Simulation Hypothesis: Optimisation, Wave-particle Duality and a New Termination Risk

Genevieve Gorrell, M.Phil (Cantab), D.Phil.
genevieve.gorrell@gmail.com
Preston, UK, May 2025

Thinking in terms of our existing within a simulation yields plausible suggestions for how our world might have come to be. For example it seems reasonable to guess our implementation is optimised. The apparent paradox of wave-particle duality may arise from optimisation. Computationally reducible material may never be run at all, but only calculated just in time. Islands of activity, complex within themselves but for some duration not affecting anything outside of themselves, might be deferred from being run, to be caught up in a separate thread only if needed. Quantum computing may be an abuse of the multithreaded nature of the simulation.

Introduction

The simulation hypothesis is the hypothesis that we are living in a computer simulation of some description. The original stimulus for simulation theory as a field was Nick Bostrom's paper (Bostrom, 2003) in which he explores the level of credence we should place in the possibility that we aren't the original 21st century Earth civilisation but a simulation thereof run by our descendents. In this way he assumes that the properties we know of our universe also apply in the hypothetical parent reality. He proposes that the tremendous level of compute power theoretically available to us would make it possible to simulate this current reality many times over.

Vazza (2025) on the other hand states that computational demands almost rule out a simulation of the complexity of the world around us. When we get into such specifics as computational requirements we must also be more specific about the nature of the simulation we are hypothesising. There are two main types in the literature, the first of which I will call the Matrix type, also known as RPG (after role-playing games, in the sense that a pre-existing human plays an avatar, see Virk (2021) and Chalmers (2016)) and the second, the Sims type, or NPC (non-player character, also from Virk (2021)); we are a character actually implemented within the game.

In the Matrix type, the conscious human is supposed to pre-exist, perhaps in the form of an envatted brain, and the simulation task is then something akin to video game rendering, most efficiently achieved "just in time", when a player actually experiences the material (Campbell *et al.*, 2017). The computational limitation then rests with the number of users, that being potentially as few as one, and not with the complexity of the world. The Matrix model is similar to Descartes' skeptical hypothesis, which mainly serves to provoke a sense of rigour in exploring what we think we know. A sense of deceit is often implied - the simulation has to be good enough to convince us.

In the Sims case, the task of the simulation is to generate the world, including the human beings living in it. Previous work (e.g. Bostrom, 2003; Beane *et al*., 2014) assumes that if the matter contained in the universe is duplicated in software, all its phenomena, including thinking human beings, will naturally be covered. This is of course much more computationally expensive. Simulating matter needn't mean that no optimisation is possible: Vazza (2025) proposes that the simulation need only cover Earth, or even a low resolution Earth. In any case, it seems likely that each simulation will be poorer than the parent reality that spawned it, and this will place some limitation on nesting realities.

In proposing matter may be implemented in software, this type of work takes an accessible step back from material realism towards the perspective that "all this is information". However, the assumptions are nonetheless materialist: the prevailing paradigm reflected in this work is that matter is the foundation of reality. We highlight the following key reasons to reject this:

- Quantum physics has amply demonstrated over the last century that classical assumptions based on local realism do not hold: locality is violated by the phenomenon of entanglement, and realism is violated by the apparent collapse of the probability wave function on observation. Matter cannot be the whole story.[1]
- The "hard problem of consciousness" is the problem we have explaining our conscious experience in terms of matter, matter being the more impoverished container. (It is much easier to explain matter in terms of consciousness.)
- All the life-like behaviour we have been able to create has been achieved through computer programming; for example cellular automata, fractals and artificial intelligence. This suggests life does not arise from matter and properties of matter, but from algorithms being run.

