Forever and Again: Necessary Conditions for “Quantum Immortality” and its Practical Implications

Under review in *Journal of Evolution and Technology*

Alexey Turchin,

Science for Life Extension Foundation,

Digital Immortality Now

March 2018

alexeiturchin@gmail.com

**Abstract**. So-called “quantum immortality” should be better called “multiverse immortality,” as it mostly depends on the existence of many copies of the observer in other worlds. We identify two necessary premises when this theory becomes true: the very large size of the Universe (not necessary because of quantum effects), and the copy-friendly theory of the personal identity. We apply decisions under uncertainty to explore the situations in which multiverse immortality could be practically used. We show that, given large uncertainty, the only practical use of the multiverse immortality is as an additive to other life extension and survival projects, because in some cases it could considerably increase their chances of success. In the case of signing for cryonics, multiverse immortality significantly increases the subjective chances of its success, and also prevents a bad form of quantum immortality (s-risk of eternal suffering). Another application of quantum immortality is the survival of a global catastrophe. We also debunk several misconceptions, such as multiverse immortality is not applicable to slow dying, or quantum immortality is similar to respawning or NDE experiences. We show that the popular objection of lowering the measure of the observer in the case of quantum immortality doesn’t work, as it compensated by merging timelines. We consider multiverse immortality as plan D of reaching immortality, where plan A is to survive until the creation of the beneficial AI via fighting aging, plan B is cryonics, and plan C is digital immortality.

**Key points:**

* “Quantum immortality” (QI) is a particular case of “multiverse immortality” (MI).
* The validity of multiverse immortality depends on the size of the Universe and the nature of personal identity.
* MI greatly increases the subjective chances of the success of cryonics.
* MI makes euthanasia impossible, but favors cryothanasia.
* MI is our last defense against existential risks.
* MI doesn’t produce decline of measure, because of merging timelines.

**Keywords**: quantum immortality, cryonics, personal identity, life extension, euthanasia

**Disclaimer**: Suicide will never give anything useful in terms of quantum immortality, only that it will probably result in a non-deadly serious injury, permanent brain damage, and infinite sufferings. The idea of quantum immortality is not an argument for suicide, but is one of the strongest arguments against it, as suicide will never work as a “suffering stopper”. If you are still interested in suicide, ask for professional psychological help.

**Memetic hazard** for people who tend to emotionally react on the play of ideas!

**Abbreviation**:

DA – Doomsday argument

QM – Quantum mechanics

QS – Quantum suicide thought experiment

QI – Quantum immortality

MI – Multiverse immortality: umbrella term for QI and big world immortality

MWI – Many worlds interpretation of quantum mechanics

Content:

[1. Introduction 3](#_Toc508042523)

[2. The nature and formalism of the multiverse immortality 5](#_Toc508042524)

[2.1 History of the idea 5](#_Toc508042525)

[2.2 Formalism of the multiverse immortality 6](#_Toc508042526)

[2.3 Atoms, infinity and combinatorial theory 6](#_Toc508042527)

[2.4 The logical indefinability of the observer’s death 7](#_Toc508042528)

[2.5 Examples of the multiverse immortality 7](#_Toc508042529)

[2.6 Why it is not easy to kill QI 8](#_Toc508042530)

[3. Necessary conditions for Multiverse-immortality: the very big size of the Universe and no-soul personal identity 8](#_Toc508042531)

[3.1 Condition 1: Repeating of the human minds in the Universe 9](#_Toc508042532)

[3.1.1 The size of the Universe and various types of the multiverse immortality 9](#_Toc508042533)

[3.1.2 Tegmark’s levels of multiverse and its additional levels 9](#_Toc508042534)

[3.1.3 The finite size of human minds 11](#_Toc508042535)

[3.1.4 Sufficient diversity of the Universe 11](#_Toc508042536)

[3.1.5 “Actuality” of the existence of the very remote copies 11](#_Toc508042537)

[3.2 Condition 2: Appropriate theory of the personal identity 12](#_Toc508042538)

[3.2.1 Various types of personal identity theories 12](#_Toc508042539)

[3.2.2 “Mars teleportation paradox” and its similarity to the multiverse immortality 12](#_Toc508042540)

[3.2.3 “Copy-friendly” and “unique-soul” approaches to the problem of the observer identity 13](#_Toc508042541)

[3.2.4. Separate types of immortality for the separate types of identity 13](#_Toc508042542)

[3.3 Personal identity theories connection with various types of multiverse immortality 13](#_Toc508042543)

[3.4 A rough probability estimation that at least some form of the multiverse immortality is true 14](#_Toc508042544)

[4. QI as a decision theory problem 14](#_Toc508042545)

[4.1 Decision theory approach to multiverse immortality 14](#_Toc508042546)

[4.2 Presidential elections as an illustration of different DTs 15](#_Toc508042547)

[4.3 Applying UDT to the multiverse immortality 15](#_Toc508042548)

[4.4 Egoistic and altruistic values 16](#_Toc508042549)

[4.5 Discounting non-existence: another thought experiment 16](#_Toc508042550)

[4.6 Multiverse-immortality and small chance of survival from decision point of view 16](#_Toc508042551)

[4.7 Value of existence and joy of QI 17](#_Toc508042552)

[4.8 Is multiverse immortality only a belief? 17](#_Toc508042553)

[4.9 “Believer against one’s will” thought experiment 17](#_Toc508042554)

[4.10 Making decisions under logical uncertainty 18](#_Toc508042555)

[4.11 Merging timelines compensate the decline of measure 18](#_Toc508042556)

[5. Possible negative effects of QI 19](#_Toc508042557)

[5.1 The idea of QI is an informational hazard 19](#_Toc508042558)

[5.2 Impossibility to die may be bad 19](#_Toc508042559)

[5.3 Long-term unescapable sufferings (s-risks) are possible 19](#_Toc508042560)

[5.4 Bad infinites and bad circles 20](#_Toc508042561)

[5.5 Marginalisation of measure, strange worlds and infinite torture 20](#_Toc508042562)

[5.6 Dangers of mistakenly ignoring QI 21](#_Toc508042563)

[5.7 Dangers of QI-euphoria 21](#_Toc508042564)

[6. Possible practical applications of the multiverse immortality 21](#_Toc508042565)

[6.1 Surviving until life extension technologies 21](#_Toc508042566)

[6.2 Multiverse-immortality significantly increases the chances of the success of cryonics 21](#_Toc508042567)

[6.3 The impossibility of euthanasia and the need for cryothanasia 22](#_Toc508042568)

[6.4 Multiverse-immortality increases the chances of the success of digital immortality, and of the acausal trade between branches 22](#_Toc508042569)

[6.5 “Universal problem solver” as an example of a bad idea 23](#_Toc508042570)

[6.6 Collective survival from a global catastrophe 23](#_Toc508042571)

[6.7 Civilizational level QS as a universal problem solver, or an instrument for calculations 23](#_Toc508042572)

[6.8 Multiverse-immortality in the age of superintelligence: Will I become AI? 24](#_Toc508042573)

[6.9 Multiverse-immortality for computer algorithms 24](#_Toc508042574)

[6.10 QI and Unfriendly AI 24](#_Toc508042575)

[Conclusion 24](#_Toc508042576)

[Appendix 1. Many-worlds immortality and probabilities 24](#_Toc508042577)

[Does survival in the quantum suicide experiment present an evidence for MW-interpretation of QM? 25](#_Toc508042578)

[Does unexpected survival present an evidence for QI? 25](#_Toc508042579)

[Uncertainty in multiverse immortality theory 26](#_Toc508042580)

[The problem of measure 26](#_Toc508042581)

[Difference in the probabilities of outcomes in the multiverse and quantum immortality 27](#_Toc508042582)

[The nature of physical experiment 27](#_Toc508042583)

[QI and presumptuous philosopher 28](#_Toc508042584)

[Marginalization: Strange outcomes will dominate after several QS experiments 28](#_Toc508042585)

[Appendix 2. Multiverse immortality and other similar probabilistic arguments 28](#_Toc508042586)

[Multiverse-immortality and the anthropic principle 28](#_Toc508042587)

[Multiverse-immortality and simulation argument 29](#_Toc508042588)

[Multiverse-immortality and Doomsday argument 29](#_Toc508042589)

[Modal realism and back causation 30](#_Toc508042590)

[Multiverse-immortality resetting and infinite timelines in the space of possible minds 30](#_Toc508042591)

[The same “me” in different worlds and superposition of interpretations 30](#_Toc508042592)

[QI and the reference class problem 30](#_Toc508042593)

[Multiverse-immortality and qualia 30](#_Toc508042594)

[Appendix 3. Next step of QI theory: Dust theory and Boltzmann brains 31](#_Toc508042595)

[Multiverse-immortality and Boltzmann brains 31](#_Toc508042596)

[Multiverse-immortality and Dust-theory universe 32](#_Toc508042597)

[Dust theory and Flux-universe 33](#_Toc508042598)

[Miracles and glitches in the observer-moments chains 34](#_Toc508042599)

[Appendix 4. Other objections to QI and wrong interpretations 34](#_Toc508042600)

[Respawning and quantum woo 34](#_Toc508042601)

[Aging and slow death 34](#_Toc508042602)

[Not-exactly-me copies 35](#_Toc508042603)

[Resetting of timeclock, respawning and reference class problem 36](#_Toc508042604)

[General anti-immortality principle based on observed age 37](#_Toc508042605)

[The impossibility of sleep 37](#_Toc508042606)

[Many futures 37](#_Toc508042607)

[QI is only artefact of thinking, not the property of the world 38](#_Toc508042608)

[Adding up to normality 38](#_Toc508042609)

[References: 38](#_Toc508042610)

# Introduction

The idea of quantum immortality (QI) is that our world is constantly branching according to the many worlds interpretation of quantum mechanics, and thus in any situation of choice between life and death, an observer will find himself in a branch of the multiverse, where he didn’t die. The idea was suggested in a form of a quantum suicide (QS) thought experiment by Tegmark (Max Tegmark, 1998) and others.

Most scientists who wrote about the idea of quantum immortality felt the obligation to disprove it (Mallah, 2009), (Aranyosi, 2012), (Randall, 2004), in a similar way to the Doomsday argument (DA) (N. Bostrom, 1999), where most articles try to demonstrate the best way to kill the DA, but do not analyze the implications of the small probability that it is true. As a result, the possible practical consequences of quantum immortality were underexplored.

In this article, we suggest that there is some probability that quantum immortality is valid; we check the conditions necessary for it to be true, and the possible practical consequences.

In other words, we transform counter-arguments into conditions which explain when quantum immortality could happen and what it will be like, and based on this, we analyze the practical consequences of the possibility that this theory is true.

QI is closely related to the effective altruism circle of problems. If quantum immortality is true, it means possible infinite future suffering for any sentient being (Aranyosi, 2012), as he will not be able to die, but will age. So it is an *s-risk* (Daniel, 2017). However, there is a way to prevent this s-risk for currently living people, by choosing positive infinite outcomes (basically, by signing to cryonics).

In Section 2 we present an overview of the history of the idea, its formalism, and its relation to the idea of death. In Section 3 we explore the difference between “quantum immortality” and “big world immortality” and will analyze the conditions necessary for them to be true. In Section 4 we will look at QI from the point of view of decision theory. In Section 5 we will explore the possible bad consequences of “natural” QI. In Section 6 we will give an overview of the practical applications of QI, including overcoming the negative effects of natural QI and cryothanasia.

Some long discussions are moved into Appendixes. In App 1 we look into a probabilities discussion connected with QI, where we show that the well-known problem of measure decrease is compensated by measure increase through merging timelines. In App 2 we will look into the connections of QI with some other weird probabilistic effects, like the Doomsday argument, the simulation argument, and Boltzmann brains. In App 3 we will look on what could be the next evolutionary step of the QI idea, which is chains of possible observer-moments, also known as “Dust theory” (Egan, 2009). In App 4 we present some counter-arguments to several well-known objections to QI, as well as misinterpretations like “respawning”.

*Use of terms*: “multiverse immortality” is the technically correct term, which is an umbrella term for two separate subtypes: “big world immortality” and “quantum immortality”. But as “quantum immortality” is the most accepted term, we use it often.



*Picture 1. Quantum suicide experiment. One can observe oneself existing only in those timelines, where one didn’t die.*

# The nature and formalism of the multiverse immortality

## 2.1 History of the idea

The earliest idea that infinite size of the universe implies some form of immortality was the idea of so-called “eternal return”, that is appearing of the same observers again and again. The early versions of the eternal return theory existed in ancient philosophy (Eliade, 1949) and still attract interest of philosophers (Bergström, 2012).

In new times, the idea that an infinite universe implies some form of immortality first came to Romantic poet Heine in the first half of the 19th century, who wrote:

“*Time is infinite, but the things in time, the concrete bodies, are finite. They may indeed disperse into the smallest particles; but these particles, the atoms, have their determinate numbers, and the numbers of the configurations which, all of themselves, are formed out of them is also determinate. Now, however long a time may pass, according to the eternal laws governing the combinations of this eternal play of repetition, all configurations which have previously existed on this earth must yet meet, attract, repulse, kiss, and corrupt each other again*” (Kaufmann, 2013).

The idea was formalized by French socialist Blanqui (Blanqui, 1872) in the book “Immortality through the stars.”

F.Nietzsche came to the same idea in 1881 when he suggested his famous “*eternal return*” theory (Nietzsche, 1883). However, Nietzsche didn’t understand (or ignored) the obvious consequences of the eternal return: *that there will be always a world, similar to our world until the moment of the observer’s death, but differ in the moment of the death in the way that the observer will not die*, which basically means immortality from the subjective view of any observer.

This is the idea of *multiverse immortality (MI)* (of which *quantum* *immortality* is an important particular case) in a nutshell, and it is a natural development of the idea of the “eternal return”.

H.Everett was probably the first to come to the idea that an infinite universe implies immortality from the subjective point of view, based on his discovery of the many worlds interpretation (MWI) of quantum mechanics (QM), but never stated this publicly (Shikhovtsev, 2003).

The idea become known as “quantum immortality” after the “quantum suicide” thought experiment that was proposed by Squires (Squires, 1994), Hans Moravec (Moravec, 1988), and Marchal (Marchal, 1991), and was popularized by Tegmark (Max Tegmark, 1998). However, the idea of MI is older than quantum mechanics and, as we will show later, doesn’t require a many worlds interpretation of QM. (If you are interested in a more layman explanation of the quantum immortality idea, see article in [Gizmondo](https://io9.gizmodo.com/5891740/quantum-suicide-how-to-prove-the-multiverse-exists-in-the-most-violent-way-possible) or see picture 1.)

