Chapter from book "Classical and quantum computation" by A. Kitaev, A. Shen and M. Vyalyi

14. The quantum analogue of NP: the class BQNP

It is possible to construct quantum analogues not only for the class P, but also for other classical complexity classes. This is not a routine process, but suitable generalizations often come up naturally. We will consider the class NP as an example. (For another example — the class IP and its quantum analogue QIP — see [**72, 38**]. We also mention that the quantum analogue of PSPACE equals PSPACE [**71**].)

14.1. Modification of classical definitions. Quantum computation, as well as probabilistic computation, is more naturally described using partially defined functions. Earlier we made do without this concept so as not to complicate matters by the inclusion of extraneous detail, but now we need it.

A *partially defined Boolean function* is a function

$$F : \mathbb{B}^n \to \{0,\ 1,\ \text{“undefined”}\}.$$

In this section it will be tacitly understood that by Boolean function we mean partially defined Boolean function.

One more comment regarding notation: we have used the symbol P both for the class of polynomially computable functions and for the class of polynomially decidable predicates; now we act analogously, using the notations P, NP, etc. for classes of partially defined functions.

P, of course, denotes the class of polynomially computable partially defined functions. We introduce a modified definition of the class NP.

Definition 14.1. A function $F : \mathbb{B}^n \to \{0,\ 1,\ \text{“undefined”}\}$ belongs to the class NP if there is a partially defined function $R \in \mathrm{P}$ in two variables such that

$$\begin{aligned} F(x) = 1 &\implies \exists\, y\ \big((|y| < q(|x|)) \wedge (R(x,y) = 1)\big) \\ F(x) = 0 &\implies \forall\, y\ \big((|y| < q(|x|)) \Rightarrow (R(x,y) = 0)\big). \end{aligned}$$

As before, $q(\cdot)$ is a polynomial.

What would change if in Definition 14.1 we replaced the condition $R \in \mathrm{P}$ by the condition $R \in \mathrm{BPP}$? First of all, we would get a different, broader, class, which we could denote by BNP. However, for this class there is another, standard, notation — MA, indicating that it falls into a hierarchy of classes defined by *Arthur-Merlin games*. We have mentioned Arthur and Merlin in connection with the definition of NP. We have also discussed games corresponding to other complexity classes (see Section 5.1). Traditionally, the term “Arthur-Merlin games” is used for probabilistic games in which Arthur is a polynomial Turing machine whereas Merlin is all-powerful;

before each move Arthur flips coins so that both players see them. The order of the letters in the symbol MA indicates the order of the moves: at first Merlin communicates y , then Arthur checks the truth of the predicate $R(x, y)$, by a polynomial probabilistic computation. The message y is sometimes called a "proof"; it may be hard to find but easy to check.

14.2. Quantum definition by analogy.

Definition 14.2. A function $F : \mathbb{B}^n \to \{0, 1, \text{"undefined"}\}$ belongs to the class BQNP if there exists a polynomial classical algorithm that computes a function $x \mapsto Z(x)$, where $Z(x)$ is a description of a quantum circuit, realizing an operator $U_x : \mathcal{B}^{\otimes N_x} \to \mathcal{B}^{\otimes N_x}$ such that

$$F(x) = 1 \implies \exists\, |\xi\rangle \in \mathcal{B}^{\otimes m_x}\ \mathbf{P}\Big(U_x|\xi\rangle \otimes |0^{N_x - m_x}\rangle,\ \mathcal{M}\Big) \geq p_1,$$

$$F(x) = 0 \implies \forall\, |\xi\rangle \in \mathcal{B}^{\otimes m_x}\ \mathbf{P}\Big(U_x|\xi\rangle \otimes |0^{N_x - m_x}\rangle,\ \mathcal{M}\Big) \leq p_0.$$

Here $\mathcal{M} = \mathbb{C}\big(|1\rangle\big) \otimes \mathcal{B}^{\otimes(N_x - 1)}$, and p_0, p_1 satisfy the condition $p_1 - p_0 \geq \Omega(n^{-\alpha})$ for some constant $\alpha \geq 0$. The quantifiers of $|\xi\rangle$ include only vectors of unit length. (We will use an analogous convention further on in this section, pushing numeric factors outside the $|\cdot\rangle$ sign.)

The vector $|\xi\rangle$ plays the role of y in the previous definition. Note that $m_x \leq N_x \leq |Z(x)| = \text{poly}(|x|)$ since the algorithm is polynomial.

In effect, the very same game of Merlin with Arthur is taking place, but now it is governed by the laws of quantum mechanics. Merlin sends a quantum message (the state $|\xi\rangle$) to Arthur, who checks it by applying the operator U_x. A suitable message will convince Arthur that $F(x) = 1$ (if this is actually so) with probability $\geq p_1$. But if $F(x) = 0$, Merlin cannot succeed in convincing Arthur to the contrary with probability higher than p_0, whatever message he sends. Instead of a pure state $|\xi\rangle$, we can allow Merlin to send an arbitrary density matrix — the maximum of the probability is achieved on a pure state anyway.

In Definition 14.2, we have the same flexibility in choosing the threshold probabilities p_0 and p_1 as in the definitions of BPP and BQP.

