i*, by letting there be multiple rounds of communications between Merlin (the all-powerful prover) and Arthur (the computationally limited verifier). That generalization, is what we call* $\mathbf{MA}(i)$*.*

Definition 7 (**IP**)**.** *Is the union of all* $\mathbf{MA}(i)$*, plus it can be both directions (i.e. both the prover and the verifier can be the one to start the first talk).*



Theorem 1. *Due to the strongly correlated nature of* **PSPACE** *and* **IP** *rooted from the motivation of their definitions, there should be at least some intuitions at this point that the following is true:*

$$\mathbf{IP} = \mathbf{PSPACE}\ [\textit{Sha92}].$$

Definition 8 (**QIP**)**.** *We give* **IP** *the additional power:*

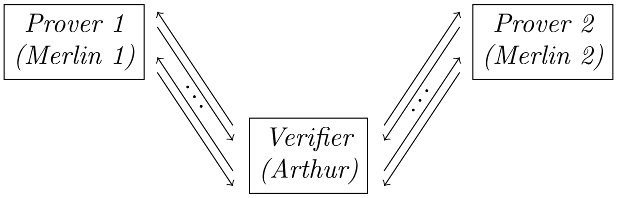
- *The verifier can be a* **BQP** *verifier.*
- *The messages sent can be quantum.*

Theorem 2. *It was actually shown that, at this point, quantum doesn't give any extra power (which makes the next result all the more surprising):*

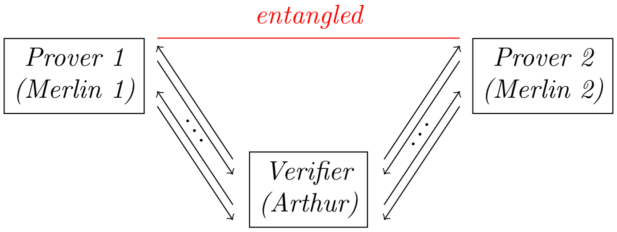
$$\mathbf{IP} = \mathbf{PSPACE} = \mathbf{QIP}\ [\textit{JJUW11}].$$

1.4.3 Multiprove Interactive Proof System (MIP)

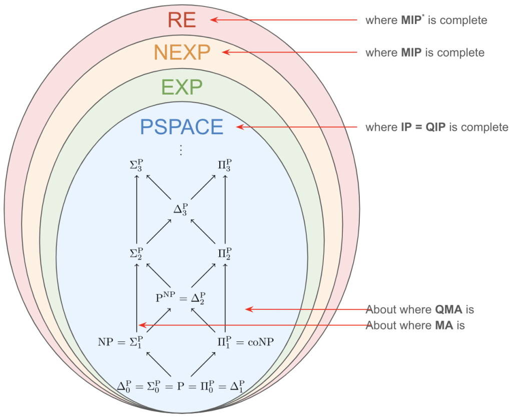
Definition 9. *The same as* **IP** *except that there can now be multiple provers:*



Definition 10 (**MIP***)**.** *:*



1.4.4 Summary



2 QMA Complete Problems

Definition 11. *The non-identity check problem. Given an n-qubit polynomial circuit C, determine whether this circuit is non-trivial up to a phase. (L_1) For all $\phi \in [0, 2\pi)$, $||C - e^{i\phi} I_{2^n}|| \geq b$. ($L_0$) There is some $\phi \in [0, 2\pi)$ such that $||C - e^{i\phi} I_{2^n}|| \leq a$. (Promise) Either L_1 or L_0 is the case and $b - a \geq 1/poly(n)$.*

