Shakespeare's Free Lunch: A Critique of the D-CTC Solution to the Knowledge Paradox

Lucas Dunlap

Abstract

In this paper I argue that the consistency condition from the Deutsch's influential model for closed timelike curves (CTCs) differs significantly from the classical consistency condition found in Lewis [15] and Novikov [16], as well as from the consistency condition found in the P-CTC model, the major rival to Deutsch's approach. Both the CCC and the P-CTC consistency condition are formulable in the context of a single history of the world. Deutsch's consistency condition relies on the existence of a structure of parallel worlds. I argue that Deutsch's commitment to realism about parallel worlds puts his solutions to the information paradox in jeopardy. I argue that, because of Deutsch's commitment to this metaphysical picture, he is committed to the existence of physical situations that are in every way indistinguishable from the paradoxes he attempts to rule out by adopting the model in the first place. Deutsch's proposed solution to the Knowledge Paradox, in particular his commitment to the actuality of the many worlds of the Everett interpretation (on which he relies to solve the paradoxes), guarantees the existence of worlds that are indistinguishable from worlds in which the genuine Knowledge Paradox arises.

1 Introduction

In this paper I will argue that because of the metaphysical commitments that underwrite Deutsch's influential model for closed timelike curves, he is committed to the existence of worlds that undermine his proposed solution to the paradox. I will suggest that there is a distinction available to Deutsch on which he can rely to alleviate some of this tension, but it too comes at a cost. Appeal to this distinction would also have served as a solution to the existence of Knowledge Paradox worlds prior to Deutsch having to develop a solution based on the D-CTC model. This calls into question the claim he makes that the features unique to the D-CTC model offer the best solutions to the paradoxes of time travel. Deutsch's time travel framework makes use of the "many worlds" of the Everett interpretation of quantum mechanics. It has been generally accepted in the literature on this topic that his model provides a a physical basis for the "multiple timelines" solution to the paradoxes of time travel well known from science fiction (see, e.g., [1]). The argument of this paper will apply to any model that grounds the existence of the "multiple timelines" or "parallel worlds" in the many worlds of the Everett interpretation of quantum mechanics.

I will discuss some general features of the concept of the *consistency condition* solution to the paradoxes of time travel. The two paradoxes on which I will be focusing my attention are the *Grandfather Paradox*, in which the effect of a time traveling system's presence in the past is to prevent that system from time traveling, and the *Knowledge Paradox*, in which a time traveling system leaves some information trace in the past which is ultimately causally responsible for its time travel.[1]

Most analyses of time travel attempt to rule out the possibilities of both of these effects. The inability to rule them out is sometimes seen as a major flaw of a proposed analysis. I will argue, however, that although the *Grandfather Paradox* is a genuine paradox—that is, it contains a contradiction—the *Knowledge Paradox* is not, and its possibility should not be seen as a fatal flaw for an account of the possibility of time travel.

In section 2, I introduce the features of Deutsch's CTC model. In the following section, I introduce the well-known *Classical Consistency Condition* (CCC) solution to the Grandfather Paradox. The principle's first appearance in the philosophical literature was in a paper by David Lewis from 1976 [15], based on a series of lectures he gave in 1971. However, Lewis is clear in his paper that he drew his inspiration from science fiction, and he believed that the consistency condition was well-known to savvy science fiction authors and fans alike.

I will go on to consider a more recent debate about the nature of consistency in the presence of CTCs, which takes place in the context of the debate between proponents of D-CTCs, and the proponents of the alternative "projective" or "post-selected" (P-CTC) model, which will be introduced below. I will argue that an analysis of Lewis's CCC will illuminate a sense in which the consistency condition in the D-CTC model is playing a very different role from the consistency condition in the P-CTC model, which is much closer to Lewis's classical version.

In light of this difference, I will argue that the consistency condition in the D-CTC model only solves the two paradoxes in conjunction with Deutsch's metaphysical commitment to the reality of the MWM and MSM worlds.[2] This commitment carries with it new problems for an analysis of time travel. I will examine one of these problematic features of time travel in the context of the Everett interpretation in detail, and I will argue that it undermines Deutsch's solution to the Knowledge Paradox from the D-CTC model.

Even though his model may have the resources to solve these problems to his

---

[1] This paradox is also variously called the *Information Paradox*, the *Unproved Theorem* paradox, the *Works of Shakespeare* paradox, and the *Uncaused Effect* paradox.

[2] See [8] for more on this distinction. In brief, Dunlap argues that the D-CTC model requires the adoption of a more general multiverse concept than found in the Everett interpretation. Dunlap calls this the "Mixed-State Multiverse" (MSM). It includes the Everett interpretation's "Many Worlds Multiverse" (MWM). The argument of this paper relies only on the quantum mechanical features of the many worlds of the Everett interpretation of quantum mechanics. I will refer to "parallel worlds" for convenience, but the argument goes through whether or not Dunlap's argument for Deutsch's model requiring the MSM is accepted.

satisfaction, it comes at a rather high cost. I'll also argue that the justification Deutsch uses for ruling out Information Paradox scenarios (he makes reference to "philosophical principles") isn't as strong as he claims. I'll give an example of another equally convincing philosophical position according to which these kinds of scenarios are possible. I'll conclude by arguing that ruling out the Information Paradox isn't a necessary part of a successful model of CTCs, and the fact that the P-CTC model doesn't rule them out shouldn't count as a strike against it.

2 Deutsch's CTC Model

A closed timelike curve is a trajectory through spacetime along which a system could travel, that would lead it into its own past, allowing it to interact with a younger version of itself. The general theory of relativity (GR) does not rule out the possibility of CTCs. Their existence is consistent with the mathematical constraints on the geometry of spacetime imposed by the theory. This fact was first pointed out to Einstein (to his great surprise) by Kurt Gödel in 1949 [18][19]. Since that time, several mathematically consistent models for CTCs have been developed (see[9] [10] [2]).

The debate about the physical possibility of, and physical constraints on, these spacetime structures has taken place largely in the context of GR, rarely taking quantum mechanics into account (see [11] for a notable exception). Deutsch's 1991 "Quantum Mechanics Near Closed Timelike Lines" represents the genesis of a different approach to analyzing CTCs [4]. Rather than engaging in the debate about the extent to which GR allows CTCs, Deutsch asked the following question: Assuming we had access to reliable CTCs, what can we do with them?