Lemma 14.1 (amplification of probabilities). *If $F \in$ BQNP, then it likewise satisfies a variant of Definition* 14.2 *in which the numbers p_0, p_1 ($p_1 - p_0 = \Omega(n^{-\alpha})$) are replaced by*

$$p_1' = 1 - \varepsilon, \quad p_0' = \varepsilon, \qquad \varepsilon = \exp(-\Omega(n^\beta)),$$

where β is an arbitrary positive constant.

Proof. The general idea of amplifying the probabilities remains as before: we consider $k = \text{poly}(n)$ copies of the circuit realizing the operator $U = U_x$.

To the results of their work we apply a variant of the majority function, with the threshold value adjusted so as to separate p_0 from p_1:

$$G(z_1,\dots,z_k)=\begin{cases}1 & \text{if } \sum_{j=1}^k z_j \geq pk,\\ 0 & \text{if } \sum_{j=1}^k z_j < pk,\end{cases} \tag{14.1}$$

where $p=(p_0+p_1)/2$. But now there appears an additional difficulty: Merlin may attempt to deceive Arthur by sending him a message that is not factorable into the tensor product.

Let us grant Merlin greater freedom, allowing him to submit any density matrix $\rho \in \mathbf{L}(\mathcal{B}^{\otimes km})$. The probability of obtaining outcomes $z_1,\dots,z_k$ by applying k copies of U to the message ρ is as follows:

$$\mathbf{P}(z_1,\dots,z_k|\,\rho)=\mathrm{Tr}\big(X^{(z_1)}\otimes\cdots\otimes X^{(z_k)}\rho\big), \tag{14.2}$$

where

$$X^{(a)}=\mathrm{Tr}_{[m+1,\dots,N]}\Big(U^\dagger \Pi_1^{(a)} U\,\big(I_{\mathcal{B}^{\otimes m}}\otimes|0^{N-m}\rangle\langle 0^{N-m}|\big)\Big). \tag{14.3}$$

Here $\Pi_1^{(a)}$ is the projection onto the subspace of states having a in the first qubit (i.e., $\mathbb{C}\big(|a\rangle\big)\otimes\mathcal{B}^{\otimes(N-1)}$).

If $F(x)=1$, Merlin can simply send the state $\rho=\rho_x^{\otimes k}$, where $\rho_x=|\xi_x\rangle\langle\xi_x|$ is the message that would convince Arthur with probability $\geq p_1$ in the original version of the game (with a single copy of U). By the general properties of quantum probability, formula (14.2) takes the form

$$\mathbf{P}(z_1,\dots,z_k|\,\rho)=\prod_{j=1}^k \mathrm{Tr}(X^{(z_j)}\rho_x)=\prod_{j=1}^k \mathbf{P}(z_j|\rho_x).$$

We will derive a lower bound for this quantity from a more general analysis given below.

In the opposite case, $F(x)=0$, we will obtain an upper bound for the probability $\mathbf{P}(z_1,\dots,z_k|\rho)$ over all density matrices ρ.

Let us select an orthonormal basis in the space $\mathcal{B}^{\otimes m}$, in which the operator $X^{(1)}$ is diagonalized (this operator is clearly Hermitian). The operator $X^{(0)}=I-X^{(1)}$ is diagonal in the same basis. We define a set of "conditional probabilities" $p(z|d)=\mathbf{P}\big(z\big||d\rangle\langle d|\big)=\langle d|X^{(z)}|d\rangle$, where $|d\rangle$ is one of the basis vectors. It is obvious that $p(z|d)\geq 0$ and $p(0|d)+p(1|d)=1$. In this notation, the quantity $\mathbf{P}(z_1,\dots,z_k|\rho)$ becomes

$$\mathbf{P}(z_1,\dots,z_k|\,\rho)=\sum_{d_1,\dots,d_k} p_{d_1,\dots,d_k}\,p(z_1|d_1)\cdots p(z_k|d_k),\qquad \sum_{d_1,\dots,d_k} p_{d_1,\dots,d_k}=1,$$

where $p_{d_1,\dots,d_k}=\langle d_1,\dots,d_k|\rho|d_1,\dots,d_k\rangle$.

This formula has the following interpretation. Consider the set of probabilities $\mathbf{P}(z_1,\dots,z_k|\rho)$ for all sequences $(z_1,\dots,z_k)\in\mathbb{B}^k$ as a vector

in a 2^k-dimensional real space; we denote this vector by $\overrightarrow{\mathbf{P}}(\rho) \in \mathbb{R}^{\mathbb{B}^k}$. We have shown that for a general density matrix ρ the vector $\overrightarrow{\mathbf{P}}(\rho)$ belongs to the convex hull of such vectors corresponding to product states, namely, $|d_1\rangle\langle d_1| \otimes \cdots \otimes |d_k\rangle\langle d_k|$. Therefore the probability of the event $G(z_1, \ldots, z_k) = 1$,

$$\mathbf{Pr}\big[G(z_1, \ldots, z_k) = 1 \big| \rho\big] = \sum_{z \in \mathbb{B}^k} G(z)\mathbf{P}(z|\rho) = \Big(\overrightarrow{G}, \overrightarrow{\mathbf{P}}(\rho)\Big),$$

achieves its maximum at a density matrix of this special type.

