# On the Relation Between Quantum Computational Speedup and Retrocausality

*Giuseppe Castagnoli*

*Elsag Bailey ICT Division and Quantum Information Laboratory, Turin, Italy. E-mail: giuseppe.castagnoli@gmail.com*

Editors: *Eliahu Cohen & Jonas Maziero*

Article history: Submitted on July 31, 2015; Accepted on January 6, 2016; Published on January 11, 2016.

**We investigate the reason for the quantum speedup (quantum algorithms require fewer computation steps than their classical counterparts). We extend the representation of the quantum algorithm to the process of setting the problem, namely choosing the function computed by the black box. The initial measurement selects a setting at random, Bob (the problem setter) unitarily changes it into the desired one. With reference to the observer dependent quantum states of relational quantum mechanics, this representation is with respect to Bob and any external observer, it cannot be with respect to Alice (the problem solver). It would tell her the function computed by the black box, which to her should be hidden. To Alice, the projection of the quantum state due to the initial measurement is retarded at the end of her problem solving action, so that the algorithm input state remains one of complete ignorance of the setting. By black box computations, she unitarily sends it into the output state that, for each possible setting, encodes the corresponding solution, acquired by the final measurement. Mathematically, we can ascribe to the final measurement the selection of any fraction $R$ of the random outcome of the initial measurement. This projects the input state to Alice on one of lower entropy where she knows the corresponding fraction of the problem setting. Given the appropriate value of $R$, the quantum algorithm is a sum over classical histories in each of which Alice, knowing in advance one of the $R$-th parts of the setting, performs the black box computations still required to identify the solution. Given a quantum algorithm, this retrocausality model provides the value of $R$ that explains its speed up; in the major quantum algorithms, $R$ is $\frac{1}{2}$ or slightly above it. Conversely, given the problem, $R = \frac{1}{2}$ always yields the order of magnitude of the number of black box computations required to solve it in an optimal quantum way.**

**Quanta 2016; 5: 34–52.**

## 1 Foreword

Consider the following problem. Bob, who is the problem setter, chooses one of the four functions $f_{\mathbf{b}}(\mathbf{a})$ shown in Table 1. Then he gives Alice, who is the problem solver, a black box (*oracle*) that, given a value of the argument $\mathbf{a}$ in the input, produces the value of $f_{\mathbf{b}}(\mathbf{a})$ in the output. Alice does not know which of the four functions is the one computed by the black box. She is to determine whether the function is constant or balanced (with the same number of zeros and ones) by performing function evaluations (*oracle queries*). Classically, Alice must perform two function evaluations, quantumly just one. Here, we refer

*Table 1: Tabular representation of the functions $f_{\mathbf{b}}(\mathbf{a})$ that are evaluated by the black box in the Deutsch's problem.*

| **a** | $f_{00}(\mathbf{a})$ | $f_{01}(\mathbf{a})$ | $f_{10}(\mathbf{a})$ | $f_{11}(\mathbf{a})$ |
|---|---|---|---|---|
| 0 | 0 | 0 | 1 | 1 |
| 1 | 0 | 1 | 0 | 1 |

to the seminal quantum algorithm that yields a *quantum computational speedup*, devised in 1985 by Deutsch [1]. Although there is a significant body of literature on the relationship between speedup and other quantum features, such as *quantum entanglement* and *discord* (see Section 3), no fundamental physical explanation nor unified mathematical mechanism is known for the speedup.

As the present issue of *Quanta* is dedicated to Richard Feynman and our subject is quantum computation, we should like to remember Feynman's pioneering contribution to the development of this new branch of science. We do this by recalling the seminal works that gave rise to the discipline of quantum computation. In 1969, Finkelstein [2] noted that computation should be possible in the quantum framework and introduced the notion of quantum unit of information, namely quantum bit or *qubit*. In 1982, Feynman [3] pointed out the essential difference between quantum and classical computation, showing that the simulation of a quantum process on a classical computer has to involve in general an amount of time × physical resources exponentially higher than that involved in the quantum process itself. This was the origin of the notion of quantum computational efficiency, now usually called speedup. In [3], Feynman also introduced the universal quantum simulator, a lattice of spin systems with freely specifiable nearest neighbor interactions that can be considered the first theoretical model of a quantum computer. The development of the notion of reversible classical computation was parallel. Bennett [4] showed in 1982 that classical computation can be ideally reversible in the limit of zero speed. His work was in the wake of the 1961 *Landauer's principle* [5] that quantifies the generation of heat necessarily consequent to the erasure of information. Still in 1982, Fredkin and Toffoli [6] developed the first theoretical model of logically reversible classical computation, in fact based on the well known Fredkin and Toffoli gates. Independently of Deutsch, Feynman produced in 1985 the quantum version of this algorithmic form of reversible computation, published in the following year [7]. The seminal idea of it was already present in [3], in fact with reference to Fredkin and Toffoli 1982 work. In 1985, Deutsch provided the first example of a quantum algorithm that requires fewer function evaluations than classically possible.

With this, the full fledged notion of quantum computation was born thanks to the insights of very few individuals. As it might happen with revolutionary science, the scientific community at large has been initially slow in expressing an interest for the new discipline. We had the fortune of contributing, with Mario Rasetti, to the organization of the first international meetings on quantum communication and computation held in Turin in the years 1992–1998. They were workshops patronized by Elsag Bailey (an Italian information and communications technology company) and the Institute for Scientific Interchange, Turin, Italy. We believe that those annual workshops have been instrumental to propagating the new discipline of quantum information throughout the scientific community. All the fathers of the discipline, with the sad exception of Feynman, the theoretical and experimental physicists and computer scientists responsible for the major developments of those years attended the workshops in question. The group pictures of the workshop participants [8] show the explosion of interest for the new science in the years 1993 through 1997.

## 2 Introduction

The usual physical representation of quantum algorithms is limited to the process of solving the problem. We extend it to the process of setting the problem, namely of choosing the function $f_{\mathbf{b}}(\mathbf{a})$ out of the set of functions. This amounts to choosing the function suffix **b**, which we call *the problem setting*, out of the set of the possible problem settings $\sigma_B \equiv \{00, 01, 10, 11\}$ (here we use Deutsch algorithm as an example).

For reasons that will soon become clear, we assume that the initial state of the quantum register $B$ that contains the setting is a mixture of all the possible settings. Its density operator is thus

$$\hat{\rho}_B = \frac{1}{4}\left(|00\rangle\langle 00|_B + |01\rangle\langle 01|_B + |10\rangle\langle 10|_B + |11\rangle\langle 11|_B\right) \tag{1}$$

At time $t_0$, Bob measures the content of register $B$, obtaining a setting at random, say $\mathbf{b} = 10$. The state of register $B$ is consequently projected on $|10\rangle\langle 10|_B$. Assume that Bob wants $\mathbf{b} = 01$. He unitarily transforms this state into $|01\rangle\langle 01|_B$, at time $t_1$.

Register $A$, meant to contain the argument of the function to be computed by the black box and eventually the solution of the problem, at time $t_1$ is in any sharp state, say $|0\rangle\langle 0|_A$. The input state of the quantum algorithm at time $t_1$ is thus $|01\rangle\langle 01|_B \otimes |0\rangle\langle 0|_A$.

Alice, with one function evaluation preceded and followed by suitable transformations, unitarily transforms it into the output state $|01\rangle\langle 01|_B \otimes |1\rangle\langle 1|_A$, at time $t_2$. The solution of the problem, 1 when the function is balanced as in the present case and 0 when it is constant, is in register $A$. Alice acquires the solution by measuring the content of $A$.

We note that this extended representation immediately calls for another extension, this time concerning the actors (observers) on the stage. We have to resort to the relational quantum mechanics of Rovelli [9], where quantum states are observer dependent. A quantum state can be sharp to an observer and a quantum superposition, or a mixture, to another one. The present representation is with respect to Bob, the problem setter, and any other observer who does not act on the problem solving process. It cannot be with respect to Alice, the problem solver. The input state of the quantum algorithm $|01\rangle\langle 01|_B \otimes |0\rangle\langle 0|_A$ would of course tell us that the content of register $B$ is 01, namely that the function chosen by Bob is $f_{01}(\mathbf{a})$. Since Alice is an observer, we assume that the state in question would tell her the same.

Throughout this work, we take for granted the legitimacy of the assumption that Alice (or Bob, or the external observer), although an abstract entity, knows what we would know in her (his) place. We assume that this is a legitimate way to take the sizes to a quantum process that necessarily involves the notion of observer.

In the present case, Alice would know that the function is balanced without performing any function evaluation. Of course the suffix of the function should be hidden to Alice because to her it is inside the black box. We physically represent this concealment by retarding the projection of the quantum state induced by the initial Bob's measurement at the end of the unitary part of Alice's problem solving action. As well known, these projections can be retarded or advanced at will along a unitary evolution that respectively follows or precedes the measurement.

The input state of the quantum algorithm to Alice, immediately after the preparation of the desired problem setting, remains

$$\frac{1}{4}\left(|00\rangle\langle 00|_B + |01\rangle\langle 01|_B + \ldots\right) \otimes |0\rangle\langle 0|_A\,. \qquad (2)$$

In fact the maximally mixed state of register $B$ remains unaltered under any unitary transformation applying to it. The two bit entropy of this state represents Alice's complete ignorance of Bob's choice.

The output state to Alice is

$$\frac{1}{4}(|00\rangle\langle 00|_B \otimes |0\rangle\langle 0|_A + |01\rangle\langle 01|_B \otimes |1\rangle\langle 1|_A + \ldots), \quad (3)$$

that is still a mixture of all the possible problem settings, each multiplied by the corresponding solution. Thus, also the solution, considered in itself, is completely undetermined. Alice's final measurement projects this state on the solution corresponding to the problem setting chosen by Bob, namely on $\frac{1}{2}\left(|01\rangle\langle 01|_B + |10\rangle\langle 10|_B\right) \otimes |1\rangle\langle 1|_A$, with probability one. In fact the solution is unpredictable to Alice but is already 1 to any other observer.

Alice's final measurement also triggers the retarded projection induced by the initial Bob's measurement, which cannot go past the unitary part of Alice's action. This further projects the above state on $|01\rangle\langle 01|_B \otimes |1\rangle\langle 1|_A$, which tells Alice both the problem setting and the solution. The two projections commute and should be considered simultaneous.

We note that either the projection of the quantum state induced by the initial Bob's measurement or that induced by the final Alice's measurement zeroes the entropy of the solution, depending on which one is performed first. This work is an exploration of the assumption that this zeroing shares in a complementary and non-redundant way between initial and final measurement.

We assume that the complete measurements behave in a contextual way, namely each would be sensitive to the other. We also assume that they reduce (in all the possible ways in quantum superposition as we will see) to partial measurements such that, together, select whatever has been selected by the complete measurements and, each by itself, reduce the entropy of the solution in a complementary and non-redundant way. For Occam razor, we exclude any redundancy between the two partial measurements, what implies that the information provided by either one is not provided by the other. In Newton's formulation, Occam razor states:

*We are to admit no more causes of natural things than such that are both true and sufficient to explain their appearances* [10].

To reconstruct the selections performed by the complete measurements, we should propagate forward in time, employing the time-forward unitary transformation, the projection of the quantum state due to the partial Bob's measurement, until it selects part of the outcome of Alice's measurement. Similarly, we should propagate backward in time, employing the inverse of the time-forward unitary transformation, the projection due to Alice's partial measurement, until it selects part of the random outcome of Bob's measurement.

We will see that everything boils down to ascribing to the final Alice's measurement the selection of part of the random outcome of the initial Bob's measurement, say the $R$-th part of the information that specifies it.

This *quantum feedback* leaves the input state of the quantum algorithm to Bob and any external observer unaltered. It projects that to Alice on a state of lower entropy where she knows the $R$-th part of the information that specifies the problem setting in advance, before performing any function evaluation.

