Essays in the Philosophy of Physics, Science, Metaphysics, and Epistemology

Francisca Sousa Pereira Cruz de Vasconcelos

Keble College

University of Oxford

Submitted in completion of the

Master of Studies in Philosophy of Physics

June 27th, 2022

Acknowledgements

There are several people and institutions I would like to thank, without whom these essays would not have been possible.

First and foremost, I would like to thank the many incredible professors, of the Oxford Faculty of Philosophy, that I had the chance to work with and learn from this year. In addition to taking their courses, I was very grateful to have several 2-hour supervisions with Prof. Owen Maroney, Prof. James Read, Dr. Sebastián Murgueitio Ramírez, and Prof. Oliver Pooley. Beyond useful feedback for improving my essays, they broadened my horizons and helped me navigate the dense philosophy literature. Largely because of their mentorship, my critical analysis skills have improved substantially and I have been exposed to some of the coolest questions at the forefront of physics and human knowledge!

Beyond the Oxford Philosophy Faculty, I would like to thank the many MIT professors, postdocs, and graduate students who previously taught and supervised me. I would not be at Oxford, nor able to pursue such an interesting degree, without their prior mentorship. I especially thank all the amazing MIT philosophy faculty, who peaked my interests in the field as an undergraduate.

Institutionally, I would like to thank the Rhodes Trust for granting me the opportunity to pursue this masters degree, through the Rhodes Scholarship. Without their monetary and general support, I never would have been able to study in the UK. I also would like to thank Keble College, for providing great accommodation and general resources in pursuing my studies.

Personally, I thank all my friends (old and new) for their constant encouragement. I would like to give a special shout-out to the OUAFC Women's Blues team for an incredible season, topped off with a great Varsity victory – score early, score often, shoe the tabs, and bleed dark blue! I also would like to acknowledge the Class of 2020 Rhodes Scholars, friends from the 2021 Statistics MSc, my degree-mates Gal·la and Yoram, my HBAC flatmates, other friends from Keble College, all the great OxCSML SIMPs, and my amazing friends from MIT undergrad. In continuing my statistics research from last year (often as a break from all the philosophy!), I would like to thank Professor Yee Whye Teh, Bobby He, and Nalini Singh for their collaboration and mentorship.

I would like to give an extra special thanks to Varsha Ramineni, for being such an amazing and supportive friend, and to Christoher McDonald for being an incredible boyfriend who provided endless love, support, encouragement, and fun times throughout this philosophy journey and my time in Oxford. In the best and worst of times, he was always there to make me a 'cuppa'.

Finalmente, quero agradecer a toda a minha família em Portugal por todo o amor e suporte. I especially would like to thank my sister and parents, for their never-ending love, daily calls, constant support, and endless encouragement. Without them, I would not be where or who I am today. Obrigada a todos. .

Abstract

This work contains a compilation of four essays submitted for assessment in completion of the Oxford Masters of Studies in Philosophy of Physics. The first essay, *Rethinking Quantum Speedup via the Quantum Singular Value Transform*, is my personal favorite – asserting that a recent breakthrough in quantum algorithms should reshape philosophical thoughts regarding the nature of quantum speedup. More traditionally in the philosophy of physics, *A Low-Entropy Big Bang is Not Needed to Fry an Egg* debates the emergence of a statistical mechanical 'arrow of time', bolstering local branch theories over the Past Hypothesis. In an ontic structural realist approach to the philosophy of science, *"Relationships All the Way Down"* draws inspiration from modern physics and the history of science to argue that only relationships are fundamental and dissolves the key concern of 'relationships without relata'. Finally, in metaphysics and theory of knowledge, *The Open Past and Many-Worlds Presentism* attacks our beliefs of a fixed, determined past and supports the Many-Worlds Presentism ontology of time. These 4 essays were selected and improved among 11 total essays written over the course of the degree: 5 in Philosophy of Physics, 3 in Philosophy of Science, and 3 in Metaphysics & Theory of Knowledge.