J. Huggo (Higgo, 1998) was the first to try to present the idea of quantum immortality coherently as a valid theory of immortality, but his analysis is not deep.

One important step towards understanding the idea is Tegmark’s four-level classification of the Multiverse (Max Tegmark, 2009).

Bevers explored how QI relates to the Born rule and apparently breaks it (Bevers, 2011).

Randall criticized QI from the point of view that most likely outcomes are neither “technological resurrection” nor “eternal decrepitude”, but chaotic random observers, similar to Boltzmann brains (Randall, 2004), but notes the usefulness of QI for cryonics. We will show that there is a self-stabilizing mechanism in the random sets of experience at the end of the article (Appendix about dust theory).

Circovic correctly mentions that the theory of personal identity is critical for QI (Cirkovic, 2004). He proceeds with proving that subjective credence should be equal to objective probabilities in the case of QI.

Almond (P. Almond, 2011) wrote a lot criticizing the idea that the quantum suicide thought experiment could be used as a proof of the many worlds interpretation of quantum mechanics. He also suggested that quantum suicide could be a “universal problem solver,” and civilizational level quantum suicide could be used to explain the Fermi paradox (P. Almond, 2008).

“Dust theory” (Egan, 2009) seems to be the next step developing the same idea, similar in significance to the step from the “eternal return” in the 19th century to the “quantum suicide” thought experiment. In the dust theory, something similar to quantum immortality is happening in every observer-moment.

## 2.2 Formalism of the multiverse immortality

There are several explanations of the quantum suicide thought experiment, and we expect that the reader already knows them. The main idea is that a quantum event is like radioactive decay triggering a powerful bomb near me. As a result, two timelines are possible, where either the bomb exploded or it did not, and if we agree with the many worlds interpretation of QM by Everett, this means that there will be two future timelines for me. In one I will exist, and in the another I will be instantly destroyed. Thus, I will be able to observe only the timeline where the bomb didn’t explode. (The experiment seems similar to the Schrodinger cat thought experiment, but here we ignore the tantalizing idea that I could be in a superposition of a dead and alive state.)

The quantum suicide experiment means that I will always observe winning in “Russian roulette”, if it is built the same way, so I will observe that I have surprising survival skills. The main question is equality of the inevitable observation of not dying to the subjective immortality.

Now we will formalize this experiment in its simplest form. Imagine that any observer consists of a discrete timer T and identity block I, which is just a constant string of numbers, not changing in time. Any observer-moment O—that is the sum of the experiences of the observer in the moment T—could be described as O(I, T).

In that case, the multiverse immorality claim is: for any O(I,T) for any T and any I exists O(I,T+1):

$∀O\left(I,T\right)\rightarrow ∃O\left(I, T+1\right), for ∀I,∀T$ (1)

The proof is the following: If the universe is sufficiently larger and random, it will always generate the string of numbers O(I,T+1) for any O(I,T), based on the Poincaré recurrence theorem (Poincaré, 1890). Tegmark showed that this recurrence is applicable even to quantum systems, despite continuity of the Schrodinger equation (M. Tegmark, 2017, also see Bocchieri & Loinger, 1957).

Note, that there is hidden claim that the existence of O(I,T+1) is enough for immortality. Exactly this claim typically raises the most objections.

## 2.3 Atoms, infinity and combinatorial theory

To apply Poincaré theorem to our universe, we need to show that our universe is infinite, and that it consists of finite atoms, combining in finite minds, and this was done by Tegmark, as we can see in the next section.

A finite universe consisting of infinitely complex minds seems to be a showstopper for the QI, but in fact, our universe is not finite in the latest version of cosmology, and quantum states do not seem to be important for mind processing, if we don’t buy the Penrose theory on the mind as an advanced quantum computer (Penrose & Gardner, 2002).

Even if the human mind is a classical quantum computer, the fact that it is able to exchange information with the world, without damage to its quantum coherency, means that the indestructibility of particular quantum states is not important.

In this case, if the MWI interpretation of QM is true, and the mind is a quantum computer, quantum immortality still works for it, but big world immortality may not work (or we should assume that personal identity in a quantum computer of mind could be restarted from scratch, if such a computer is restarted after being turned off, but after getting the same classical information input.)

## 2.4 The logical indefinability of the observer’s death

There is a well-known conjecture: “All humans are mortal, I am a human, so I will die.” However, it has several flaws. One of them is the “fallacy of four terms” (Copi, n.d.) as the word “death” has different meanings in the premises and in the conclusion.

In the premises, “death” means “death-for-others”, that is *observation of death by the outside observers*, that is, cessation of the living activity of the body. In the conclusion of the syllogism, the term “death-for-me” is used, which means *the death of the observer from the observer’s point of view*, which is typically represented by the idea of “nothingness after death” or cessation of the process of the observation. However, nobody is able to experience and report anything about nothingness after death (by definition), so it is not possible to prove that it will actually happen. (Another fallacy is that not all humans are currently dead, so we can’t claim that all are mortal.)

The multiverse immortality means that death—from the observer’s point of view—is impossible. This impossibility means that, for any state of any observer with time-stamp T0, there is a possible next state with T0+1, and as everything possible exists, such an observer exists. Thus, a last moment of experience is impossible, and death as the cessation of the process of experiencing, is impossible.

Almond attacks this idea in (Paul Almond, 2011c), where he noted that if MWI is not true, the death of observer is a consistent explanation or reality, but if we assume that MWI is true, death in some branches immediately becomes an inconsistent explanation. This jump between varying descriptions of death is not logical, and undermines the idea of QI, according to Almond. In other words, according to Almond, if we agree that death is theoretically possible, because of a small world, we should also agree that the death is possible in the branching world, as the death of some branches.

This is not a question of fact, but one of interpretation: should we regard the existence of my copy as an immortality or not. The answer of the problem raised by Almond is connected with the nature of personal identity, which will be analyzed in the next section. If identity is connected with measure of existence, then lowering this measure is equal to death, with some probability. If identity depends only on information, then changes in the number of copies is irrelevant.

I will illustrate this on the following examples:

1. imagine many parallel running copies, like a flat world with some thickness (Yudkowsky, 2008), like pages with the same text in a book. If I cut the copies in half, they will die completely, but the book will be just half thinner. This interpretation is based on idea of “fluid of reality” (Yudkowsky, 2009a), and is applicable if the mind is not a classical computer. The loss of “reality fluid” doesn’t necessary mean death, which happens only if the “reality fluid” is not inter-mixing and consists of layers. If it is intermixing, its loss will be like a loss of blood, i.e., partial death, but survivable.
2. Imagine that in the world there are many copies of the novel “War and Peace.” If some of these copies are deleted, it will not affect the plot of the book. This interpretation is based on the idea of the mind as a pure classical computer.

The nature of death depends on nature of personal identity and could be described as the end of identity.

Another counter argument is that if I am not currently on Mars, that doesn’t mean that I am dead on Mars. In any given moment, there are lot of branches which are not happening with me, but it doesn’t mean that I die every moment—and if we assume non-existence of branches as deaths, I am dying every second in millions of ways.

## 2.5 Examples of the multiverse immortality

There is nothing strange in the multiverse immortality, if we look on simpler objects and beings:

**Numbers are immortal.** All numbers will appear over and over again in multiple worlds, and we could say that any number is “immortal”. The same is true for number series. The number 27 will appear again and again, but we are not surprised by this, as the number is small compared to the size of our world, where we can find 27 ships and 27 sheep. However, repetition of longer numbers becomes more and more surprising, but only because our world seems to be too small for such repetition.

**Simple molecules are immortal. A** molecule of water will appear again and again, because it is very simple, and consists of very abundant components. But more complex compounds are rare. Some short organic molecules are likely to repeat rather often. But a molecule with around 100 randomly connected atoms will be so rare, that it will appear once in our Hubble volume. That is, the simpler the object, and the more generative is its environment, the more likely it will have something like multiverse immortality.

**Biological viruses**. Even some living beings could enjoy this type of immortality in an observable way. Biological viruses of one species are all equal to each other (excluding mutations) and the death of one copy is not the death of the virus.

**Computer viruses’** copies are exactly equal to each other, and the virus is not interested in preserving any exact copy, it is only interested in the dissemination of copies. So, a virus could use reasoning similar to the multiverse immortality, that even if it disappears, the Internet is large enough for another copy of the virus exist somewhere.

 The main question is, if it this be applied to the human mind. Is the universe sufficiently large, and does the nature of human identity prevent such a form of immortality?

## 2.6 Why it is not easy to kill QI theory

There are a lot of objections to QI, and the idea seems to be too complex to be true. However, any attempt to kill the idea comes with a price, that we assume some properties of the world that are even more weird. For example, if we assume a small universe, we assume the existence of some strange force which is limiting its growth, and also fine-tune the remaining small part for life—which sounds similar to God.

If we assume that human identity is fragile, and its informational copy is not enough, we also should postulate some weird theory of personal identity based on the existence of either a soul, or at least non-informational carrier of identity, which may be qualia or quantum mind. These theories may open possibilities for other forms of immortality.

# 3. Necessary conditions for multiverse-immortality: the very big size of the Universe and no-soul personal identity

For the multiverse immortality condition (1) from section 2.2 to be valid for human minds, that is to represent actual immortality, two preconditions need to be true:

1. *The universe must include all possible variants of human minds*, that is, it should be large enough for human minds to be repeated in it and minds must be of finite complexity. This means that O(I,T+1) observer-moment should always exist.
2. *Human personal identity theory must support immortality through copies:* from the existence of similar moments O(I,T), and O(I,T+1) should follow continuity of the personal identity O(I,T), ⇒ O(I,T+1). In other words, the identity solution should be “no-soul” and “copy-friendly”.

In this section, we will take an overview of both conditions.

## 3.1 Condition 1: Repeating of the human minds in the Universe

### 3.1.1 The size of the Universe and various types of the multiverse immortality

While the popular name of the “multiverse immortality” is “quantum immortality,” as it is connected with the many worlds interpretation of quantum mechanics, multiverse immortality in fact doesn’t require a MW-interpretation of QM to be true, and doesn’t use any quantum effects for work. So, the name is misleading, and results in attracting some quantum woo.

The only thing it needs is a sufficiently large and diverse universe for the existence of my copies. But how large, and how diverse?

Universe must have all possible minds and all possible their observer moments, only in this case for any O(I,T) will exist O(I,T+1). For this to happens, 3 conditions must be true:

**Very large Universe size**: The universe should be large enough to include all the possible combinations of atoms inside a human skull. The condition of the “infinity” of the Universe’s size is not necessary, but would suffice.

**Finite size of minds.** Humanminds should consist of finite combinations of finite elements.

**Diversity of the Universe**: The universe should be also diverse enough to allow even very rare combinations of atoms (or at least mind states) to appear in it.

We will prove all three conditions in the next subsections.

### 3.1.2 Tegmark’s levels of multiverse and its additional levels

If the Universe is really-really big, it is difficult to kill the idea of multiverse immortality. Max Tegmark identified four levels which independently create a world—large enough to include all possible observer’s minds (at least of human mind size). The best description, and his own comment is in (M. Tegmark, 2017):

1. *Level 1. Cosmological Inflation:* Our bubble universe is very large. Supported by observations of anisotropy. One estimate is 10^{10^{10^{122}}} megaparsecs (Page, 2007).
2. *Level 2. Chaotic inflation*: quantum foam during the Big Bang creates many bubbles with differing laws of physics. Supported by ideas that fine tuning of physical laws is explained by the observer selection effect, also known as anthropic principle.
3. *Level 3. Many worlds interpretation of quantum mechanics*. The most simple of the QM explanations.
4. *Level 4. Mathematical universe*: all possible mathematical structures exist. The most simple explanation of the fact that everything exists at all.

Besides original Tegmark’s levels, several other levels of multiverse are possible:

1. *Other causally independent Big Bangs*. If one was possible, there is a mechanism to create something from nothing, and such a mechanism can’t be exhausted, as the mechanism can’t have causal power over causally disconnected other Big Bangs. Such other Big Bangs are not the same as other bubbles in a quantum foam: they are completely independent “pieces of foam.”
2. *Simulations and multilevel simulation*. Bostrom stated that only local galactic supercluster could support 10E46 simulated humans every century (N. Bostrom, 2003). There is also the idea of nested multilevel simulation (Torres, 2014).
3. *Infinite time*. In most contemporary cosmological theories, there is no infinite time, except some cyclic universe models.
4. *Other branes in the many-dimensional space.* Many branes appear in some variants of string theory.
5. *Boltzmann brains*. Calculations mimicking an observer-moment of a human mind could appear as fluctuations in vacuum.
6. *Dust theory*. Calculations mimicking an observer-moment appearing as random causal chains inside thermodynamic objects.
7. *Fecund universes.* This is the idea that the universes could replicate and self-tune via black holes (Smolin, 1992). And if this process exists, it could create infinitely many universes. Some fecund universe theories even include the need for the existence of human intelligence (perhaps to create a black hole with the required properties). (evodevouniverse, 2018)
8. *Full blown modal realism.* Everything possible exists.
9. *String theory landscape (14)* universe (Kachru, Kallosh, Linde, & Trivedi, 2003), (B. Greene, 2011). String theory landscape multiverse explains why we have these exact laws of physics by postulating something like 10E500 separate types of universe, with individual numbers of dimensions, fields etc.

These nine additional levels are more hypothetical, while the first levels are supported by the data of contemporary science, especially Levels 1 and 2, as explained by M. Tegmark, (2017). So, “big world immortality” in a very large universe receives stronger scientific support than even “quantum immortality,” which is based on currently untestable MWI.

One attractive figure of Levels 1 and 2 is that they are causally disconnected only currently, but come from the same causal process of the same Big Bang, and could become connected again after an infinitely long time in the future. So, Level 1 and 2 copies surely have the same ontological status as any object in the observable universe, and even your neighbor next door (more about actuality in the next section).

Surely, some other levels of the big world are possible, which we do not know about yet.

Is it possible to disprove somehow big world theory? Maybe if God exists, he could limit the number of existing worlds?

Another possible objection to the big world is idea that only “me-now” actually exists(similar to the “moving spotlight theory” of now (Deasy, 2015), “actual now” (Bitbol, 1994) and “ontological privilege” of now (Frischhut & Skiles, 2013)), and everything unobservable doesn’t exist. But even this will not kill multiverse immortality, as in that case, consciousness becomes an ontological necessity which helps to create now-moment and thus can’t be turned off.

Given all this, I estimate that big universe hypothesis has a 0.8 probability of being true. (It is not actual probability, but the level of my credence in the theory, or logical uncertainty.)