David Deutsch had already established himself a chief founding figure in the information-theoretic approach to quantum mechanics and in quantum computation.[3] The approach he takes to analyzing CTCs mirrors the QIT approach to the analysis of quantum mechanics. The focus is on what can be practically achieved in a world that gives us access to these resources.

Deutsch has expressed the belief that the proper way to understand physical processes is in terms of information flow [5]. Our fundamental analyses of the evolutions and interactions of physical systems should be in terms of the manipulation and exchange of information. It was only natural, then, that Deutsch's attempt to answer the question of the power of CTCs was formulated in terms of computation and information transmission.

Most importantly, however, the move away from the GR regime was motivated by what Deutsch saw as a major failing of all previous analyses of CTCs: They ignored the fact that quantum mechanics allows systems to exist in states that are classically impossible.[4]

---

[3] For example, he developed the first quantum computational algorithm which demonstrated a significant speed–up over any classical counterpart. [3][12]

[4] Even Hawking's [11], in which his argument is formulated in the semi–classical gravity

In particular, Deutsch was interested in the power of quantum mechanics to solve the paradoxes of time travel. All previous attempts to solve the two major time travel paradoxes—the Grandfather Paradox and the Knowledge Paradox—had proceeded from the assumption that the state of the system traveling though time must be a definite classical state. Part of the power of quantum mechanics is that it allows for the existence of superposed and quantum mixed states. In the former case, the system is not in a definite state with respect to the basis of interest, but rather in some linear combination of its eigenstates. The latter case is even more general, where the system may not even be in a definite linear combination of eigenstates of the measurement basis. The important feature of these states is that the are in a sense "in between" the states allowed in classical physics.

This is relevant because the classical solutions to the paradoxes of time travel have the feature of ruling out certain initial experimental setups, since the propagation of a system in that definite state along a CTC would yield a contradiction. This feature of the classical solutions is often referred to as *superdeterminism*, since it puts constraints on the initial conditions of an experiment, that go over and above the constraints imposed by the deterministic theory itself.[5] Deutsch has expressed serious discomfort with this feature of the classical solutions (among others). His D-CTC model is an attempt to show that quantum mechanics can solve the paradoxes of time travel without ruling out any initial experimental condition.

The Grandfather paradox is a physical situation in which a time–traveling system's presence in the past prevents itself from time traveling in the first place. The Knowledge Paradox is a situation where a time–traveling physical system's presence in the past is causally responsible for its having time traveled to the past in the first place. A simple example of a Grandfather Paradox is going back in time to kill yourself as a baby. If you don't survive childhood, who comes back in time to assassinate you? A simple example of a Knowledge Paradox is using the plans for a time machine given to you by your time–traveling future self to go back in time and give your past self the plans. Who designed the time machine? It seems to exist in a causal loop.

The classical consistency condition (CCC) proposed independently by David Lewis and Igor Novikov, states that the history of the world must be self–consistent. This entails that trajectories that would take physical systems to the past to enact a Grandfather Paradox are impossible, because they would lead to a physical contradiction (the baby both survives toddlerdom, and doesn't). However, CCC doesn't rule out the closed causal loop of the Knowledge Paradox.

---

regime only takes quantum effects into account in predicting the existence of certain kinds of fundamental particles on the interior of a wormhole. Quantum considerations are not applied to the possible states in which systems find themselves while traversing the CTC.

[5] "Superdeterminism" is a slight misnomer in this context, given that, in a chronology–violating region, all events are in the past of all others, and therefore *determinism* itself is difficult to define. The idea being expressed by the term, however, is clear: in a chronology–violating region, all events need to be self–consistent, meaning that certain sequences of events are ruled out, even though they would not be inconsistent with the dynamics of a chronology–respecting region.

After all, being uncaused is not inconsistent. This permissiveness with respect to uncaused effects is the other major feature of CCC to which Deutsch objects.

Deutsch's analysis of CTCs is formulated in terms of quantum computational circuits. In order to be able to present the D-CTC model completely, it is necessary to introduce the basic concepts of QIT and quantum computation.

In his well known 1991 paper [4], Deutsch introduced a model for the analysis of the physical behavior of CTCs. Prior to his work, the standard way of analyzing the physical effects of chronology-violating regions of spacetime was in terms of their underlying geometry. Deutsch considered this approach to be insufficient because it fails to take quantum mechanical effects into account. He proposed an alternative approach which involves analyzing the behavior of CTCs in terms of their information processing capabilities.

He begins his account by defining a notion of equivalence between spacetime-bounded networks containing chronology-violating regions. A network in this context is to be understood as a spacetime geometry which takes as input the initial state of a physical system and outputs the system's final state. Two networks are *denotationally equivalent* if their outputs are the same function of their inputs. That is to say, regardless of whether two networks have differing spacetime geometries, if the function that maps their initial states to their final states is the same, they are denotationally equivalent.

Next he introduces the idea that the transformation between any two denotationally equivalent networks is trivial. Insofar as we are interested in analyzing CTCs in terms of their physical effects (that is, their output given a certain input), we are free to use the simplest model available in the denotational equivalence class of a particular network for the purpose of our analysis of the information flow through a CTC.

The final step of his proposal is to introduce a simple standard form into which any spacetime-bounded network can be trivially transformed for the purpose of analysis. The simple standard form involves translating all spacetime-bounded networks into circuits in which each particle traveling in the original network is replaced by sufficiently many carrier particles, each of which have a single 2-state internal degree of freedom (a bit). The regions in which the particles interact are localized (by denotationally trivial transformations) into gates, such that the states of the particles so not evolve while traveling between them. And finally, all chronology-violating effects of the network are localized to sufficiently many carrier particles on closed loops, which only interact with chronology-respecting particles in gates.

Deutsch points out that chronology violation itself makes no difference to the behavior of a network unless there is a closed loop of information. In the original network, this closed information path could potentially not be confined to the trajectory of any single particle (since the carriers can interact with each other), but for any such network, there is a denotationally trivial transformation which will localize the closed loop of information on sufficiently many carriers on closed paths.

The real innovation of this approach is that it can very easily accommodate quantum mechanical effects by relaxing the requirement that the carrier particles be in a well-defined classical state after interactions. If viewed classically, networks containing chronology violations can lead to paradoxes that seem to put unnaturally strong constraints on possible initial conditions of physical systems (e.g. you are somehow prohibited from getting in the time machine that would take you back to kill your grandfather). Deutsch uses his model to argue that, when quantum mechanics is taken into account, these unnatural constraints on initial states disappear. Deutsch's fixed point theorem states that CTCs "place no retrospective constraints on the state of a quantum system" [4]. That is to say, for any possible input state, there will be a paradox-free solution.