In the case where G is the threshold function (14.1),

$$p_{\max} \stackrel{\text{def}}{=} \max_{\rho} \mathbf{Pr}\big[G(z_1, \ldots, z_k) = 1 \big| \rho\big] = \sum_{j \geq l} \binom{k}{j} p_*^j (1 - p_*)^{k-j},$$

where $p_* = \max_{|\xi\rangle} \langle \xi | X^{(1)} | \xi \rangle$. The number $p_{\max}$ equals the probability of getting $\geq pk$ "heads" for k coins tossed, $\mathbf{Pr}\Big[k^{-1} \sum_{j=1}^k v_j \geq p\Big]$, where $\mathbf{Pr}\big[v_j = 1\big] = p_*$. This probability can be estimated using Chernoff's inequality (13.4). Thus we obtain[11]

$$\begin{aligned} p_{\max} &\leq \exp(-2(p - p_*)^2 k) \qquad \text{if } p \geq p_*, \\ p_{\max} &\geq 1 - \exp(-2(p - p_*)^2 k) \quad \text{if } p \leq p_*. \end{aligned}$$

According to the assumptions of the lemma, $p_* \leq p_0$ if $F(x) = 0$, and $p_* \geq p_1$ if $F(x) = 1$. We have chosen p so that $p - p_0 \geq \Omega(n^{-\alpha})$ and $p_1 - p \geq \Omega(n^{-\alpha})$. Choosing $k = cn^{2\alpha+\beta}$ (for a suitable constant c), we get exactly the estimate which is required,

$$\begin{aligned} p_{\max} &\leq \exp(-\Omega(n^\beta)) \qquad \text{if } F(x) = 0, \\ p_{\max} &\geq 1 - \exp(-\Omega(n^\beta)) \quad \text{if } F(x) = 1. \end{aligned}$$

□

Remark 14.1. An important point in the proof was the fact that $X^{(0)}$ and $X^{(1)}$ are diagonalized over the same basis. In general, the amplification of probability for nontrivial complexity classes (both classical and quantum) is a rather subtle thing.

14.3. Complete problems. Similarly to the class NP, the class BQNP has complete problems. Completeness is understood with respect to the same polynomial reduction that we considered earlier (i.e., Karp reduction; see Definition 3.4). Here is the simplest example.

[11] We have omitted the factor 2 from (13.4) because it includes both cases that are now considered separately (see the proof of Chernoff's inequality in the solution to Problem 13.3). This is an unimportant factor anyway.

PROBLEM 0. Consider a function F which is defined on a subset of the words of this form:

$$z = \big((\text{description of a quantum circuit } U),\, p_0, p_1\big),$$

where by description of a circuit we mean its approximate realization in the standard basis, and p_0, p_1 are such that $p_1 - p_0 \geq \Omega(n^{-\alpha})$ (n is the size of the circuit, $\alpha > 0$ is a constant). The function F is defined as follows:

$F(z) = 1 \iff$ there exists a vector $|\xi\rangle$, on which we get 1 in the first bit with probability greater than p_1;

$F(z) = 0 \iff$ for all $|\xi\rangle$ the probability of getting 1 in the first bit is smaller than p_0.

The completeness of Problem 0 is obvious: by saying that the problem is complete we just rephrase Definition 14.2.

We now consider more interesting examples. To start, we define a quantum analog of 3-CNF — the local Hamiltonian (locality is the analogue of the condition that the number of variables in each clause is bounded).

Definition 14.3. An operator $H : \mathcal{B}^{\otimes n} \to \mathcal{B}^{\otimes n}$ is called a *k-local Hamiltonian* if it is expressible in the form

$$H = \sum_j H_j[S_j],$$

where each term $H_j \in \mathbf{L}(\mathcal{B}^{\otimes |S_j|})$ is a Hermitian operator acting on a set of qubits S_j, $|S_j| \leq k$.

In addition, we put a normalization condition, namely, $0 \leq H_j \leq 1$, meaning that both H_j and $I - H_j$ are nonnegative.

PROBLEM 1: THE LOCAL HAMILTONIAN. Let

$$z = \big(\text{description of a } k\text{-local Hamiltonian } H,\, a, b\big),$$

where $k = O(1)$, $0 \leq a < b$, $b - a = \Omega(n^{-\alpha})$ ($\alpha > 0$ is a constant). Then

$F(x) = 1 \iff$ H has an eigenvalue not exceeding a,

$F(x) = 0 \iff$ all eigenvalues of H are greater than b.

Proposition 14.2. *The problem* LOCAL HAMILTONIAN *belongs to* BQNP.

Proof. At the outset we describe the general idea. We construct a circuit W that can be applied to a state $|\eta\rangle \in \mathcal{B}^{\otimes n}$ so as to produce a result 1 or 0 ("yes" or "no"): it says whether Arthur accepts the submitted state or not. The answer "yes" will occur with probability $p = 1 - r^{-1}\langle\eta|H|\eta\rangle$, where r is the number of terms in the Hamiltonian H. If $|\eta\rangle$ is an eigenvector

corresponding to an eigenvalue $\lambda \le a$, then the probability of the answer "yes" is

$$p = 1 - r^{-1}\langle\eta|H|\eta\rangle = 1 - r^{-1}\lambda \ge 1 - r^{-1}a,$$

and if every eigenvalue of H exceeds b, then

$$p = 1 - r^{-1}\langle\eta|H|\eta\rangle \le 1 - r^{-1}b.$$

At first we construct such a circuit for a single term. This will be just a realization of the POVM measurement corresponding to the decomposition $I = H_j + (I - H_j)$. We could use the general result about POVM measurements (see Problem 11.8), but let us give an explicit construction from scratch.