There are still some scientists who do not buy the multiverse theories, like (Hossenfelder, 2018).

The main argument of the critic is that as other universes are non-observable, it is difficult to observe them, and thus they are beyond real science. However, it is not an argument for a small world.

Tipler showed that eternal recurrence can’t happened in closed universe described by general relativity (Tipler, 1980).

It is difficult to prove that the Universe is infinite, but to prove that it is finite is impossible, so a multiverse immortality can’t be disproved.

### 3.1.3 The finite size of human minds

Not only does universe have to be infinitely large, but human minds need to be finite for multiverse immortality to work.

As the human brain consists of a finite number of atoms (around 5x10E26), it seems to be finite. Tegmark shows that quantum effects do not prevent the number of states from finite, as any quantum system in the finite space could only be in the finite number of internal states.

Moreover, personal conscious memory size was estimated around 2.5 Gygabites (Carrigan Jr, 2006), and this means that only 10E(10E10) separate people are possible, and as most of them will have immeasurably small differences, the actual number is smaller.

If we look on the observer-moment size, that is the current size of human experience, it is dominated by the visual field, which could be roughly estimated by the size of the signal which the eye sends to the brain. It uses 1 million nerves in optic nerve fibers https://en.wikipedia.org/wiki/Retina, each probably presenting one pixel, so the actual size of an observer-moment is around 1 megabyte, and given a lot of non-significant peripheral noise, it is probably even smaller. Thus, there are 10E1000000 separate visual observer-moments, most of which are noise.

Surely, the most important part of an observer-moment is not what I see, but what I feel and think about it, but these parts are probably less complex than the visual field. For example, assuming that the size of the observer-moment is one sentence, and such a sentence can’t have more than 20 words, most of which are simple and frequent, so they are selected from just 1000 of the most typical words, we could conclude that there are 10E60 possible sentences, which is less than the number of particles in the visual universe (and more like the number of small meteors).

### 3.1.4 Sufficient diversity of the Universe

But the infinite size of the universe and finite size of human minds is not enough to prove multiverse immortality. The universe must be diverse enough for any possible mind states to appear, because we need more than just repetition of events, we need an escape from even the worst possible situations.

Imagine that my spacecraft is falling into the Sun. At the first glance, it seems that there is no chance for my survival. However, in a sufficiently diverse universe, an alien space craft could rescue me, or a rogue black hole moving at near-light speed could remove the Sun—or, most likely, it is just a dream.

### 3.1.5 “Actuality” of the existence of the very remote copies

One important condition is that a big universe must exist *actually* (Menzel, 2017).

In short, “actuality” is the idea that some remote thing exists in the same way as I exist now. In a sense, “actuality” is actual “existence”, which implies that there is no second-rate form of existence. Existence can’t be discounted.

 The difference becomes clear if we look at the modal status of events in the past. Depending on the theory of time used, one would say that past events are actual (that is in a timeless universe), or not-actual, that is, they existed but do not exist now, so they are more than just possible events, but less than actually existing events.

There is an open question in the philosophical ontology if non-causally connected regions which could be claimed to exist actually. However, in our case, in some theories they are connected, as in the cases of MW-interpretation of QM, the many simulation universe, and the eternal universe; or they were connected, or they will be connected.

For multiverse immortality, my copies must exist actually, or they should at least become actual in the future. If they are not actual, they do not exist.

There are two solutions to the actuality problem:

1. Modal realism, that is, everything possible exists
2. Actuality is somehow connected with the “now” moment and is “irradiated” from it.

Both solutions imply multiverse immortality, as we show above, and thus, for the case of multiverse immortality, it doesn’t matter.

There is also the idea of “relative actuality”, wherein actuality appears only if two objects are causally connected. Thus, unobservable regions of the universe are not actual. The theory has some difficulties.

In the case of quantum immortality, actuality is provided by the fact that next observer-moments exist in the future, and are causally connected with me-now. But if we start to think about my copies in the multiple Big Bangs, it is easy to ask: if they are so remote, and so disconnected from my universe, how do they could help in terms of my immortality? This question is an intuitive expression of the doubts in the actuality of remote copies.

Intuitively, it seems that the more potential for causal connection between copies, the better it is for multiverse immortality, but it is not clear how grounded this intuition is.

There is even a point of view that if all moments of time are actually existing, there should not be fear of death, as each moment continue to exist (Deng, 2015).

## 3.2 Condition 2: Appropriate theory of the personal identity

### 3.2.1 Various types of personal identity theories

The theory of personal identity is notoriously difficult, despite seeming simple, and attempts to describe it will quickly produce controversies and spill off-topic.

We don’t know what human identity is, and one of the reasons is that we don’t know the nature of consciousness and qualia (Chalmers, 1996), and without it, personal identity can’t be solved.

According to Parfit (Parfit, 1984), there are two types of identity. One is based on the continuity of the personal memory, and another is based on the continuity of observations. For any real human, his identity consists of several complexly intertwined types of identity, and discussion of it is not simple, as it also includes his image of self, his social identity etc. See my map of the types of identity <http://lesswrong.com/lw/nuc/identity_map/>. For simplification, we will focus on two main types of identity.

We will call them *informational identity* and *observer identity*. Informational identity is rather straightforward and less controversial, as it simply explains the measure of similarities of two minds. If two minds have the same memories M, they are informationally identical, of just each other copies.

Observational identity answers the question: what I will experience in the next moment of time?

### 3.2.2 “Mars teleportation paradox” and its similarity to the multiverse immortality

In the Mars teleportation paradox, a person is scanned, and his copy is sent and recreated on Mars, and his original is instantly destroyed on Earth. If one agrees to such a type of transportation, he is asked what if the destroying part of the system does not work, resulting in two copies, one on Earth, and the other on Mars. Which of the copies will he be?

If humans were computer programs, there would be no contradiction. A program could have two copies, and it would not be surprised by it. But in the case of a human being, human intuitions about his personal identity contradict the idea of copying.

The Mars teleportation paradox is similar to the multiverse immortality, with just one difference: no data is actually sent to a remote celestial body, as my copy is already there.

### 3.2.3 “Copy-friendly” and “unique-soul” approaches to the problem of the observer identity

There are three main views on the nature of the *observer identity* (*informational identity* theory also has several levels, depending on the estimations of the amount of similarity that must be preserved for two minds to be regarded as copies):

1. Copy-friendly view. It postulates that there is no way to measure observer identity, it is just based on the fear of a human being losing his body. The main problem of this approach is that it could completely dissolve the idea of identity, which will contradict the ways humans feel and plan their actions.
2. Unique-soul view. This point of view is that some unique substrate of identity exists, which can’t be measured from outside, but which dictates what I will experience in the next moment. Such substrate is either “soul,” or causal continuity of the human consciousness between subsequent observer-moments.
3. Another popular theory of identity is open individualism, which denies the existence of the “identity” and postulates that all sentient beings are the same in core.

“Causal continuity as the nature of the human identity” is rather popular view, but it has some difficult problems. For example, if a person has a narcosis, his brain activity almost stops, should that be regarded as the death of personality? And what about the abrupt end of a night dream?

Also, causal continuity doesn’t mean exclusivity of the next state of consciousness, as two new states could follow smoothly on from the “parent” state, and if quantum immortality is true, this is exactly what happens within it: minds smoothly branch into two (and more) minds every moment in time.

I would add that “human identity” is a complex socio-biological adaptation, which is not very similar to the identity of minds in general, typically explored by philosophers.

Identity also could be described through values and *choices*, and this approach will be discussed in Section 4.

Later we will show that, despite the illusion that *informational identity* and *observer identity* could become two independent timelines in the case of QI, they will actually merge soon.

### 3.2.4. Separate types of immortality for the separate types of identity

Immortality could be defined as *preserving ones’ identity without end.* (And death is the end of one’s identity.)

“Without end” doesn’t mean “infinitely long”, as time-like curves are possible (more about them later).

But the definition of personal immortality depends on the definition of personal identity. This means that there are (at least) two types of immortality, one for the informational identity and another for observer identity, and they may be realized in completely different timelines.

In other words, multiverse immortality for *information identity* means that for any person with memories M about events (T0… Tn), there exists a person with memories M’ for the events (T0…. Tn, Tn+1).

And multiverse immortality for the *observational identity* means that for any observer-moment O(I,T), there exists an observer-moment O(I, T+1). That is, for any observer-moment there is a next observer-moment, which logically follows from the previous observer-moment.

## 3.3 Personal identity theories connection with various types of multiverse immortality

Various combinations of the *personal identity theory* and the big world theory produce various answers about the possibility of multiverse immortality.

For example, if causal connection is required, AND the big world interpretation of QM is true, then multiverse-immortality will work.

In addition, if only informational identity is true AND only a cosmological inflation based big world is true, multiverse-immortality will also work.

The relation between big world theories and identity theories is presented in table 1. This table shows how multiverse immortality depends on these two variables. YES means that multiverse immortality will work, NO means that it will not work.

Both variables are unknown to us currently. Simply speaking, multiverse-immortality will not work if an actually existing world is small or if personal identity is very fragile.

Table 1. The relation between big world theories and personal identity theories

|  |  |  |  |
| --- | --- | --- | --- |
| **Big world theory** *⇒***Personal identity theory**⇓ | MWI interpretation of quantum mechanics | Big world (cosmological inflation) | Only visible universe actually exists |
| I am only information: copy-friendly, no-soul view on personal identity | Yes | Yes | No |
| Non-informational identity carrier exists: soul or continuity of consciousness is required | Yes | No | No |
| Zero-identity, I am everybody | Yes | Yes | yes |

## 3.4 A rough probability estimation that at least some form of the multiverse immortality is true

Here we see that only in two thirds of the cases, the combination of personal identity theory and big world theory is right for multiverse immortality. To get a priori probability of multiverse immortality, we need to multiply this estimate by a previously given estimate that at least one multiverse theory is true, for which we get 0.6. Certainly, the numerical estimates are arbitrary, but they show that uncertainty is very large.

Thus, very rough estimates of the veracity of multiverse immortality gives around even chances, that is p=0.5. I remind you, that this is not classical probability, but my current bet based on some vague Bayesian estimations of the logical uncertainty of the several propositions.

It means that any application of the multiverse-immortality where we need it to work for sure are dangerous, as there is almost half a chance that it will not work. However, we could still use it as an additive to other life extension and problem-solving instruments—or be afraid of negative forms of multiverse immortality.

# 4. QI as a decision theory problem

## 4.1 Decision theory approach to multiverse immortality

We could approach QI as decision theory problem.

The formulation of a decision problem is rather simple: Were one to play Russian Roulette for money, where it has probability P of not firing, and brings utility U (like winning 1 million USD), and probability 1-P of causing death, which has negative utility Ud (neg.utility of death). We will call this thought experiment as Paid *Russian Roulette.*

In a normal word model, the condition looks simple from a utilitarian perspective:

PU 〉 (1-P)Ud (2)

If many-worlds immortality is true, it seems that the condition is

PU 〉 0

But if we assume that there is an agent who has credence X in the validity of the multiverse immortality theory, then:

PU 〉 (1-X)(1-P)Ud

However, such calculations assume implementation of some form of decision theory (DT). There are two main DTs: causal DT and evidential DT (Soares & Levinstein, 2017).

Evidential decision theory seems to favor quantum immortality, as it recommends actions based only on existing experiences, and as I could experience only not being killed in Russian Roulette, I will win (and non-experience is not experience of nothing.)

In the case of causal DT the situation is more tricky, as killing myself is a causal event affecting winning, and thus should be counted as actually happening, i.e., non-receiving of the prize.

I remind the reader that decision theories are not something that actually exist, but just the best ways of choosing a winning strategy (P. Greene, 2017).

## 4.2 Presidential elections as an illustration of different DTs

Imagine, that there are presidential elections with millions of voters, and there are two main candidates, Trump and Sanders. Trump has higher popularity, and Sanders is not very popular. I want Sanders to win. And if everybody who wants to vote for Sanders does so, he will win.

Causal decision theory implies that my vote is very unlikely to change the election result, especially, as I know that candidates are not even in polls. Nothing will change, whether I vote or not.

Evidential decision theory says that if I vote for Trump, I will win.

However, if everybody who wants to vote for Sanders actually votes, he will win. So, if I think that my decision to go to vote is based on the same lines of reasoning as all other Sanders supporters, if I go to vote, all of them will also go, and we will win. Reasoning like this motivated Wei Dai to create updateless decision theory. “Updateless,” in this case, means that I should not update my decision based on the fact that my vote is insignificant, I should still vote for my favorable candidate.

The formulation of Updateless DT: “UDT specifies that the optimal agent is the one with the best algorithm—the best mapping from observations to actions—across a probability distribution of all world-histories” (LessWrong Wiki, 2018), (Armstrong, 2011). In the case of our presidential election example, “all world-histories” are all instances of Sanders supporters, so the best algorithm here is that all of them will go to the election and vote.

## 4.3 Applying UDT to the multiverse immortality

In example above of paid Russian roulette, an observer would expect all his friends to start playing the same game, so many of them would die soon. So, he will soon be living in an empty world where he will not be able to spend his money, and where there will be many other forms of sufferings.

More practical example of Paid Russian Roulette is when a person decides not to invest in personal life extension, hoping that multiverse-immortality will help him to survive until the appearance of powerful life extension technologies (some people on *LessWrong* expressed this opinion as a reason for not signing for cryonics). But if everybody were to behave this way, you would observe more deaths around you, and few of your friends would return to live.

However, if everybody were to sign up for cryonics, as its probability of working is increased by QI (more later), in this world there would be more funding for cryonics, and this would be win-win situation for all.

## 4.4 Egoistic and altruistic values

Another way to connect multiverse-immortality with decision theory is to introduce the notions of egoistic and altruistic values.

* *Egoistic values* care only about my future experiences, about pleasure, or about escaping pain.
* *Altruistic values* care about what will happen to other people. Imagine a person who cares about his family.

As we show above, even in the case of purely egoistic values, QI may be not the best choice in situation where other people may start to play paid Russian roulette, or similar games.

Surprisingly, altruistic values don’t always exclude QI, as in some cases it could be used to escape infinite future suffering for any human being, for example in the case of cryothanasia.

We could also divide values, as values about experiences, and values about the actual state of the world. Most values about the state of the world are altruistic.

 For example, if an ecologist wants to preserve the elephant population, he probably wants to preserve it in all actual future timelines, so he cares about the integral of the elephant population in all futures, weighted by the probabilities of these futures. This integral is independent of his survival in these timelines (if he is not the main activist). So, he is not interested in surviving in just one unique time-line, where there are a lot of elephants, but in the fact that in all timelines there will be a lot of elephants, whether he is alive or dead.