This is the result of a consistency condition implied by the quantum mechanical treatment of time-traveling carrier particles interacting with later versions of themselves. If we let $|\psi\rangle$ be the initial state of the "younger" version of the carrier particle, and let $\hat{\rho}$ be the density operator of the "older" version of the carrier particle, then the joint density operator of the two particles entering the region of interaction is

$$|\psi\rangle\langle\psi| \otimes \hat{\rho} \tag{1}$$

and the density operator of the two carrier particles after the interaction is

$$U(|\psi\rangle\langle\psi| \otimes \hat{\rho})U^{\dagger} \tag{2}$$

where U is the interaction unitary. The consistency condition requires that the density operator of the younger version of the carrier particle as it leaves the region of interaction is the same as that of the older version as it enters the region of interaction.

This makes intuitive sense, because it is the interaction that causes the earlier version of the carrier particle to become the later version. When translated via a denotationally trivial transformation to a network in which the chronology-violating behavior is localized to a single particle on a CTC that interacts with a chronology-respecting (CR) carrier particle, the consistency condition for the CTC system is

$$\rho_{\mathrm{CTC}} = \mathrm{Tr}_{\mathrm{CR}}[U(|\psi\rangle\langle\psi| \otimes \rho_{\mathrm{CTC}})U^{\dagger}]. \tag{3}$$

This requirement says that, after tracing out the CR qubit, the density operator of the system on the CTC *after* the interaction is the same as it was *before* the interaction. That is to say, after the interaction, the carrier particle on the CTC enters the "future mouth" of the CTC, and exits the "past mouth" of the CTC *before* the interaction. The state of the particle that comes out of the past mouth must be the same as the system that enters the future mouth.[6] Furthermore, ρ_{CTC} depends on $|\psi\rangle$, so the input state on the causality-respecting carrier particle has an effect on the state of the particle it will interact with.

The output of the circuit (i.e. the final state of the CR qubits) depends on the input of system $|\psi\rangle$ and ρ_{CTC}. And, as we see in the previous equation,

---

[6] Because the procedure involves taking the partial trace of the system, and requiring consistency for only the state of the system bound to the CTC, any entanglement with systems in CR region is broken when the CTC-bound quibit exits the past mouth of the CTC.

ρ_{CTC} itself depends on $|\psi\rangle$. Therefore, the evolution of the CR qubit is nonlinear with respect to the input $|\psi\rangle$.

$$\rho_{\mathrm{output}} = \mathrm{Tr}_{\mathrm{CTC}}[U(|\psi\rangle\langle\psi| \otimes \rho_{\mathrm{CTC}})U^{\dagger}]. \tag{4}$$

Whatever the physical situation is, its information flow can be redescribed in a form that has the following features: There are a finite number of qubits bound to a CTC. These interact via unitaries with a finite number of qubits that follow an ordinary chronology–respecting trajectory. The CR qubits are measured after their interaction with the CTC qubits, and their state is the final state of the system. In the region of interaction, the CTC qubits behave according to ordinary quantum mechanics, and interact with the CR qubits via unitary interactions. The CTC qubits do not evolve in any way while traveling back along the CTC. The nonlinearity of the systems overall evolution is entirely due to the consistency conditions nonlinearity.

This means that the closed information loops of chronology violation can be isolated into localized regions of spacetime. The effects that can be generated by interaction with the CTC can range over all of space, of course. But they must be the result of entanglement, or prior causal interaction, with systems in the region of interaction with the chronology violating qubits.

In light of the model's reliance on this nonlinear consistency condition, Deutsch's claim that CTCs, when properly understood, place no constraints on the possible states of the quantum system may be stronger than is warranted. While it is true that, unlike the classical analysis of time travel paradoxes, his model places no constraints on the input state of the causality-respecting system, it *does* constrain the possible states of the system confined to the CTC.

While Deutsch's solution seems more intuitively plausible than the constraint on initial conditions that prevents the occurrence of classical time travel paradoxes, it is nonetheless puzzling. In the classical case, it is somehow forbidden that I get in the time machine that will take me back to kill my grandfather. There isn't necessarily any obvious causal mechanism that prevents me. It is simply impossible, to avoid paradox, that I ever actually carry out my mission. This constraint is often described as *superdeterministic*, since it is something above and beyond simple determinism that rules out the possibility of me getting into the time machine. David Lewis's influential formulation of the classical consistency condition from his [15] alleviates some of this tension by redescribing the time travel narrative as a single, self-consistent history. The drawback of this approach is that it seriously undermines the notion that the time traveler has free will.

Deutsch characterizes his problem with the classical solutions as stemming from the fact that they violate what he calls the *principle of autonomy.*

> According to this principle, it is possible to create in our immediate environment any configuration of matter the laws of physics permit locally, without reference to what the rest of the universe may be doing. [7]

He claims that classical solutions to the Grandfather Paradox, which impose global consistency, violate this principle.

> Under this principle, the world outside the laboratory can physically constrain our actions inside, even if everything we do is consistent, locally, with the laws of physics. Ordinarily we are unaware of this constraint, because the autonomy and consistency principles never come into conflict. But classically, in the presence of CTCs, they do. [7]

Although Deutsch's reference to the principle of autonomy makes it seem as though his problem with the classical solutions is that they always rule out certain initial conditions of an experiment involving a CTC, this can't be quite right. For most initial setup conditions, including those trajectories that seem to entail a Grandfather Paradox, there is a fixed–point solution which shows that there is a self-consistent sequence of events that will avoid the paradoxical outcome. Often these solutions involve a self-interaction that changes the state of the younger system heading toward the future mouth of the CTC, such that when it exists the past mouth, it is no longer on a trajectory that will prevent it from time traveling. Novikov explains this class of solutions in terms of billiard ball entering the future mouth of a CTC on a trajectory that will lead to a collision with itself in the past, preventing it from entering the future mouth.

> If we take into account the collision from the very beginning, then the collision is very weak, just a slight touch between the two balls that nudges the younger ball only slightly. The younger ball then moves along a trajectory slightly different from our expectation, but still enters mouth B. It reappears from mouth A in the past and continues along its motion, still on a trajectory that differs only slightly from the trajectory it would have traveled on had it not suffered a collision. The result of the slight difference in trajectory is that the collision with the younger version of itself is not a strong collision, but rather a weak collision, a glancing blow. Therefore we have a consistent solution. [16]

There are some situations with which Deutsch is concerned that he claims do not have a classical fixed point. For example, the classical bit in the system depicted in Figure 1. After exiting the past mouth of the CTC, the system encounters a NOT gate, which flips its state. It then enters the future mouth, and repeats the process. Classically, this system would oscillate between the two values allowed.[7]

In the quantum version of this circuit, however, there is a unique fixed point.