Let $H_j = \sum_s \lambda_s |\psi_s\rangle\langle\psi_s|$. This operator acts on a bounded number of qubits, $|S_j| \le k$. Therefore we can realize the operator

$$W_j : |\psi_s, 0\rangle \mapsto |\psi_s\rangle \otimes \left(\sqrt{\lambda_s}|0\rangle + \sqrt{1-\lambda_s}|1\rangle\right)$$

by a circuit of constant size. It acts on the set of qubits $S_j \cup \{\text{"answer"}\}$, where "answer" denotes the qubit that will contain the measurement outcome.

We compute the probability that the outcome is 1. Let $|\eta\rangle = \sum_s y_s|\psi_s\rangle$ be the expansion of $|\eta\rangle$ in the orthogonal system of eigenvectors of H_j. We have, by definition of the probability,

$$\begin{aligned}
\mathbf{P}_j(1) &= \langle\eta, 0|\, W_j^\dagger (I \otimes \underbrace{|1\rangle\langle 1|}_{\text{answer}}) W_j\, |\eta, 0\rangle \\
&= \left(\sum_s y_s^* \langle\psi_s, 0|\right) W_j^\dagger (I \otimes \underbrace{|1\rangle\langle 1|}_{\text{answer}}) W_j \left(\sum_t y_t\, |\psi_t, 0\rangle\right) \\
&= \sum_{s,t} \sqrt{1-\lambda_s}\; y_s^*\, \sqrt{1-\lambda_t}\; y_t\; \langle\psi_s|\psi_t\rangle \;=\; \sum_s (1-\lambda_s) y_s^* y_s \\
&= 1 - \langle\eta|H_j|\eta\rangle.
\end{aligned}$$

The general circuit W chooses the integer j randomly and uniformly, after which it applies the corresponding operator W_j. This procedure can be realized by the measuring operator $\sum_j |j\rangle\langle j| \otimes W_j$, applied to the initial vector $\left(\frac{1}{\sqrt{r}}\sum_j |j\rangle\right) \otimes |\eta, 0\rangle$. (Here $|j\rangle$ denotes a basis vector in an auxiliary r-dimensional space.) The probability of getting the outcome 1 is

$$\mathbf{P}(1) = \sum_j \frac{1}{r}\mathbf{P}_j(1) = \sum_j \frac{1}{r}\left(1 - \langle\eta|H_j|\eta\rangle\right) = 1 - r^{-1}\langle\eta|H|\eta\rangle.$$

□

14.4. Local Hamiltonian is BQNP-complete.

Theorem 14.3. *The problem* LOCAL HAMILTONIAN *is* BQNP*-complete with respect to the Karp reduction.*

The rest of this section constitutes a proof of this theorem. The main idea goes back to Feynman [**24**]: replacing a unitary evolution by a time independent Hamiltonian (i.e., transition from the circuit to a local Hamiltonian).

Thus, suppose we have a circuit $U = U_L \cdots U_1$ of size L. We will assume that U acts on N qubits, the first m of which initially contain Merlin's message $|\xi\rangle$, the rest being initialized by 0. The gates U_j act on pairs of qubits.

14.4.1. ***The Hamiltonian associated with the circuit.*** It acts on the space

$$\mathcal{L} = \mathcal{B}^{\otimes N} \otimes \mathbb{C}^{L+1},$$

where the first factor is the space on which the circuit acts, whereas the second factor is the space of a step counter (clock). The Hamiltonian consists of three terms which will be defined later,

$$H = H_{\text{in}} + H_{\text{prop}} + H_{\text{out}}.$$

We are interested in the minimum eigenvalue of this Hamiltonian, or the minimum of the *cost function* $f(|\eta\rangle) = \langle\eta|H|\eta\rangle$ over all vectors $|\eta\rangle$ of unit length. We will try to arrange that the Hamiltonian has a small eigenvalue if and only if there exists a quantum state $|\xi\rangle \in \mathcal{B}^{\otimes m}$ causing U to output 1 with high probability. In such a case, the minimizing vector $|\eta\rangle$ will be related to that $|\xi\rangle$ in the following way:

$$|\eta\rangle = \frac{1}{\sqrt{L+1}} \sum_{j=0}^{L} U_j \cdots U_1 |\xi, 0\rangle \otimes |j\rangle.$$

In constructing the terms of the Hamiltonian, we will try to "enforce" this structure of the vector $|\eta\rangle$ by imposing "penalties" that increase the cost function whenever $|\eta\rangle$ deviates from the indicated form.