Most real humans have a mix of egoistic and altruistic values.

Also, from a biological evolution point of view, we should try to have children in all branches.

## 4.5 Discounting non-existence: another thought experiment

Typically, humans make plans only for the periods of time when they expect to be alive with high probability. But QI suggests ignoring complete non-existence probabilities.

For example, A is a writer who is writing a book, but he is doing it just for fun, not for publishing, so it is an egoistic value. To finish the book, he needs 50 years, and he will be 90 years old at the time. So, he could have two lines of reasoning:

1. As my life expectancy is 70 years, there is only a five percent chance that I will survive to 90, so it is unlikely that I will finish the book. So maybe I should try a shorter book, which has only 0.1 utility for me.
2. But because of QI, I will survive to 90 from my observational point of view, so it is better to start working on the longer book.

## 4.6 Multiverse-immortality and small chance of survival from decision point of view

We could imagine the following thought experiment as a counterargument for multiverse immortality. Imagine that I am going to be killed with 99 percent probability, but my only chance of survival is to jump from the cliff to another cliff where I have only one percent chance of a successful jump. Multiverse-immortality says that it is rational to jump anyway, as I will surely survive. But ordinary rationality says that jumping for a one percent chance is also rational, as it is only way to survive. And, as the decisions are the same, the theories don’t have any meaningful difference.

Certainly, the probability of death has an emotional impact on a human being, as we used to think that a one percent chance of survival is extremely bad, and it would affect a person’s ability to jump. Believing in multiverse-immortality, that is anticipating the experience of only time-lines where one will survive, could add courage.

Imagine that a doctor suggests to you either surgery which will kill you with a probability of 99 per cent, or several months of slow dying, but you will have time to say goodbye to your family etc. In that case, multiverse-immortality could move odds in decision making, as the one per cent chance of survival could turn into 50 (because of uncertainty of will it work or not).

## 4.7 Value of existence and joy of QI

V.Kosoy, in a personal communication, suggested that the idea of QI should not affect our choices from the decision theory point of view, if we correctly calculate expected utilities. But there is (at least) one important exception. Many people have an extremely negative preference for non-existence, also known as *fear of death*. QI claims that there will be always some form of survival, maybe not good, not so proven, with lower measure, but survival. There is a difference between having a one percent chance of survival, and surviving with one percent of measure. This difference typically disappears in the expected utility calculations. If someone has a preference for existence, QI has a big utility gain for him.

I personally remember that when I came to the idea of the multiverse immortality in 1990 while reading Nietzsche, I felt a large stone over my head disappear—this stone was fear of death. The relief from fear of death could be an actual and measurable benefit of QI. In my case, I felt euphoria for one or two months until I recognized penitential negative consequences of QI, and later I stopped having any emotional impact from QI; the fear of death returned and disappeared again for some time after I signed up for cryonics.

## 4.8 Is multiverse immortality only a belief?

One could say that pruning of the branches in QI is equal to death, let’s call him Deather. Another one could say that only existing future copies should be counted, non-Deather.

They could even undergo a collective QS experiment, but it will not change the beliefs of both. Deather will say that it is just a random survival event, which is not proving anything, and non-Deather will say that it is evidence for survival.

The main difference between these points of view is that they ignore pure rational decision making and human decision making, and this difference will be illustrated in the following example.

QI only postulates the fact that the next observer-moment O(I, T+1) always actually exists, but does not give any interpretation of it ,and humans don’t have any practical intuition about what to do with the existence of your copies (at least for now, brain uploading will make the problem sharper). To check this belief in an experiment is risky, expensive and mathematically controversial.

## 4.9 “Believer against one’s will” thought experiment

After Mary survived several rounds of QS, she could say: “I still don’t believe that MWI is true, and I think that my survival is pure randomness. I also think that I will most likely completely disappear after the next rounds. I want to drink water now, but I will not drink water, as the next several rounds of QS will happen before I will able to get to the bottle, and I will be dead in 999 of 1000 of futures.” After several rounds she becomes even more thirsty, and decides to go for the water bottle anyway.

Here her subjective surprise of survival will decrease with each round, and she may even start making plans which include survival after several QS, and which put zero utility to the dead ends.

That is, after several survivals, subjective experience and planning will be in contrast to theoretical understanding of Mary. This is similar to the *survivor fallacy* or *gambler fallacy* by its psychological effects. However, the gambler will always be punished by his overconfidence in his luck— and Mary may become a true believer, but will ultimately be punished by what Almond called “low measure marginalization”, that is, she will survive in unexpected ways (Paul Almond, 2011d).

## 4.10 Making decisions under logical uncertainty

Not only do we not know if the multiverse theory is true, or the true nature of human identity, but we also don’t know the decision theory to apply correctly here. Our rationality is bounded, and we can’t spend an infinitely long time on discussing very complex problems.

So here I suggest a *conservative approach* to the tough problems: make decisions which don’t have extremely negative consequences in all plausible theories of reality and decision making, and do things which have a small cost in all theories, but a very high positive payoff in at least one theory.

This is close to the advice given by Taleb (Taleb, 2007) who discussed how to use black swans for profit.

Applying this conservative approach will help you not to lose much in our uncertain world, and give yourself a chance to win, so it is “winning first” (P. Greene, 2017) meta-level decision theory.

Applying this conservative decision theory to QI, we should:

* Not commit suicide in the paid quantum suicide thought experiment.
* Try to escape bad forms of quantum immortality, like eternal aging.
* Choose actions through which the probability of success is strongly increased by QI, but which are still not bad if it is not true, like cryonics
* Choose actions, which have both egoistic and altruistic value, and do not take actions which will damaging, if everybody else were to do them.

In Section 6 we will analyze this in details.

## 4.11 Merging timelines compensate the decline of measure

One of the main arguments against QI is that my measure declines while branching continues. However, as my mind state becomes simpler when I am closer to death, I am de facto merged with “other me” who also has a simpler mind now, but before, had some differences.

For example, there is me with memories (M1, M2… Mt) and another person with memories (N1, M2…Mt), so we have a different first memory. Maybe my first cup was green and his cup was yellow. If, because of the early start of Alzheimer’s, we both forget the color of our first cup, we are now the same person with memories (M2…Mt), which basically means that my measure grew two times! If we look at the observer-moments complexity, the result will be the same: the closer the death, the more simple my observer-moment, and the more people I merge with.

This simplification could compensate loss of measure because of dying. In our example, after the measure of the person (M2…Mt) doubles, one of them could die, and the measure will be again 1. Losing every 1 bit of complexity produces the doubling of measure, so the simplification process could generate dramatic jumps of measure. For example, forgetting just 1 kilobite of personal information is equal to a 10E300 jump of measure— which is enough to run QS 1000 times.

If someone really cares about his measure, he may try to forget as much information as possible in order to increase his measure. More about this idea will be said in the appendix about dust theory and flux universe.

But this seems to work only in the case of the big world immortality, where copies are not causally connected. Whether the timelines merge in the case of the multiverseinterpretation of QM, is not clear. One view is that in QM, a single event may have multiple histories (Feynman, Hibbs, & Styer, 2010). In that case, merging seems to be possible.

Pereira (2017) suggested that the fact that my current observer-moment is rather simple, means that superintelligent observer-moments are rare, and thus superintelligence is impossible. But in this section, we demonstrated that simpler observer-moments get an incredible burst of measure, and this should compensate their implausibility. For example, simple molecules are more numerous in the universe than in the mind, so if we peak at a random molecule, it is more likely to be simple.

# 5. Possible negative effects of QI

## 5.1 The idea of QI is an informational hazard

It seems that some people were motivated to object to QI based on their ethical position. QI may lead an unstable person to suicide (as it seems already happened with Everett’s daughter), so we should claim that Quantum suicide is false. However, the genie as out of the bottle, and thus we should explain that quantum immortality has nothing to do with “suicide” – this was a very incorrect name, which could lead an unstable person to think that suicide will somehow will make him immortal.

But if quantum immortality is true, any attempt at suicide will result in a failure, and some health damage, so it will only increase suffering. This should be explained to compensate for the damage of the QS idea.

Similarly, the idea of the existence of the soul may result in a much stronger temptation into suicide, as it at least promises another better world, but I have never heard that the idea was hidden for fear of resulting in suicides. Religions try to stop suicide (which is logical on their premises) by adding additional rules against it. So, QI itself is not promoting suicide, and personal instability may be the main course of suicidal ideation.

Euthanasia seems a much more dangerous idea because of multiverse-immortality, to a person who already suffers.

To be buried alive also becomes more probable because of Multiverse-immortality.

## 5.2 Impossibility to die may be bad

Big world immortality means that death is impossible. It doesn’t mean that I, for sure, will have a linear infinite life span, as timer resets and memory loss are possible (see Appendix).

But it means that euthanasia is impossible, and one can’t choose death, even if he really wants it. There is no escape from suffering and there is no way to go into “another” world (at least if we exclude more exotic ideas about multiverse-immortality in the multilevel simulation discussed below).

## 5.3 Long-term unescapable sufferings (s-risks) are possible

If death is impossible, someone could be locked into a very bad situation, where he can’t die, but also can’t become healthy again. But I don’t think it will be too long, as when the probability of survival becomes very small, strange ways of survival will dominate, for example, aliens could arrive with a cure for the illness (“low measure marginalization”, (Paul Almond, 2011d), but more likely, the suffer will find himself in a simulation, or resurrected by superintelligence in our world, perhaps because of cryonics.

Aranyosi summed the problem: “David Lewis’s point that there is a terrifying corollary to the argument, namely, that we should expect to live forever in a crippled, more and more damaged state, that barely sustains life. This is the prospect of eternal quantum torment” (Aranyosi, 2012), (Lewis, 2004).

The idea of outcomes infinitely worse than death was explored by Daniel (2017), and he called them s-risks (but for the whole of humanity). In our case, everybody will have his personal hell.

Aranyosi suggested *comforting сorollary*, based on the idea that multiverse-immortality requires not remaining in the “alive state”, but remaining in a conscious state, and thus damage to the brain should not be very high. It means, according to Aranyosi, that being in the nearest vicinity of death is less probable than being in just “the vicinity of the vicinity”: the difference is akin to the difference between constant agony and short-term health improvement.

However, we all know that very bad constant states of health which, nevertheless, don’t affect consciousness, are possible, which include cancer, whole body paralysis, Alzheimer’s, depression, blindness, and lock-in syndrome.

We argue that for 21st century people these outcomes become less probable, as medical technology development increases the number of possible futures where any decease will be cured, or where a person will be put in cryostasis, or wake up in the next level of nested simulation.

Aranyosi suggested several other reasons why eternal sufferings are less probable:

1. *Early escape from bad situation*: “According to my line of thought, you should rather expect to always luckily avoid life-threatening events in infinitely many such crossing attempts, by not being hit (too hard) by a car to begin with. That is so because according to my argument the branching of the world, relevant from the subjective perspective, takes place earlier than it does according to Lewis. According to him it takes place just before the moment of death, according to my reasoning it takes place just before the moment of losing consciousness” (Aranyosi, 2012).
2. *Limits of sufferings.* “The more damage your brain suffers, the less you are able to suffer” (Aranyosi, 2012).
3. *Inability to remember the sufferings.* “Emergence from coma or the vegetative state is followed by amnesia is not an eternal life of suffering, but rather one extremely brief moment of possibly painful self-awareness – call it the ‘Momentary Life’ scenario.” (Aranyosi, 2012).

## 5.4 Bad infinites and bad circles

Multiverse-immortality may cause one to be locked into very stable but improbable world—think about the episode “White Christmas” of the “Black Mirror” series, where a character is locked into a simulation of a room for a subjective 30 million years.

Another bad option is *circular chain of observer-moments*. Multiverse-immortality does not require that the “next” moment will be in the actual future (especially in a timeless universe, where all moments are equally actual). Thus ground-hog day becomes possible.

It could be very short, like several seconds, where dying consciousness repeatedly comes to the same state as several seconds ago, and as it doesn’t have any future moments it resets to the last similar moment. Surely it could happen only in a very narrow state of consciousness, where internal clock and memory are damaged.

## 5.5 Marginalisation of measure, strange worlds and infinite torture

Alomd suggests to the term “marginalisation of measure” to the situation where improbable survival requires existence of very improbable worlds.

Some very improbable worlds may be extremely bad. For example, a world where future AI decided to torture human beings eternally has a priori low probability, but it could be one of the most obvious ways to survive almost eternally in a human form. This is one of the types of s-risks connected with superintelligence (Sotala & Gloor, 2017).

## 5.6 Dangers of mistakenly ignoring QI

These several extremely bad outcomes could become real possibilities for any observer, if QI is true. That is why trying to dismiss QI as a weird probabilistic trick may have almost infinite negative cost.

However, we could use properties of QI to replace bad outcomes for good outcomes, in the case of currently living people.

## 5.7 Dangers of QI-euphoria

A person who strongly emotionally believes in QI may choose more risky actions, such as not fastening seat belts in cars. But the same argument is even more applicable to a religious person who believes in fate. Fatalism is associated with higher level of accidents (Şimşekoğlu et al., 2013), and the highest world level is in Iran.

# 6. Possible practical applications of the multiverse immortality

## 6.1 Surviving until life extension technologies

Multiverse-immortality is just an observation selection effect which does not have any positive or negative value per se.

It may help to escape death, but not injuries, and thus it means very long aging, and continuous sufferings in the “ordinary” 21st century world. As we assumed above, MWI-immortality has no more than an even chance of working, and even if it works, it will result in infinitely long aging full of suffering. Thus, there is nothing good in it.

But we could still use multiverse-immortality if we see it as a probability shift in the required direction, and do not put ourselves at needless risk by attempting some “suicide experiments”.

For example, as we live in the 21st century, a sufficiently long aging process will help us to survive until the appearance of powerful life extension technologies. These technologies will produce rejuvenation, and will help us to live almost forever, without help of multiverse-immortality. The needed additional survival time may be just several decades, and could require a person not dying at 50 years old, but surviving until 90. This rather small life expectancy shift, which has probability of around 10 percent could be not observed as something surprising by the person.

By trying to live as long as possible, we use the additional probability shift provided by multiverse-immortality, but even if it doesn’t exist, we still get useful results, as we are trying to do everything possible for life extension.

Multiverse-immortality also implies that it is better to choose medical treatments similar to Russian roulette, for example, surgery over chemo, as an immediate death outcome does not exist, but slow dying still exists.