Definition 12. *The k-local matrix consistency problem. Consider $m \leq poly(n)$ density matrices $\{\rho_i\}_{i=1}^m$, where ρ_i depends only on the set of qubits Q_i with $|Q_i| \leq k$. Denote $Q = \{1, \cdots, n\}$ as the set of all qubits. (L_1) There is a consistent n-qubit density matrix ρ, meaning for all $i \in \{1, \cdots, m\}$, the partial trace $tr_{Q \backslash Q_i}(\rho) = \rho_i$. ($L_0$) All matrices ρ have a significant non-consistency, meaning there exists some $i \in \{1, \cdots, m\}$ such that $|tr(tr_{Q \backslash Q_i}(\rho) - \rho_i)| \geq b$. (Promise) Either L_1 or L_0 is the case and $b \geq 1/poly(n)$.*

2.1 The Local Hamiltonian Problem

2.1.1 The Hamiltonian

In a molecular/cluster system involving n electrons, the energy of these electrons is the sum of their kinetic energy $\hat{T}$, their electric potential energy from atomic nuclei $\hat{V}$, and electron-electron repulsion energy $\hat{U}$. The Hamiltonian for this system is

$$\hat{H} = \hat{T} + \hat{V} + \hat{U}. \tag{1}$$

In general, the Hamiltonian is the sum of all the energy operators of a system.

All observables $\hat{O}$ (position, momentum, energy, etc.) of a state $|\psi\rangle$ are described by operators in quantum mechanics. The observed value is given by the expectation value of the operator with the state

$$\langle\psi| \hat{O} |\psi\rangle . \tag{2}$$

When $|\lambda\rangle$ is a normalized eigenvector of $\hat{O}$ with eigenvalue λ,

$$\langle\lambda| \hat{O} |\lambda\rangle = \lambda \tag{3}$$

Since all observables are real-valued in the physical world, we constrain all eigenvalues of observables to be real. This can be done by enforcing $\hat{O}$ to be Hermitian ($\hat{O} = \hat{O}^\dagger$, where † represents the conjugate transpose). Another neat consequence of making observables Hermitian is that the eigenvectors of $\hat{O} : \mathbb{C}^d \to \mathbb{C}^d$ span the full rank d of the space $\mathbb{C}^d$ (see spectral theorem of Hermitian matrices). This extends to $d \to \infty$. More specifically, we can diagonalize the Hamiltonian is

$$\hat{H} = \sum_n E_n |n\rangle \langle n| , \tag{4}$$

where $|n\rangle$ are the eigenvectors of $\hat{H}$ with eigenvalue E_n. The energy of an arbitrary state can thus be written as the following expectation value

$$\begin{aligned} \langle\psi| \hat{H} |\psi\rangle &= \langle\psi| \left(\sum_n E_n |n\rangle \langle n| \right) |\psi\rangle && (5) \\ &= \sum_n E_n \langle\psi|n\rangle \langle n|\psi\rangle && (6) \\ &= \sum_n E_n || \langle\psi|n\rangle ||^2. && (7) \end{aligned}$$

To maximize entropy per the second law of thermodynamics, at thermal equilibrium, a system distributes its states according to the probability distribution

$$Pr(|\psi\rangle = |n\rangle) \propto \exp\left\{ -\frac{E_n}{k_B T} \right\} \tag{8}$$

meaning the lowest-energy state E_0, often called the ground state, has the highest probability of occurring. If we cool the system such that $T \to 0$, it is the only state that will exist.

In physics, we generally care about time-independent eigenstates, though knowing these states allows us to determine the time-evolution of any arbitrary state. By Schrödinger's equation,

$$\hat{H} \left|\psi(t, \vec{x})\right\rangle = i\hbar \frac{\partial}{\partial t} \left|\psi(t, \vec{x})\right\rangle . \tag{9}$$

the time-evolution of a system depends on the Hamiltonian. The Hamiltonian is generally has no explicit time dependence (e.g. the electric potentials between two charged objects depends only on their position), so it makes some sense to separate the state into time and position components.

$$\psi(t, \vec{x}) = \phi(t)\psi(\vec{x}). \tag{10}$$

Let $\left|n(\vec{x})\right\rangle$ be an eigenvector of $\hat{H}$ with corresponding eigenvalue E.

$$\hat{H}\psi(t, \vec{x}) = E\phi_n(t) \left|n(\vec{x})\right\rangle . \tag{11}$$