Contents

List of Figures

x

List of Abbreviations

BDV . . . . Bub's Disjunctive View

BH . . . . Branching Hypothesis

GBT . . . . Growing Block Theory

GPU . . . . Graphical Processing Unit

GR . . . . General Relativity

IGUS . . . . Information Gathering and Utilizing System

NISQ . . . . Noisy Intermediate-Scale Quantum

OFPAA . . . . . . Oblivious Fixed-Point Amplitude Amplification

OSR . . . . Ontic Structural Realism

PH . . . . Past Hypothesis

QFT . . . . Quantum Field Theory

QPP . . . . Quantum Parallelism Process

QPT . . . . Quantum Parallelism Thesis

QSP . . . . Quantum Signal Processing

QSVT . . . . Quantum Singular Value Transformation

RQM . . . . Relational Quantum Mechanics

SIV . . . . Steane's Interpretational View

SVD . . . . Singular Value Decomposition

TA . . . . Turning Argument

The discovery and development of quantum computers have raised a number of interesting philosophical issues. Perhaps the most important issue is how to explain why quantum computers appear to be faster than classical computers for some computational tasks.

— Duwell (2018)

1

Rethinking Quantum Speedup via the Quantum Singular Value Transform

Contents

1.1 Introduction

Since the 1990s, quantum computer fabrication has progressed exponentially[1]. Unfortunately, quantum algorithm design has not been as fruitful. Quantum computer scientists, lacking fundamental knowledge of the source(s) of quantum speedup, have discovered only a handful of quantum algorithms providing substantial speedup over classical counterparts. While compelling evidence exists for a quantum speedup (Preskill, 2018), there is no rigorous proof that the class of efficiently solvable quantum computational problems exceeds those efficiently solvable on classical computers. Precise identification of quantum speedup could enable a proof of quantum computational advantage[2] and generalized framework for efficient quantum algorithm design.

In this work, I present and expose the limitations of three predominant theories for quantum speedup: the Quantum Parallelism Thesis, Steane's Interpretational View, and Bub's Disjunctive View. I also present a recent breakthrough in quantum algorithms – the Quantum Singular Value Transform, or so-called "grand unification of quantum algorithms" – and argue that future philosophy regarding quantum speedup must incorporate this algorithm. I conclude with a discussion of the Quantum Singular Value Transform's implications for existing philosophical theories of quantum speedup. Specifically, I argue that it generalizes Bub's Disjunctive View, confirming Bub's intuition for quantum speedup.

1.2 Potential Sources of Quantum Speedup

I begin by presenting three main philosophical arguments for quantum speedup: (1) the Quantum Parallelism Thesis, (2) Steane's Interpretational View, and (3) Bub's Disjunctive View. These views are contradictory and face distinct challenges in reconciling different observed forms of quantum speedup.

1.2.1 The Quantum Parallelism Thesis

The earliest and most common intuition for quantum speedup is the **Quantum Parallelism Thesis (QPT)**. Inspired by their development of the first quantum algorithm, Deutsch and Jozsa (1992) argued that *quantum computers are more efficient than their classical counterparts because they can compute multiple values of a function in a single step* (Deutsch, 1998; Ekert and Bouwmeester, 2000). In the original arguments, Deutsch held that the QPT is consistent only with the many-worlds interpretation of quantum mechanics. However, this view has since been dismissed by most (Duwell, 2007). For the remainder of this work I will focus on a modern, interpretation-independent presentation of the QPT by Duwell (2018).

Duwell claims that evidence for the QPT comes from the **Quantum Parallelism Process (QPP)**, which is associated with several traditional, efficient quantum algorithms. Consider a system of $n + m$ qubits, where the first n qubits serve as an input register and the last m qubits serve as an output register. Let $|x\rangle$ describe the state of the n input qubits and $|y\rangle$

[1] Schoelkopf's Law predicts a Moore's Law-like exponential scaling in quantum processor performance, which has been empirically observed to-date (Devoret and Schoelkopf, 2013).

[2] I do not mean to suggest this would prove $\mathsf{P} \neq \mathsf{NP}$, but a weaker claim that quantum computation provides guaranteed speedup over classical computation for certain problems.

describe the state of the m output qubits. Recall that applying a Hadmard, $\hat{H}$, to every qubit in $|x\rangle$ achieves the uniform superposition state,

$$|x\rangle \xrightarrow{\hat{H}^{\otimes n}} \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \,. \tag{1.1}$$

Now, suppose there exists unitary operator $\hat{U}_f$, such that

$$|x\rangle\,|y\rangle \xrightarrow{\hat{U}_f} |x\rangle\,|f(x) \oplus y\rangle \,, \tag{1.2}$$

where $f : \{0, ..., 2^n - 1\} \to \{0, ..., 2^m - 1\}$ and $\oplus$ denotes addition modulo 2. Creating a uniform superposition over $|x\rangle$, setting $|y\rangle = |0\rangle$, and applying $\hat{U}_f$ computes $f(x)$ for every x as

$$\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle\,|0\rangle \xrightarrow{\hat{U}_f} \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle\,|f(x)\rangle \,. \tag{1.3}$$

Thus, in the QPP, a single operation evaluates a function over all possible inputs of interest.