Given a value of $R$ different from 0 and 1, there are many ways of taking the $R$-th part of the information that specifies the problem setting. There is thus the need of reconciling the notion of Alice's advanced knowledge with this multiplicity. We also need to provide an operational interpretation of this notion. We kill two birds with one stone by applying to quantum computation Feynman's path integral formulation of quantum mechanics [11]. Given the appropriate choice of the value of $R$, it turns out that the quantum algorithm can be seen as a sum over classical histories in each of which Alice knows in advance one of the possible $R$-th parts of the problem setting and performs the function evaluations still necessary to find the solution of the problem. The sum is over all the possible ways of taking this $R$-th part.

Incidentally, let us notice the many ways the present interpretation of the speedup draws on Feynman's findings, including his and Wheeler's work on the symmetry of physical phenomena under time-reversal [12]. As things are now, the interplay between the path integral formulation of quantum mechanics and the current form of quantum retrocausality is at the complex systems level; investigating whether it can be brought to a fundamental level would seem to be an interesting prospect.

Given an oracle problem and a value of $R$, the present retrocausal interpretation of the speedup yields a number of function evaluations required to solve the problem. Conversely, given a quantum algorithm, it yields the value of $R$ that explains its speedup.

We have compared the present interpretation of the speedup with the major quantum algorithms discovered so far. In all the quantum algorithms that solve the problem with a single function evaluation, as that of Deutsch, we have $R = \frac{1}{2}$. This also applies to Grover quantum search algorithm for database size 4, Deutsch–Jozsa algorithm, and the algorithms of Simon and the Abelian hidden subgroup. The latter algorithm [13] has unified about ten historical algorithms, among which the famous Shor's factorization algorithm. In Grover algorithm, when database size goes past 4, first $R$ goes slightly above $\frac{1}{2}$ then it goes back to $\frac{1}{2}$ for database size tending to infinity.

In the corresponding sample of problems, $R = \frac{1}{2}$ always corresponds to an existing quantum algorithm and yields the order of magnitude of the number of function evaluations required to solve the problem in an optimal quantum way. If this held in general, we would have a very powerful tool, the way of assessing the order of magnitude of the number of function evaluations (oracle queries) required to solve a generic oracle problem in an optimal quantum way.

## 3 Positioning the work

The present work is the further development of the approach followed in [14–16]. We have further clarified the retrocausal interpretation of the speedup and developed a procedure for computing the number of function evaluations required to solve a generic oracle problem with quantum retrocausality $R = \frac{1}{2}$.

This retrocausal interpretation is in line with the tenet of time-symmetric quantum mechanics of Aharonov and collaborators [17–19], which states that the complete description of the quantum process between initial and final measurement requires knowledge not only of the outcome of the initial measurement, but also of that of the final one. This naturally implies that the latter outcome has back in time implications on the upstream process. As a matter of fact, the form of quantum retrocausality utilized in the present work has been inspired by the work of Dolev and Elitzur [20] on the non-sequential behavior of the wave function highlighted by partial measurement.

More specifically, the notion of advanced knowledge is in line with the finding that apparently random events in quantum mechanics are caused by events in the future [21]. Aharonov and collaborators analyzed a sequence of weak and strong measurements performed on an Einstein–Podolsky–Rosen pair, and concluded that "the most reasonable resolution seems to be that of the two-state vector formalism, namely, that the choice of the experimenter has been encrypted within the weak measurement's outcomes, even before the experimenters themselves know what their choice will be" [21]. Extending the notion of advanced knowledge to the two-state vector formalism would be an interesting prospect.

The work has points of contact with those of Morikoshi who highlighted the problem-solution symmetry of Grover's and the phase estimation algorithms, and noted that it may be relevant for the explanation of the speedup [22]. He further showed that Grover algorithm violates a temporal Bell inequality [23]. There should be a connection between this violation and the form of quantum retrocausality we are dealing with.

Besides [14–16], we are not aware of literature relating the speedup to quantum retrocausality. There are of course other approaches to the problem of finding a common reason for the variety of speedups found until now. An exemplification is as follows.

Jozsa and Linden showed that the presence of multipartite entanglement with number of parties increasing unbounded with problem size is necessary for achieving exponential speedup in pure state quantum computing. They also conjectured that there could be exponential speedup in the absence of entanglement in mixed state quantum computing [24].

The notion of *quantum discord* was introduced independently in [25] and [26]. Discord is a measure of non-classical correlations between two subsystems of a quantum system that are not necessarily entangled. It coincides with entanglement in pure state quantum computing. Discord could be of high practical interest, since it shows the possibility of achieving a speedup in mixed state quantum computing, which is the realistic form of computation in the presence of noise.

Gross and collaborators showed that, contrary to the topical thought at the time, quantum states can be too entangled to be useful for the purpose of computation [27].

At present, no single reason behind the speedup was found from the standpoint of entanglement or discord. The speedup appears to always depend on the exact nature of the problem while the reason for it varies from problem to problem [26].

With respect to the above approaches, the novelty of the present one is replacing entanglement by the retrocausality measure $R$, a possibly more fundamental quantum feature as it might explain entanglement [28]. Here the unifying feature is the permanence of $R$ in a right surrounding of $\frac{1}{2}$, where the number of queries required to solve the oracle problem varies a little. If this were true in general, beyond the sample of quantum algorithms examined, given an oracle problem we would know the order of magnitude of this number. This should be the object of further research, the present work is an exploration.

One of the main approaches to the study of the speedup is that of quantum computer science, which classifies the hardness of computational problems given a mathematical abstraction of the physical mechanism that performs the computation, typically the universal quantum Turing machine. For example, Aaronson studied the class of problems efficiently solvable with a quantum computer given a capability of postselecting measurement outcomes, and demonstrated that this class coincides with the well known classical complexity class PP (Probabilistic Polynomial-time), as well as the closure of either class under intersection [29]. Physically, the ability to postselect on a measurement yielding a specific outcome means that the computation process is to satisfy a constraint placed in its future, what can be seen as a form of retrocausality. One might wonder whether the retrocausal interpretation of the speed up could provide a more direct physical interpretation of complexity classes.

We eventually cite tree size complexity [30] and contextually based [31] arguments. In the former, a measure of the complexity of the multiqubit state is shown to be related to the speedup of a variety of quantum algorithms. The latter addresses the relation between speedup and the contextual character of quantum mechanics. It has led to identifying a form of fault tolerant quantum computation (by *magic states*) that is specially resilient to noise. Also the present retrocausal interpretation of the speedup could be considered a contextually based argument. The reduction of the initial and final measurements of a quantum process to partial non-redundant measurements is of course contextual in character.

## 4 The seminal Deutsch algorithm

Let us review the usual representation of Deutsch algorithm, limited to the process of solving the problem. We need two quantum registers: $A$, of basis vectors $|0\rangle_A$ and $|1\rangle_A$, and $V$, of basis vectors $|0\rangle_V$ and $|1\rangle_V$. We use ket vectors instead of density operators as in the original Deutsch algorithm.

Bob chooses one of the four functions in Table 1, say $f_{01}(\mathbf{a})$, and gives Alice the black box that computes it. Alice knows Table 1 but does not know which is the function chosen by Bob. She is to find whether it is constant or balanced through function evaluations. She prepares register $A$ with the value of $\mathbf{a}$ for which she wants to perform function evaluation. The black box computes the value of $f_{01}(\mathbf{a})$ and adds it modulo two to the former content of register $V$. Being logically reversible, modulo two addition can be implemented unitarily. In the introduction we omitted register $V$ because transformations are unitary also without it, but they are more difficult to explain.

For reasons that will soon become clear, the input state of the quantum algorithm is

$$|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle_A(|0\rangle_V - |1\rangle_V). \tag{4}$$

Alice applies to register $A$ the Hadamard transform $\hat{H}_A$, which transforms $|0\rangle_A$ into $\frac{1}{\sqrt{2}}(|0\rangle_A + |1\rangle_A)$ and $|1\rangle_A$ into $\frac{1}{\sqrt{2}}(|0\rangle_A - |1\rangle_A)$, producing the state

$$\hat{H}_A|\psi\rangle = \frac{1}{2}(|0\rangle_A + |1\rangle_A)(|0\rangle_V - |1\rangle_V), \tag{5}$$

then asks the black box to compute the value of the function. Let $\hat{U}_f$ be the corresponding unitary transformation (defined in the Hilbert space of all registers). We have

$$\hat{U}_f\hat{H}_A|\psi\rangle = \frac{1}{2}(|0\rangle_A - |1\rangle_A)(|0\rangle_V - |1\rangle_V). \tag{6}$$

Function evaluation is performed in quantum parallelism for each term of the input state superposition. It leaves the term $|0\rangle_A(|0\rangle_V - |1\rangle_V)$, appearing in the input state 5, unaltered. In fact, here the argument of the function, the content of register $A$, is 0. The computation of $f_{01}(0)$ yields 0 that module two added to the former content of register $V$ leaves everything unaltered. Function evaluation instead changes the term $|1\rangle_A (|0\rangle_V - |1\rangle_V)$ into $|1\rangle_A (|1\rangle_V - |0\rangle_V) = -|1\rangle_A (|0\rangle_V - |1\rangle_V)$. In fact now we have to module two add $f_{01}(1) = 1$ and this changes $|0\rangle_V$ into $|1\rangle_V$ and $|1\rangle_V$ into $|0\rangle_V$.

Then Alice applies a second time the Hadamard transform to register $A$, obtaining the output state

$$\hat{H}_A \hat{U}_f \hat{H}_A |\psi\rangle = \frac{1}{\sqrt{2}} |1\rangle_A (|0\rangle_V - |1\rangle_V). \qquad (7)$$

Eventually she measures the *content* of register $A$, namely the observable $\hat{A}$ of eigenstates $|0\rangle_A$ and $|1\rangle_A$ and eigenvalues respectively 0 and 1. She reads the eigenvalue 1, which tells her that the function is balanced (the final content of register $A$ is 0 when the function is constant and 1 when it is balanced).

Thus the problem of checking whether the function given by Bob is constant or balanced is always solved with just one function evaluation quantumly, against two classically.

The mathematics of the speedup in this quantum algorithm is obvious, in the sense that we have it under the eyes. However, the mathematics of different quantum algorithms is different from one algorithm to another as there is no known universal scheme. The *mechanism* of the speedups, provided there is such one, is not known.

## 4.1 Time-symmetric and relativized representations

To start with, we extend the representation of Deutsch algorithm to the process of choosing the black box. To this end, we should add an imaginary quantum register $B$ of basis vectors $|00\rangle_B$, $|01\rangle_B$, $|10\rangle_B$, and $|11\rangle_B$. This register contains the problem setting, namely the suffix **b** of the function chosen by Bob. The previous black box, which computed $f_{\mathbf{b}}(\mathbf{a})$ for a well determined value of **b** and any value of **a**, is replaced by a universal one that computes $f_{\mathbf{b}}(\mathbf{a})$ for any values of **b** and **a**. Registers $A$ and $V$ have the same role as before.

For reasons that will soon become clear, we assume that register $B$ is initially in the maximally mixed state

$$\hat{\rho}_B = \frac{1}{4}(|00\rangle\langle 00|_B + |01\rangle\langle 01|_B + |10\rangle\langle 10|_B + |11\rangle\langle 11|_B). \qquad (8)$$

As we will need a detailed representation of quantum states and operators, for reasons of encumbrance we represent all states as ket vectors, not matrices. To this end, we move to the random phase representation [32] of the maximally mixed state of register $B$ that is

$$|\psi\rangle_B = \frac{1}{2}(e^{\imath\varphi_0}|00\rangle_B + e^{\imath\varphi_1}|01\rangle_B + e^{\imath\varphi_2}|10\rangle_B + e^{\imath\varphi_3}|11\rangle_B), \qquad (9)$$

where the $\varphi_i$ are independent random phases with uniform distribution in $[0, 2\pi]$. We will be dealing with a trivial application of the random phase representation: we can always think that the quantum state evolves as a pure state with the $\varphi_i$ fixed phases. Only when we have to compute its von Neumann entropy, we should remember that the $\varphi_i$ are random variables. The von Neumann entropy of state $|\psi\rangle_B$, as that of $\hat{\rho}_B$, is 2 bits.