The idea that the world arises computationally (AKA digital physics, see Wolfram (1997), Wheeler (2018), Fredkin (2003), Vopson (2023)) can be quite easily grasped in the form of the simulation hypothesis, since many people now have experience of programming or of using software such as computer games. The simulation hypothesis is suggestive of some common-sense inferences we can make about the requirements of our existence. We cannot make a simulation without:

- An energy supply
- Some form of switch architecture, such as a silicon chip
- Algorithms/parameters
- Data, albeit possibly just random seed data

1
https://www.scientificamerican.com/article/the-universe-is-not-locally-real-and-the-physics-nobel-prize-winners-proved-it/

Thinking in terms of computational generation, rather than in terms of matter obeying natural laws, flags up to any programmer the amount of coding that would have to go into even the most basic set-up (Vopson, 2023, argues similarly). The idea that no programming at all would be involved seems somewhat incredible, like supposing that your computer, given long enough, will eventually program itself. We can certainly try to figure out the most minimal algorithm set that would explain what we see around us, in accordance with Occam's Razor, but nonetheless, algorithms don't normally write themselves. Furthermore, programmers aren't often maximally elegant in their formulations. In fact, we might consider a perfect creator rather less likely than an imperfect one. If we recognise that it is likely that there is a degree of somewhat complex programming underlying our experience, that opens doors in understanding physical anomalies such as action at a distance. As far as the origin of "base reality" is concerned, the absolute origin of all things, we have nothing on which to base any inference. But *this* reality appears to be computed.

In this paper we suggest that it is worth considering that our implementation is optimised in order to maximise usage of computational processing power. The implications of anything short of infinite processing resource are that optimisation is relevant and likely. We suggest three possible areas of optimisation: not simulating material that can be simply calculated; not necessarily simulating isolated material; and reusing simulated material where possible. We relate these to quantum mechanics; namely, the role of observation in collapsing the wave function, and the curious status of Schrodinger's cat. Finally, we outline a new potential termination risk. Firstly, however, we introduce relevant findings from quantum physics, and previous research that looks to simulation theory for an explanation thereof.

Review: Quantum Physics and Simulation Theory

In wave-particle duality we see an apparent paradox. Where laser light is shone through two slits, an interference pattern appears on the screen behind, suggesting light is a wave. However, if individual photons are fired towards the two slits and detected at the slits, the photon behaves as a particle, and over many trials, two bars appear on the screen behind, one for each slit, with no interference pattern. In short, the photon behaves as a probability wave unless it is observed. When you consider it is possible to only decide whether or not to measure the photon *after* it has been fired, which potentially was much earlier, for example if the particle was fired from another planet, it is hard to escape the conclusion that the world, and even the past, is rendered for our benefit, and offers no substantial reality for our materialist scientists to investigate.

Schrodinger's Cat is a thought experiment in which the emission of a particle causes poison to be released into a box in which a cat is isolated. However, the emission of the particle is indeterminate, meaning that whether the cat is alive or dead is also indeterminate. Within the classical view of physics, this is regarded as absurd. Modern experiments have confirmed that given sufficient isolation it is possible for quite large objects to remain indeterminate (Thomas, 2021) and indeed quantum computing, whilst error-prone, has demonstrated that quantum indeterminacy can be deliberately maintained for a complex of material. In quantum computing, the state of indeterminacy is used to allow a bit (qubit) to simultaneously hold different values, making for efficient parallelized calculation, where a conventional computer would have to work through each possible value separately.

A small body of research makes the connection between simulation theory and quantum physics. Simulation theorist Marcus Arvan (e.g. Arvan 2014) proposes that we exist in a peer-to-peer simulation with a node for each individual. The difference between our individual realities, which must ultimately be made congruent, explains quantum uncertainty and a variety of other phenomena. An attraction of the work is that it draws on our actual experience of implementing P2P online role-playing games (MMORPGs) such as Nine Chronicles[2] to yield insights about how multiple realities are merged in practical terms. Quantum phenomena are also addressed by Campbell et al (2017) in the context of a Matrix-style simulation. Beane et al (2014) explore the limitations and quantum observations we might expect to see in a numerical simulation of the matter of the universe. The meta-message from this plurality of research may be that as soon as we embrace the idea of reality as a computational implementation, a great many options become available to us for resolving the issues that our current entrenchment in materialism, despite all counter evidence, is causing for us.