---

[7] Deutsch admits that discreteness in the classical domain is an approximation, so perhaps an argument could be made that the fixed point of such a system would involve a failure of the gate to operate properly. But assuming everything works as advertised, there is no classical fixed point.

Figure 1: A classical bit bound to a CTC with a *NOT* gate. In this example, there is no classical fixed point solution.

If the qubit is in the mixed state

$$\frac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right) \tag{5}$$

the consistency is satisfied.[8]

The classical fixed point solutions show that in most cases, you are not constrained in terms of the initial setup conditions of your experiment. For most input states, there is at least one self-consistent sequence of events that can follow. I believe Deutsch is concerned with the fact that we are not free, even in light of the classical fixed–point solutions, to send a system in any state we choose *into the CTC*. The classical fixed–point solutions involve a self-interaction before the input system enters the future mouth of the CTC. This, I believe, is at the heart of Deutsch's discomfort with the classical solutions to the paradoxes of time travel.

3 The Consistency Condition in the Philosophy Literature

David Lewis, in his 1976 paper "The Paradoxes of Time Travel", argued that time travel was in fact logically possible, and that time travel seems to entail the existence of Grandfather Paradox scenarios was not enough to show that time travel narratives were inconsistent.

He introduced the CCC, which states that scenarios in which a time traveler kills his own grandfather cannot occur because they contain a contradiction. There are not two different versions of the past—one in which the time traveler wasn't present, and his grandfather lived, and one in which the time traveler was present and his grandfather was killed. Rather, there is just a single "run" of time, and it must be self-consistent.

---

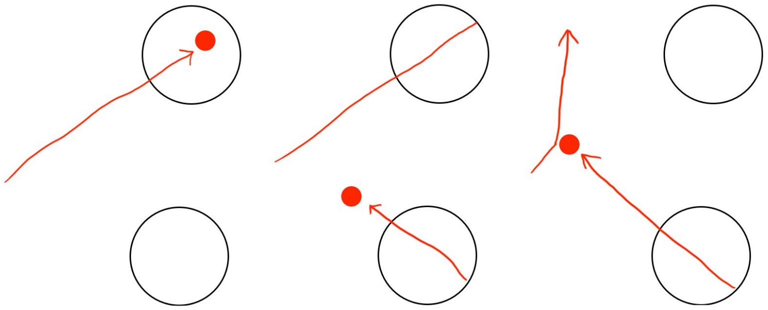
[8] As is true in the classical case, most fixed–point solutions in the D-CTC model are not unique. However, the information flow in this circuit is at the heart of the Grandfather Paradox, and is therefore especially important for Deutsch's goal to show that his model can solve the paradoxes of time travel.

> Either the events of 1921 timelessly do include Tim's killing of Grandfather, or else they timelessly don't. We may be tempted to speak of the 'original' 1921 which lies in Tim's personal past, many years before his birth, in which Grandfather lived; and of the 'new' 1921 in which Tim now finds himself waiting in ambush to kill Grandfather. But if we do speak so, we merely confer two names on one thing. The events of 1921 are doubly located in Tim's (extended) personal time, like the trestle on the railway, but the 'original' 1921 and the 'new' 1921 are one and the same. If Tim did not kill Grandfather in the 'original' 1921, then if he does kill Grandfather in the 'new' 1921, he must both kill and not kill Grandfather in 1921—in the one and only 1921, which is both the 'new' and 'original' 1921. [15]

Lewis's version of the consistency condition is the requirement that the single history in which all events take place be self-consistent. This self-consistency rules out all histories in which a time traveler somehow prevents himself from time traveling, since those histories would contain a contradiction.

This is identical to what physicists call the "Novikov Principle", though Lewis's statement of the principle predated Novikov's in print. Novikov's version of the CCC is formulated purely in terms of the behavior of simple physical systems. He begins by noting that, for every CTC, there is a trajectory along which one could send a billiard ball, such that the older version of the ball exiting the past mouth of the wormhole would interfere with the trajectory of the younger version heading towards the future mouth. Setting the ball along this initial trajectory enacts a Grandfather Paradox, since the ball is preventing itself from time traveling.

Figure 2: A billiard ball on a Grandfather Paradox trajectory. First the younger version enters the future mouth of the wormhole. Then the older version exits the past mouth of the wormhole. Then the older version interferes with the original trajectory of the younger version, preventing it from entering the future mouth.



Most of the trajectories that involve self–interaction are unproblematic—a slight nudge in one direction or the other from the older version ensures that the younger version is on a slightly altered trajectory, so that when it exits the past mouth of the wormhole as the older version, it only just nudges its younger counterpart. These are the consistent solutions to the problem of self–interaction.

However, certain initial trajectories do entail a Grandfather Paradox interaction. These initial conditions are ruled out by Novikov's consistency principle. Like Lewis, Novikov requires that the history of events be self-consistent. Therefore, certain initial states cannot obtain. In this sense, the two version of CCC both impose superdeterminism.

This is the constraint that Deutsch believes violates the autonomy principle. CCC violates this principle because it rules out the possibility of certain experimental setups, namely those which would send the billiard ball on a Grandfather Paradox trajectory. The characteristic feature of both of these formulations of CCC is that they rely only on a single history of the world to solve the Grandfather Paradox. They require simply that the single history of the universe contain no contradictions—that it be self-consistent. It is this feature that is shared by the P-CTC consistency condition, and which differs in the D-CTC model.

4 P-CTC Consistency v. D-CTC Consistency

The versions of the consistency condition which appear in the P-CTC model and the D-CTC model are formulated rather differently than Lewis's CCC. And although the two quantum consistency conditions appear to be very similar, I will argue that the P-CTC consistency condition is in fact closer in an important way to CCC, and that the D-CTC consistency condition is very different.