By the way, $\hat{\rho}_B$ is the mathematical average over all the $\varphi_i$ of the outer product $|\psi\rangle_B \langle\psi|_B$; reading state $|\psi\rangle_B$ is also simple: it is a mixture of pure states with the phases $\varphi_0$, $\varphi_1$, $\varphi_2$, $\varphi_3$ all different, in fact a dephased quantum superposition.

The overall initial state of the three registers, at time $t_0$, is thus

$$|\psi\rangle = \frac{1}{\sqrt{8}}(e^{\imath\varphi_0}|00\rangle_B + e^{\imath\varphi_1}|01\rangle_B + e^{\imath\varphi_2}|10\rangle_B + e^{\imath\varphi_3}|11\rangle_B)|0\rangle_A(|0\rangle_V - |1\rangle_V). \qquad (10)$$

In order to prepare register $B$ in the desired problem setting, at time $t_0$ Bob measures its content, namely the observable $\hat{B}$ of eigenstates the basis vectors $|00\rangle_B, |01\rangle_B, \ldots$ and eigenvalues respectively $00, 01, \ldots$ Note that $\hat{B}$ commutes with $\hat{A}$. The measurement outcome is completely random. Say it comes out the eigenvalue $\mathbf{b} = 10$. The state immediately after measurement is

$$\hat{P}_B |\psi\rangle = \frac{1}{\sqrt{2}} |10\rangle_B |00\rangle_A (|0\rangle_V - |1\rangle_V), \qquad (11)$$

where $\hat{P}_B$ is the projection of the quantum state induced by Bob's measurement. Then Bob applies to register $B$ a unitary transformation $\hat{U}_B$ that changes the random measurement outcome into the desired problem setting, say $\mathbf{b} = 01$. At time $t_1$ we will have

$$\hat{U}_B \hat{P}_B |\psi\rangle = \frac{1}{\sqrt{2}} |01\rangle_B |0\rangle_A (|0\rangle_V - |1\rangle_V). \qquad (12)$$

State 12 is the input state of the quantum algorithm in the representation extended to the process of setting the problem. There are of course many $\hat{U}_B$ that change $|10\rangle_B$ into $|01\rangle_B$. For simplicity of exposition, we choose the one that bit by bit changes zeros into ones and ones into zeros

$$\hat{U}_B \equiv |11\rangle\langle 00|_B + |10\rangle\langle 01|_B + |01\rangle\langle 10|_B + |00\rangle\langle 11|_B. \qquad (13)$$

The output state of the extended representation of the quantum algorithm is

$$\hat{H}_A \hat{U}_f \hat{H}_A \hat{U}_B \hat{P}_B |\psi\rangle = \frac{1}{\sqrt{2}} |01\rangle_B |1\rangle_A (|0\rangle_V - |1\rangle_V). \qquad (14)$$

Of course, input and output states are the same as in the usual representation of the quantum algorithm up to the presence of the ket $|01\rangle_B$.

We note that this extension immediately calls for another one, this time concerning the actors (observers) on the stage. We resort to relational quantum mechanics where quantum states are observer-dependent. State 12 is *with respect to* Bob, the problem setter, and any other observer who does not act on the problem solving process. It cannot be with respect to Alice, the problem solver. The sharp state $|01\rangle_B$ would tell her, before she starts her search for the solution, that the function chosen by Bob is $f_{01}(\mathbf{a})$. She would know that it is balanced without performing any function evaluation. The suffix of the function should be hidden to Alice because to her it is inside the black box. To physically represent this fact, it suffices to retard the projection $\hat{P}_B$ until the end of the unitary part of Alice's action, at time $t_2$.

To her, the state of register $B$ in the input state of the quantum algorithm is still maximally mixed. In fact $\hat{U}_B$ leaves state 10 unaltered up to an irrelevant permutation of the independent random phases. Thus, disregarding the permutation, state 10 is the input state to Alice.

We started with register $B$ in a maximally mixed state to represent the fact that, to Alice, the problem setting is physically hidden.

Summing up, states 10 through 14 are the representation of the quantum algorithm with respect to Bob. In the representation with respect to Alice, the input state, which coincides with the initial state, is

$$\hat{U}_B|\psi\rangle = |\psi\rangle = \frac{1}{\sqrt{8}}(e^{\imath\varphi_0}|00\rangle_B + e^{\imath\varphi_1}|01\rangle_B + e^{\imath\varphi_2}|10\rangle_B \\ +e^{\imath\varphi_3}|11\rangle_B)|0\rangle_A(|0\rangle_V - |1\rangle_V). \qquad (15)$$

The two bit entropy of the state of register $B$ represents Alice's complete ignorance of the problem setting. The output state is

$$\hat{H}_A\hat{U}_f\hat{H}_A\hat{U}_B|\psi\rangle = \frac{1}{\sqrt{8}}\Big[(e^{\imath\varphi_0}|00\rangle_B + e^{\imath\varphi_3}|11\rangle_B)|0\rangle_A \\ +(e^{\imath\varphi_1}|01\rangle_B - e^{\imath\varphi_2}|10\rangle_B)|1\rangle_A\Big](|0\rangle_V - |1\rangle_V), \qquad (16)$$

We can see that, for each possible problem setting (value of $\mathbf{b}$ contained in register $B$), Alice has built the corresponding solution of the problem $s(\mathbf{b})$ in register $A$.

Eventually, at time $t_2$, she acquires the solution by reading the content of register $A$, namely by measuring $\hat{A}$. We should keep in mind that the output state 16 is with respect to Alice. The same state with respect to Bob and any other observer is $\frac{1}{\sqrt{2}}|01\rangle_B|1\rangle_A(|0\rangle_V - |1\rangle_V)$. The measurement outcome is unpredictable to Alice, it is already 1 to any other observer. Thus Alice's measurement must select the eigenvalue 1 with probability one, projecting state 16 on

$$\hat{P}_A\hat{H}_A\hat{U}_f\hat{H}_A\hat{U}_B|\psi\rangle = \frac{1}{\sqrt{2}}(e^{\imath\varphi_1}|01\rangle_B - e^{\imath\varphi_2}|10\rangle_B) \\ \otimes|1\rangle_A(|0\rangle_V - |1\rangle_V), \qquad (17)$$

where $\hat{P}_A$ is the projection induced by the final Alice's measurement. State 17 is further projected on

$$\frac{1}{\sqrt{2}}|01\rangle_B|1\rangle_A(|0\rangle_V - |1\rangle_V) \qquad (18)$$

by the retarded projection induced by the initial Bob's measurement. We note that inverting the order of the two projections leaves the end result unaltered. As a matter of fact, since the projection due to Bob's measurement cannot be retarded beyond the unitary part of Alice's action, we should see the two projections as simultaneous. In this way Alice, by measuring $\hat{A}$, also acquires the content of register $B$. In fact state 18, with register $B$ in the sharp state $|01\rangle_B$, tells Alice that the problem setting chosen by Bob is $\mathbf{b} = 01$. This state is common to the representation with respect to Alice and to that with respect to Bob.

In view of what will follow, we note that Alice's measurement of $\hat{A}$ in the output state with respect to her is equivalent to the measurement of $\hat{B}$. In fact either measurement projects state 16 on state 18, where the sharp states of registers $B$ and $A$ tell Alice both the setting and the solution of the problem.

## 4.2 Quantum feedback

The random phase representation of the reduced density operator of register $A$ in the output state 16 is

$$|\psi\rangle_A = \frac{1}{\sqrt{2}}\left(e^{\imath\Phi_0}|0\rangle_A + e^{\imath\Phi_1}|1\rangle_A\right), \qquad (19)$$

where $\Phi_0$ and $\Phi_1$ are independent random phases with uniform distribution in $[0, 2\pi]$. The usual density matrix representation is

$$\hat{\rho}_A = \frac{1}{2}(|0\rangle_A\langle 0|_A + |1\rangle_A\langle 1|_A). \qquad (20)$$

Here, $\mathcal{E}_A$, the entropy of $|\psi\rangle_A$ or $\hat{\rho}_A$, is 1 bit. The zeroing of $\mathcal{E}_A$ can be due to either the projection of the quantum state associated with the measurement of $\hat{B}$ in the initial state 10, retarded at the end of the unitary part of Alice's action, or that associated with the measurement of $\hat{A}$ in the output state 16 (we have seen that the two projections should be considered simultaneous). The present work is an exploration of the assumption that the zeroing of $\mathcal{E}_A$ shares between the two measurements.

To this end, we assume that the two complete measurements reduce to partial measurements that obey the following two *Occam conditions*:

(1) together, they select whatever was selected by the complete measurements, and

(2) each performed alone, contribute in a complementary and non-redundant way to the zeroing of $\mathcal{E}_A$. By this we mean that no information provided by either partial measurement is provided by the other.

The assumption that the two partial measurements contribute equally to the zeroing of $\mathcal{E}_A$, namely that $R = \frac{1}{2}$, explains the speedup of the present quantum algorithm.

We should reduce the initial Bob's measurement and the final Alice's measurement to two partial measurements submitted to (1) and (2), and the condition of equally contributing to the zeroing of $\mathcal{E}_A$.

We have seen that the measurement of $\hat{A}$ in the relativized output state 16 is equivalent to that of $\hat{B}$. Thus we can move to the problem of reducing two measurements of $\hat{B}$, one performed by Bob in the initial state 10 and the other by Alice in the output state 16, to two partial measurements, say of $\hat{B}_i$ and $\hat{B}_j$, satisfying Occam conditions. In the most general terms, $\hat{B}_i$ and $\hat{B}_j$ are Boolean functions of $\hat{B}$, such as: $\hat{B}_0$, the content of the left cell of register $B$, $\hat{B}_1$, the content of the right cell, $\mathrm{XOR}\left(\hat{B}_0, \hat{B}_1\right)$, the exclusive or between the two former contents, etc.

We provide an example of reduction of the complete measurements to such partial measurements. We keep the assumption that the initial measurement of $\hat{B}$ randomly selects the eigenvalue $\mathbf{b} = 10$ and that Bob, by $\hat{U}_B$, changes it into $\mathbf{b} = 01$. Let $\mathbf{b} \equiv b_0 b_1$; we assume that the eigenvalue $b_0 = 1$ is selected at time $t_0$ by the measurement of $\hat{B}_0$ in the initial state and that the eigenvalue $b_1 = 1$ is selected at time $t_2$ by the measurement of $\hat{B}_1$ in the output state.

To reconstruct the selections performed by the complete measurements, we should propagate forward in time, by $\hat{H}_A \hat{U}_f \hat{H}_A \hat{U}_B$, the projection induced by the former measurement and backward in time, by its inverse, the projection induced by the latter measurement. The two propagations can be performed in any order, the reconstruction is the same.

Let us perform the backward propagation first. The measurement of $\hat{B}_1$ in the output state 16, which assumedly selects $b_1 = 1$, projects this state on

$$|\chi\rangle = \frac{1}{2} \left(e^{\imath\varphi_1} |01\rangle_B |1\rangle_A + e^{\imath\varphi_3} |11\rangle_B |0\rangle_A\right) \left(|0\rangle_V - |1\rangle_V\right). \tag{21}$$

We advance at time $t_0$ the two ends of this projection. The result is the projection of the initial state 10 on

$$\hat{U}_B^\dagger \hat{H}_A^\dagger \hat{U}_f^\dagger \hat{H}_A^\dagger |\chi\rangle = \frac{1}{2} \left(e^{\imath\varphi_3} |00\rangle_B + e^{\imath\varphi_1} |10\rangle_B\right) \otimes |0\rangle_A \left(|0\rangle_V - |1\rangle_V\right). \tag{22}$$

The permutation of the independent random phases is irrelevant. At this point the measurement of $\hat{B}_0$ in state 22, which assumedly selects $b_0 = 1$, projects it on

$$|\xi\rangle = \frac{1}{\sqrt{2}} |10\rangle_B |0\rangle_A \left(|0\rangle_V - 1_V\right). \tag{23}$$

Of course, the state 23, under $\hat{H}_A \hat{U}_f \hat{H}_A \hat{U}_B$, evolves into state 18, the final state common to both representations (to Bob and to Alice). We have reconstructed the selections performed by the complete measurements. Furthermore, the reduction of $\mathcal{E}_A$ induced by either partial measurement, performed alone, is half bit and no information acquired by either partial measurement is acquired by the other. Conditions (1) and (2) are satisfied.