Principle 1: Don't Simulate if you can Calculate

Wolfram (1997) defines computational reducibility as a property of a system in which outcomes can be predicted. For example, in the case of a binary star system, if the properties of the orbit are known, the position at any given time point can be calculated. Computationally irreducible behaviour is that where the outcome cannot be predicted. For example, in a three star system, chaotic behaviour results, and the only way to find out what is going to happen is to simulate it. I suggest that the reason we are being simulated at all is because we are not computationally reducible, and the only way to find out what happens is to run us.

Yet large chunks of our physical world are either computationally reducible or might be considered so for practical purposes, such as simple combinations of inert matter. If our Cosmic Programmer seeks to optimise, we might expect that these are never run at all, but simply calculated just in time, if and only if they feed into some complex thread that is not reducible and therefore must be run. This is reminiscent of the video rendering suggestions for optimising a matrix-style simulation with the difference that no deceit is intended, nor even a specific focus on observation, but merely making the necessary information available to the *complex process* whilst not wasting computational resource.

With regards wave-particle duality, computing particle final positions and even history, either singly or in aggregate, can be done just-in-time, as required by an actual simulated element. Electrons are never run and only notionally exist at all. When final positions are computed in aggregate the probability distribution produces an interference pattern. When they are computed singly, they don't. This is simply the way the simulation is implemented.

It might be said that the physical world appears to be somewhat of an "object-orientated" interface on top of the algorithm set responsible for generating the underlying complex data flow, in the sense that the essence of the material world is to contain objects. It would be

---

[2] https://medium.com/nine-chronicles/nine-chronicles-a-new-paradigm-in-gaming-973768988093

interesting if the property of being reducible or irreducible were a property on classes of object. It would be interesting if the things in our life were of two (or more) types on this basis, and some of them, for example perhaps some astronomical bodies, exist therefore "outside of time".

This suggestion differs from Arvan's (2014) in which the probability distribution arises from modelling uncertainty in a peer-to-peer simulation, where multiple observers have their own version, and these must ultimately be reconciled. It is possible that both his and my ideas are simultaneously in operation in some form. I don't focus on separate observers, but similar issues are raised in the next section in the context of islands of complexity.

Principle 2: Rank Islands of Complexity

An obvious optimisation, as described above, is to calculate simple, computationally reducible phenomena "just in time". Where a simple calculation meets a complex, computationally irreducible one, it must be calculated in order that the simulation can proceed. Could it be that the wave function collapses as soon as it intersects with any complex phenomenon? In that case, Schrodinger's cat constitutes a complex phenomenon in itself, so would live or die at the mundane moment on that basis. Alternatively, a further optimisation would be that the cat, having been isolated, is taken out of the equation until the box opens, and literally ceases to exist in any concrete form, since it affects nothing. It's a notional cat. Once the box is opened and the cat must be discovered dead or alive, the gap in its timeline might be filled in as necessary in a separate thread, since complex phenomena must be run to determine their outcome. The cat is run up to the present in a separate thread, and in the mean time, the scientists' thread waits, and as far as the cat was concerned, nothing unusual happened there.

Running the irreducible parts of the entire simulation can be made more computationally tractible as follows. An island is a unit within which computational complexity is occurring (and which therefore must be run) in isolation from other parts of the simulation. For example, the cat is isolated and is therefore an island. The scientists may also be an island, but a larger one.

First, all islands are ranked according to some measure of significance, or else randomly. Then the first island is run until such a time as it interconnects with another, for example when the scientist opens the box. At that point, the cat is run in a separate thread up to the present time and merged into the island. So the indeterminate status of the cat arises from the fact that the fate of the cat hasn't happened yet, and may never happen if no-one looks in the box, thus avoiding running unnecessary islands.