The term H_{in} corresponds to the condition that, at step 0, all the qubits but m are in state $|0\rangle$. Specifically,

$$H_{\text{in}} = \left(\sum_{s=m+1}^{N} \Pi_s^{(1)} \right) \otimes |0\rangle\langle 0|, \tag{14.4}$$

where $\Pi_s^{(\alpha)}$ is the projection onto the subspace of vectors for which the s-th qubit equals α. The second factor in this formula acts on the space of the counter. (Informally speaking, the term $\Pi_s^{(1)} \otimes |0\rangle\langle 0|$ "collects a penalty" by

adding 1 to the cost function whenever the s-th qubit is in state $|1\rangle$ while the counter being in state $|0\rangle$.)

The term H_{out} corresponds to the final state and equals

$$H_{\text{out}} = \Pi_1^{(0)} \otimes |L\rangle\langle L|. \tag{14.5}$$

Here we assume that the bit of the result has number 1. (That is, at step L the first qubit should be in state $|1\rangle$, or a penalty will be imposed.)

And, finally, the term H_{prop} describes the propagation of a quantum state through the circuit. It consists of L terms, each of which corresponds to the transition from $j-1$ to j:

$$H_{\text{prop}} = \sum_{j=1}^{L} H_j, \tag{14.6}$$

$$H_j = -\frac{1}{2} U_j \otimes |j\rangle\langle j-1| - \frac{1}{2} U_j^\dagger \otimes |j-1\rangle\langle j| + \frac{1}{2} I \otimes \big(|j\rangle\langle j| + |j-1\rangle\langle j-1|\big).$$

Each term H_j acts on two qubits of the space $\mathcal{B}^{\otimes N}$, as well as on the space of the counter (the latter is not represented by qubits yet).

14.4.2. ***Change of basis.*** We effect the change of basis given by the operator

$$W = \sum_{j=0}^{L} U_j \cdots U_1 \otimes |j\rangle\langle j|.$$

(It is worth mentioning that W is a measuring operator with respect to the the value of the counter j.) The change of basis means that we represent the vector $|\eta\rangle$ in the form $|\eta\rangle = W|\widetilde{\eta}\rangle$; from now on, we are dealing with $|\widetilde{\eta}\rangle$ instead of $|\eta\rangle$. Under such a change, the Hamiltonian is transformed into its conjugate, $\widetilde{H} = W^\dagger H W$. We consider how the conjugation by the operator W acts on the terms of H.

On the term H_{in} the conjugation has no effect:

$$\widetilde{H}_{\text{in}} = W^\dagger H_{\text{in}} W = H_{\text{in}}. \tag{14.7}$$

The action on the term H_{out} is:

$$\widetilde{H}_{\text{out}} = W^\dagger H_{\text{out}} W = \left(U^\dagger \Pi_1^{(0)} U\right) \otimes |L\rangle\langle L|. \tag{14.8}$$

Each operator H_j in (14.6) is the sum of three terms. Let us write the action of the conjugation on the first of them:

$$\begin{aligned} &W^\dagger\bigl(U_j \otimes |j\rangle\langle j-1|\bigr)W \\ &\quad = \sum_{p,t} \bigl(U_p \cdots U_1 \otimes |p\rangle\langle p|\bigr)^\dagger \bigl(U_j \otimes |j\rangle\langle j-1|\bigr) \bigl(U_t \cdots U_1 \otimes |t\rangle\langle t|\bigr) \\ &\quad = \Bigl((U_j \cdots U_1)^\dagger U_j \bigl(U_{j-1} \cdots U_1\bigr)\Bigr) \otimes \Bigl(\bigl(|j\rangle\langle j|\bigr)^\dagger |j\rangle\langle j-1| \bigl(|j-1\rangle\langle j-1|\bigr)\Bigr) \\ &\quad = I \otimes |j\rangle\langle j-1|. \end{aligned}$$

Conjugation of the two other terms proceeds analogously, so that we obtain

$$\begin{aligned} \widetilde{H}_j &= W^\dagger H_j W \\ &= I \otimes \frac{1}{2}\Bigl(|j-1\rangle\langle j-1| - |j-1\rangle\langle j| - |j\rangle\langle j-1| + |j\rangle\langle j|\Bigr) = I \otimes E_j, \end{aligned}$$

$$\widetilde{H}_{\mathrm{prop}} = W^\dagger H_{\mathrm{prop}} W = I \otimes E, \tag{14.9}$$

where

$$E = \sum_{j=1}^{L} E_j = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & & & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & & \\ & -\frac{1}{2} & 1 & -\frac{1}{2} & \\ & & -\frac{1}{2} & \ddots & \ddots \\ 0 & & & \ddots & \ddots \end{pmatrix}.$$

14.4.3. ***Existence of a small eigenvalue if the answer is "yes".*** Suppose that the circuit U gives the outcome 1 ("yes") with probability $\geq 1 - \varepsilon$ on some input vector $|\xi\rangle$. This, by definition, means that

$$\mathbf{P}(0) = \langle \xi, 0 | U^\dagger \Pi_1^{(0)} U | \xi, 0 \rangle \leq \varepsilon.$$

We want to prove that in this case $\widetilde{H}$ (and so also H) has a small eigenvalue. For this, it is sufficient to find a vector $|\widetilde{\eta}\rangle$ such that $\langle \widetilde{\eta} | \widetilde{H} | \widetilde{\eta} \rangle$ is small enough (the minimum of this expression as a function of $|\widetilde{\eta}\rangle$ is attained at an eigenvector).