One can see that, eventually, everything boils down to ascribing the selection of one of the two bits (the right one in present assumptions) of the random outcome of the initial measurement to the final measurement. We are not sending a message backward in time. Each of the bits that specify the outcome of the initial measurement is independently and randomly selected. We are just ascribing half of these random selections to the final rather than the initial measurement.

We note that sharing between Bob's and Alice's measurements the zeroing of $\mathcal{E}_A$ does not affect Bob's freedom of choosing the function computed by the black box. We should keep in mind that the probability that Alice's measurement of $\hat{B}$ in state 16 selects $\mathbf{b} = 01$, or that the measurement of $\hat{B}_1$ selects $b_1 = 1$ (the right digit of 01), is one. This means that the measurement of $\hat{B}_1$ just reads the right digit of the problem setting $\mathbf{b} = 01$ freely chosen (determined) by Bob, without possibly altering it, or affecting Bob's freedom of choosing it. This goes along with the fact that the backward propagation of the projection due to the measurement of $\hat{B}_1$ in the output state does not determine any part of Bob's choice, but the right digit of the random outcome of Bob's measurement $\mathbf{b} = 10$, which is before that choice.

The kind of retrocausation discussed above has been invoked in various articles to explain Einstein–Podolsky–Rosen non-locality, see for example [28]. It has no consequence in the representation of the quantum algorithm with respect to Bob and any external observer. To them, it leaves the input state of the algorithm, namely state 12, unaltered. It just tells that, say, the left digit of the random outcome of Bob's measurement $\mathbf{b} = 10$ has been randomly selected by Bob's measurement and the right digit has been randomly selected back in time by the future Alice's measurement, which is an apparently inconsequential fact.

Things change dramatically in the representation with respect to Alice who is the problem solver.

We have seen that the projection induced by Alice's measurement of $\hat{B}_1$ in the output state 16 must propagate backward in time through the inverse of $\hat{H}_A \hat{U}_f \hat{H}_A \hat{U}_B$ until $t_0$, where it selects the right digit of the random outcome of Bob's measurement 10. Let us see the value of this backward propagation at time $t_1$, immediately after the application of $\hat{U}_B$ and before that of $\hat{H}_A \hat{U}_f \hat{H}_A$. This time we should advance the two ends of the projection of state 16 on state 21 by the inverse of $\hat{H}_A \hat{U}_f \hat{H}_A$. The result is the projection of state 15, the input state of the quantum algorithm with respect to Alice, on

$$\hat{H}_A^\dagger \hat{U}_f^\dagger \hat{H}_A^\dagger \left|\chi\right\rangle = \frac{1}{2}\left(e^{\imath\varphi_1}\left|01\right\rangle_B + e^{\imath\varphi_3}\left|11\right\rangle_B\right) \otimes \left|0\right\rangle_A \left(\left|0\right\rangle_V - \left|1\right\rangle_V\right). \qquad (24)$$

This is an outstanding consequence. State 24, the input state to Alice under the assumption that the selection of the solution equally shares between Bob's and Alice's measurements, tells her, before she performs any function evaluation, that the suffix of the function chosen by Bob is either $\mathbf{b} = 01$ or $\mathbf{b} = 11$, namely that $\mathbf{b} \in \{01, 11\}$. We can say that Alice *knows in advance* that $\mathbf{b} \in \{01, 11\}$, since this knowledge comes from the projection of the quantum state induced by her future measurement.

We are at a fundamental level where knowing is doing (David Finkelstein, private communication). Alice is the problem solver, her knowing in advance that $\mathbf{b} \in \{01, 11\}$ would simply mean that the quantum algorithm requires the number of function evaluations classically required to identify the solution starting from that knowledge. This interpretation seems to exactly fit a Feynman's sum over classical histories representation of the quantum algorithm. The quantum algorithm can be seen as a sum over classical histories in each of which Alice knows in advance one of the possible halves of the information that specifies the problem setting and performs the function evaluations classically required to find the solution. This also accounts for there being a plurality of ways of taking half of the information: all these possibilities are taken in quantum superposition.

In the present case, the number of function evaluations required to discriminate between $f_{01}(\mathbf{a})$ and $f_{11}(\mathbf{a})$ is just one. The value of the function for the argument $\mathbf{a} = 0$ does the job as can be seen from the tables of the two functions in question in Table 1. Since it is 0, the function must be $f_{01}(\mathbf{a})$, what implies that it is balanced.

Let us see this in more detail. A classical history is a classical trajectory of the quantum registers, namely a causal sequence of sharp register states. For example

$$e^{\imath\varphi_1}\left|01\right\rangle_B \left|0\right\rangle_A \left|0\right\rangle_V \overset{\hat{H}_A}{\rightarrow} e^{\imath\varphi_1}\left|01\right\rangle_B \left|0\right\rangle_A \left|0\right\rangle_V \overset{\hat{U}_f}{\rightarrow} e^{\imath\varphi_1}\left|01\right\rangle_B \left|0\right\rangle_A \left|0\right\rangle_V \overset{\hat{H}_A}{\rightarrow} e^{\imath\varphi_1}\left|01\right\rangle_B \left|1\right\rangle_A \left|0\right\rangle_V. \qquad (25)$$

The left-most state is one of the elements of the input state superposition 15. The state after each arrow is one of the elements of the superposition generated by the unitary transformation of the state before the arrow; the transformation in question is specified above the arrow.

In history 25, the problem setting is $\mathbf{b} = 01$. Alice performs function evaluation for $\mathbf{a} = 0$ (second and third state). This behavior is justifiable by two instances of Alice's advanced knowledge. One is $\mathbf{b} \in \{01, 11\}_B$, the other $\mathbf{b} \in \{01, 10\}_B$. The value of the function for $\mathbf{a} = 0$ in either case tells that the function in the black box is $f_{01}(\mathbf{a})$ and thus that it is balanced.

## 5 Generalization

We have shown that $R = \frac{1}{2}$ explains the speedup of Deutsch algorithm. Implicit in this demonstration is a capability to compute the number of function evaluations required to solve Deutsch's problem with retrocausality $R = \frac{1}{2}$. Now we make this capability explicit, while extending it to the generic oracle problem. In other words, given a generic oracle problem (with no need to know the quantum algorithm that solves it), we show how to compute the number of function evaluations required to solve it with retrocausality $R = \frac{1}{2}$ according to the present model.

We focus on $R = \frac{1}{2}$ for the conjecture that this value of $R$ always yields the order of magnitude of the number of function evaluations required to solve the problem in an optimal quantum way. This is the case of all the quantum algorithms examined in this work. We will further discuss the plausibility of this conjecture in Section 9.

A generic oracle problem can be formulated as follows. We have a set of functions $f_{\mathbf{b}} : \{0, 1\}^n \to \{0, 1\}^m$ with $m \le n$. The suffix $\mathbf{b}$ ranges over the set of all the problem settings $\sigma_B$. Bob chooses one of these functions (a value of $\mathbf{b}$) and gives Alice the black box (oracle) that computes it. Alice knows the set of functions but does not know which is the function chosen by Bob. She is to find a certain feature of the function (for example, whether it is constant or balanced in the algorithm of Deutsch, or its period in that of Shor) by performing function evaluations (oracle queries). We call the feature in question, which is the solution of the problem and a function of $\mathbf{b}$, $s(\mathbf{b})$.

### 5.1 Time-symmetric representation to Alice

Provided that a register $B$ contains the problem setting $\mathbf{b}$ and a register $A$ will eventually contain the solution $s(\mathbf{b})$, the most general form of the input and output states of the unitary part $\hat{U}$ of Alice's problem-solving action, in the representation of the quantum algorithm to her, is

$$|\mathrm{in}\rangle_{BAW} = \frac{1}{\sqrt{c}} \left( \sum_{\mathbf{b}\,\in\,\sigma_B} e^{\imath\varphi_{\mathbf{b}}} |\mathbf{b}\rangle_B \right) |00\ldots\rangle_A |\psi\rangle_W , \quad (26)$$

$$\begin{aligned} |\mathrm{out}\rangle_{BAW} &= \hat{U} |\mathrm{in}\rangle_{BAW} \\ &= \frac{1}{\sqrt{c}} \sum_{\mathbf{b}\,\in\,\sigma_B} e^{\imath\varphi_{\mathbf{b}}} |\mathbf{b}\rangle_B |s(\mathbf{b})\rangle_A |\varphi(\mathbf{b})\rangle_W , \end{aligned} \quad (27)$$

where $c$ is the cardinality of $\sigma_B$, $|\psi\rangle_W$ and $|\varphi(\mathbf{b})\rangle_W$ are normalized states of a register $W$, which stands for any other register or set of registers.

$\hat{U}$ should not change the problem setting. It suffices that register $B$ is the control register of all function evaluations, what means that the content of register $B$ affects the output of function evaluation while remaining unaltered through it, and the unitary transformations before and after each function evaluation do not apply to $B$. Correspondingly, $\hat{U}$ sends the input into the output independently term by term and keeping the value of **b** unaltered

$$\forall \mathbf{b} : \ \hat{U} |\mathbf{b}\rangle_B |00\ldots\rangle_A |\psi\rangle_W = |\mathbf{b}\rangle_B |s(\mathbf{b})\rangle_A |\varphi(\mathbf{b})\rangle_W . \quad (28)$$

Given the oracle problem, namely all the pairs **b** and $s(\mathbf{b})$, and provided that one is free to add suitable *garbage qubits* to register $W$, it should not be difficult to put the input and output states in a form compatible with the existence of such a $\hat{U}$ between them. In the following, we assume that states 26 and 27 are of this form. We will see that we do not need to know the form of $\hat{U}$ to the end of ascertaining the number of function evaluations required to solve the oracle problem with quantum retrocausality $R = \frac{1}{2}$; it suffices to know all the pairs **b** and $s(\mathbf{b})$.

Note that, for Eq. 28, the projection of the quantum state induced by any measurement on the content of register $B$ in the output state, advanced by $\hat{U}^\dagger$, becomes the projection induced by performing the same measurement in the input state. Conversely, the projection induced by any measurement on the content of $B$ in the input state, retarded by $\hat{U}$, becomes the projection induced by performing the same measurement in the output state. This goes along with the fact that the reduced density operator of register $B$ remains the same throughout $\hat{U}$. Its random phase representation is

$$|\psi\rangle_B = \frac{1}{\sqrt{c}} \sum_{\mathbf{b}\,\in\,\sigma_B} e^{\imath\varphi_{\mathbf{b}}} |\mathbf{b}\rangle_B , \quad (29)$$

and the usual density matrix representation is

$$\hat{\rho}_B = \frac{1}{c} \sum_{\mathbf{b}\,\in\,\sigma_B} |\mathbf{b}\rangle \langle \mathbf{b}|_B . \quad (30)$$

## 5.2 Quantum feedback

Given Eqs. 26 and 27, we show how to share the selection of the solution between initial and final measurements and derive the corresponding Alice's advanced knowledge in the case $R = \frac{1}{2}$.

It is simpler to assume that $\hat{U}_B$ is the identity. In this way we can think that the initial Bob's measurement is performed in state 26. Of course, its selection of a value of **b** also determines that of $s(\mathbf{b})$.

We reformulate Occam conditions (1) and (2) for the particular case $R = \frac{1}{2}$. We should reduce in all the possible ways the two measurements of $\hat{B}$, one on the part of Bob in the input state and the other on the part of Alice in the output state (see Section 4.2), to two partial measurements, of $\hat{B}_i$ and $\hat{B}_j$, such that:

(1′) together, they select whatever is selected by the complete measurements, and

(2′) each performed alone, they contribute in an equal and non-redundant way to the selection of the solution.