Original arguments for the QPT (Deutsch, 1998; Ekert and Bouwmeester, 2000; Duwell, 2007) attributed quantum speedup to simultaneous computation over all possible states. However, a significant objection was raised by Steane (2003) and Bub (2010). Although the final state in Equation 1.3 encodes every input x and its corresponding function evaluation $f(x)$, at measurement time the state will collapse to state $|x'\rangle\,|f(x')\rangle$ for a single input x' and function evaluation $f(x')$. Obtaining the full set of function evaluations still requires $O(2^n)$ system measurements, which is on the same order of operations as classical computation[3]. Thus, QPP computed functions are *inaccessible* on the basis of a single system measurement.

In light of Steane and Bub's criticism, it appears unclear how the QPT would provide any source of speedup. Duwell (2018), however, defends the QPT via the notion of *distinguishability*. Given the algorithmic focus of this work, I present a reformulation of Duwell's argument based on standard quantum algorithm design intuition. Although exponential measurements are needed to obtain the full set of function evaluations $f(x)$ $\forall x \in \{0, ..., 2^n - 1\}$, in practice, efficient quantum algorithms are not designed to need all $f(x)$ evaluations. Instead, quantum algorithms leverage *quantum interference* to constructively interfere desired outcome amplitudes and destructively interfere undesired outcome amplitudes. This can be achieved by the new unitary operation $\hat{U}'_f$ that computes

$$\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle\,|0\rangle \xrightarrow{\hat{U}'_f} \sum_{x=0}^{2^n-1} \alpha_x\,|x\rangle\,|f(x)\rangle \,, \tag{1.4}$$

where α_x are amplitude coefficients such that $\sum_{x=0}^{2^n-1} |\alpha_x|^2 = 1$. Thus, the two key challenges in designing efficient quantum algorithms are: (1) formulating the problem such that the desired algorithm output can be encoded in $O(1)$ states and (2) designing a $\hat{U}'_f$ which amplifies desired output state amplitudes, α_x, significantly more than undesired state amplitudes.

However, if speedup results from simultaneous computation, there is no clear distinction between quantum computation and, say, bounded-error probabilistic computing or massively

[3]In the analogous classical computing setup, one would simply apply the function f to each $x' \in \{0, ..., 2^n - 1\}$, which is 2^n operations of $f(x)$.

parallelized classical computation. In fact, it would be disappointing if the only computational advantage provided by quantum computers was parallelization, because quantum computers would amount essentially to specialized hardware, similar to classical graphical processing units (GPUs). Steane and Bub argued that sheer parallelization does not enable quantum speedup, to which Duwell responded that clever algorithm design must strategically leverage quantum interference. However, neither Duwell nor the QPT provide concrete guidelines or insight for achieving this crucial quantum interference – resulting in a lack of explanatory power.

1.2.2 Steane's Interpretational View

Beyond the previously presented inaccessability argument, Steane (2003) argues that quantum computers cannot acquire or manipulate classical information more efficiently than classical computers. Specifically, Equation (1.3) gives a false impression of the performed amount of computation, since computation cannot be measured by directly comparing numbers of operations between distinct methods. For example, each step in binary search is not $\frac{N}{2}$ steps because of the number of operations needed for linear search over the same N-element list. Steane also posits that simultaneous computation of 2^N operations should amount to $O(\frac{1}{\text{Exp}(N)})$ error rates, while quantum computers empirically obtain $O(\frac{1}{\text{Poly}(N)})$ error rates. Finally, he argues that the unitary evolution of quantum processes must imply that different quantum computational paths are not independent.

With all these criticisms of the QPT, Steane presents his own theory of quantum speedup: **Steane's Interpretational View (SIV)**. Unlike the QPT's view of quantum computers as large processes exploiting massive parallelism, SIV argues that quantum computers are *small systems, which can exploit entanglement-generated correlations*. Specifically, SIV proposes that speedup is achieved for certain computational tasks because quantum computers do not need to represent entities, but instead use entanglement to manipulate correlations between entities. For example, in Equation (1.3), the correlation between x and $f(x)$ is fully represented, even though the values of $f(x)$ are not (since at measurement only one $f(x')$ is sampled). Further, in Shor (1994)'s algorithm, the extracted period is a correlation between values of $f(x)$ – no physical record of x remains once the algorithm terminates.