In the space of the counter we choose the vector

$$|\psi\rangle = \frac{1}{\sqrt{L+1}} \sum_{j=0}^{L} |j\rangle. \tag{14.10}$$

We set $|\widetilde{\eta}\rangle = |\xi, 0\rangle \otimes |\psi\rangle$ and estimate $\langle \widetilde{\eta} | H | \widetilde{\eta} \rangle$.

It is clear that $E|\psi\rangle = 0$. Therefore

$$\langle \widetilde{\eta} | \, \widetilde{H}_{\mathrm{prop}} \, | \widetilde{\eta} \rangle = 0 = \langle \widetilde{\eta} | \, \widetilde{H}_j \, | \widetilde{\eta} \rangle.$$

Since all auxiliary qubits are initially set to 0, we immediately obtain from the defining formula (14.4) that

$$\langle\widetilde{\eta}|\widetilde{H}_{\text{in}}|\widetilde{\eta}\rangle = 0.$$

It remains to estimate the last term

$$\langle\widetilde{\eta}|\,\widetilde{H}_{\text{out}}\,|\widetilde{\eta}\rangle = \langle\widetilde{\eta}|\Big(U^{\dagger}\Pi_1^{(0)}U\otimes|L\rangle\langle L|\Big)|\widetilde{\eta}\rangle = \mathbf{P}(0)\,\frac{1}{L+1} \le \frac{\varepsilon}{L+1}.$$

Thus we have proved that

$$\langle\widetilde{\eta}|\,\widetilde{H}\,|\widetilde{\eta}\rangle \le \frac{\varepsilon}{L+1},$$

so that H itself has an eigenvalue with the very same upper bound.

14.4.4. ***Lower bound for the eigenvalues if the answer is "no".*** Suppose that for any vector $|\xi\rangle$ the probability of the outcome 1 does not exceed ε, i.e.,

$$\langle\xi,0|U^{\dagger}\Pi_1^{(0)}U|\xi,0\rangle \ge 1-\varepsilon.$$

We will prove that, in this case, all eigenvalues of H are $\ge c\,(1-\sqrt{\varepsilon})\,L^{-3}$, where c is some constant.

The proof is rather long, so we will outline it first. We represent the Hamiltonian in the form $\widetilde{H} = A_1 + A_2$, where $A_1 = \widetilde{H}_{\text{in}} + \widetilde{H}_{\text{out}}$, and $A_2 = \widetilde{H}_{\text{prop}}$. Both terms are nonnegative. It is easy to show that the null subspaces of A_1 and A_2 have trivial intersection (i.e., $\mathcal{L}_1 \cup \mathcal{L}_2 = \{0\}$), hence the operator $A_1 + A_2$ is positive definite. But this is not enough for our purpose, so we obtain lower bounds for nonzero eigenvalues of A_1 and A_2, namely, 1 for A_1, and $c'L^{-2}$ for A_2. In order to estimate the smallest eigenvalue of $A_1 + A_2$, we also need to know the angle between the null subspaces. The angle $\vartheta(\mathcal{L}_1, \mathcal{L}_2)$ between subspaces $\mathcal{L}_1$ and $\mathcal{L}_2$ with trivial intersection is given by

$$\cos\vartheta(\mathcal{L}_1,\mathcal{L}_2) = \max_{\substack{|\eta_1\rangle\in\mathcal{L}_1\\ |\eta_2\rangle\in\mathcal{L}_2}} \big|\langle\eta_1|\eta_2\rangle\big|, \qquad 0 < \vartheta(\mathcal{L}_1,\mathcal{L}_2) < \frac{\pi}{2}. \tag{14.11}$$

Lemma 14.4. *Let A_1, A_2 be nonnegative operators, and $\mathcal{L}_1$, $\mathcal{L}_2$ their null subspaces, where $\mathcal{L}_1 \cap \mathcal{L}_2 = \{0\}$. Suppose further that no nonzero eigenvalue of A_1 or A_2 is smaller than v. Then*

$$A_1 + A_2 \ge v \cdot 2\sin^2\frac{\vartheta}{2},$$

where $\vartheta = \vartheta\,(\mathcal{L}_1, \mathcal{L}_2)$ is the angle between $\mathcal{L}_1$ and $\mathcal{L}_2$.

The notation $A \ge a$ (A an operator, a a number) must be understood as an abbreviation for $A - aI \ge 0$. In other words, if $A \ge a$, then all eigenvalues of A are greater than, or equal to, a.