Let $\mathcal{E}_A$ be the von Neumann entropy of the solution, namely of the trace over registers $B$ and $W$ of state $|\mathrm{out}\rangle_{BAW}$. Point (2′) implies the following two conditions

$$\Delta\mathcal{E}_A(\hat{B}_i) = \Delta\mathcal{E}_A(\hat{B}_j), \quad (31)$$

where $\Delta\mathcal{E}_A(\hat{B}_i)$ is the reduction of $\mathcal{E}_A$ due to the measurement of $\hat{B}_i$, $\Delta\mathcal{E}_A(\hat{B}_j)$ is the reduction of $\mathcal{E}_A$ due to the measurement of $\hat{B}_j$, and

$$\boxed{\text{no partial measurement outcome provides enough information to select the solution}} \quad (32)$$

In fact the cases are two: if both outcomes provided enough information, then there would be redundant information, what is forbidden by the no-redundancy condition. If only one did, then the two partial measurements would not contribute equally to the selection of the solution, what is forbidden by the equality condition. Condition 32 is redundant when **b** is an unstructured bit string as in Deutsch algorithm, it is not when **b** is structured.

Alice's measurement of $\hat{B}_j$ (as any measurement of a Boolean function of $\hat{B}$), performed alone, must induce a projection of the output state 27 on a state of the general form

$$|\chi\rangle = \frac{1}{\sqrt{c'}} \sum_{\mathbf{b}\,\in\,\sigma'_B} e^{\imath\varphi_{\mathbf{b}}} |\mathbf{b}\rangle_B |s(\mathbf{b})\rangle_A |\varphi(\mathbf{b})\rangle_W , \quad (33)$$

where $\sigma'_B$ is a subset of $\sigma_B$ of cardinality $c'$. Alice's advanced knowledge is obtained by advancing by $\hat{U}^\dagger$ the two ends of this projection at the input of the quantum algorithm, at time $t_1$ immediately after the preparation of the problem setting (which here is the outcome of Bob's measurement). Even without knowing $\hat{U}^\dagger$, we know that, for Eq. 28, this projects the input state 26 on

$$\frac{1}{\sqrt{c'}}\left(\sum_{\mathbf{b}\,\in\,\sigma'_B} e^{\imath\varphi_{\mathbf{b}}}\,|\mathbf{b}\rangle_B\right)|00\ldots\rangle_A\,|\psi\rangle_W\,. \tag{34}$$

In particular, it projects the maximally mixed state 29 of register $B$ on the state of lower entropy

$$\frac{1}{\sqrt{c'}}\sum_{\mathbf{b}\,\in\,\sigma'_B} e^{\imath\varphi_{\mathbf{b}}}\,|\mathbf{b}\rangle_B\,, \tag{35}$$

which represents Alice's advanced knowledge, namely Alice knows in advance that $\mathbf{b}\in\sigma'_B$. For short we say that Alice's measurement of $\hat{B}_j$ projects $\sigma_B$ on $\sigma'_B$.

Still for Eq. 28, the same projection can be obtained by measuring $\hat{B}_j$ in the input state. We also note that, mathematically, nothing changes if we assume to start with that the two complete measurements reduce to two partial measurements, of $\hat{B}_i$ and $\hat{B}_j$, both performed in the input state. Conditions (1′) and (2′) define the same pairs of partial observables $\hat{B}_i$ and $\hat{B}_j$ no matter whether $\hat{B}_j$ is measured in the input or output state. In fact moving the measurement from the output to the input state leaves all selections and reductions of the entropy of the solution unaltered.

This latter way of assessing Alice's advanced knowledge highlights a symmetry hidden in the former one. We are left with two partial measurements of the content of register $B$ that satisfy conditions (1′) and (2′), both performed in the input state. We can lose the memory of which partial measurement is performed by Alice and which by Bob. Evidently, either partial measurement can be the one performed by Alice. Therefore, given a pair of partial measurements, of $\hat{B}_i$ and $\hat{B}_j$, in the input state 26 that satisfy conditions (1′) and (2′), either partial measurement performed alone projects the maximally mixed state 29 of register $B$ on an instance of Alice's advanced knowledge. By the way, in this sense we can say that, with quantum retrocausality $R=\frac{1}{2}$, Alice knows *half* of the problem setting in advance.

It is important to note that register $W$, which we have considered for generality and could be necessary to construct the quantum algorithm, is not involved in the definition of the pairs $\hat{B}_i$ and $\hat{B}_j$. Let us recall the conditions their measurements (which can be both performed in the input state 26) are submitted to: (i) together, they select a value of $\mathbf{b}$, (ii) the information acquired by either measurement is not acquired by the other, (iii) they satisfy Eq. 31, and (iv) they satisfy requirement 32. Conditions (i), (ii) and (iv) only involve the input state of register $B$, namely state 29. Also condition (iii) does not involve register $W$, as the reductions of the entropy of the solution $\Delta\mathcal{E}_A(\hat{B}_i)$ and $\Delta\mathcal{E}_A(\hat{B}_j)$ concern the trace of the output state 27 over registers $B$ and $W$.

Therefore, to the end of determining $\hat{B}_i$ and $\hat{B}_j$, we can work with $|\text{in}\rangle_{BA}$ and $|\text{out}\rangle_{BA}$, the traces over $W$ of $|\text{in}\rangle_{BAW}$ and $|\text{out}\rangle_{BAW}$; it suffices to drop $|\psi\rangle_W$ and the $|\varphi(\mathbf{b})\rangle_W$. States $|\text{in}\rangle_{BA}$ and $|\text{out}\rangle_{BA}$, in turn, can be written solely on the basis of the pairs $\mathbf{b}$ and $s(\mathbf{b})$, namely of the oracle problem.

Since the quantum algorithm can be seen as a sum over classical histories in each of which Alice knows in advance one of the possible halves of the problem setting and performs the function evaluations still necessary to identify the solution, given an oracle problem, we can know the number of function evaluations required to solve it with quantum retrocausality $R=\frac{1}{2}$.

### 5.2.1 Example of application

We apply the present procedure to Deutsch's problem. Of course, we should ignore Deutsch algorithm.

Given the problem, namely all the pairs $\mathbf{b}$ and $s(\mathbf{b})$, we write down $|\text{in}\rangle_{BA}$ and $|\text{out}\rangle_{BA}$ (of course we obtain the traces over register $V$ of states 15 and 16). We do not need to know $\hat{U}$. It suffices to know that there can be a unitary transformation between input and output that satisfies Eq. 28. Under conditions (1′) and (2′), $|\text{in}\rangle_{BA}$ and $|\text{out}\rangle_{BA}$ define the pairs of partial observables $\hat{B}_i$ and $\hat{B}_j$ we are looking for, in particular it is easier to think that they are both measured in $|\text{in}\rangle_{BA}$. It is not a constructive definition, however finding the pairs in question will be easy in all the cases examined in this work. In the case of Deutsch's problem they are any two of the three partial observables: $\hat{B}_0$, $\hat{B}_1$, and $\hat{B}_X\equiv\text{XOR}\left(\hat{B}_0,\hat{B}_1\right)$. These partial observables are Boolean functions of $\hat{B}$ and the measurements of any two of them satisfy conditions (1′) and (2′) with $\Delta\mathcal{E}_A(\hat{B}_i)=\Delta\mathcal{E}_A(\hat{B}_j)=\frac{1}{2}$ bit.

Given a problem setting, say $\mathbf{b}=01$, either partial observable, $\hat{B}_i$ or $\hat{B}_j$, corresponds to an instance of Alice's advanced knowledge as follows. We should assume that its measurement selects the eigenvalue that matches with the problem setting. With problem setting $\mathbf{b}\equiv b_0b_1=01$, this implies that the measurement of $\hat{B}_0$ selects $b_0=0$, that of $\hat{B}_1$ selects $b_1=1$, and that of $\hat{B}_X$ selects $\text{XOR}(b_0,b_1)=1$. The corresponding projections of $\sigma_B$ are respectively on $\{00,01\}_B$, $\{01,11\}_B$, and $\{01,10\}_B$. Thus the instances of Alice's advanced knowledge are $\mathbf{b}\in\{01,00\}_B$, $\mathbf{b}\in\{01,11\}_B$, and $\mathbf{b}\in\{01,10\}_B$, as obvious in hindsight. For any of these instances, Alice can solve the problem with a single function evaluation.

We call the present procedure *the advanced knowledge rule*. Given a generic oracle problem, this rule defines the number of function evaluations required to solve it with quantum retrocausality $R=\frac{1}{2}$. The importance of this rule depends on the confidence that can be placed in the assumption that retrocausality $R = \frac{1}{2}$ is always attainable. This is the case in all the quantum algorithms examined in the present work. Whether it is the case in general should be the object of further work, the present one is an exploration.

# 6 Grover Algorithm

Bob hides a ball in one of $N = 2^n$ drawers (equivalently, he marks an item in an unstructured database of size $N$). Alice is to locate it by opening drawers. In the classical case, to be *a priori* certain of locating the ball, Alice should plan to open $O(N)$ drawers, in the case of Grover [33] quantum search algorithm $O\left(\sqrt{N}\right)$.

The problem, an oracle one, is formalized as follows. Let $\mathbf{b}$ and $\mathbf{a}$, belonging to $\{0,1\}^n$, be respectively the number of the drawer with the ball and that of the drawer that Alice wants to open. Checking whether the ball is in drawer $\mathbf{a}$ amounts to evaluating the function $f_{\mathbf{b}}(\mathbf{a}) : \{0,1\}^n \rightarrow \{0,1\}$, which is 1 if $\mathbf{a} = \mathbf{b}$ and 0 otherwise.

Bob chooses one of the functions $f_{\mathbf{b}}(\mathbf{a})$ (that is a value of $\mathbf{b}$) and gives Alice the black box that computes it. Alice is to find the value of $\mathbf{b}$ chosen by Bob by performing function evaluations for appropriate values of $\mathbf{a}$.

We will distinguish between $n = 2$ and $n > 2$. The speedup of Grover's algorithm with $n = 2$ is explained by $R = \frac{1}{2}$. When $n$ goes past 2, $R$ slightly goes above $\frac{1}{2}$, to go back to $\frac{1}{2}$ for $n \rightarrow \infty$.

## 6.1 Grover algorithm with $n = 2$

### 6.1.1 Time-symmetric representation to Alice

The input and output states of the quantum algorithm to Alice are respectively

$$\hat{U}_B|\psi\rangle = |\psi\rangle = \frac{1}{\sqrt{8}}(e^{\imath\varphi_0}|00\rangle_B + e^{\imath\varphi_1}|01\rangle_B + e^{\imath\varphi_2}|10\rangle_B + e^{\imath\varphi_3}|11\rangle_B)|00\rangle_A(|0\rangle_V - |1\rangle_V), \qquad (36)$$

$$\hat{\Im}_A\hat{U}_f\hat{H}_A\hat{U}_B|\psi\rangle = \frac{1}{\sqrt{8}}(e^{\imath\varphi_0}|00\rangle_B|00\rangle_A + e^{\imath\varphi_1}|01\rangle_B|01\rangle_A + e^{\imath\varphi_2}|10\rangle_B|10\rangle_A + e^{\imath\varphi_3}|11\rangle_B|11\rangle_A)(|0\rangle_V - |1\rangle_V). \quad (37)$$

The function of registers $B$, $A$, and $V$ is as in Deutsch algorithm. $\hat{U}_B$ unitarily transforms the random outcome of Bob's measurement into the desired problem setting, $\hat{H}_A$ is the Hadamard transform on register $A$, $\hat{U}_f$ is function evaluation, and $\hat{\Im}_A$ is a unitary transformation on register $A$ that is called *inversion about the mean*. Note that we could write the input and output states of registers $B$ and $A$ only on the basis of the pairs $\mathbf{b}$ and $s(\mathbf{b})$, and without knowing Grover algorithm. The state of register $V$ is irrelevant for the determination of Alice's advanced knowledge.

Measuring $\hat{A}$ in the output state 37 yields the number of the drawer with the ball chosen by Bob.