While appealing, there are problems with Steane's view. Steane claims quantum interference is only possible if the superposition terms are part of a single, isolated coherent-state (i.e. the quantum computer is not entangled with the environment). In fact, this justifies why quantum computers perform a single process, rather than many distinct simultaneous computations:

> When we examine an efficient quantum algorithm such as Shor's, we find that it is indeed essential to the working of the algorithm that the evaluations of $f(x)$ in superposition do not individually have any subsequent influence on other parts of the universe. If they did, the resulting entanglement would prevent the algorithm from working. (Steane, 2003)

However, it is unreasonable to expect that experimental quantum processors will ever be fully isolated from the environment. For example, cryogenic cooling, which prevents thermal excitations in superconducting quantum computers, can never reach 0K. Furthermore, fault tolerance only mandates an error threshold, not elimination of error altogether. If Steane's claim were correct, it would be impossible to experimentally implement any algorithm and quantum computing would be doomed to the same fate as analog computing – a belief few currently hold. In fact, small noisy quantum processors, of the present Noisy Intermediate-Scale Quantum (NISQ) era (Preskill, 2018), have proven competitive with the world's largest classical supercomputers for specific computational tasks (Arute et al., 2019). Thus, Steane's argument does not seem fully valid in light of modern experimental results. Furthermore, like the QPT, SIV faces a lack of explanatory power. While arguing that entanglement is key to quantum speedup, SIV does not provide specific guidance for designing quantum algorithms which leverage said entanglement to outperform classical ones.

1.2.3 Bub's Disjunctive View

Similar to Steane and unlike the QPT, Bub (2010) argues that quantum speedup is achieved via *fewer* computations than classical computers. Unlike Steane, however, he believes quantum speedup lies in the difference between classical and quantum logic. Specifically, **Bub's Disjunctive View (BDV)** posits that, while classical disjuncts can only be `True` or `False` in virtue of the disjuncts' truth values, quantum disjuncts can be `True` or `False` irrespective of the disjuncts' values. Throughout the paper, Bub demonstrates how information processing in hidden-subgroup type problems – i.e. Deutsch (1985), Simon (1997), and Shor (1994)'s algorithms – exploit the non-Boolean logic represented by their Hilbert spaces' projective geometry and subspace structure. For the sake of space, I will only briefly explain BDV's application to Simon's and Shor's algorithm, instead elaborating upon the simplest algorithm – Deutsch's.

1.2.3.1 BDV at a High-Level

Period-finding algorithms, such as Simon's and Shor's, split the domain of a periodic function, via a period, into mutually exclusive and collectively exhaustive domain partitions. Thus, distinguishing a period from all other periods becomes the problem of distinguishing one partition from all other partitions. To determine a desired partition, a classical algorithm must evaluate the function for a subset of input values – the number of computation steps growing exponentially with input size.

From the perspective of BDV, quantum period-finding algorithms represent alternative domain partitions as Hilbert-space subspaces that are predominantly orthogonal, but may contain small overlap regions. Each Hilbert subspace is spanned by the 2^n quantum computational basis states, a superposition of which is input to the quantum algorithm. After the period-finding algorithm performs a suitable transformation, a computational basis state measurement can identify the Hilbert subspace containing the state, identifying the desired partition and period. No function evaluation is needed, but the algorithm may need to run multiple times if the state collapses to an overlap region, making the measurement inconclusive.

1.2.3.2 BDV for Deutsch's Algorithm

Deutsch's XOR problem asks whether a Boolean function[4] $f : \mathcal{B} \to \mathcal{B}$, where $\mathcal{B} \in \{0, 1\}$, is "constant" (i.e. returns the same output for all inputs: $f(0) = 0$ and $f(1) = 0$ *or* $f(0) = 1$ and $f(1) = 1$) or "balanced" (i.e. returns 0 for half of the input domain and 1 for the other half: $f(0) = 0$ and $f(1) = 1$ *or* $f(0) = 1$ and $f(1) = 0$). Classically, this requires querying the oracle with both 0 and 1, then comparing the results. Meanwhile, Deutsch's algorithm can

[4]Computed via a "black-box" oracle.

determine the function nature, 50% of the time, via a single measurement. Although traditionally not considered a period-finding algorithm, Deutsch's algorithm can be mapped into the same BDV framework as Simon's and Shor's. As we will see, function nature (whether constant or balanced) is mapped into Hilbert-space subspaces.