In our case we will get the estimates 1 and $c'L^{-2}$ for the nonzero eigenvalues of A_1 and A_2 (as already mentioned), and $\sin^2\vartheta \geq (1-\sqrt{\varepsilon})/(L+1)$ for the angle. From this we derive the desired inequality

$$H \geq c\,(1-\sqrt{\varepsilon})\,L^{-3}.$$

Proof of Lemma 14.4. It is obvious that $A_1 \geq v(I-\Pi_{\mathcal{L}_1})$ and $A_2 \geq v(I-\Pi_{\mathcal{L}_2})$, so it is sufficient to prove the inequality $(I-\Pi_{\mathcal{L}_1})+(I-\Pi_{\mathcal{L}_2}) \geq 2\sin^2(\vartheta/2)$. This, in turn, is equivalent to

$$\Pi_{\mathcal{L}_1}+\Pi_{\mathcal{L}_2} \leq 1+\cos\vartheta. \tag{14.12}$$

Let $|\xi\rangle$ be an eigenvector of the operator $\Pi_{\mathcal{L}_1}+\Pi_{\mathcal{L}_2}$ corresponding to an eigenvalue $\lambda > 0$. Then

$$\Pi_{\mathcal{L}_1}|\xi\rangle = u_1|\eta_1\rangle, \quad \Pi_{\mathcal{L}_2}|\xi\rangle = u_2|\eta_2\rangle, \qquad u_1|\eta_1\rangle + u_2|\eta_2\rangle = \lambda|\xi\rangle,$$

where $|\eta_1\rangle \in \mathcal{L}_1$ and $|\eta_2\rangle \in \mathcal{L}_2$ are unit vectors, and u_1, u_2 are nonnegative real numbers. From this we find

$$\lambda = \langle\xi|\,(\Pi_{\mathcal{L}_1}+\Pi_{\mathcal{L}_2})\,|\xi\rangle = u_1^2+u_2^2,$$

$$\lambda^2 = \big(u_1\langle\eta_1| + u_2\langle\eta_2|\big)\big(u_1|\eta_1\rangle + u_2|\eta_2\rangle\big) = u_1^2+u_2^2+2u_1u_2\mathrm{Re}\langle\eta_1|\eta_2\rangle.$$

Consequently,

$$(1+x)\lambda-\lambda^2 = x(u_1\pm u_2)^2 \geq 0, \quad \text{where } x = \big|\mathrm{Re}\langle\eta_1|\eta_2\rangle\big|.$$

Thus $\lambda \leq 1+x \leq 1+\cos\vartheta$. □

We will now obtain the above-mentioned estimates. The subspaces A_1 and A_2 can be represented in the form

$$\begin{aligned}\mathcal{L}_1 = \Big(\mathcal{B}^{\otimes m}\otimes|0^{N-m}\rangle\otimes|0\rangle\Big) \oplus \Big(\mathcal{B}^{\otimes N}\otimes\mathbb{C}\big(|1\rangle,\dots,|L-1\rangle\big)\Big) \\ \oplus \Big(U^\dagger\big(|1\rangle\otimes\mathcal{B}^{\otimes(N-1)}\big)\otimes|L\rangle\Big)\end{aligned} \tag{14.13}$$

(the last factor in all three terms pertains to the counter), and

$$\mathcal{L}_2 = B^{\otimes N}\otimes|\psi\rangle, \tag{14.14}$$

where the vector $|\psi\rangle$ was defined by formula (14.10).

For the estimate

$$A_1|_{\mathcal{L}_1^\perp} \geq 1 \tag{14.15}$$

it suffices to note that A_1 is the sum of commuting projections, so that all eigenvalues of this operator are nonnegative integers.

For the estimate of $A_2|_{\mathcal{L}_2^\perp}$ we need to find the smallest positive eigenvalue of the matrix E. The eigenvectors and eigenvalues of E are

$$|\psi_k\rangle = \alpha_k\sum_{j=0}^{L}\cos\Big(q_k\big(j+\frac{1}{2}\big)\Big)\,|j\rangle, \qquad \lambda_k = 1-\cos q_k,$$

where $q_k = \pi k/(L+1)$ $(k = 0, \ldots, L)$. From this it follows that

$$A_2|_{\mathcal{L}_2^\perp} \geq 1 - \cos\left(\frac{\pi}{L+1}\right) \geq c' L^{-2}. \tag{14.16}$$

Finally, we need to estimate the angle between the subspaces $\mathcal{L}_1$ and $\mathcal{L}_2$. We will estimate the square of the cosine of this angle,

$$\cos^2 \vartheta = \max_{\substack{|\eta_1\rangle \in \mathcal{L}_1 \\ |\eta_2\rangle \in \mathcal{L}_2}} \left|\langle \eta_1 | \eta_2 \rangle\right|^2 = \max_{|\eta_2\rangle \in \mathcal{L}_2} \langle \eta_2 | \Pi_{\mathcal{L}_1} | \eta_2 \rangle. \tag{14.17}$$

Since the vector $|\eta_2\rangle$ belongs to $\mathcal{L}_2$, it can be represented in the form $|\eta_2\rangle = |\xi\rangle \otimes |\psi\rangle$ (cf. (14.14)). According to formula (14.13), the projection onto $\mathcal{L}_1$ breaks into the sum of three projections. It is easy to calculate the contribution of the second term; it equals $(L-1)/(L+1)$. The first and third terms add up to

$$\frac{1}{L+1} \langle \xi | \left(\Pi_{\mathcal{K}_1} + \Pi_{\mathcal{K}_2}\right) | \xi \rangle \leq \frac{1 + \cos\varphi}{L+1},$$

where $\mathcal{K}_1 = \mathcal{B}^{\otimes N}$, $\mathcal{K}_2 = U^\dagger \left(|1\rangle \otimes \mathcal{B}^{\otimes (N-1)}\right)$ and φ is the angle between these two subspaces. (Here we have used inequality (14.12), obtained in the course of the proof of Lemma 14.4.)