## 6.2 Multiverse-immortality significantly increases the chances of the success of cryonics

Multiverse-immortality significantly increases the chances of cryonics. If multiverse-immortality is true, the biggest share of timelines where I survive until 2100 all include cryopreservation, as was mentioned by (Randall, 2004).

For example, if, in the normal world, cryopreservation’s chances are 0.1 per cent, and my chances to live until the year 2100 naturally is 1 in a million (here we exclude the effects of new life extension technologies described above, and only look at typical human life expectancy distribution), it means that I have 1000 times more chances to survive to 2100, because of cryonics, compared with my personal longevity.

As multiverse-immortality means that I will survive to 2100 anyway, it implies that my chances of cryopreservation success grow to 50 percent (which is our estimation that the multiverse will work), or 500 times. In other words, no matter how small cryonics chances, multiverse-immortality increases them up to the level of its own success probability.

You just have to sign up for cryonics, and multiverse will do the rest. However, if you want that your friends to survive too, you will still have to invest in the quality and probability of the success of cryotechnologies.

Signing up for cryonics replaces the default outcome of the “bad immortality” of infinitely long aging, with a good “resurrection” outcome, only in the worlds where you needed it, and where all life extension technologies exist.

Signing up for cryonics is also good from the UDT decision point of view, as it means that other people, who have similar lines of thought as you, will also sign up. And if more people sign up, cryocompanies will have more money for research in improving technology and lowering its price.

## 6.3 The impossibility of euthanasia and the need for cryothanasia

Euthanasia assumes that the voluntary death of the terminally ill patient will have two practical results: the cessation of personal suffering, and alleviation of the burden on the relatives. The first one seems to be the most important reason for the euthanasia, but if QI is true, the patient in his subjective timeline will only experience the failure of the procedure, which most likely means that his suffering will continue, and will be even worse, as he will have the traumatic experience of something akin to regaining consciousness in a morgue, and starting to feeling the pain again without the help of painkillers (Plus whatever damage may have been done by the euthanasia attempt method). Also, his hopes of stopping his suffering will be ruined, and he will witness the shock of his relatives.

However, if one chooses a combination of euthanasia and cryopreservation, called cryothanasia (Istvan, 2014), (Minerva & Sandberg, 2017), his most probable line of survival will be regaining consciousness in the future, where he will be cured from his illness.

Surprisingly, this is not a popular idea. One of the reasons is that cryonics is mostly legal in US, yet euthanasia is legal only in Europe, so combining the two procedures looks legally difficult.

## 6.4 Multiverse-immortality increases the chances of the success of digital immortality, and of the acausal trade between branches

Similar logic to that of cryopreservation is applicable to *digital immortality*, that is, the idea that I will be reconstructed by future AI, based on my information traces. Digital immortality will also become a very likely outcome from all possible ways of my survival, thanks to multiverse immortality. It could be an even more likely outcome than cryonics, as a lot of people could be reconstructed, even if they didn’t sign up for cryonics.

However, there is a problem of information loss. One possible solution is to replace the lost pieces of information with random information (P. Almond, 2006). In some universes, this random data will be equal to my data, so an effect similar to multiverse-immortality will be used to reconstruct exactly me. However, what to do with copies of me, which are not exactly similar?

In fact, this is not a big problem, as these not-similar minds will be exact resurrections of minds from different universes, and these universes in turn will resurrect me. So, the total measure of all resurrections will be the same for me, and as high as if I were to be resurrected without the use of random noise.

For example, the information of my apartment number is lost and can’t be recovered or reconstructed. Future AI creates a copy of me and generates a quantum random number X as my apartment number, and puts in my mind. Let’s say it is 73. In another universe, another AI does the same, and generates X=68. As a result, both me(68) and me(73) are exactly resurrected, but in separate universes. Me(68) lives in the universe where the actual number of the apartment was 73, and me(73) lives in the 68-universe. Thus, both branches help each other, and don’t need to run many copies, and the total “measure” of each copy is not declining. This could be called “acausal trade” between branches.

Some have suggested (P. Almond, 2006) that completely random code could be created once only, and that will be enough to resurrect all humans, as for every possible human, there will be a branch of the quantum multiverse. However, any resurrection will happen in some context, for example, some relatives will be alive, and they have to recognize the reconstructed person as their kin. In the case of partial random noise filling, information from relatives is already taken into account. Also, it means over relating on untestable multiverse-immortality.

In the case of only filling the small gaps with random noise, there will be no over-relation on the multiverse-immortality, as it will be only an additive help.

## 6.5 “Universal problem solver” as an example of a bad idea

Multiverse-immortality could be used to create “universal problem solver”. Imagine a bomb, connected to a computer, which will kill me only if a certain condition is not met. In that case, I will survive only in the worlds where the condition happens, and the system could work as a “universal problem solver”.

However, this is exactly the wrong application of the multiverse-immortality, as the bomb might not work, or might work partly, and only injure me.

But the main reason is that if this practice becomes widespread, I (and everybody else) will soon find myself in an empty world, where everybody kills themselves for some reason, and it will be similar to a global catastrophe and human extinction. It is unlikely that I will be first and only one who uses the idea, so the practice will be catastrophic for my world, and any personal gains will be outweighed by the damage.

## 6.6 Collective survival from a global catastrophe

There are several hypothetical situations where the death of a person is connected by high probability causal connection with the death of his group of people.

For example, the collapse of a submerged submarine will probably kill everybody on board. In that case, my personal survival will mean that the submarine collapse will not happen, and it will appear that multiverse immortality helps the survivability of the ship. I explored nuclear submarines as possible survival refuges in the case of a global catastrophe, and this could help increase their survivability (Turchin & Green, 2017). The same is true for a space ship.

If I sit near the potential target of a nuclear strike, it also could lessen the observed probability of nuclear war, so this could be used as a protection against other global risks.

Some have suggested that the fact that we have survived without nuclear war could be explained by QI. However, it is unlikely that WW3 will result in guaranteed human extinction, so there most likely will be some survivors, and this explanation doesn’t work.

Some even suggested that CERN’s collider could end the world via some physical experiment disaster, but we survive by luck only in the timelines where LHC has some technical difficulties—see (Sandberg, 2008) which also has some nice Bayesian calculations.

Even if a global catastrophe (like a pandemic or asteroid impact) happens, multiverse-immortality ensures that I will be the only survivor. Thus, it will be not technically human extinction, but it is unlikely I will be able to restart the civilization.

## 6.7 Civilizational level QS as a universal problem solver, or an instrument for calculations

Almond (P. Almond, 2008) suggested that alien civilization could go through ages of collective QS, which they could use for problem solving, or even for complex calculations. It could be a solution for the Fermi paradox, Doomsday argument and simulation argument.

I think that it is unlikely that *all* civilization will choose this path, as it is clearly a very risky idea, thus it is not as good an explanation of the Fermi paradox as it is for the DA and the SA.

## 6.8 Multiverse-immortality in the age of superintelligence: Will I become AI?

As I said above, if superintelligence appears, it will able to reconstruct me based on my traces. It selects the worlds where AI will be interested in such reconstruction, hopefully, because it will be benevolent.

If we look at longer timelines, multiverse-immortality implies that I will exist for billions and trillions of years. The most likely way this could happen is that I will become AI, or will somehow integrate my personality with it.

## 6.9 Multiverse-immortality for computer algorithms

Yudkowsky wrote “The Anthropic Trilemma” <http://lesswrong.com/lw/19d/the_anthropic_trilemma/>, about manipulation of observed probabilities by creation of large numbers of virtual copies, and manipulation of this number. But any computer program could exist in many copies, and the Internet is a big world for it.

## 6.10 QI and Unfriendly AI

If AI uses QI to reach some of its goals, it will disappear from our world.

If AI decided to kill all humans, we could suggest some form of deal (Turchin, 2017), where AI kills all humanity instantly with 0.99 probability, and completely preserve us, and even plays benevolent AI with a 0.01 probability. In that case, AI gets 0.99 of its expected utility of killing humans, like their atoms, but humanity enjoys collective quantum immortality (P. Almond, 2008) and gains complete survival. As a result, values of both entities will be satisfied.

# Conclusion

In this article, we explored one of the weird probabilistic arguments connected with our place in the big world, that is, so called “quantum immortality,” and we found that it has some probability of being valid, not enough to be used for sure, but enough to be afraid of its bad consequences or to use at as possible multiplier of some already existent good processes like cryonics.

# Appendix 1. Multiverse immortality and probabilities

The questions about probability is notoriously difficult as they always come to the question “what is probability”, especially in the situation with multiverse and many observer’s copies. See great short overview of problem by Wei Dai: <http://lesswrong.com/lw/1iy/what_are_probabilities_anyway/>

My view on the “probabilities” is that it is incorrect to ask “what are they”, as they are not real objects, but instruments which we constructed to predict the world and act in it. So, we can’t pretend that they actually exist.

As “probabilities” are prediction instruments sometimes we could use different prediction instruments, for example, expected utility, which is combination of the probability and our preferences.

It is not easy to discuss QI without applying probabilities to it, and here we explore some probability related questions.

Many authors prove or attack QI exactly based on some probabilistic calculations, like (Mallah, 2009), and (P. Almond, 2011).

## Does survival in the quantum suicide experiment present an evidence for MW-interpretation of QM?

The main problem arising around Multiverse-immortality is how to define probabilities in the situation where many copies of observer exist. One of the solutions is idea that the probabilities are not important per se but only after their multiplication on expected utility of action.

Defining probabilities in the case where many copies exist is known to be notoriously difficult. Some suggested that calculating probabilities is useless intermediate step and we could get better results if we calculate expected utilities.

Tegmark (Max Tegmark, 1998) suggested that quantum suicide experiment could be used to test many worlds interpretation of quantum mechanics. Almond (P. Almond, 2011) and others invest a lot to demonstrate that this will not work.

But survival in the quantum suicide experiment could happen because of the multiverse immortality, not quantum immortality, that is just because level 1 and level 2 in Tegmarks’ multiverse. Thus, even if such survival is the evidence for immortality, it can’t be used to distinguish what kind of multiverse provided needed my copies, so it can’t be used as a proof of Multiverse interpretation.

However, the real question is – could be immortality be tested via set of experiments at all?

## Does unexpected survival present an evidence for QI?

If I survive plane crash, is it an evidence that I will survive the next plane crash?

Mallah thinks that it is not: “Suppose there are 10 billion people, and 200 of them decide to try QS, so about 100 of those survive. The effective probability of a person being any one of those QS survivors is about 100 in 10 billion. This is true in either the single-world or MWI case, so seeing that you are a QS survivor does not provide evidence either for or against the MWI”. Almond share the similar view.

It is similar to the Doomsday argument where the fact that I am surprisingly early in the human history is taken as an evidence for short live expectancy of the human civilization – and this idea is often rejected on the same grounds, as the Mallah did for Multiverse-immortality, that where will be always be a small group of people who is very biased because of their very rare position in time.

However, it is also well known that it is not easy to kill Doomsday Argument, as it requires to choose one of two self-locations assumptions (self-indication or self-sampling) and choose a side in the Sleeping Beauty problem. We will not try to solve the problem here, as there is immense literature on the topic, but we just will tell that we can’t be sure. (In short, moving from 3-perspective into being inside the situation gives me additional information – more Bostrom’s book (Nick Bostrom, 2013)).

This position I call Meta-Doomsday argument, which gives credence 0.5 to both DA solutions. However, as DA is itself a probabilistic argument, it produces multiplication of of our credence on its prediction, which keeps the prediction almost untouched!

In our case of MWI immortality, it means that we should take unexpected survival evidence with 0.5 discount. That is, as a priory probability in Mallah’s example is also 0.5, total credence to the Multiverse-immortality in case of unexpected very low probability survival should be 0.5x0.5=0.75.

Almond invested a lot in proving that survival in QS is not evidence of MWI (P. Almond, 2011), and thus not evidence that survivals will continue, so that MW-immortality will continue to work, the same way as it happens in gambler fallacy where series of wins is not evidence that wins will continue. However, it is still the evidence that the coin is biased.

It may be also interested to note that this discussion is similar to Presumptuous Philosopher paradox (Nick Bostrom & Cirković, 2003). However, all such arguments require very large, but finite universe. In the infinite universe where every possible observer exists, its existence is not evidence for anything.

We could use another thought experiment. Imagine that someone is playing in coin tossing and get 5 heads in a row. Is it an evidence of his magic powers? No, as 5 heads in row is rather often event with probability 1 in 32. Imagine, however, that someone got 100 heads in a row. This event has probability around 10E-30 and the most probable its explanation is that the coin is strongly biased. The same way if I survive until 100 years old, it is not evidence of any type of immortality, as there are thousands of people of this age. However, if I survive until 150 years old (without any technological interventions) – it will be surprising even from the point of the outside viewer as such random event has probability 10E-30 to happen on Earth (given that survival probability halves every 1 years after 100), so it would imply existence 10E30 “other earths”.

The question, “Does survival until 100 years present a Bayesian evidence for quantum immortality”, is more complicated. From the outsider’s point of view, it is not surprising, as there will be many such survivors. But for me, becoming such survivor has probability around 1 in 1000, so I will find this surprising and may try to explain via quantum immortality.

Almond wrote: “If MWI is correct, Mary should be just as surprised to find herself in such an observer moment [after 1000 QS experiments with probability 10E-300], out of all the observer moments she could be in, as she would be to find that she had survived a long sequence of experiments if MWI were incorrect. The situation is no more or less implausible with MWI being true than it was without MWI being true. Any implausibility associated with an unlikely sequence of events happening has merely been replaced by implausibility associated with being located in a low measure part of the Everett multiverse” (Paul Almond, 2011d).

 The logic of Almond here is flawed as he assumed that quantum immortality doesn’t work, but this is exactly what he is trying to prove, so he uses conclusion as premises and gets circular logic.

By the way, there is no need to repeat this experiment, as nature already did it in the form of anthropic principle. The fact that our universe looks like fine-tuned for human existence implies existence of 10E500 different universes with different random parameters, according to explanation of the anthropic principle via string theory landscape. If we agree that our existence is an evidence for very large universe, we basically use the same logic as if we will try to prove QI vie unexpected survival.

## Uncertainty in multiverse immortality theory

There are two lines of uncertainty theoretical – we don’t know will it work, and practical – we don’t know which outcome is the most probable – may be injuries or time reset is more probable.

## The problem of measure

“Measure” is rather obscure term used to describe quantity of existence. It is roughly equal to the number of similar copies, but takes into account quantum nature of them and is measured by the amplitude of probability.