The main alternative approach to giving a quantum mechanical analysis of CTCs is the P-CTC approach. The approach begins by conceiving of the CTC as a quantum communication channel to the past. For this reason, the quantum teleportation protocol is taken as the starting place for the P-CTC model. Seth Lloyd, one of the prominent proponents of this approach says

> [...] If quantum teleportation is combined with post-selection, then the result is a quantum channel to the past. The entanglement occurs between the forward- and backward- going parts of the curve, and post-selection replaces the quantum measurement and obviates the need for classical communication, allowing time travel to take place. [13]

Consistency in the P-CTC model is achieved via post-selection. In the conventional teleportation protocol, in order for Alice to teleport the (potentially unknown) state $|\psi\rangle$ to Bob, they must first share a pair of maximally entangled particles. Alice makes a joint measurement on $|\psi\rangle$ and her half of the entangled pair. Based on the outcome she gets, she will communicate to Bob (via a classical channel) which one of four operations he could perform on his half of the entangled pair to produce $|\psi\rangle$ for himself. One of the four operations Alice might instruct Bob to undertake is to *do nothing* to his particle. That is to say, in 25% of the cases, Bob will already have a particle in the state $|\psi\rangle$ before he receives Alice's communication.

If we select out just these cases, then we can conceive of this as a situation where the state $|\psi\rangle$ is instantaneously teleported to Bob. In an experimental setting, by limiting our attention to the subset of results that have this feature, a P-CTC can be effectively simulated. It's important to note that, since the teleportation protocol preserves entanglement, the P-CTC model has the same feature. The state $|\psi\rangle$ may be entangled with other systems, and a P-CTC would displace this entanglement into the past.

One of the major differences between the D-CTC model and the P-CTC model is the form the consistency condition takes. The P-CTC consistency condition, as described by Lloyd as follows:

> [...] A generalized measurement made on the state entering the curve should yield the same results, including correlations with other measurements, as would occur if the same measurement were made on the state emerging from the curve. The CTC should behave like an ideal quantum channel [...].[14]

What's required to be consistent in this framework is a measurement on the state of the system at the future mouth of the CTC must be the same as a measurement on the state that emerges from the past mouth. This entails that the single history of the world in which all events take place be self-consistent. In this respect it is similar to the CCC.

D-CTC consistency does not share this feature. Recall Deutsch's mathematical statement of the consistency condition:

$$\rho_{\mathrm{CTC}} = \mathrm{Tr}_{\mathrm{sys}}\left[V\left(|\psi\rangle\langle\psi| \otimes \rho_{\mathrm{CTC}}\right)V^{\dagger}\right]$$

It is only the state of the system bound to the CTC that must be the same at both mouths. This does not guarantee that the correlations with other systems that would have obtained had we measured the CTC-bound system prior to its entering the future mouth of the CTC be preserved when it exists that past mouth.

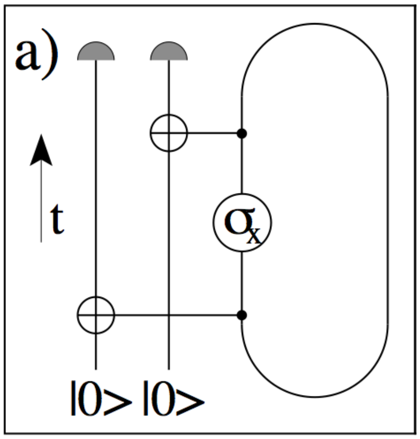
Deutsch's interpretation of the mixed state is crucial for the model to make the predictions he needs (see [8]). The relevant features of that interpretation are that the mixed state that enters the CTC represents a collection of distinct worlds, in half of which the quibit is in state $|0\rangle$, and in half of which the qubit is in state $|1\rangle$, and that the CTCs are "'gateways' between Everett universes" [4]. In the D-CTC model, ρ_{CTC} represents a collection of worlds, which are connected with each other via the CTCs. Deutsch's consistency condition applies to *this* object. The collection of worlds picked out by ρ_{CTC} is required to be self-consistent overall, but each individual world may undergo changes.

The main difference between the P-CTC and D-CTC consistency conditions is that the state of the system exiting the past mouth of a P-CTC must give rise to the same measurement outcomes as the state that entered the future mouth.

Notably, this includes correlations of outcomes between the system on the CTC and CR systems. If the system traveling around the CTC is entangled with one of the CR qubits, that entanglement would be preserved when the system is sent to the past. Lloyd et al. impose this condition because they interpret it as being more true to the CCC. As Lloyd et al. say, "our mechanism embodies in a natural way the Novikov principle that only logically self-consistent sequences of events occur in the universe" [14]. Deutsch's consistency condition does not have this feature. The D-CTC model breaks entanglement between the CTC-bound qubit and any system in the CR region.[9]

Lloyd uses the following Grandfather Paradox circuit to distinguish between the two approaches.

Figure 3: The Grandfather Paradox circuit from [14]



The system bound to the CTC is only stable in the state $\rho = \frac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right)$, because each "time" it goes around the CTC (in pseudotime) its state is flipped from $|1\rangle$ to $|0\rangle$, or from $|0\rangle$ to $|1\rangle$.

Deutsch interprets this as representative of real physical systems in the definite states $|1\rangle$ and $|0\rangle$, that, when taken together, constitute the mixed state on the CTC. As Lloyd says

> The strange aspect of Deutsch's solution comes when one attempts to follow the state of the time-traveller through the CTC. To preserve self-consistency, the 1 component (time traveller alive) that enters the loop emerges as the 0 component (time traveller dead), while the 0 component (time traveller dead) that enters the loop emerges as the 1 component (time traveller alive). Thus, the CTC preserves the overall mixed state, but not the identity of the components [...]. [14]

---

[9] This is true provided that the physical situation being modeled warrants the inclusion of D-CTC in the simplified version of the circuit, i.e., that there is a closed loop of information. See [8] for a discussion of this point.

This aspect of the D-CTC model is explained by the fact that Deutsch conceives of these time travel scenarios as playing out across parallel worlds. On Deutsch's model, what is required to be consistent is the make-up of the collection of worlds which realize the mixed state on the CTC. For each individual world, however, there is no need for a prohibition against killing grandfathers. There will be worlds in which time travelers appear and kill people who look and act very much like their own grandfathers, but this is not inconsistent because in addition to being time travelers, these murderous adventurers are also *trans-world* travelers. They came from a world in which their own grandfather was alive, and traveled to a different world in which they kill a *counterpart* of their own grandfather. Each time traveler can kill the grandfather they find, because he isn't actually their own, and killing someone else's grandfather does not create a paradox.

On the P-CTC model, however, a different outcome is predicted.