There are three steps in Deutsch's algorithm: (1) initialize to $|0\rangle|0\rangle$ in the computation basis, (2) apply a Hadamard to the first qubit, and (3) apply an oracle unitary transformation, $\hat{U}_f : |x\rangle|y\rangle \to |x\rangle|y \oplus f(x)\rangle$, which implements Boolean function f. Dependent on the values of $f(0)$ and $f(1)$,

$$|0\rangle|0\rangle \xrightarrow{\hat{H}\otimes\hat{I}} \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big)|0\rangle \xrightarrow{\hat{U}_f} \frac{1}{\sqrt{2}}\big(|0\rangle|f(0)\rangle + |1\rangle|f(1)\rangle\big), \tag{1.5}$$

the system could be in one of four possible states:

$$|c_1\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle|0\rangle + |1\rangle|0\rangle\big) \tag{1.6}$$

$$|c_2\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle|1\rangle + |1\rangle|1\rangle\big) \tag{1.7}$$

$$|b_1\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle|0\rangle + |1\rangle|1\rangle\big) \tag{1.8}$$

$$|b_2\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle|1\rangle + |1\rangle|0\rangle\big), \tag{1.9}$$

where constant f achieves $|c_1\rangle$ or $|c_2\rangle$ and balanced f achieves $|b_1\rangle$ or $|b_2\rangle$[5].

Constant states $|c_1\rangle$, $|c_2\rangle$ and balanced states $|b_1\rangle$, $|b_2\rangle$ span two planes in the Hilbert space $\mathcal{H}^2 \otimes \mathcal{H}^2$. Recall that Von Neumann (1932)'s probability calculus of quantum logic defines disjunction ($\vee$) as the closed span of the Hilbert subspaces' union and conjunction ($\wedge$) as the intersection. Define the constant plane $\mathcal{P}_c$ as the quantum logical disjunction of constant state projection operators,

$$\mathcal{P}_c = \hat{P}_{|c_1\rangle} \vee \hat{P}_{|c_2\rangle}, \tag{1.10}$$

where $\hat{P}_{|c_1\rangle} = |c_1\rangle\langle c_1|$ and $\hat{P}_{|c_2\rangle} = |c_2\rangle\langle c_2|$. Similarly, define balanced plane $\mathcal{P}_b$ as

$$\mathcal{P}_b = \hat{P}_{|b_1\rangle} \vee \hat{P}_{|b_2\rangle}, \tag{1.11}$$

where $\hat{P}_{|b_1\rangle} = |b_1\rangle\langle b_1|$ and $\hat{P}_{|b_2\rangle} = |b_2\rangle\langle b_2|$. Although the states $|c_1\rangle$, $|c_2\rangle$ are not orthogonal to the states $|b_1\rangle$, $|b_2\rangle$, their disjunctions – the planes $\mathcal{P}_c$ and $\mathcal{P}_b$ – are orthogonal except for an intersection ray,

$$\vec{\mathcal{I}}_{c,b} = \mathcal{P}_b \wedge \mathcal{P}_c = \frac{1}{2}\big(|00\rangle + |01\rangle + |10\rangle + |11\rangle\big) = \frac{1}{\sqrt{2}}\big(|c_1\rangle + |c_2\rangle\big) = \frac{1}{\sqrt{2}}\big(|b_1\rangle + |b_2\rangle\big). \tag{1.12}$$

This geometry is illustrated in Figure 1.1.

Moving to the Hadamard basis,

$$|+\rangle = \hat{H}|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \tag{1.13}$$

$$|-\rangle = \hat{H}|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \tag{1.14}$$

[5] I assume no errors, i.e. bit-flips, in all computations presented in this work.

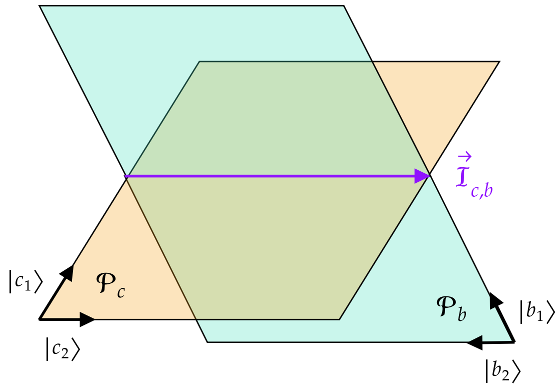

Figure 1.1: A geometrical representation of BDV for Deutsch's algorithm. The constant plane $\mathcal{P}_c$ is mostly orthogonal to the balanced $\mathcal{P}_b$, except for an intersection ray $\vec{\mathcal{I}}_{c,b}$.