### 6.1.2 Quantum feedback

We apply the advanced knowledge rule to Grover's problem with $n = 2$. This yields the number of function evaluations required to solve the problem with quantum retrocausality $R = \frac{1}{2}$. The pairs of partial observables are the same as in Deutsch algorithm: all the pairs among $\hat{B}_0$, $\hat{B}_1$, and $\hat{B}_X$. One can see that they satisfy conditions (1′) and (2′) with $\Delta\mathcal{E}_A(\hat{B}_i) = \Delta\mathcal{E}_A(\hat{B}_j) = 1$ bit.

Suppose that the problem setting chosen by Bob is $\mathbf{b} = 01$, namely Bob hides the ball in drawer 01. The instances of Alice's advanced knowledge are: $\mathbf{b} \in \{01, 00\}_B$, $\mathbf{b} \in \{01, 11\}_B$, and $\mathbf{b} \in \{01, 10\}_B$. In other words, Alice knows in advance that the ball is in one of a pair drawers (one of which with the ball in it). This allows her to locate the ball by opening either drawer (that is by performing just one function evaluation).

All the above could be derived solely from $|\mathrm{in}\rangle_{BA}$ and $|\mathrm{out}\rangle_{BA}$, the traces over register $V$ of states 36 and 37, which can be written solely on the basis of the pairs $\mathbf{b}$ and $s(\mathbf{b})$. One does not need to know Grover algorithm. However, it is of course in agreement with the $n = 2$ instance of Grover algorithm. This means that the speedup of this instance is explained by quantum retrocausality $R = \frac{1}{2}$.

The present instance of Grover algorithm can be seen as a sum over classical histories in each of which Alice knows in advance that the ball is in a pair of drawers and locates it by opening either drawer. An example history is

$$e^{\imath\varphi_1}|01\rangle_B|00\rangle_A|0\rangle_V \xrightarrow{\hat{H}_A} e^{\imath\varphi_1}|01\rangle_B|11\rangle_A|0\rangle_V \xrightarrow{\hat{U}_f} e^{\imath\varphi_1}|01\rangle_B|11\rangle_A|0\rangle_V \xrightarrow{\hat{\Im}_A} e^{\imath\varphi_1}|01\rangle_B|01\rangle_A|0\rangle_V. \quad (38)$$

The problem setting is $\mathbf{b} = 01$. Alice performs function evaluation for $\mathbf{a} = 11$ (second and third state). Therefore, we must assume that Alice's advanced knowledge is $\mathbf{b} \in \{01, 11\}_B$. Since the output of function evaluation is zero (the content of register $V$ remains unaltered), she finds that the number of the drawer with the ball is $\mathbf{b} = 01$.

## 6.2 Grover algorithm with $n > 2$

We should make a clarification to start with. If $n > 2$, the original Grover algorithm does not provide the solution of the problem with absolute certainty. For this, one has to resort to the revision of Grover algorithm made by Long [34] (see also [35]). Long's algorithm can be tuned to provide the solution of Grover's problem with certainty with any number of function evaluations provided it is above the minimum number required by the optimal quantum algorithm, which is

$$K = \frac{\pi}{4\arcsin(2^{-\frac{n}{2}})} \approx \frac{\pi}{4}2^{\frac{n}{2}}. \tag{39}$$

Incidentally, this is also the number required by Grover algorithm, which however does not provide the solution with certainty when $n > 2$.

With $R = \frac{1}{2}$, the number of function evaluations required by the present retrocausality model would be $2^{\frac{n}{2}} - 1 \approx 2^{\frac{n}{2}}$. In fact, Alice knows in advance $Rn$ of the $n$ bits that specify the number of the drawer with the ball, thus $\frac{n}{2}$ bits for $R = \frac{1}{2}$. This means that she must open in the worst case $2^{\frac{n}{2}} - 1$ drawers (if all were empty, then she would know that the ball is in the only drawer left).

We note anyhow that also the number of function evaluations foreseen by the advanced knowledge rule, for $R = \frac{1}{2}$, is that of an existing quantum algorithm, which is in fact Long's algorithm tuned on $2^{\frac{n}{2}} - 1$ function evaluations.

When $n$ goes past 2, Alice's advanced knowledge should increase over the $\frac{n}{2}$ bits of the case $R = \frac{1}{2}$, so that the problem can be solved with $\approx \frac{\pi}{4}2^{\frac{n}{2}}$ function evaluations rather than $\approx 2^{\frac{n}{2}}$. This increase must be slight: an increase of just one bit would halve the required number of function evaluations. Correspondingly, $R$ should slightly go above $\frac{1}{2}$. It should also be noted that, for $n \to \infty$, we have $R = \frac{1}{2}$ again.

## 7 Deutsch–Jozsa algorithm

Deutsch–Jozsa [36] algorithm is a generalization of the seminal Deutsch algorithm that yields an exponential speedup. In the respective problem, the set of functions is all the constant and *balanced* functions (with the same number of zeroes and ones) $f_{\mathbf{b}} : \{0,1\}^n \to \{0,1\}$. Four of the eight functions, for $n = 2$, are shown in Table 2.

The bit string $\mathbf{b} \equiv b_0 b_1 \dots b_{2^n-1}$ is both the suffix and the table of the function $f_{\mathbf{b}}(\mathbf{a})$, which is the sequence of function values for increasing values of the argument. Alice is to find whether the function chosen by Bob is constant or balanced by computing $f_{\mathbf{b}}(\mathbf{a})$ for appropriate values of $\mathbf{a}$. Classically, this requires in the worst case a number of function evaluations exponential in $n$. It requires just one function evaluation in the quantum case.

***Table 2:*** *Tabular representation of the functions $f_{\mathbf{b}}(\mathbf{a})$ that are evaluated by the black box in the Deutsch–Jozsa problem.*

| **a** | $f_{0000}(\mathbf{a})$ | $f_{1111}(\mathbf{a})$ | $f_{0011}(\mathbf{a})$ | $f_{1100}(\mathbf{a})$ | ... |
|---|---|---|---|---|---|
| 00 | 0 | 1 | 0 | 1 | ... |
| 01 | 0 | 1 | 0 | 1 | ... |
| 10 | 0 | 1 | 1 | 0 | ... |
| 11 | 0 | 1 | 1 | 0 | ... |

### 7.1 Time-symmetric representation to Alice

The input and output states of the quantum algorithm to Alice are respectively

$$\hat{U}_B|\psi\rangle = |\psi\rangle = \frac{1}{4}(e^{\imath\varphi_0}|0000\rangle_B + e^{\imath\varphi_1}|1111\rangle_B + e^{\imath\varphi_2}|0011\rangle_B + e^{\imath\varphi_3}|1100\rangle_B + \dots)|00\rangle_A(|0\rangle_V - |1\rangle_V), \tag{40}$$

$$\hat{H}_A\hat{U}_f\hat{H}_A\hat{U}_B|\psi\rangle = \frac{1}{4}\Big[(e^{\imath\varphi_0}|0000\rangle_B - e^{\imath\varphi_1}|1111\rangle_B)|00\rangle_A + (e^{\imath\varphi_2}|0011\rangle_B - e^{\imath\varphi_3}|1100\rangle_B)|10\rangle_A + \dots\Big](|0\rangle_V - |1\rangle_V). \tag{41}$$

Registers $B$, $A$, and $V$ and the unitary transformation $\hat{U}_B$ have the same function as in the previous quantum algorithms. $\hat{H}_A$ is the Hadamard transform on register $A$ and $\hat{U}_f$ is function evaluation. Note that we could have written the input and output states of registers $B$ and $A$ only on the basis of the pairs $\mathbf{b}$ and $s(\mathbf{b})$.

Measuring $\hat{A}$ in the output state 41 says that the function is constant if the measurement outcome is all zeros, balanced otherwise.

### 7.2 Quantum feedback

We apply the advanced knowledge rule to Deutsch–Jozsa problem. Given the problem setting of a balanced function, there is only one pair of partial measurements of the content of register $B$ compatible with conditions (1′) and (2′). With problem setting, say, $\mathbf{b} = 0011$, $\hat{B}_i$ must be the content of the left half of register $B$ and $\hat{B}_j$ that of the right half. The measurement of $\hat{B}_i$ yields all zeros, that of $\hat{B}_j$ all ones.

In fact, a partial measurement yielding both zeroes and ones would violate condition 32: it would provide enough information to identify the solution, namely the fact that $f_{\mathbf{b}}$ is balanced. Given that either partial measurement must yield all zeroes or all ones, it must concern the content of half register. Otherwise either Eq. 31 would be violated or the problem setting would not be completely determined, as readily checked.

One can see that, with $\mathbf{b} = 0011$, the measurement of $\hat{B}_i$, performed alone, projects $\sigma_B$ on the subset $\{0011, 0000\}_B$, that of $\hat{B}_j$ on $\{0011, 1111\}_B$. Either subset represents the part of the problem setting that Alice knows in advance. Eq. 31 is satisfied with $\Delta\mathcal{E}_A(\hat{B}_i) = \Delta\mathcal{E}_A(\hat{B}_j) = 1$ bit.

The case of the problem setting of a constant function is analogous. The only difference is that there are more pairs of partial measurements that satisfy the above said conditions. Say that the problem setting is $\mathbf{b} = 0000$. The measurements of the content of the left and right half of register $B$ (each performed alone) projects $\sigma_B$ on respectively $\{0000, 0011\}_B$ and $\{0000, 1100\}_B$, the measurements of the content of even and odd cells (say the leftmost one is odd) on respectively $\{0000, 0101\}_B$ and $\{0000, 1010\}_B$, etc.

There is a shortcut to finding the subsets in question. Here the problem setting, namely the bit string $\mathbf{b}$, is the table of the function chosen by Bob. For example $\mathbf{b} = 0011$ is the table $f_{\mathbf{b}}(00) = 0$, $f_{\mathbf{b}}(01) = 0$, $f_{\mathbf{b}}(10) = 1$, $f_{\mathbf{b}}(11) = 1$. We call *good half table* any half table in which all the values of the function are the same. One can see that good half tables are in one-to-one correspondence with the subsets of $\sigma_B$ in question. For example, the good half table $f_{\mathbf{b}}(00) = 0$, $f_{\mathbf{b}}(01) = 0$ corresponds to the subset $\{0011, 0000\}_B$, is the identical part of the two bitstrings in it. Thus, given a problem setting (that is an entire table), either good half table, or identically the corresponding subset of $\sigma_B$, is a possible instance of Alice's advanced knowledge.

Because of the structure of tables, given the advanced knowledge of a good half table, the entire table and thus the solution can be identified by performing just one function evaluation for any value of the argument $\mathbf{a}$ outside the half table.

Summing up, the advanced knowledge rule says that Deutsch–Jozsa problem can be solved with just one function evaluation. This is in agreement with Deutsch–Jozsa algorithm, what also means that the speedup of this algorithm is explained by quantum retrocausality $R = \frac{1}{2}$.

The present instance of Deutsch–Jozsa algorithm can be seen as a sum over classical histories in each of which Alice knows in advance that Bob has chosen one of a pair of functions and discriminates between the two with just one function evaluation. An example history is

$$e^{\imath\varphi_2} |0011\rangle_B |00\rangle_A |0\rangle_V \overset{\hat{H}_A}{\rightarrow} e^{\imath\varphi_2} |0011\rangle_B |10\rangle_A |0\rangle_V \overset{\hat{U}_f}{\rightarrow} e^{\imath\varphi_2} |0011\rangle_B |10\rangle_A |1\rangle_V \overset{\hat{H}_A}{\rightarrow} e^{\imath\varphi_2} |0011\rangle_B |10\rangle_A |1\rangle_V . \quad (42)$$

Since the problem setting is $\mathbf{b} = 0011$ and Alice performs function evaluation for $\mathbf{a} = 10$, her advanced knowledge must be $\mathbf{b} \in \{0011, 0000\}_B$; if it were $\mathbf{b} \in \{0011, 1111\}_B$, she would have performed function evaluation for either $\mathbf{a} = 00$ or $\mathbf{a} = 01$. The result of function evaluation, $f_{\mathbf{b}}(10) = 1$, tells that the function chosen by Bob is $f_{0011}(\mathbf{a})$, hence that it is balanced.

One can see that the present analysis, like the notion of good half table, holds unaltered for $n > 2$.