What if the scientists' thread was frozen, waiting for the cat to observe them? There could be some heuristic that favours the larger island, or maybe the cat's thread was run first and found to result in an inert state - if the scientists had been frozen, they wouldn't have been able to open the box. The cat one way or another became computationally reducible and the simulation reached a cul-de-sac. So the next island is tried.

Suppose there's an alien civilisation in some other location in the Milky Way. They aren't run unless they come into contact with us, because Earth is island A to their island B. Eventually, we make some form of contact with them, so they need to be caught up to the present in order to merge. However, in the course of catching them up, it turns out they contacted Earth earlier. Now we must all revert to that moment and merge.

Of course, there may be some more optimal way to prioritise islands. Several may be run in step. The coding around synchronising islands might be similar to Arvan's (2014) proposal of a peer-to-peer simulation, in which different perspectives on reality must be reconciled.

For the most part, however, not running islands - the proverbial tree that falls in the forest that no-one is watching - saves compute power, and splitting the problem into islands means they can be run sequentially on limited computing resources.

Could it be that we ourselves will one day be pruned out of existence? We may eventually be irrelevant to the main thread(s) of the simulation, or placed far down the queue and never got to. The code lingers, but we aren't run. Eventually we are garbage-collected. The ultimate painless death might be to make yourself completely irrelevant.

Principle 3: Reuse Simulated Material

It seems plausible that the Cosmic Programmer might optimise by avoiding simulation repetition. Where a piece of simulation differs only in certain ways from another, why not reuse it, and use a shortcut or heuristic to make necessary edits? The human mind is remarkably flexible in filling in the gaps in for example an incomplete memory or an action on our part that we can't explain (e.g. Gazzaniga, 1995), so any change to our history needn't be perfect. Where simulation has produced a game state that includes a set of individuals with mental states that are correct aside from certain self-contained edits, it might well go undetected to hack the edits in *post hoc*.

For example, in the scenario outlined above, the entire history of planet B was run before realising that planet A affected it some time earlier. It must be re-run from the contact point, but large sections might be reused. The personal history of an individual on planet B may be reused, but their knowledge of the date that alien contact occurred must be edited.

Rizwan Virk discusses the Mandela Effect in his book "The Simulated Multiverse". The Mandela Effect refers to the phenomenon of a significant number of people having clear memories of Nelson Mandela dying in prison in the 1980s, though in fact he died in 2013. A variety of further evidence of differing quality has accreted under the heading of the Mandela Effect, and can easily be found online. Virk proposes that the Mandela Effect can be explained as a glitchy merge between timelines in a simulation in which alternative realities are run in parallel. This is somewhat echoed here, but my emphasis is on *not* running extra material if it's avoidable, and I don't include any concept of choice points. My simulations make their mind up and stick to it.

The reader may explore for themselves the quality of evidence for the Mandela Effect, but the Missing Thunderbird Photograph is a colourful example. There is also interesting first-hand report data from Philip K. Dick[3] in which he describes his experience of remembering alternative timelines.

Termination Risk

The optimistic hope surrounding the advent of quantum computing, implicitly, is that the machine the universe is run on, however literally we take that, is inherently probabilistic, and our quantum computing leverages that to enable efficient computing. However, it is also possible that the probabilism is not due to some fundamental structural difference, the nature of which is currently beyond out ken, but merely implemented in a conventional computing style using multithreading. It appears magical to us because *our thread waits*.

One idea about how quantum computing works would be that in putting a particle in an indeterminate state, and indeed potentially creating a complex calculation based on these indeterminate particles, we have created an island, A, which isn't run until it our island B observes it. Indeed we pile much material into this island A. When we observe the complex, it needs to be run up to the present, which is done in a separate thread. From our point of view we get a fast result, but in fact our thread may have had to wait for island A to complete before it could resume. It saved us waiting, but if we have any concern about a limitation on the resources available to our simulated universe, we might not want to think of it as a freebee.