The letter E stands for energy. Plugging this into Schrödinger's equation gives

$$i\hbar \frac{\partial}{\partial t} \phi_n(t) \left|n(\vec{x})\right\rangle = E\phi_n(t) \left|n(\vec{x})\right\rangle \tag{12}$$

$$\phi_n(t) = \exp\left\{-i\frac{E}{\hbar}t\right\}. \tag{13}$$

Thus, the time-evolution of the eigenstate is described as a simple oscillatory function. At any moment in time $t = T$, any arbitrary state $\left|\Psi(t = T, \vec{x})\right\rangle$ can be represented as a sum of the eigenstates of the Hamiltonian as the eigenvectors span the full rank. If at $t = 0$,

$$\left|\Psi(0, \vec{x})\right\rangle = \sum_n \alpha_n \left|n(\vec{x})\right\rangle \tag{14}$$

for arbitrary weights α_n, the time-evolved state is just

$$\left|\Psi(t, \vec{x})\right\rangle = \sum_n \alpha_n \phi_n(t) \left|n(\vec{x})\right\rangle \tag{15}$$

$$= \sum_n \alpha_n \exp\left\{-i\frac{E_n}{\hbar}t\right\} \left|n(\vec{x})\right\rangle . \tag{16}$$

2.1.2 The Classical-Ising Model

Minimization of Hamiltonians can be used to solve classical problems.

Let $G = (V, E)$ be a graph with n verticies in set V and pairs of verticies in the set E of edges. Define Z_i to be the Pauli z-matrix acting on qubit i. Then, the Max-Cut Hamiltonian is

$$\hat{H} = \sum_{e=(u,v)\in E} Z_u \otimes Z_v \tag{17}$$

$$= \sum_{(u,v)\in E} \left|00\right\rangle \left\langle 00\right|_{u,v} - \left|01\right\rangle \left\langle 01\right|_{u,v} - \left|10\right\rangle \left\langle 10\right|_{u,v} + \left|11\right\rangle \left\langle 11\right|_{u,v} \tag{18}$$

$$= \sum_{x\in\{0,1\}^n} (m - 2c(x)) \left|x\right\rangle \left\langle x\right| , \tag{19}$$

where $c(x)$ is the number of satisfied constraints (number of edges that will be cut) and $m = |E|$.

This is similar to the real-world quantum Heisenburg model subject to a transverse magnetic field on the x-axis. For some number of qubits arranged in a ring, the Hamiltonian for this model is

$$H(J_x, J_y, J_z, g) = J_x \sum_i X_i \otimes X_{i+1} + J_y \sum_i Y_i \otimes Y_{i+1} + J_z \sum_i Y_i \otimes Y_{i+1} + g \sum_i X_i, \tag{20}$$

where $J_x, J_y, g = 0$ gives the Max-Cut problem Hamiltonian and $J_x = J_y = J_z, g = 0$ gives something called the quantum Max-Cut problem. Minimizing the 2D extension to the Heisenburg Hamiltonian is **QMA**-complete.

2.2 The Local Hamiltonian Problem

Definition 13. *A k-local n-qubit Hamiltonian $\hat{H}$ acts trivially on only $k \leq n$ qubits. In other words, the $(2^n \times 2^n)$-dimensional $\hat{H}$ can be fully described as the tensor product of a $(2^k \times 2^k)$-dimensional tensor and identity for the other $n - k$ qubits.*

Definition 14. *The k-local Hamiltonian problem for a n-qubit Hamiltonian is a promise problem where (i) the Hamiltonian satisfies*

$$\hat{H} = \sum_{i=1}^{m} \hat{H}_i, \tag{21}$$

where each $\hat{H}_i$ are k'-local where $k' \leq k$ and $0 \preccurlyeq \hat{H}_i \preccurlyeq \hat{I}_{2^n}$, and (ii) given $a, b \in [0, m]$ with $b - a \geq poly(n)$, decide whether the minimum eigenvalue λ_0 of $\hat{H}$ is $\leq a$ or $\geq b$ promised it is in one of the two categories.