> In any real-world situations, the σ_x transformation in not perfect. Then, replacing σ_x with $e^{-i\theta\sigma x} = \cos\frac{\theta}{2}\mathbb{1} - i\sin\frac{\theta}{2}\sigma_x$ (with $\theta \simeq \pi$), the non-linear post-selection amplifies fluctuations of θ away from π. This eliminates the histories plagued by the paradox and retains only the self-consistent histories in which the time traveler fails to kill her grandfather (the unitary in the curve is 1 instead of σ_x), and the two output qubits have equal value: P-CTCs fulfill our self-consistency condition. [14]

This is consistent with the predictions of CCC. The probability that a time traveler will succeed in killing the man they find in the past is zero.

It is in this sense that the D-CTC consistency condition is radically different from the P-CTC consistency condition and CCC. The latter two require a single self-consistent history in which all events take place. The former requires consistency in the make-up of the collection of distinct parallel worlds which realize the mixed state ρ_{CTC}.

This also gives a plausible answer to why entanglement isn't preserved through a D-CTC. The CTC-bound system that is potentially entangled with CR systems in its environment before it enters the future mouth of the wormhole is not numerically identical to the system that exits the past mouth of the wormhole. The qubit that emerges in the past has traveled from a different world. Preserving entanglement relations does not make sense in the context of the MSM model of time travel.

5 The D-CTC Knowledge Paradox Solution

In addition to giving a relatively compelling solution to the Grandfather Paradox, the MWI analysis of time travel also provides a solution to the Knowledge Paradox, about which Deutsch is particularly concerned. He emphasizes the importance of a solution to the Knowledge Paradox repeatedly throughout his writings on CTCs:

> Knowledge paradoxes violate the principle that knowledge can come into existence only as a result of problem–solving processes, such as biological evolution or human thought. Time travel appears to allow knowledge to go from the future to the past and back, in a self-consistent loop, without anyone or anything ever having to grapple with the corresponding problems. What is philosophically objectionable here is not that knowledge-bearing artifacts are carried into the past—it is the "free lunch" element. The knowledge required to invent the artifacts must not be supplied by the artifacts themselves. [7]

> The real problem with closed timelike lines under classical physics is that they could be used to generate knowledge in a way that conflicts with the principles of the philosophy of science, specifically with the evolutionary principle. [4]

> It is a fundamental principle of the philosophy of science that the solutions of problems do not spring fully formed into the Universe, i.e., as initial data, but emerge only through evolutionary or rational processes. [4]

> This "near inconsistency", forcing a violation of the evolutionary principle, is a far more serious paradox than the "actual" inconsistencies of paradoxes 1–3. Those inconsistencies merely indicate that the initial data have one set of values rather than another, something which is true anyway, and starting from those values the subsequent evolution, though strange, does not contradict the philosophy of science. But because of the "near inconsistency" of [the knowledge paradox] the only permitted initial data cause an evolution that does contradict the philosophy of science. [4]

It is important to note that Deutsch emphasizes the fact that he considers the possibility of the Knowledge Paradox to be in conflict with principles from the philosophy of science. The evolutionary principle, which states that the existence of knowledge must be the result of problem-solving processes, is the most directly relevant principle. But why does Deutsch believe in it? Is it really a fundamental principle of the philosophy of science?

I will argue that Deutsch has a more fundamental principle in mind, which serves as a justification for the evolutionary principle. He is basing his justification of the evolutionary principle on the idea that uncaused effects are impossible. Everything must be explainable. Explanation plays an important role in his unified picture of science, articulated in *The Fabric of Reality.*

> Science seeks better explanations. A scientific explanation accounts for our observations by postulating something about what reality is like and how it works. We deem an explanation to be better if it leaves fewer loose ends (such as entities whose properties are

> themselves unexplained), requires fewer and simpler postulates, is more general, meshes more easily with good explanations in other fields and so on. [6]

Though he goes on to say that there is no necessary connection between explanatory power and truth, insofar as our aim is to develop a scientific understanding of the world, explanation is a necessary guide.

The existence of a Knowledge Paradox would represent unexplainable knowledge. This is in tension with a thoroughly scientific analysis of the possible existence of CTCs, and should therefore be avoided. Any model that will eliminate the unexplained presence of knowledge should be preferred. This is why Deutsch devotes so much effort to showing that D-CTCs solve the Knowledge Paradox.

It will be useful to have an example of a Knowledge Paradox scenario in hand to see how Deutsch's solution is supposed to work. A famous example involves a Shakespeare scholar building a time machine to bring his favorite edition of the great author's *Complete Works* back to Elizabethan London to have it autographed by Shakespeare himself at a time before his first play was staged. When he arrives, he asks where he can find Shakespeare, and is directed to the gutter in the alley behind the pub. He finds an illiterate inebriate passed out in a puddle, and throws his book to the ground in disgust before returning to his own time. When Shakespeare wakes, he finds the book, and goes on to plagiarize his life's work from it.

Deutsch adopts the position that the structure of parallel MSM worlds in his D-CTC model solves the problem of Knowledge Paradox. Since time travel into the past necessarily involves time travel into a different parallel world, the knowledge that the time traveler brings with him (in the form of the book) *also* comes from another world. In his original world, Shakespeare wrote his plays and sonnets. But in the world into which he travels, the *counterpart* of Shakespeare that he encounters has the luxury of simply copying them.

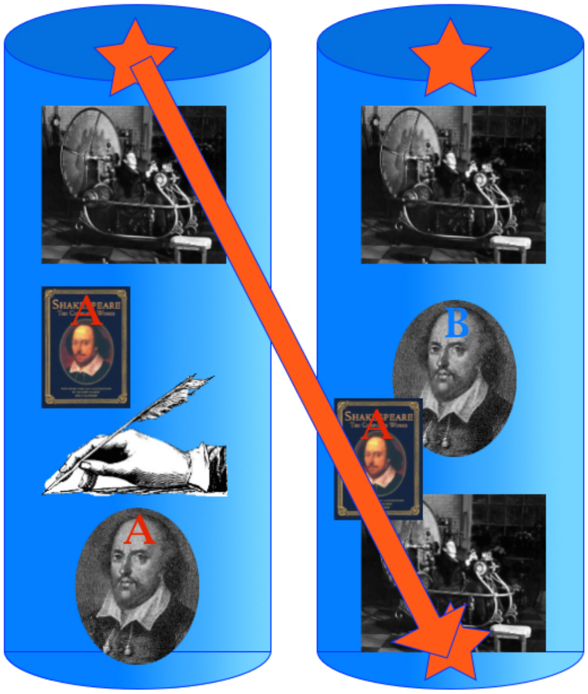
Deutsch argues that this avoids the paradox of getting "something for nothing" because, even though the counterpart of Shakespeare in the second universe didn't have to put in work to produce the plays for which he took credit, they were nonetheless the result of "genuine creative effort, albeit in another universe" [7].