My measure in case of Multiverse-immortality is declining. There are two views on it. One is that declining measure is bad, as it means less existence. However, it is not clear why exactly it is bad. My measure compared to all other objects in the universe is already very small, and it is not clear what is bad if it becomes even smaller, but I can’t observe it. (If we exclude ideas of “quantum magic” where I could trade my measure on ability to affect observed probabilities by some form of orchestrated reduction, which is unsurprisingly similar to controlled quantum suicide). More about trading measure and probabilities in “The anthropic trilemma” by Yudkowsly (Yudkowsky, 2009b).

Another view is that there is no problem in measure, until at least one my copy exists, in the same way, as there is no problem how many copies of a novel exist if at least one survives.

The possible solution resolving contradiction of these two views is the idea of *relative measure.* It is the number of observer-moments O(I,t) in the situation X compared to all observer-moments O(I,t). That is, it is the proportion of me-tomorrow who will win in a lottery compared with the number of all me-tomorrow observer-moments. In the QS experiment with lottery the total measure is declining, the measure of winnings is constant, and the relative measure of wins is growing.

Measure is similar to the “fluid of reality” and to “actuality” which was mentioned above.

Almond argue that measure is needed – or we should experience not typical but bizarre observer-moment, as they are more numerous (but see our discussion about Boltzmann brains and their expected randomness). Measure is quantified amount of “reality fluid” or actuality. If we say that measure of something is low, we mean that it exists in less extent then we.

If we accept the “measure view” on survival, a person, who survived a potentially deadly accident should be still unhappy as he just lost half of his measure. There is no way to feel that you lost part of your “measure”.

Almond states that if there is no measure, and all observer moments counts by numbers, I will find most likely myself in the end of my life – literary in last second. But this assumes that QI is not true. If it is true, for any “last moment” there is the next moment.

New way to calculate the measure: it is as shares of all worlds where the moment O(I, T+1) exist: for example, if 1 such moment exist in real world and 9 in simulation, the real one has 0.1 measure. So, measure is relative, not absolute.

Also, if one values having higher measure, he must strive to creating as much his copies as possible, as it will increase measure.

## Difference in the probabilities of outcomes in the multiverse and quantum immortality

One interesting thing is that in the multiverse immortality the number of “very rare survival” observers could be higher than in the quantum immortality. For example, you are falling on Sun. The probability that some quantum effects temporary change gravity and move you away is extremely small, like the number with millions zeros. However, in the other part of the big universe there are another copies of you, which have an illusion that they are falling on Sun, but actually not. In that case big immortality provides more probable way of survival.

## The nature of physical experiment

Classical idea of the physical experiment presumes existence of the pure objective observer. If someone decline to accept Multiverse-immortality, he has to rule out principle role of the observer.

If we said that unexpected survival is not the argument for Multiverse-immortality as did Almond, we create too strong argument that kills power of evidences.

If all possible observer-moments exist, any observation is not evidence for anything. If we measure g which is expected to be g=10, where will be always an observer who find that g=7, and thus finding that g=7 is not evidence for g=10. Or, we have to said, that probability of me-now to be a given observer is somehow depending on the measure of the observers in the universe, as if observer-moments are boxes and my consciousness is a ball which fall in the box.

## QI and presumptuous philosopher

We don’t know if multiverse immortality works. But if I will survive several severe accidents or will live until very long age, I may take it as a Bayesian evidence for the theory, but it is still questionable, see presumptuous philosopher (PP).

In our case, the fact that I exist according to PP-logic is an evidence to the existence of all possible observers and infinite universe.

## Marginalization: Strange outcomes will dominate after several QS experiments

Almond called “Low Measure Marginalization”, that is, survival in an unexpected ways (Paul Almond, 2011d).

There are several possible types of marginalization:

* Saved by aliens,
* In a dream,
* In a short-term hallucination
* Data stamping mistake
* In a simulation with afterlife
* As a Boltzmann brain
* An error in estimation of risk or in firing of QS weapon
* A jump to past observer-moment
* Respawning

Marginalization means appearing in less probable worlds which are also less likely to be ordered. It has surprising consequence that QI is more likely to work in them! Because the size of set O(I, T+1) is not much smaller than O(I, T) set. That is, “measure” is declining slowly after marginalization. The reason for it is that in the strange non-ordered world all possible observer-moments are more or less equally frequent, so frequency is not declining.

For example, in a bizarre dream you are much more likely to be saved by “aliens”.

# Appendix 2. Multiverse immortality and other similar probabilistic arguments

## Multiverse-immortality and the anthropic principle

I also think that it is nothing extraordinary in Multiverse-immortality idea, and it adds up to normality (in immediate surroundings). We all already witness to the examples of the similar ideas. That is the anthropic principle and the fact that we found ourselves on habitable planet while most planets are dead. Similarly, the fact that I was born, but not my billions potential siblings. Survivorship bias could explain finding oneself in very improbable conditions and QI is the same idea projected in the future.

In a sense, anthropic principle is the same idea, but turned to the past.

Almond argues against such interpretation: “It should be noted that this does not apply to anthropic arguments that suggest that it should have been unlikely for the universe to develop in a form that allowed life, or that the universe should have been too unstable to allow life, and that MWI has dealt with this in the past by ensuring that a path was available through the multiverse that allowed our universe to become a stable one, suitable for life, before conscious observers existed. Without observers, or observer moments in the universe before the measure reduction occurs, there is no reason not to think that observer moments after the measure reduction are places in which observers should think it quite reasonable to find themselves. To put this another way, the thought experiment about waking on Monday or Tuesday, in 2.1, is irrelevant if you do not have any observer moments appearing before Tuesday” (Paul Almond, 2011d).

## Multiverse-immortality and simulation argument

Simulation argument contribute powerfully to the big world idea and if we buy it, most my copies (and me now) are in some forms of simulations which are created by advanced civilizations for whatever reason. The number of my copies in simulations is much larger than expected number of my copies in any other big world model.

If my simulation is turned off or I die in it, where will be always another simulation where my copy is still running, so soon I will find myself in the simulations which are designed to preserve life of characters.

If I am not in a simulation, the simulation probability (that is I am in a simulation) increases after number of “deaths”, because in a simulation even weirdest ways of survival are possible. If I fail in 1 km abyss I will likely to survive only in the simulation where miracle ability to fly is turned off – or I will wake up from the dream (as dreams are in fact most widespread type of simulations).

However, dying is also increasing the probability that I will “wake up” on a higher level of the simulation, as if I was a player on computer game. Surely, it is possible that our world is a computer game for a higher-level character, the same way as we play in RPG (or have dreams). So, after waking up I will “remember” who I actually am, but there are many different variants of it (see subsection below): the same game-character could be played by different players. These players could be played by even higher level players in the nested simulation, so the idea of multiverse immortality stacked well with it (Torres, 2014).

There is probably also a share of simulations which simulate “afterlife” according to human believes about it. This share will also increase after a number of my “deaths” as I will be able to find myself only in the simulations where there is afterlife, miracle survival or awakening, and will not find myself into simulations which simply shutdown or terminate my file.

The phenomenon of respawning often reported by people (where they find themselves before a deadly accident and live this moment the second time without dying) may be also explained by the fact we are in the simulation.

## Multiverse-immortality and Doomsday argument

We already discussed connection of DA and Multiverse-immortality above.

 As my “measure” decreases, it is less likely that I will find myself in the later stages of “immortal timeline” if I think that current moment is random.

Almond on DA and QI: “The Doomsday argument, if correct, suggests that our civilization will end soon, based on the idea that our own position in time should not be too unusual, but in an MWI context, the Doomsday argument could be interpreted as meaning that our civilization will have a huge and ongoing decrease of measure in the future: that our civilization will continue in some futures, but that its measure will decrease all the time as something like the quantum suicide experiment – whether contrived by our descendants or imposed on them – occurs on a continual basis. In such a situation, observers should expect to find themselves in relatively high measure parts of the civilization’s history – maybe where we are now” (Paul Almond, 2011d).

## Modal realism and back causation

If everything possible exists, I should exist. However, if I exist, there should exists most probable way of my appearing, and Darwinian evolution seems to be the one. This need to create “me” create something resembling back-causation pressure on the past events, the same way as my father and mother has to meet no matter how difficult it was.

## Multiverse-immortality resetting and infinite timelines in the space of possible minds

Another possible type or resetting could happen in the case of “slow death”. In that case, my state of mind will be simpler and I will forget part of my personal identity as well as my timestamp. This partial simpler observer-moment could be equal to many more observer-moments of larger groups of people, which also experience these narrow states on consciousness.

This could produce something similar to “reincarnation”. For example, I forget my surname and my age. Now I am just Alexey and my state of consciousness could be equal to a state of any child with the same name.

However, it creates frightening possibility that every night when my state of mind narrows, I could become everyone else and randomly jump inside any person with similar narrow state of mind. One guy on LessWrong had panic attacks because of it.

## The same “me” in different worlds and superposition of interpretations

*Universal reference class* by Almond: “In previous articles, the author gave the term ultimate reference class to the set of all possible descriptions of your situation – all possible situations in which you could be right now” (P. Almond, 2011). *Ultimate reference class* includes all possible situations which creates given observer-moment, and includes all the same observer-moments in these situation.

It means that questions about actual nature of reality is useless, as the same observer-moment could be Boltzmann-brain, in different simulations and in different real worlds, and even belong to different peoples.

The next step in this logic is to suggest that these different interpretations of the same observed reality do not just linearly sum up, but could interfere with each other in the similar way as different paths of a quantum particle interfere. In that case we will find ourselves in the reality which resists attempts to any final interpretation and this will fundamentally limit our ability to predict future outcomes.

## QI and the reference class problem

Almond suggested that QS is not working as an argument for MWI (Paul Almond, 2011d), as more observer-moment will be before QS test. However, there are two reference classes: O(I, T) and O(I,T+1). If I know that I am in O(I,T+1) I know that I already survive QS.

## Multiverse-immortality and qualia

Observer-moments consist of *atoms of experience* that is qualia. That is qualia is minimal living parts of consciousness. In the moment of death consciousness dissipate, but individual qualia can’t cease to exist, because *green* can’t stop to be *green*. Simple qualia could exist in may be in nature, that is like dust universe, or in other people’s minds. So, qualia are themselves immortal, but also don’t have anything personal.

# Appendix 3. Next step of QI theory: Dust theory and Boltzmann brains

## Multiverse-immortality and Boltzmann brains

There is a dangerous thought – what if I just a Boltzmann brain? There is popular objection to this idea, that most of BBs have very random experiences, but I am not, so I am not a BB. However, it doesn’t work, for two reasons.

First is similar to the reasons why QS experiment is not evidence of MWI: that is, non-random BBs also exists and the fact that I am nonrandom one is not the evidence for that I am not BB. It is also similar to the Sleeping beauty, which is told that she is nonrandom mind and has to estimate the probability that she is BB. For one-theirder solution of the Sleeping beauty, she can’t update its estimation to be a BB based on non-randomness of her environment.

Another reason that randomness argument doesn’t work is that a BB is not able to make coherent conclusions about its own randomness. So, BB could think that he has non-random experiences, but still have random experiences, as most BBs are not logical.

So where is no way for BB to prove that he is not BB, and it seems to bad as he will disappear after a microsecond. However, Multiverse-immortality gives a solution. For any BB-observer-moment exists another next observer-moment in the real life, and if we agree with Multiverse-immortality, my mind will jump from BB to the real me.

It is not hard to see that “real me” is not necessary state here, as the next observer-moment could be also BB (lets call it BB2), if it has memories about last moment equal to BB1 and the continuous line of experiences.

This idea is known as flux-universe and will be discussed in the next subsection.

"What we can do, however, is recognize that it’s no way to go through life. The data that an observer just like us has access to includes not only our physical environment, but all of the (purported) memories and knowledge in our brains. In a randomly-fluctuating scenario, there’s no reason for this “knowledge” to have any correlation whatsoever with the world outside our immediate sensory reach. In particular, it’s overwhelmingly likely that everything we think we know about the laws of physics, and the cosmological model we have constructed that predicts we are likely to be random fluctuations, has randomly fluctuated into our heads. There is certainly no reason to trust that our knowledge is accurate, or that we have correctly deduced the predictions of this cosmological model.” <https://arxiv.org/pdf/1702.00850.pdf>

EY: <http://lesswrong.com/lw/17d/forcing_anthropics_boltzmann_brains/> and my comments

"Why Boltzmann Brains Are Bad" by Sean M. Carroll <https://arxiv.org/pdf/1702.00850.pdf>

Two excepts: " The data that an observer just like us has access to includes not only our physical environment, but all of the (purported) memories and knowledge in our brains. In a randomly-fluctuating scenario, there’s no reason for this “knowledge” to have any correlation whatsoever with the world outside our immediate sensory reach. In particular, it’s overwhelmingly likely that everything we think we know about the laws of physics, and the cosmological model we have constructed that predicts we are likely to be random fluctuations, has randomly fluctuated into our heads. There is certainly no reason to trust that our knowledge is accurate, or that we have correctly deduced the predictions of this cosmological model.” - my thought in <https://arxiv.org/pdf/1702.00850.pdf>

"If we discover that a certain otherwise innocuous cosmological model doesn’t allow us to have a reasonable degree of confidence in science and the empirical method, it makes sense to reject that model, if only on pragmatic grounds”

My opinion: I agree with idea that BB can’t know is he BB or not, and wrote about it on LessWrong, but it is not the basis to conclude that BB-theory has zero probability. We can’t put zero probability to theories if we don’t like them, because it is great way to start to ignore any cognitive biases.

My position: There is no problem to be BB:

1) If nothing else exist, different BB states are connected with each other like digits in natural set, and this way of their connection create almost normal world, and it may have some testable predictions. (Dust theory)

2) If special type of BB, called BB-AIs exist and dominate landscape, such BB-AIs create simulations which are full of human minds, so we are probably in one of them. (The idea is that superintelligent computers are more probable than messy human minds and so are more often type of BB; Or if any BB-AI create more human simulations than random BB appear)

3) If real world exist and BB exist, each BB correspond to some state in real world. As any observer should think as of all sets of similar observers under UDT, it means that I can’t be BB, but I am number of BB plus some real me. And I could ignore BB-part of me, because some form of “quantum immortality”, every second transfer dead BBs into the “real me”. In short: “Big world immortality” completely neutralise BB problem.

##  Multiverse-immortality and Dust-theory universe

The idea of dust universe comes from the novel of Greg Egan *Permutation city*. The idea is that if the physical world (including consciousness) is just a succession of states, then why would it matter that these states occur sequentially in time and at the same place. If any state of mind is just some information, that is sequences of numbers, the same sequences of numbers could appear as random sequences of atoms inside physical bodies.