Every programmer will have had the experience of accidentally running an infinite loop or running a program that we then realise will take a very long time to complete. In this case, the program must be interrupted. We are now at the stage where many people are experimenting with quantum computing. Is it possible that creating an infinite loop in quantum computing would lead to our thread hanging indefinitely? We might take some comfort in the fact that if it can happen, it probably has already, and if we have crashed our simulation, clearly the Cosmic Programmer is prepared to restart us or has implemented a catch.

Yet if we abuse the multithreaded nature of our simulation too extremely, that might be considered a form of outcome, and bring the experiment to an end. The most immediate danger of this seems to lie in Bitcoin mining, that being a bottomless pit for futile computation. For further termination risks in a simulated universe, see Greene (2020), Turchin et al. (2019) and Braddon-Mitchell and Latham (2024).

Might the Cosmic Programmer prefer us to spend more time focused on personal and interpersonal integration and organisation, and less time hacking the fabric of the simulation? Is it possible that we'll break the simulation before we achieve some plausible intended goal, such as becoming an advanced civilisation capable of avoiding self-annihilation? We conclude on that cautionary note.

---

[3] "If you find this world bad, you should see some of the others", AKA the Metz speech, https://www.youtube.com/watch?v=RkaQUZFbJjE

Bibliography

Arvan, M. (2014). A unified explanation of quantum phenomena? The case for the peer-to-peer simulation hypothesis as an interdisciplinary research program.

Beane, S. R., Davoudi, Z., & J. Savage, M. (2014). Constraints on the Universe as a Numerical Simulation. *The European Physical Journal A*, *50*(9), 148.

Bostrom, N. (2003). Are we living in a computer simulation?. *The philosophical quarterly*, *53*(211), 243-255.

Braddon-Mitchell, D., & Latham, A. J. (2024). Ancestor simulations and the Dangers of Simulation Probes. *Erkenntnis*, *89*(3), 1257-1267.

Campbell, T., Owhadi, H., Sauvageau, J., & Watkinson, D. (2017). On testing the simulation theory. *arXiv preprint arXiv:1703.00058*.

Chalmers, D. J. (2016). The Matrix as metaphysics. *Science Fiction and Philosophy: From Time Travel to Superintelligence*, 35-54.

Fredkin, E. (2003). An introduction to digital philosophy. *International journal of theoretical physics*, *42*, 189-247.

Gazzaniga, M. S. (1995). Principles of human brain organization derived from split-brain studies. *Neuron*, *14*(2), 217-228.

Greene, P. (2020). The termination risks of simulation science. *Erkenntnis*, *85*(2), 489-509.

Thomas, R. A., Parniak, M., Østfeldt, C., Møller, C. B., Bærentsen, C., Tsaturyan, Y., ... & Polzik, E. S. (2021). Entanglement between distant macroscopic mechanical and spin systems. *Nature Physics*, *17*(2), 228-233.

Turchin, A., Batin, M., Denkenberger, D., & Yampolskiy, R. (2019). Simulation typology and termination risks. *arXiv preprint arXiv:1905.05792*.

Vazza, F. (2025). Astrophysical constraints on the simulation hypothesis for this Universe: why it is (nearly) impossible that we live in a simulation. *Frontiers in Physics*, *13*, 1561873.

Virk, R. (2021). *The Simulated Multiverse: An MIT Computer Scientist Explores Parallel Universes, the Simulation Hypothesis, Quantum Computing and the Mandela Effect* (Vol. 2). Bayview Books, LLC.

Vopson, M. M. (2023). The second law of infodynamics and its implications for the simulated universe hypothesis. *AIP Advances*, *13*(10).

Wheeler, J. A. (2018). Information, physics, quantum: The search for links. *Feynman and computation*, 309-336.

Wolfram, S. (1997). New kind of science.