Deutsch is very concerned about the possibility of time travel allowing for knowledge to exist absent the kind of dynamic or rational processes we usually think give rise to it. However, with the resources afforded him by his version of the Everett interpretation of quantum mechanics, he seems to have a satisfactory solution.

6 Problems for Deutsch's Solution

I argue that this reliance on a metaphysical position to solve the Knowledge Paradox poses a unique problem for Deutsch. Committing himself to realism

Figure 4: Deutsch's solution to the Information Paradox. There is no "free lunch" because the work that Shakespeare_B in the second world plagiarizes was written by Shakespeare_A in the first world.



about the MWM means that he is committed to the existence of worlds in which any physically possible events happen, however improbable. This is an unusual, but unproblematic feature of MWI in most cases. However, I argue it presents a potentially fatal problem for the solution to the Knowledge Paradox offered by Deutsch.

The problem arises when we recognize that among the very improbable world histories MWI guarantees are actual are worlds which are in every way indistinguishable from worlds in which there is a genuine violation of the fundamental evolutionary principle for the existence of knowledge. Deutsch's commitment to the actuality of the many worlds of the Everett interpretation includes worlds that are in every way indistinguishable from the Knowledge Paradox worlds he is trying to avoid. In these worlds, the Shakespeare counterpart is genuinely getting something for nothing: a free lunch. Consider the following two cases:

In the first, a time traveler named Tim loves the works of Shakespeare, and decides to go visit him before he's written his first play. He follows a CTC back to late 16th century England. Of course, because of the way time travel works, Tim actually disappears from his own universe A, and appears in universe B. The man that he meets in London is not Shakespeare_A, but is Shakespeare_B. So there is no problem that Tim leaves his copy of the *Collected Works* and returns to his own time. Among the future possible timelines of universe B is one that is in every other way (other than the fact that Shakespeare$_B$ is plagiarizing the plays and taking credit for them) indistinguishable from universe A. In 2121 Tim$_B$ is born, and at the exact same time that Tim$_A$ left *his* universe in a time machine, Tim$_B$ steps into his own time machine, destined for 16th century England (in universe C, of course).

In the second case, in universe X as the result of a random fluctuation, a man claiming to be Tim the time traveler appears in the alleyway behind a pub in 16th century England, looking for Shakespeare. He carries with him a book with the words "Complete Works of Shakespeare" legible on the front (though not *printed* per se, since they were the result of a random fluctuation). The man has a conversation with the layabout Shakespeare and leaves a book with him. Everything then proceeds in a way totally indistinguishable from the history of universe B. Shakespeare copies the plays out of the book, staging them periodically, and takes credit for them. In the year 2121, a boy named Tim$_X$ is born, and at the exact same time that Tim$_A$ and Tim$_B$ left *their* universes in a time machine, Tim$_X$ steps into his own time machine, destined for 16th century England. In this case, though, as a result of a fluctuation, he disappears into nothing, paying back the energy debt that allowed for the initial fluctuation.

In the first story, however many iterations happen, the *Complete Works* were, at some point in their history, *actually* written by someone (Shakespeare$_A$). There is no free lunch. Deutsch's evolutionary principle is not violated.

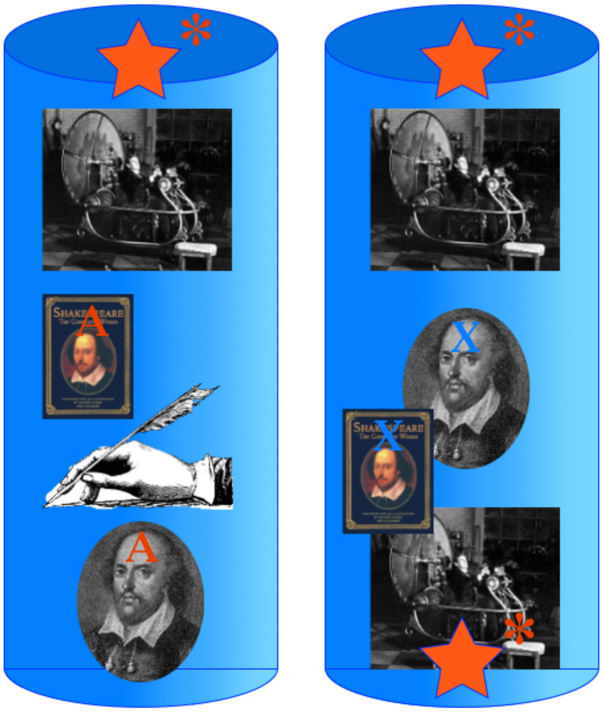
However, in the second story, *no one* wrote the *Complete Works.* Yet the two worlds are indistinguishable in every way. Tim the time traveler in universe X, though he was the result of a fluctuation, could be imagined to have all of the memories that Tim$_X$ accrues over his life before he steps into the time machine.

The only difference between the two stories is that in the first, the time traveler is stipulated to be identical with an individual from another universe. However, there is no way that this could ever be verified.

Any experiments to test for the presence of a CTC would necessarily send our test particles into other universes, never to be seen again. There is a possibility that we could receive test particles *from* other universes, though it's not clear what that evidence would establish, since we would have no access to the conditions under which the test particles were sent. And furthermore, for any conceivable data we could get that truly is the result of travel through a "gateway" from another universe, there is guaranteed to be another world in which indistinguishable evidence is gathered that is merely the result of a fluctuation.

Deutsch is careful to refer to the paradox under consideration as the "Knowledge Paradox", and not by the alternative, "Information Paradox". This potentially gives him some room to avoid this objection by claiming that the genuine paradoxical situations involve *knowledge*, whereas *The Complete Works of Shakespeare* that genuinely has no author, as in the worlds I point out, contains only *information*, and does not contain *knowledge.* We can come to have knowledge *of* that information (by studying the *Complete Works* in universe

Figure 5: A pair of worlds guaranteed to exist by Deutsch's metaphysical picture. The two worlds are unconnected, but the sequence of events in them is entirely indistinguishable from those represented in Figure 2. The *Complete Works of Shakespeare*$_X$ was not authored by anyone, but was the result of a random fluctuation.