The one difference of state of minds in dust theory with BBs is that they are not vacuum fluctuations but a random information sequences inside actually existing physical bodies. They could even include the same causal connections, as in thoughts in human mind, as atoms hit each other because of thermal movement, and these hits create different causal structures, which could be similar to causal stricture of neuronal exchange in mind or to calculation inside a computer.

In other words, complex random calculation in nature could be exactly the same as calculations which create human experiences, and some observer-moments could exist inside a stone in your garden.

This difference – quantum fluctuations create random minds or random thermal processes create random observer-moments – is not important, as the result is the same. What is really interesting in the dust theory, is the idea that close to each other observer-moments could produce subjectively continuous chain of experiences without having any causal connection with each other, which is exactly the same way as multiverse immortality works. Multiverse-immortality connects each state of consciousness with the next one in the dust theory.

If the dust theory is enough to explain the existence of my line of consciousness, it means that existence of any actual universe is not needed, or at least principally unknowable from the inside such line. Thus, we don’t know what kind of the “dusts speckles” creates needed combinations. However, pure “mathematical dust” is possible, if each state of consciousness is just a random number or is hidden inside Pi number.

This may have some observable consequences, if “primordial randomness” will able to break inside the stream of the subsequent states of minds. But this speculation is outside our topic.

## Dust theory and Flux-universe

A poster on LessWrong expressed strong emotional concern about the consequence of the dust theory which he called *flux-universe*. Every night, he goes to sleep and his consciousness “simplifies”, that is he has much simpler observer-moments, during which he forgets most information about his life. Many other people also could have the same simple observer-moments, and, if dust theory is true, only current observer-moment defines next observer-moment – thus if he forgets all his life and everything he loves – he could “jump” into another observer-moment of another person and never return to his original self! This thought terrified him, as it was like death of his original personality and he was afraid to sleep.

I suggested several counterarguments:

1. If he “jumps” into another person after infinitely many jumps he will eventually return back and will not remember anything, so he is not loosing his life forever.
2. It is better to think about the same situation from the point of view that he has many “futures”
3. *It also not a bug, but a feature,* as it creates possibility to perform some form of “*magic*”. If a person is in bad but very rare situation, for example, he had just lost his win in negative lottery, he could put his mind in very simple state, may be meditation or full concentrating on very common activity – so simple, that he completely forgets everything about bad luck event. In that case, much larger number of observers, who didn’t have this bad lack event, will also meditate in the same way, and he will jump into them and and then randomly become one of them. And this is exactly the same situation he feared about, but now used as instrument to change observed reality. (Surely, this magic, even if it works, fails altruistic test of updateless decision theory, as someone else will be on his place.) Also, if it works, much more complex magic is possible, in which I consciously manipulate my uncertainties about some aspect of the observed world in order to get some control over it. In a long run, it seems to be zero sum game, as simple observer states will be exploitable limited resource by those who want to escape bad life-lines.
4. Another reason, why we should not be afraid that dust theory throws us in a kaleidoscopic *flux universe* is that if lines of similar observer-moments exist in the space of all possible minds, such lines will probably come to some “strange attractors”, where some form of stabilization mechanism exists. Stabilization mechanism means that observer-moments rather coherently follows one another. There are several candidates for such mechanism: a) nature of human dreams means that we never have really simple *blank* observer moments, we start to have some form of random thoughts and images (hypnogogia) as soon as we become sleepy, and they are very individual b) human observer-moments consist of qualia as elements and qualia could be very individual: I can’t jump into a person with different red color if I see now my red color. c) we constantly have perceptive memory of last moment. d) even more weird stabilization mechanisms are possible (like Man in Black), or *counterglitches*, which glitch in the way to make the world more coherent. These stabilization mechanisms could fight expected randomness of Boltzmann brains (discussed above) – but not completely, and some strange things like *glitches in the matrix* (but there are no matrix) could be observed.

One reason why all this is related to the Multiverse-immortality is that dust theory is about constant jumping between “worlds” and copies, and in the Multiverse-immortality it happens only in the moment of death.

## Miracles and glitches in the observer-moments chains

(This section does not belong to this article per se, but I leave it here for a while. I do not believe in the following, but it has high enough a priory probability to be worse be discussed)

If we suggest that human life is just as a line of similar observer-moments, like 4567, 4568, 4569, it has the following consequences:

1. For any random observer-moment there is another observer-moment, which follows from it, and makes it less random. If you see static, the next observer-moment will be just you looking on TV screen.
2. If from A observer moment follows B moment, and from B follows C moment, it doesn’t mean that from A follows C, as only near observer-moments must be similar, but A may be not coherent to C. So, observer-moments have not transitive relation. The consequences of it is fluidity of the world, when we compare A and C – we could find some disagreements.
3. For observer-moments chains there is no time, and past observer-moments could be somehow similar to the future moments, which could take form of “precognition”
4. Here should be some connection with quantum theory, that is any observer-moment may result from infinitely many different paths of past global wave function, or in other words, any observer-moment is surrounded by uncollapsed world. But is seems to be different level of description of the same phenomena.

# Appendix 4. Other objections to QI and wrong interpretations

## Respawning and quantum woo

In the internet, the term of “quantum immortality” is extrapolated on the other type of events which veracity can’t be checked. It is the stories typically about automotive accident, where a person has a crash, almost die, and when find himself 1-2 minutes before the event in the car and escape an accident. Typical example is here: <https://www.reddit.com/r/Glitch_in_the_Matrix/comments/7frtxt/i_literally_diedor_so_i_thought/>

The people in comments call the story a “quantum immortality event”, but this is wrong use of term, no matter if the story is fictional or not. I suggested to call such types of stories “respawning”.

This mix of terms results from general woo regarding quantum things, so anything miracle will be easily called “quantum”.

## Aging and slow death

Tegmark suggested that the main counterargument for multiverse immortality is that it is not applicable to the situation of slow death, like death of aging.

In these situations, I slowly lose most of my identity which is based on memories. At the end rather simple mind remains, but it is not me. This mind maybe still able to suffer, but it lost most of my valuable traits.

One outcome of it is that I will become equal to so many other simple minds that I could become almost anyone else, and it is similar to “reincarnation” – we will discuss it later.

Here I explore another outcome, based on the situation where I will regain most of my memories.

Imagine a situation that I am gradually losing memories (and memories is all that is needed for personal identity). For example, my memories consisted of 10 numbers, like 1234543210, but now I remember only first 3, that is 123. However, there is still some possible future timelines where I will regain my memories, where my brain function improves, and I will again remember 1234543210. In that case, I could percept all the story as jumping from the last moment where I had full memories to the first moment when I again have all full memories, similar to the situation where will lost our memories during night sleep but without problems regain them in the morning and often feel it as if we jumped from the evening to the morning, and forgetting almost all period of narrow consciousness during night dream.

It means that if our identity-memory will fall below some threshold, we will jump over period of narrow consciousness to the next observer-state where we will regain our memory, and thanks to the multiverse immortality where will be always such next me, who regain the memories.

Almond suggested “Quantum Brain Damage”: “A more general version of the thought experiment, the quantum brain damage thought experiment, shows problems with such a simplistic view. In this thought experiment, the quantum outcome determines a degree of damage to your brain, ranging from no damage at all to complete destruction, with many intermediate degrees of damage”. In that case arises the problem of personal identity and death definition, as it is not clear how to draw the line between “me” and not-me (Paul Almond, 2011b). But existing of timelines where I will remember, who I am, makes the problem of the threshold irrelevant: I will always emerge above any threshold.

Another thought experiment by Almond: “A more general version of the thought experiment, the delayed quantum suicide thought experiment, shows problems with such a simplistic view. In this thought experiment, there is a variable time delay between the quantum event and your death occurring, ranging from a very long time delay to a negligibly short time delay, with many intermediate time delays being possible” (Paul Almond, 2011a).

## Not-exactly-me copies

Almond suggested the following thought experiment: “We can also construct a thought experiment that seems to force an advocate of the idea that surviving quantum suicide suggests that MWI is true into dealing with more expansive reference classes than the one proposed here. Suppose we run a quantum suicide experiment that is less clean. Some series of quantum events, will generate a number between 0 and 10E10. That number will then determine the extent to which your brain is damaged. If the number is 10E10 your brain will not be damaged at all. If the number is 10E10-1 your brain will be damaged just a little bit. (This hardly matters: your brain could get more random damage in a second of normal life anyway.) If the number is 10E10-2 it will be damaged a bit more severely, and so on. The worse case scenario is if the number is 0. In this situation, your brain is damaged to such an extent that it is just a random mess: something that is no more apparently a human brain than it is anything else. The question is: how intelligent do you expect to be after such a process? If you think that P(Survival) = 1 for standard quantum suicide, you should think that some of the really damaged brains that result here should be excluded: they will effectively be dead versions of you, but where do you draw the line, or do you assign probabilities in some variable way? This situation seems to force you into some kind of reference class consideration that is worse than the one presented in the previous argument, so the case here is that the reference class issue in the previous argument is fairly minor in comparison with some more expansive reference classes that need to be dealt with to discuss quantum suicide” (P. Almond, 2011).

It is clear that even inside one observer-moment some parts of the experience are more important than the other. Human attention gradually decline to the sides of the visual field, and something which is part of an observer-moment gradually become part of its environment.

All this rise important question: if I die, and in the another universe my copy exist (so it is not quantum immortality, but big world immortality based on existence of my copies in causally disconnected regions), which has difference with me, equal to several bits, could I find my next moment in this non-exact copy?

It seems that answer “yes” is correct if the difference is small. Firstly, because I even will not be able to consciously observe and account for small one-pixel difference. Second, is that our states of mind are always changing from the second to second, and so it is normal that next my moment will be the somewhat different than previous moment.

It is incorrect to say that I somehow “jump” from here to there. I am already “there”, and just stop to exist here.

Not-exactly-me my copies dominate total number of copies. For example, if I go out and have a fatal accident, there is a lot my copies in other worlds which didn’t go out at all in that day. The next day, the fact of did they go out or not will completely vanish from their memory and will not become a part of their personal identity. So, timelines (went out on Monday and not hit by a car) and (didn’t go out on Monday) will merge on Tuesday, and the first one could dominate over time as more and more timelines will forget going out. As a result, the most of timelines where I survive the car accident will be not the ones where I miraculously dodge the car or survive the crash, but those where I didn’t remember going out at all. See the similar lines of reasoning in (Aranyosi, 2012).

It is especially true for older people, whos personal identity is very rigid and is not evolving every moment and whos memory is not very strong. They could be regarded and fell themselves as the same person even after years between two moments in life. So they could even “jump over” the whole deadly illness. For example, John (t=60) gets deadly cancer. For John (t=61) the most likely outcome to be alive is that he didn’t become ill at all.

## Resetting of timeclock, respawning and reference class problem

As we said before, multiverse immortality could be formulated in form that for any observer-moment 1 (I, T) exists observer-moment 2 (I, T+1).

However, Almond showed (P. Almond, 2011) that the “measure” (roughly speaking, the number of copies in quantum world) is decreasing in the transition from 1 to 2, and if a subject can’t measure time T properly, he will be in the situation similar to Sleeping Beauty experiment, and (with some assumption how we solve Sleeping beauty) he will more likely find himself before the QS experiment.

For example, if he doesn’t know if the quantum suicide experiment already happened or not in the moment T0, he should think that is doesn’t happened yet, as the number of copies before experiment is higher.

However, there is another possible solution of the QS experiment. That the experimenter dreamed about being in the moments T and T+1, but in fact he is still in the moment T-10. For example, he had a false awakening in the morning before the experiment and dreamed that he performed the experiment, but after experiment he awake again in his bad in the moment T-10, but having memories about realistic dream where he has the experiment.

In the real settings, probably 1 of 100 experimenters will have dreams about the planed QS experiment the night before it. If the probability to survive the QS suicide is set 1 in 1000, it means that from observational point of view the most probable outcome of experiment us to find one self before the experiment with false memories about the experiment! See more about reports of such experiences in the subsection 7.1.

However, the person could awake from the dream about QS experiment not in his timeline. I mean that after the experimenter actually performed QS experiment, he will find in the most probable moment (I, T+1) in the universe, and many such moments will be the false awakening dreams of another experimenters in another worlds, who are not yet started the experiment. That is, actual timeline could continue in someone’s dream.

Remember, that it still not enough to explain reports about respawning, as classical big world immortality doesn’t allow any information transfer or premonition visions, which often described in the respawning cases. Only if some form of evolutionary evolved ways of orchestrating quantum collapse (Penrose & Gardner, 2002) exists for probability manipulation (Yudkowsky, 2009b), it could become possible. However, rigorous thinking on this topic is not easy as it is field of quantum woo and out of scope of this article.

The less likely is survival in QS experiment, the more likely outcome of it is that the next observer state will be in the world where he (or his copy) only hallucinated about the experiment.

Almond about time resetting: “if you had somehow lost track of how old you were – say you had severe amnesia – you should find it unlikely that you were in one of the observer moments in a category with a low combined measure, if these exist, corresponding to being, for example, millions of years old” (Paul Almond, 2011d)

## General anti-immortality principle based on observed age

The idea is that if immortality will be possible, including Multiverse-immortality, when a random observer should find his age infinitely higher that median age of people around himself. (Mallah, 2009).

His idea is exactly the Doomsday argument, applied to individual human life expectancy. It is surprising that after all attempts to disprove it as invalid, it now was revived inly to kill another weird probabilistic idea.

However, all anti Doomsday argument arguments works here, like that my position is not random, as I still don’t know is multiverse immortality is applicable to me.

Also, the fact that my measure decreasing may play the role here.

## The impossibility of sleep

One objection to multiverse immortality is that the same logic could be also applied to the losing consciousness during night and thus sleep becomes impossible for me tonight as there will be always a world there I didn’t sleep in the next moment. The same way waking up from a dream forgetting also becomes impossible.

However, human dreaming is not working this way, as we still have some consciousness and constantly have some forms of dreams even in the deep state. But in case of deep coma and narcosis the problem arises again.

The main thing here is point of application of the multiverse immortality: is it applicable to the observer-moments, or to human minds which include all their memories, including those which I do not remember now but is able to extract from my memory?

It seems that it could be applied to both, and it is not the big problem, as we, humans, are adapted to having two levels minds: observer-moment pure subjective experiences and larger personality which we never observer the whole, but which consists of all our life memories and skills.

## Many futures

In the MW interpretation of quantum mechanics, any agent faces many futures which will actually happen with him, so if we don’t have problems with many futures idea before the QS experiment, we will not have problems with the idea after the experiment.