The quantity $\cos^2 \varphi$ equals the maximum probability for the initial circuit to produce the outcome 1. By hypothesis this probability is not greater than ε. So we can continue the estimate (14.17):

$$\langle \eta_2 | \Pi_{\mathcal{L}_1} | \eta_2 \rangle \leq \frac{L-1}{L+1} + \frac{1+\sqrt{\varepsilon}}{L+1} = 1 - \frac{1-\sqrt{\varepsilon}}{L+1}.$$

Consequently, $\sin^2 \vartheta = 1 - \cos^2 \vartheta \geq (1 - \sqrt{\varepsilon})/(L+1)$ as asserted above.

14.4.5. ***Realization of the counter.*** We wrote a nice Hamiltonian almost satisfying the required properties. It has but one shortcoming — the counter is not a qubit. We could, of course, represent it by $O(\log L)$ qubits, but then the Hamiltonian would be only $O(\log L)$-local, not $O(1)$-local.

This shortcoming can be removed if we embed the counter space in a larger space. We take L qubits, enumerated from 1 to L. The suitable embedding $\mathbb{C}^{L+1} \to \mathcal{B}^{\otimes L}$ is

$$|j\rangle \mapsto |\underbrace{1, \ldots, 1}_{j}, \underbrace{0, \ldots, 0}_{L-j}\rangle.$$

The operators on the space $\mathbb{C}^{L+1}$ used in the construction of the Hamiltonian H are replaced in accordance with the following scheme:

$$
\begin{aligned}
&|0\rangle\langle 0| \text{ on } \Pi_1^{(0)}, && |0\rangle\langle 1| \text{ on } \big(|0\rangle\langle 1|\big)_1 \Pi_2^{(0)},\\
&|j\rangle\langle j| \text{ on } \Pi_j^{(1)}\Pi_{j+1}^{(0)}, && |j-1\rangle\langle j| \text{ on } \Pi_{j-1}^{(1)}\big(|0\rangle\langle 1|\big)_j \Pi_{j+1}^{(0)},\\
&|L\rangle\langle L| \text{ on } \Pi_L^{(1)}, && |L-1\rangle\langle L| \text{ on } \Pi_{L-1}^{(1)}\big(|0\rangle\langle 1|\big)_L.
\end{aligned}
\tag{13.23}
$$

Now they are 3-local (and the Hamiltonian itself, acting also on the qubits of the initial circuit, is 5-local).

To be more precise, we have replaced the Hamiltonian H, acting on the space $\mathcal{L} = \mathcal{B}^{\otimes N} \otimes \mathbb{C}^{L+1}$, by a new Hamiltonian H_{ext}, defined on the larger space $\mathcal{L}_{\text{ext}} = \mathcal{B}^{\otimes N} \otimes \mathcal{B}^{\otimes L}$. The operator H_{ext} maps the subspace $\mathcal{L} \subseteq \mathcal{L}_{\text{ext}}$ into itself and acts on it just as H does.

Now a new problem arises: what to do with the extra states in the extended space of the counter? We will cope with this problem by adding still another term to the Hamiltonian H_{ext}:

$$H_{\text{stab}} = I_{\mathcal{B}^{\otimes N}} \otimes \sum_{j=1}^{L-1} \Pi_j^{(0)} \Pi_{j+1}^{(1)}.$$

The null subspace of the operator H_{stab} coincides with the old working space $\mathcal{L}$, so that the supplementary term does not change the upper bound for the minimum eigenvalue for the answer "yes".

For the answer "no" the required lower bound for the eigenvalues of the operator $H_{\text{ext}} + H_{\text{stab}}$ can be recovered in the following way. Both terms leave the subspace $\mathcal{L}$ invariant, so that we can also examine the action of $H_{\text{ext}} + H_{\text{stab}}$ on $\mathcal{L}$ and on its orthogonal complement $\mathcal{L}^{\perp}$ independently. On $\mathcal{L}$ we have $H_{\text{ext}} \geq c(1 - \sqrt{\varepsilon})L^{-3}$ and $H_{\text{stab}} = 0$, and on $\mathcal{L}^{\perp}$ we have $H_{\text{ext}} \geq 0$ and $H_{\text{stab}} \geq 1$. (Here we use the fact that each of the terms of the Hamiltonian, (14.4), (14.5) and (14.6), remains nonnegative for the change (13.23).) In both cases

$$H_{\text{ext}} + H_{\text{stab}} \geq c(1 - \sqrt{\varepsilon})L^{-3}.$$

This completes the proof of Theorem 14.3.

14.5. The place of BQNP among other complexity classes. It follows directly from the definition that the class BQNP contains the class MA (and so also BPP and NP). Nothing more definitive can be said at present about the strength of "nondeterministic quantum algorithms".[12]

Nor can much more be said about its "weakness".

[12]Caveat: in the literature there is also a different definition of "quantum NP", for which a complete characterization in terms of classical complexity classes can be obtained (cf. [**2, 73**]).