X), but Tim the time traveler did not possess knowledge of Tim$_X$'s life in the mid 2100s, and the book he gave to Shakespeare$_X$ contained merely information (and did so only accidentally). He draws this distinction between information and knowledge in [4]:

> "Knowledge is not the same thing as information, nor is it any function of information alone. There is as yet no quantitative measure of "knowledge" that could be incorporated into physics. However, it is reasonable to suppose that the requirement that a system contain no independent information (which is what the maximum entropy rule effectively says) might also imply that the system contains no independent knowledge." [4, 3204]

Deutsch may have the resources to distinguish between these two situations, but it seems to me that it comes at a significant cost. One of the primary goals of his model of CTCs was to eliminate the possibility of this paradox.

However, the metaphysical position that he takes on to help build the foundation of his CTC model *necessitates* the belief that there are completely empirically indistinguishable histories from the ones he has tried to eliminate, in which there genuinely is a "free lunch". He can dispense with them by claiming that they do not contain real knowledge, and they are in principle distinct from the histories which *do*.

This can be seen as being motivated by the principle of the importance of explanation in a scientific theory. The existence of the *Complete Works of Shakespeare* in world X is not the result of an evolutionary process, and is therefore merely information. Knowledge requires an evolutionary process, and so there must be an explanation for its existence. Deutsch's solution to the Knowledge Paradox offers an explanation for the existence of the *Complete Works of Shakespeare* in universe B—it's there because it was carried over by Tim_A from universe A. Since there is no explanation for the existence of the *Complete Works of Shakespeare* in universe X, it does not count as knowledge.

7 Knowledge vs. Information

If the above is the solution to the present challenge, it is unclear why he didn't just use this distinction at the outset to solve the puzzle of the Knowledge Paradox. He could easily have said that artifacts which are the result of a Knowledge Paradox situation contain only *information*, and therefore are unproblematic.

He would not have needed to rely on the trans–world explanation for the existence of future artifacts in the past. He could simply have claimed that any history in which a Knowledge Paradox exists, the relevant artifact contains only *information*, and not *knowledge* of how to create it.

It has been pointed out (e.g. [17]) that the P-CTC model does not rule out the Knowledge Paradox. In the context of the debate between the proponents of the two CTC models, this is seems to be used as a mark against P-CTCs. However, in light of the trouble that the D-CTC model has handling these scenarios, perhaps we should reassess whether we should require our CTC models to rule them out.

In the very same paper where Lewis introduced CCC, he writes of the possibility of Knowledge Paradoxes.

> But where did the information come from in the first place? Why did the whole affair happen? There is simply no answer. The parts of the loop are explicable, the whole of it is not. Strange! But not impossible, and not too different from inexplicabilities we are already inured to. Almost everyone agrees that God, or the Big Bang, or the entire infinite past of the universe, or the decay of a tritium atom, is uncaused and inexplicable. Then if these are possible, why not also the the inexplicable causal loops that arise in time travel? [15]

Given the extra metaphysical structure Deutsch needs to take on to solve the Grandfather Paradox, given that the additional benefit of the solution to the

Knowledge Paradox is now somewhat cast into doubt, and given that there is a serious philosophical position that advocates for the acceptance of Knowledge Paradox effects, perhaps proponents of the D-CTC model should rethink his insistence on ruling out a free lunch.

It is perfectly consistent with a well-developed philosophy of science to believe that there are uncaused effects, as Lewis shows us. It's true that it would be a very strange world in which such a thing occurred. But that is not an argument against the possibility of a Knowledge Paradox obtaining.

References

[1] Arntzenius, F., Maudlin, T. Time travel in modern physics. *The Stanford Encyclopedia of Philosophy* Zalta, E. (ed.) (Winter 2013 edition).

[2] Davies, P. *How to Build a Time Machine.* Penguin Books, New York, New York, paperback edition, 2003.

[3] Deutsch, D. Quantum theory, the Church–Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London A*, 400(97), 1985.

[4] Deutsch, D. Quantum mechanics near closed timelike lines. Physical Review D **44**(10), 3197–3217 (1991)

[5] Deutsch, D. The structure of the multiverse. In *Proceedings of the Royal Society of London A*, volume 458, pages 2911–23, 2002.

[6] Deutsch, D. *The Fabric of Reality.* Penguin Books, New York, New York, paperback edition, 1998.

[7] Deutsch, D, Lockwood, M. The quantum physics of time travel. *Scientific American*, pages 50–56, March 1994.

[8] Dunlap, L. The metaphysics of D-CTCs: on the underlying assumptions of Deutsch's quantum solution to the paradoxes of time travel (2016). Studies in the History and Philosophy of Modern Physics **56** 39–47

[9] Gott, J.R. Closed timelike curves produced by pairs of moving cosmic strings: Exact solutions. *Physical Review Letters*, 66(9):1126–1129, 1991.

[10] Gott, J.R. *Time Travel in Einstein's Universe.* Houghton Mifflin Company, New York, NY, 2001.

[11] Hawking, S.W. Chronology protection conjecture. *Physical Review D*, 46(2):603, 1992.

[12] Lloyd, S. Quantum information science. 2009. Online Preprint. Retrieved from http://web.mit.edu/2.111/www/notes09/spring.pdf.

[13] Lloyd, S., Maccone, L., Garcia-Patron, R., Giovannetti, V., Shikano, Y. The quantum mechanics of time travel through post-selected teleportation. 2010. arXiv:1007.2615v2 [quant-ph].

[14] Lloyd, S., Maccone, L., Garcia-Patron, R., Giovannetti, V., Shikano, Y., Pirandola, S., Rozema, L.A., Darabi, A., Soudagar, Y., Shalm, L.K., Steinberg, A.M. Closed timelike curves via post-selection: Theory and experimental demonstration. 2010. arXiv:1005.2219v1 [quant-ph].

[15] Lewis, D.: The paradoxes of time travel. American Philosophical Quarterly **13**(2), 145–152 (1976)

[16] Novikov, I.: Can we change the past? In: The Future of Spacetime, pp. 57–86. Norton, New York, NY (2002)

[17] Ralph, T.C. Problems with modeling closed timelike curves with post-selected teleportation. 2011. arXiv: 1107.4675v1 [quant-ph].

[18] Thorne, K.S. *Black Holes & Time Warps.* W. W. Norton & Company, New York, New York, 1994.

[19] Yourgrau, P. *Gödel Meets Einstein: Time Travel in the Gödel Universe.* Open Court, Chicago, Illinois, 1999.