- *The accepted languages L_1 are the set of Hamiltonians where $\lambda_0 \leq a$.*
- *The rejected are those with $\lambda_0 \geq b$, and we are promised we do not get L_*, where $a < \lambda_0 < b$.*

In general, $\hat{H}_i$ may satisfy the condition $0 \preccurlyeq \hat{H} \preccurlyeq \hat{I}_{2^n}$. However, we can always normalize $\hat{H}$ such that the $-I_{2^n} \preccurlyeq \hat{H}_i \preccurlyeq I_{2^n}$. Then, we can define a new positive-semidefinite Hamiltonian $\hat{H}'_i = (\hat{H} + I_{2^n})/2$ that satisfies $0 \preccurlyeq \hat{H}'_i \preccurlyeq I_{2^n}$. Under this transformation, the eigenvalue λ'_0 of $\hat{H}' = \sum_{i=1}^{m} \hat{H}'_i$ is given by $\lambda'_0 = (\lambda_0 + m)/2$.

Theorem 3. *There is a version of the k-local Hamiltonian problem that is* **QMA***-complete.*

Proof. It is easy to verify that the k-local Hamiltonian problem is in **QMA**. To show completeness, for a promise problem L in **QMA**, we can construct a map from all instances $x \in L$ to a Feynman-Kitaev Hamiltonian. For a given verifier circuit C_n, the input is the n-qubit string x, the s-qubit ancilla, and the $w(n) \leq \text{poly}(n)$ verifier. An additional $T + 1 \leq \text{poly}(n)$ time-keeping states are added for each gate $\{U_t\}_{t=1}^{T}$ where $C_n = U_T \cdots U_1$.

At the initial state, a penalty Hamiltonian

$$\hat{H}_{init} = \sum_{i=1}^{s} |1\rangle \langle 1|_{A,i} \otimes |0\rangle \langle 0|_C \tag{22}$$

checks to make sure the ancilla (denoted A) is initialized to zero and the proper input x (denoted X) is given. The clock (denoted C) state is $|t = 0\rangle$.

Then, an additional T Hamiltonians ($t \in \{1, \cdots, T\}$) assign penalties for deviating from the circuit

$$\hat{H}_t = \frac{1}{2} \left(I \otimes (|t-1\rangle \langle t-1|_C + |t\rangle \langle t|_C) - U_t \otimes |t\rangle \langle t-1|_C - U_t^* \otimes |t-1\rangle \langle t|_C \right). \tag{23}$$

Finally, a penalty is assigned for a 0-measurement in the output (treat a 1-measurement as accepting)

$$\hat{H}_{fin} = |0\rangle\langle 0|_1 \otimes |T\rangle\langle T|_C . \tag{24}$$

The total Feynman-Kitaev (penalty) Hamiltonian is just

$$\hat{H} = \hat{H}_{init} + \sum_{t=1}^{T} \hat{H}_t + \hat{H}_{fin}. \tag{25}$$

Completeness and soundness can be shown for this Hamiltonian with well-defined a and b. Furthermore, as each quantum gate acts on at most 2 qubits and the clock register uses $\lceil \log(T+1) \rceil$ bits, we have k-locality where $k = \lceil \log(T+1) \rceil + 2 = O(\log(n))$. This can be further reduced to a constant. [dW23] ■

References

[dW23] Ronald de Wolf. Quantum computing: Lecture notes. 2023.

[JJUW11] Rahul Jain, Zhengfeng Ji, Sarvagya Upadhyay, and John Watrous. Qip = pspace. *J. ACM*, 58(6), dec 2011.

[MW05] Chris Marriott and John Watrous. Quantum arthur–merlin games. *computational complexity*, 14(2):122–152, 2005.

[Sha92] Adi Shamir. Ip = pspace. *J. ACM*, 39(4):869–877, oct 1992.