## 8 Simon and hidden subgroup algorithms

In Simon's [37] problem, the set of functions is all the $f_{\mathbf{b}} : \{0,1\}^n \rightarrow \{0,1\}^{n-1}$ such that $f_{\mathbf{b}}(\mathbf{a}) = f_{\mathbf{b}}(\mathbf{c})$ if and only if $\mathbf{a} = \mathbf{c}$ or $\mathbf{a} = \mathbf{c} \oplus \mathbf{h}(\mathbf{b})$; $\oplus$ denotes bitwise modulo 2 addition. The bit string $\mathbf{h}(\mathbf{b})$, depending on $\mathbf{b}$, is a sort of period of the function.

Table 3 shows four of the six functions, for $n = 2$. The bit string $\mathbf{b}$ is both the suffix and the table of the function. We note that each value of the function appears exactly twice in each table; thus 50% of the rows plus one always identify $\mathbf{h}(\mathbf{b})$.

Bob chooses one of these functions. Alice is to find the value of $\mathbf{h}(\mathbf{b})$ by performing function evaluation for appropriate values of $\mathbf{a}$.

In present knowledge, a classical algorithm requires a number of function evaluations exponential in $n$. The quantum part of Simon algorithm solves with just one function evaluation the hard part of this problem, namely finding a string $\mathbf{s}_j(\mathbf{b})$ *orthogonal* [37] to $\mathbf{h}(\mathbf{b})$. There are $2^{n-1}$ such strings. Running the quantum part yields one of these strings at random. The quantum part is iterated until finding $n-1$ different strings. This allows Alice to find $\mathbf{h}(\mathbf{b})$ by solving a system of modulo 2 linear equations. Thus, on average, finding $\mathbf{h}(\mathbf{b})$ requires $O(n)$ iterations of the quantum part; in particular $O(n)$ function evaluations. Moreover, if we put an upper bound to the number of iterations, *a priori* there is always a non-zero probability of not finding $n-1$ different strings.

We apply the advanced knowledge rule directly to the complete Simon's problem of finding $\mathbf{h}(\mathbf{b})$ through function evaluations. This is not the problem solved by the quantum part of Simon algorithm, which is finding at random one of the $\mathbf{s}_j(\mathbf{b})$ orthogonal to $\mathbf{h}(\mathbf{b})$. The value of $R$ that explains the speedup of the quantum part of Simon algorithm will be a by-product of applying the advanced knowledge rule to Simon's problem.

*Table 3: Tabular representation of the functions $f_{\mathbf{b}}(\mathbf{a})$ that are evaluated by the black box in the Simon's problem.*

| | $\mathbf{h}(0011) = 01$ | $\mathbf{h}(1100) = 01$ | $\mathbf{h}(0101) = 10$ | $\mathbf{h}(1010) = 10$ | $\dots$ |
|---|---|---|---|---|---|
| $\mathbf{a}$ | $f_{0011}(\mathbf{a})$ | $f_{1100}(\mathbf{a})$ | $f_{0101}(\mathbf{a})$ | $f_{1010}(\mathbf{a})$ | $\dots$ |
| 00 | 0 | 1 | 0 | 1 | $\dots$ |
| 01 | 0 | 1 | 1 | 0 | $\dots$ |
| 10 | 1 | 0 | 0 | 1 | $\dots$ |
| 11 | 1 | 0 | 1 | 0 | $\dots$ |

## 8.1 Time-symmetric representation to Alice

Knowing all the pairs $\mathbf{b}, \mathbf{h}(\mathbf{b})$ from Table 3 we can write $|\text{in}\rangle_{BA}$ and $|\text{out}\rangle_{BA}$ as

$$|\text{in}\rangle_{BA} = \frac{1}{\sqrt{6}}(e^{\imath\varphi_0}|0011\rangle_B + e^{\imath\varphi_1}|1100\rangle_B + e^{\imath\varphi_2}|0101\rangle_B + e^{\imath\varphi_3}|1010\rangle_B + \dots)|00\rangle_A, \quad (43)$$

$$|\text{out}\rangle_{BA} = \frac{1}{\sqrt{6}}\Big[(e^{\imath\varphi_0}|0011\rangle_B + e^{\imath\varphi_1}|1100\rangle_B)|01\rangle_A + (e^{\imath\varphi_2}|0101\rangle_B + e^{\imath\varphi_3}|1010\rangle_B)|10\rangle_A + \dots\Big]. \quad (44)$$

We must assume that there can be a unitary transformation between the un-traced states $|\text{in}\rangle_{BAW}$ and $|\text{out}\rangle_{BAW}$.

## 8.2 Quantum feedback

The analysis is similar to that of Deutsch–Jozsa algorithm. This time a good half table should not contain a same value of the function twice, what would provide enough information to identify the solution of the problem, namely the *period* $\mathbf{h}(\mathbf{b})$, thus violating condition 32 of the advanced knowledge rule.

With $\mathbf{b} = 0011$, that is $f_{\mathbf{b}}(00) = 0$, $f_{\mathbf{b}}(01) = 0$, $f_{\mathbf{b}}(10) = 1$, $f_{\mathbf{b}}(11) = 1$, one way of sharing the table into two good halves is $f_{\mathbf{b}}(00) = 0$, $f_{\mathbf{b}}(10) = 1$ and $f_{\mathbf{b}}(01) = 0$, $f_{\mathbf{b}}(11) = 1$. The corresponding subsets of $\sigma_B$ are respectively $\{0011, 0110\}_B$ and $\{0011, 1001\}_B$; one can check that each half table is the identical part of the two bit-strings in the corresponding subset of $\sigma_B$. Either good half table or identically either subset is a possible instance of Alice's advanced knowledge. Eq. 31 is satisfied with $\Delta\mathcal{E}_A(\hat{B}_i) = \Delta\mathcal{E}_A(\hat{B}_j) \approx 0.585$ bit (entropy reduction from $\log_2(3)$ bit to 1 bit).

We note that sharing each table into two halves is peculiar to Deutsch–Jozsa and Simon algorithms. In the quantum part of Shor's [38] factorization algorithm (finding the period of a periodic function), taking two shares of the table that do not contain a same value of the function twice implies that each share is less than half table if the domain of the function spans more than two periods.

Given the advanced knowledge of a good half table, the entire table and then $\mathbf{h}(\mathbf{b})$ can always be identified by performing just one function evaluation for any value of the argument $\mathbf{a}$ outside the half table. Thus, the advanced knowledge rule says that, with $R = \frac{1}{2}$, Simon's problem is solved with just one function evaluation. Under the assumption that $R = \frac{1}{2}$ is always attainable, Simon algorithm, which requires $O(n)$ function evaluations, would be suboptimal. The above also shows that the speedup of the quantum part of Simon algorithm is explained by $R = \frac{1}{2}$. In fact, once $\mathbf{h}(\mathbf{b})$ is known (with just one function evaluation in the case of quantum retrocausality $R = \frac{1}{2}$), generating at random the $\mathbf{s}_j(\mathbf{b})$'s requires no further function evaluations.

We give the simplest instance, $n = 2$, of the quantum algorithm that finds $\mathbf{h}(\mathbf{b})$ with just one function evaluation. Register $W$ reduces to the usual register $V$ that contains the result of function evaluation modulo 2 added to its previous content. The input and output states of $V$ are both $\frac{1}{\sqrt{2}}(|0\rangle_V - |1\rangle_V)$. We have $\hat{U} = \hat{\mathcal{P}}_A\hat{H}_A\hat{U}_f\hat{H}_A$, where $\hat{H}_A$ is Hadamard on register $A$, $\hat{U}_f$ is function evaluation, and $\hat{\mathcal{P}}_A$ is the permutation of the basis vectors $|01\rangle_A$ and $|10\rangle_A$. Checking whether there is the similar algorithm for $n > 2$ should be the object of further work.

The sum over histories representation can be developed as in Deutsch–Jozsa algorithm. If, for example, Alice's advanced knowledge is $\mathbf{b} \in \{0011, 0110\}_B$, she can identify the value of $\mathbf{h}(\mathbf{b})$ by performing a single function evaluation for either $\mathbf{a} = 01$ or $\mathbf{a} = 11$, see Table 3, etc.

The fact that Alice knows in advance a good half table, and can thus identify the entire table and hence the solution with just one function evaluation, clearly holds unaltered for $n > 2$. It should also apply to the generalized Simon's problem and to the Abelian hidden subgroup problem. In fact the corresponding algorithms are essentially Simon algorithm. In the hidden subgroup problem, the set of functions $f_{\mathbf{b}} : G \to W$ map a group $G$ to some finite set $W$ with the property that there exists some subgroup $S \leq G$ such that for any $\mathbf{a}, \mathbf{c} \in G$, $f_{\mathbf{b}}(\mathbf{a}) = f_{\mathbf{b}}(\mathbf{c})$ if and only if $\mathbf{a} + S = \mathbf{c} + S$. The problem is to find the hidden subgroup $S$ by computing $f_{\mathbf{b}}(\mathbf{a})$ for the appropriate values of $\mathbf{a}$. Now, a large variety of problems solvable with a quantum speedup can be re-formulated in terms of the hidden subgroup problem [13]. Among these are the seminal Deutsch's problem, finding orders, finding the period of a function (thus the problem solved by the quantum part of Shor's factorization algorithm), discrete logarithms in any group, hidden linear functions, self shift equivalent polynomials, Abelian stabilizer problem, and the graph automorphism problem [39, pp. 146–147].

## 9 Discussion

To highlight the simplicity of the present interpretation of the quantum speedup, we refer to Grover algorithm where the problem setting and the solution of the problem coincide with one another. We also assume that $\hat{U}_B$ is the identity, so that the problem setting is directly the random outcome of Bob's measurement.

To the end of selecting the setting or identically the solution of the problem, the initial Bob's measurement (required to select the problem setting) and the final Alice's measurement (required to read the solution), are redundant with one another. From the mathematical standpoint, the selection of any part of the information that specifies the random outcome of the initial measurement can be ascribed to the final measurement. We should naturally assume that part of the selection performed by the final measurement propagates backward in time, by the inverse of the time-forward unitary transformation, until selecting the part of the outcome of the initial measurement whose selection is ascribed to the final measurement. We have seen that, in the representation of the quantum algorithm relativized to Alice, this backward propagation tells Alice the part of the outcome of the initial measurement in question (which is identically a part of the solution), before she starts her search for the solution.

We have called $R$ the fraction of information whose selection is ascribed to Alice's measurement. Given a value of $R$, we need to reconcile the notion of Alice's advanced knowledge of an $R$-th part of the information with the fact that such a part can be taken in a plurality of ways. Moreover, we need an operational interpretation of the advanced knowledge notion. We kill two birds with one stone by resorting to Feynman's path integral formulation of quantum mechanics. Given an appropriate value of $R$, the quantum algorithm can be seen as a sum over classical histories in each of which Alice knows in advance one of the possible $R$-th part of the information that specifies the problem setting (or the solution) and performs the oracle queries still necessary to find the solution. The sum covers all the possible ways of taking that part. An immediate consequence is that the number of oracle queries required to solve the problem with retrocausality $R$ is that of a classical algorithm that knows in advance the $R$-th part of the problem setting.

Put in this simplified form, the present interpretation of the speedup would seem to be self-evident. What we have until now is an exact relation between $Q$, the number of queries required to solve an oracle problem on the one side and advanced knowledge of part of the problem setting, or the corresponding retrocausality measure $R$, on the other side. $R = 0$ means classical computation, whereas $R > 0$ quantum computation and speedup.

We have applied this retrocausal interpretation of the speedup to the major quantum algorithms known today. $R$ is exactly $\frac{1}{2}$ in all the quantum algorithms requiring a single oracle query, slightly above $\frac{1}{2}$ in Grover algorithm with database size $N > 4$ (more than one query is required), whereas it goes back to $\frac{1}{2}$ for $N \to \infty$. Moreover, $R = \frac{1}{2}$ always corresponds to an existing quantum algorithm and gives at least the order of magnitude of the number of queries required by the optimal one.