In case of big world acausal immortality, a person has only one future in his normal life, but after the moment of his death, there are infinitely many exact and non-exact copies of him in very different situations. Thus, a reasonable question arises: where exactly his timeline would jump, or, if we rightly decline the idea of “jumping”, the question is what to do with infinitely many futures, which will actually happen simultaneously.

The question, however, is not factual, but normative, the same way as all questions about personal identity.

Multiverse-immortality is not answering to the question *what to expect?* It answers to the question *what you should not expect,* and it tells that *non-existence is impossible*: there always will be next observer-moment. However, correctly predicting observations about the next observer-moment may be easy or difficult depending on many circumstances.

“Many futures” explains that there will be many infinite future timelines with different topology. Some will be circular, some will include infinite growth inside superintelligent AI.

Carl Shulman: “I think this sidesteps the underlying intuitions too quickly. We have cognitive mechanisms to predict "our next experience," memories of this algorithm working well, and preferences in terms of "our next experience." If we become convinced by the data that this model of a unique thread of experience is false, we then have problems in translating preferences defined in terms of that false model. We don't start with total utilitarian-like preferences over the fates of our future copies (i.e. most aren't eager to lower their standard of living by a lot so as to be copied many times (with the copies also having low standards of living)), and one needs to explain why to translate our naive intuitions into the additive framework (rather than something more like averaging). <http://lesswrong.com/lw/6op/preference_for_many_future_worlds/4iuy>

## QI is only artefact of thinking, not the property of the world

Nesov suggested that QI appears only in thinking in the world, but not a real think in the world: “Note that this statement talks about a ritual of cognition, not about the world: it talks about what one can *remember*, but obviously it's possible to *infer* that in some circumstances you'll die, or that in counterfactuals following different past events you've died. So this kind of "immortality" is an artefact of an artificial limitation on the ways of conceptualizing real world, one that can easily be lifted and thus shown to be not about an actual property of the world”. <http://lesswrong.com/lw/87o/a_pessimistic_view_of_quantum_immortality/548k>

However, actual statement of QI is about the state of the world, not about the cognition: It states that O(I, T+1) moment always actually exists.

## Adding up to normality

One of meta-level objections to this type of arguments is that “its all should come to normality”. However, the correct presentation of this principle is that “its all should look like normality in the near vicinity” of us, and Multiverse-immortality doesn’t contradict it. This addition of *near vicinity* is based on the fact that as we love in normally looking world, all theories which contradicts my current experiences should be false. However, not near events may be more and more strange. For example, black holes are possible, but just not near my planet.

# References:

1. Almond, P. (2006). *Many-Worlds Assisted Mind Uploading: A Thought Experiment*. Retrieved from https://web.archive.org/web/20110513092111/http://www.paul-almond.com/ManyWorldsAssistedMindUploading.htm
2. Almond, P. (2008). Civilization-Level Quantum Suicide. Retrieved from https://web.archive.org/web/20120310105846/http://www.paul-almond.com/CivilizationLevelQuantumSuicide.htm
3. Almond, P. (2011a). *Delayed Quantum Suicide: Assuming you will survive quantum suicide may be simplistic.* Retrieved from https://web.archive.org/web/20130503122611/http://www.paul-almond.com/
4. Almond, P. (2011b). *Quantum Brain Damage: Assuming you will survive quantum suicide may be simplistic.* Retrieved from https://web.archive.org/web/20130503122611/http://www.paul-almond.com/
5. Almond, P. (2011c). Quantum Suicide and Inconsistency: Assuming you will survive quantum suicide may be simplistic. Retrieved from https://web.archive.org/web/20120309233248/http://www.paul-almond.com/QuantumSuicideInconsistency.pdf
6. Almond, P. (2011d). *Repeated quantum suicide would not suggest that MWI is correct – even to the person doing it.* Retrieved from https://web.archive.org/web/20120310205434/http://www.paul-almond.com/QuantumSuicide.pdf
7. Almond, P. (2011). Repeated quantum suicide would not suggest that MWI is correct: a stronger version of the argument. Retrieved from https://web.archive.org/web/20120310023143/http://www.paul-almond.com/QuantumSuicide2.pdf
8. Aranyosi, I. (2012). Should we fear quantum torment? *Ratio*, *25*(3), 249–259.
9. Armstrong, S. (2011). Anthropic decision theory. *ArXiv Preprint ArXiv:1110.6437*.
10. Bergström, L. (2012). Death and eternal recurrence. *The Oxford Handbook of Philosophy of Death*.
11. Bevers, B. (2011). Can Many-Worlds Survive a Quantum Doomsday.
12. Bitbol, M. (1994). *IS NOW A MOMENT IN TIME? A Discussion of McTaggart’s Argument Against the Reality of Time From a Transcendental Idealist Standpoint*.
13. Blanqui, A. (1872). Eternity Through the Stars by Auguste Blanqui. Retrieved December 29, 2017, from https://www.marxists.org/reference/archive/blanqui/1872/astres.htm
14. Bocchieri, P., & Loinger, A. (1957). Quantum recurrence theorem. *Physical Review*, *107*(2), 337.
15. Bostrom, N. (1999). The doomsday argument is alive and kicking. *Mind*, *108*(431), 539–551.
16. Bostrom, N. (2003). Astronomical waste: The opportunity cost of delayed technological development. *Utilitas*, *15*(3), 308–314.
17. Bostrom, N. (2013). *Anthropic bias: Observation selection effects in science and philosophy*. Routledge.
18. Bostrom, N., & Cirković, M. M. (2003). The Doomsday Argument and the Self–Indication Assumption: Reply to Olum. *The Philosophical Quarterly*, *53*(210), 83–91. https://doi.org/10.1111/1467-9213.00298
19. Carrigan Jr, R. A. (2006). Do potential SETI signals need to be decontaminated? *Acta Astronautica*, *58*(2), 112–117.
20. Chalmers, D. (1996). *The Conscious Mind*. Oxford University Press, New York.
21. Cirkovic, M. M. (2004). Is Quantum Suicide Painless? On an Apparent Violation of The Principal Principle. *ArXiv:Quant-Ph/0412147*. Retrieved from http://arxiv.org/abs/quant-ph/0412147
22. Copi, I., & Cohen, M. (1990). *Introduction to Logic*. New York: Macmillan. Google Scholar.
23. Daniel, M. (2017). S-risks: Why they are the worst existential risks, and how to prevent them (EAG Boston 2017). Retrieved from https://foundational-research.org/s-risks-talk-eag-boston-2017/
24. Deasy, D. (2015). The moving spotlight theory. *Philosophical Studies*, *172*(8), 2073–2089.
25. Deng, N. (2015). On Whether B-Theoretic Atheists Should Fear Death. *Philosophia*, *43*(4), 1011–1021. https://doi.org/10.1007/s11406-015-9638-y
26. Egan, G. (2009). Dust Theory FAQ. Retrieved from http://www.gregegan.net/PERMUTATION/FAQ/FAQ.html
27. Eliade, M. (1949). The Myth of the Eternal Recurrence. *Translation Feb Srkaraty,*.
28. evodevouniverse. (2018). Cosmological natural selection (fecund universes) - Evo Devo Universe. Retrieved February 8, 2018, from http://evodevouniverse.com/wiki/Cosmological\_natural\_selection\_(fecund\_universes)
29. Feynman, R. P., Hibbs, A. R., & Styer, D. F. (2010). *Quantum mechanics and path integrals*. Courier Corporation.
30. Frischhut, A. M., & Skiles, A. (2013). Time, Modality, and the Unbearable Lightness of Being. *Thought: A Journal of Philosophy*, *2*(3), 264–273.
31. Greene, B. (2011). *The hidden reality: Parallel universes and the deep laws of the cosmos*. Vintage.
32. Greene, P. (2017). Success-First Decision Theories. *Newcomb’s Problem. Cambridge University Press, Forthcoming*.
33. Higgo, J. (1998). Does the “many-worlds” interpretation of quantum mechanics imply immortality? Retrieved from https://web.archive.org/web/20050203192203/http://higgo.com:80/quantum/qti.htm
34. Hossenfelder, S. (2018). Scientific Theory And The Multiverse Madness. *NPR.Org*. Retrieved from https://www.npr.org/sections/13.7/2018/01/22/579666359/scientific-theory-and-the-multiverse-madness
35. Istvan, Z. (2014). Should Cryonics, Cryothanasia, and Transhumanism Be Part of the Euthanasia Debate? *Huffington Post Tech: The Blog*, *23*.
36. Kachru, S., Kallosh, R., Linde, A., & Trivedi, S. P. (2003). De Sitter vacua in string theory. *Physical Review D*, *68*(4), 046005.
37. Kaufmann, W. A. (2013). *Nietzsche: Philosopher, psychologist, antichrist*. Princeton University Press.
38. LessWrong Wiki. (2018). Updateless decision theory - Lesswrongwiki. Retrieved February 8, 2018, from https://wiki.lesswrong.com/wiki/Updateless\_decision\_theory
39. Lewis, D. (2004). How many lives has Schrödinger’s cat? *Australasian Journal of Philosophy*, *82*(1), 3–22.
40. Mallah, J. (2009). Many-Worlds Interpretations Can Not Imply “Quantum Immortality.” *ArXiv:0902.0187 [Quant-Ph]*. Retrieved from http://arxiv.org/abs/0902.0187
41. Marchal, B. (1991). Mechanism and personal identity. In *Proceedings of WOCFAI* (Vol. 91, pp. 335–345).
42. Menzel, C. (2017). Actualism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (2014th ed.). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/win2017/entries/actualism/
43. Minerva, F., & Sandberg, A. (2017). Euthanasia and cryothanasia. *Bioethics*, *31*(7), 526–533.
44. Moravec, H. (1988). *Mind children: The future of robot and human intelligence*. Harvard University Press.
45. Nietzsche, F. (1883). *Thus spoke zarathustra*. Jester House Publishing.
46. Page, D. N. (2007). Susskind’s Challenge to the Hartle-Hawking No-Boundary Proposal and Possible Resolutions. *Journal of Cosmology and Astroparticle Physics*, *2007*(01), 004–004. https://doi.org/10.1088/1475-7516/2007/01/004
47. Parfit, D. (1984). *Reasons and persons*. OUP Oxford.
48. Penrose, R., & Gardner, M. (2002). *The Emperor’s New Mind: Concerning Computers, Minds, and the Laws of Physics* (1 edition). Oxford: Oxford University Press.
49. Pereira, T. (2017). An Anthropic Argument against the Future Existence of Superintelligent Artificial Intelligence. *ArXiv:1705.03078 [Cs]*. Retrieved from http://arxiv.org/abs/1705.03078
50. Poincaré, H. (1890). Sur le probleme des trois corps et les équations de la dynamique. *Acta Mathematica*, *13*(1), A3–A270.
51. Randall, A. F. (2004). Quantum miracles and immortality. *Quantum*, *5*, 8.
52. Sandberg, A. (2008). Bayes, Moravec and the LHC: Quantum Suicide, Subjective Probability and Conspiracies. Retrieved from http://www.aleph.se/andart/archives/2008/09/bayes\_moravec\_and\_the\_lhc\_quantum\_suicide\_subjective\_probability\_and\_conspiracies.html
53. Shikhovtsev, E. (2003). Biographical Sketch Of Hugh Everett, III. Retrieved from http://space.mit.edu/home/tegmark/everett/everett.html#contents
54. SHOULD WE FEAR QUANTUM TORMENT? - Aranyosi - 2012 - Ratio - Wiley Online Library. (n.d.). Retrieved January 2, 2018, from http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9329.2012.00540.x/abstract
55. Şimşekoğlu, Ö., Nordfjærn, T., Zavareh, M. F., Hezaveh, A. M., Mamdoohi, A. R., & Rundmo, T. (2013). Risk perceptions, fatalism and driver behaviors in Turkey and Iran. *Safety Science*, *59*, 187–192. https://doi.org/10.1016/j.ssci.2013.05.014
56. Smolin, L. (1992). Did the universe evolve? *Classical and Quantum Gravity*, *9*(1), 173.
57. Soares, N., & Levinstein, B. A. (2017). *Cheating death in damascus*. Technical report, Machine Intelligence Research Institute, 2017. https://intelligence. org/files/DeathInDamascus. pdf. Retrieved from https://intelligence.org/files/DeathInDamascus.pdf
58. Sotala, K., & Gloor, L. (2017). Superintelligence as a Cause or Cure for Risks of Astronomical Suffering.
59. Squires, E. J. (1994). *The Mystery of the Quantum World* (2 edition). Bristol ; Philadelphia: CRC Press.
60. Taleb, N. N. (2007). *The black swan: The impact of the highly improbable* (Vol. 2). Random house.
61. Tegmark, M. (1998). The Interpretation of Quantum Mechanics: Many Worlds or Many Words? *Fortschritte Der Physik*, *46*(6–8), 855–862. https://doi.org/10.1002/(SICI)1521-3978(199811)46:6/8<855::AID-PROP855>3.0.CO;2-Q
62. Tegmark, M. (2009). The multiverse hierarchy. *ArXiv Preprint ArXiv:0905.1283*.
63. Tegmark, M. (2017). The Universes of Max Tegmark. Retrieved from http://space.mit.edu/home/tegmark/crazy.html
64. Tipler, F. J. (1980). 3 - General Relativity and the Eternal Return†. In *Essays in General Relativity* (pp. 21–37). Academic Press. https://doi.org/10.1016/B978-0-12-691380-4.50009-5
65. Torres, P. (2014). Why Running Simulations May Mean the End is Near. Retrieved from https://ieet.org/index.php/IEET2/more/torres20141103
66. Turchin, A. (2017). Messaging future AI. Retrieved from https://goo.gl/YArqki
67. Turchin, A., & Green, B. . (2017). Aquatic refuges for surviving a global catastrophe. *Futures*, *89*, 26–37. https://doi.org/10.1016/j.futures.2017.03.010
68. Yudkowsky, E. (2008). Where Physics Meets Experience - Less Wrong. Retrieved February 8, 2018, from http://lesswrong.com/lw/ps/where\_physics\_meets\_experience/
69. Yudkowsky, E. (2009a). Eliezer\_Yudkowsky comments on Newcomb’s Problem and Regret of Rationality - Less Wrong. Retrieved February 8, 2018, from http://lesswrong.com/lw/nc/newcombs\_problem\_and\_regret\_of\_rationality/126j
70. Yudkowsky, E. (2009b). The anthropic trilemma. Retrieved from http://lesswrong.com/lw/19d/the\_anthropic\_trilemma/