Relating $Q$ to $R$ should be interesting from the foundational standpoint and might have practical applications. The following prospects seem to be liable of further study:

First, there should be a relationship between $R$ and the violation of the temporal Bell inequalities. For example, we have $R \approx \frac{1}{2}$ in Grover algorithm whereas this algorithm violates a Bell inequality [23]. It is reasonable to think that any quantum algorithm with $R > 0$ does the same. If this were proved, given an optimal quantum algorithm that yields the solution with $Q$ queries and $R > 0$, this would mean that no classical algorithm could make it with the same number of queries. Often, as in the case of Shor's algorithm, one does not know whether the speedup is an essential quantum feature or is just due to our ignorance of an equally efficient classical algorithm.

Second, we have seen that relating the speedup to a quantitative fundamental quantum feature, entanglement in [24] and discord in [25, 26], is common to other attempts to unification. These attempts have not yet managed to identify a unifying element that explains the different speedups [26]. In the present work, entanglement is replaced by the retrocausality measure $R$, a perhaps more fundamental feature as far as it can explain entanglement [28]. The possible unifying element, for the time being, is the fact that in all the quantum algorithms examined the value of $R$ is in a right surrounding of $\frac{1}{2}$ where the value of $Q$ has little variance. All this might allow the conjecture that $R = \frac{1}{2}$ is always attainable in quantum computation and always gives the order of magnitude of the number of queries required by the optimal quantum algorithm. If this conjecture were true, we would have a very powerful tool. Given a generic oracle problem, we would know the order of magnitude of the number in question (Section 5).

Here, let us make a distinction between entangling and non-entangling unitary evolutions. The unitary evolution under consideration is that between initial and final measurement. An example of non-entangling evolution would be the identity transformation of the initial (sharp) measurement outcome. In this case $R$ can assume any value between zero and one, including the extremes where everything is selected by only the initial measurement (in which case $R = 0$) or only the final measurement (in which case $R = 1$). Of course, quantum algorithms instead create maximal entanglement between the setting and the solution of the problem. In this case, there seems to come out a distinction between past and future. In fact, while $R = 0$ is always possible, as it corresponds to classical computation, $R = 1$ seems to be obviously impossible. It would mean solving an oracle problem without oracle queries. Then, in the case of a unitary evolution that produces maximal entanglement, there must be an upper bound to the value of $R$ that can be attained. As things are now, one could think that this bound is reached in quantum search, because of the fundamental character of this algorithm. Then it would be $R = \frac{1}{2}$, or $R$ slightly above $\frac{1}{2}$. Conversely, a value of $R$ that does not overcome this bound should always be attainable. This idea is also supported by the fact that, in all the quantum algorithms examined, $R$ always attains at least the value $\frac{1}{2}$. These considerations seem to authorize the conjecture that $R = \frac{1}{2}$ is always attainable for fundamental reasons, independently of the specificity of the quantum algorithm. By the way, the idea that, at the microscopic level, the arrow of time is that of growing entanglement dates back to 1988 [40] and underwent further developments starting from 2006 [41].

To better assess the trust that can be placed in this conjecture, there can be the following directions of research: checking the value of $R$ on other classes of quantum algorithms (such as quantum random walks); verifying whether there exists a quantum algorithm that solves Simon's problem with certainty with a single oracle query for $n > 2$, as foreseen by the present interpretation of the speedup for $R = \frac{1}{2}$ (see Section 8); and, learning more about the relation between $R$ and quantum entanglement.

It could also be interesting to look for possible cross fertilizations between the present work and other works on time-symmetric quantum mechanics or the speedup. The present notion of quantum retrocausality might apply to: some possibly related findings of time-symmetric quantum mechanics, as in [20, 21]; the violation of temporal Bell inequalities on the part of quantum computation [23]; and, a more direct physical interpretation of quantum computation complexity classes [29].

## Acknowledgments

I would like to thank David R. Finkelstein for useful discussions.

## References

[1] Deutsch D. Quantum theory, the Church-Turing principle and the universal quantum computer. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 1985; 400 (1818): 97–117. doi:10.1098/rspa.1985.0070

[2] Finkelstein D. Space-time structure in high energy interactions. In: Fundamental Interactions at High Energy. Gudehus T, Kaiser G, Perlmutter A (editors), New York: Gordon and Breach, 1969, pp. 324–338.

[3] Feynman RP. Simulating physics with computers. International Journal of Theoretical Physics 1982; 21 (6–7): 467–488. doi:10.1007/bf02650179

[4] Bennett CH. The thermodynamics of computation—a review. International Journal of Theoretical Physics 1982; 21 (12): 905–940. doi:10.1007/bf02084158

[5] Landauer R. Irreversibility and heat generation in the computing process. IBM Journal of Research and Development 1961; 5 (3): 183–191. doi:10.1147/rd.53.0183

[6] Fredkin E, Toffoli T. Conservative logic. International Journal of Theoretical Physics 1982; 21 (3–4): 219–253. doi:10.1007/bf01857727

[7] Feynman RP. Quantum mechanical computers. Foundations of Physics 1986; 16 (6): 507–531. doi:10.1007/bf01886518

[8] Castagnoli G. Turin Elsag Bailey-ISI international workshops on quantum communication and computation, 1993, 1995 and 1997 group pictures. Retrieved: January 7, 2016. http://www.giuseppecastagnoli.com/images

[9] Rovelli C. Relational quantum mechanics. International Journal of Theoretical Physics 1996; 35 (8): 1637–1678. doi:10.1007/bf02302261

[10] Hawking SW. On the Shoulders of Giants. The Great Works of Physics and Astronomy. Philadelphia: Running Press, 2003.

[11] Feynman RP, Hibbs AR. Quantum Mechanics and Path Integrals. New York: McGraw-Hill Companies, 1965.

[12] Wheeler JA, Feynman RP. Interaction with the absorber as the mechanism of radiation. Reviews of Modern Physics 1945; 17 (2–3): 157–181. doi:10.1103/RevModPhys.17.157

[13] Mosca M, Ekert A. The hidden subgroup problem and eigenvalue estimation on a quantum computer. Lecture Notes in Computer Science 1999; 1509: 174–188. arXiv:quant-ph/9903071, doi:10.1007/3-540-49208-9_15

[14] Castagnoli G, Finkelstein DR. Theory of the quantum speed-up. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 2001; 457 (2012): 1799–1806. doi:10.1098/rspa.2001.0797

[15] Castagnoli G. Quantum correlation between the selection of the problem and that of the solution sheds light on the mechanism of the quantum speed-up. Physical Review A 2010; 82 (5): 052334. arXiv:1005.1493, doi:10.1103/PhysRevA.82.052334

[16] Castagnoli G. Probing the mechanism of the quantum speed-up by time-symmetric quantum mechanics. Presented at the 92nd Annual Meeting of the AAAS Pacific Division, Quantum Retrocausation: Theory and Experiment, 2011. arXiv:1107.0934v9

[17] Aharonov Y, Bergmann PG, Lebowitz JL. Time symmetry in the quantum process of measurement. Physical Review B 1964; 134 (6): 1410–1416. doi:10.1103/PhysRev.134.B1410

[18] Aharonov Y, Albert DZ, Vaidman L. How the result of a measurement of a component of the spin of a spin-$\frac{1}{2}$ particle can turn out to be 100. Physical Review Letters 1988; 60 (14): 1351–1354. doi:10.1103/PhysRevLett.60.1351

[19] Aharonov Y, Popescu S, Tollaksen J. A time-symmetric formulation of quantum mechanics. Physics Today 2010; 63 (11): 27–32. doi:10.1063/1.3518209

[20] Dolev S, Elitzur AC. Non-sequential behavior of the wave function. 2001; arXiv:quant-ph/0102109

[21] Aharonov Y, Cohen E, Elitzur AC. Can a future choice affect a past measurement's outcome? Annals of Physics 2015; 355: 258–268. arXiv:1206.6224, doi:10.1016/j.aop.2015.02.020

[22] Morikoshi F. Problem-solution symmetry in Grover's quantum search algorithm. International Journal of Theoretical Physics 2011; 50 (6): 1858–1867. doi:10.1007/s10773-011-0701-6

[23] Morikoshi F. Information-theoretic temporal Bell inequality and quantum computation. Physical Review A 2006; 73 (5): 052308. arXiv:quant-ph/0602011, doi:10.1103/PhysRevA.73.052308

[24] Jozsa R, Linden N. On the role of entanglement in quantum-computational speed-up. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 2003; 459 (2036): 2011–2032. arXiv:quant-ph/0201143, doi:10.1098/rspa.2002.1097

[25] Ollivier H, Zurek WH. Quantum discord: a measure of the quantumness of correlations. Physical Review Letters 2001; 88 (1): 017901. doi:10.1103/PhysRevLett.88.017901

[26] Henderson L, Vedral V. Classical, quantum and total correlations. Journal of Physics A: Mathematical and General 2001; 34 (35): 6899–6905. arXiv:quant-ph/0105028, doi:10.1088/0305-4470/34/35/315

[27] Gross D, Flammia ST, Eisert J. Most quantum states are too entangled to be useful as computational resources. Physical Review Letters 2009; 102 (19): 190501. doi:10.1103/PhysRevLett.102.190501

[28] Price H, Wharton KB. Disentangling the quantum world. Entropy 2015; 17 (11): 7752–7767. doi:10.3390/e17117752

[29] Aaronson S. Quantum computing, postselection, and probabilistic polynomial-time. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 2005; 461 (2063): 3473–3482. arXiv:quant-ph/0412187, doi:10.1098/rspa.2005.1546

[30] Cai Y, Le HN, Scarani V. State complexity and quantum computation. Annalen der Physik 2015; 527 (9–10): 684–700. arXiv:1503.04017, doi:10.1002/andp.201400199

[31] Howard M, Wallman J, Veitch V, Emerson J. Contextuality supplies the 'magic' for quantum computation. Nature 2014; 510 (7505): 351–355. arXiv:1401.4174, doi:10.1038/nature13460

[32] Bohm D, Pines D. A collective description of electron interactions: III. Coulomb interactions in a degenerate electron gas. Physical Review 1953; 92 (3): 609–625. doi:10.1103/PhysRev.92.609

[33] Grover LK. A fast quantum mechanical algorithm for database search. In: Proceedings of the 28th Annual ACM symposium on Theory of Computing. New York: Association for Computing Machinery Press, 1996, pp. 212–219. arXiv:quant-ph/9605043, doi:10.1145/237814.237866

[34] Long GL. Grover algorithm with zero theoretical failure rate. Physical Review A 2001; 64 (2): 022307. arXiv:quant-ph/0106071, doi:10.1103/PhysRevA.64.022307

[35] Toyama FM, van Dijk W, Nogami Y. Quantum search with certainty based on modified Grover algorithms: optimum choice of parameters. Quantum Information Processing 2013; 12 (5): 1897–1914. doi:10.1007/s11128-012-0498-0

[36] Deutsch D, Jozsa R. Rapid solution of problems by quantum computation. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 1992; 439 (1907): 553–558. doi:10.1098/rspa.1992.0167

[37] Simon DR. On the power of quantum computation. Proceedings of the 35th Annual IEEE Symposium on Foundations of Computer Science, November 20-22, 1994, pp. 116–123. doi:10.1109/sfcs.1994.365701

[38] Shor PW. Algorithms for quantum computation: discrete logarithms and factoring. Proceedings of the 35th Annual IEEE Symposium on Foundations of Computer Science, November 20–22, 1994, pp. 124–134. doi:10.1109/sfcs.1994.365700

[39] Kaye P, Laflamme R, Mosca M. An Introduction to Quantum Computing. Oxford: Oxford University Press, 2007.

[40] Lloyd S. Pure state quantum statistical mechanics and black holes. Chapter 3 in: Black holes, demons, and the loss of coherence: how complex systems get information and what they do with it. PhD Thesis, Rockefeller University, New York, 1988. arXiv:1307.0378

[41] Popescu S, Short AJ, Winter A. Entanglement and the foundations of statistical mechanics. Nature Physics 2006; 2 (11): 754–758. doi:10.1038/nphys444