the intersection ray is simply $|{+}{+}\rangle$, the constant plane $\mathcal{P}_c = \hat{P}_{|++\rangle} \vee \hat{P}_{|+-\rangle}$ is spanned by

$$|{+}{+}\rangle = \frac{1}{\sqrt{2}}(|c_1\rangle + |c_2\rangle) \tag{1.15}$$

$$|{+}{-}\rangle = \frac{1}{\sqrt{2}}(|c_1\rangle - |c_2\rangle), \tag{1.16}$$

and the balanced plane $\mathcal{P}_b = \hat{P}_{|++\rangle} \vee \hat{P}_{|--\rangle}$ is spanned by

$$|{+}{+}\rangle = \frac{1}{\sqrt{2}}(|b_1\rangle + |b_2\rangle) \tag{1.17}$$

$$|{-}{-}\rangle = \frac{1}{\sqrt{2}}(|b_1\rangle - |b_2\rangle). \tag{1.18}$$

If an observable with eigenstates $\{|{+}{+}\rangle, |{+}{-}\rangle, |{-}{+}\rangle, |{-}{-}\rangle\}$[6] measures the algorithm output, the system will collapse into the 3D subspace orthogonal to $|{-}{+}\rangle$. Constant f obtains $|{+}{+}\rangle$ or $|{+}{-}\rangle$, with 50% probability of each outcome. Similarly, balanced f obtains $|{+}{+}\rangle$ or $|{-}{-}\rangle$, with 50% probability each. If $|{+}{+}\rangle$ obtains, we cannot distinguish whether f is constant or balanced. However, if $|{+}{-}\rangle$ obtains, f must be constant and if $|{-}{-}\rangle$ obtains, f must be balanced. Thus, with 50% probability, Deutsch's algorithm determines whether f is constant or balanced, via a single measurement.

1.2.3.3 BDV Limitations

BDV is arguably the most complex, most sophisticated, and least intuitive of the three presented theories for quantum speedup. This is promising, since a useful theory of quantum speedup is expected to be complex! BDV's analysis of Deutsch's algorithm demonstrates that quantum speedup is explainable by the projective geometry of the system's subspaces, defined using quantum logic. BDV's explanatory power is arguably greater than that of the QPT or SIV, especially in application to Simon's and Shor's algorithms. However, it is still a limited theory.

Bub's original BDV proposal primarily explains speedup for period-finding algorithms – i.e. Shor's and Simon's. As described in Section 1.2.3, these algorithms require a domain partitioning,

[6] With a computational basis of $\hat{Z}$-eigenstates, this would amount to measurement of the $\hat{X}$-observable. Alternatively, in the computational basis, this is achieved by applying $\hat{H}$, then measuring the $\hat{Z}$-observable.

lending naturally to the geometric interpretation of BDV. Thus, one might wonder whether BDV's speedup explanation is too closely tied to the structure of period-finding problems and inapplicable to other algorithm types. Bub attempts to bolster BDV's merits by claiming that Deutsch's algorithm is distinct from Simon's and Shor's, but can still be mapped into the BDV framework. However, Deutsch's algorithms is one of the simplest quantum algorithms, yet still requires non-trivial analysis to map into the BDV framework, as demonstrated in Section 1.2.3.2. Therefore, Bub's failure to mention or apply BDV to complex algorithms such as search or Hamiltonian simulation is unsurprising.

While BDV seemingly has more explanatory power than the high-level QPT and SIV, Bub appears unable to generalize BDV to all classes of quantum algorithms. For the remainder of this work, I will discuss the Quantum Singular Value Transform algorithm and argue that it generalizes BDV by mathematically encoding all quantum algorithms into Bub's desired projective Hilbert subspaces.

1.3 Inspiration from a Recent Algorithms Breakthrough

The **Quantum Singular Value Transformation (QSVT)** algorithm (Gilyén et al., 2019) instantiates all known efficient quantum algorithms, via its tuneable parameterization, and was deemed *The Grand Unification of Quantum Algorithms* (Martyn et al., 2021). In this section, I will (1) discuss the state of quantum algorithms prior to the QSVT, (2) present the QSVT algorithm, (3) apply the QSVT to search, and (4) explain how the QSVT unifies quantum algorithms.

1.3.1 Prior to QSVT

The first quantum algorithms – Deutsch and Jozsa (1992) and Bernstein and Vazirani (1997) – solved toy problems to demonstrate a quantum computational advantage. More recent quantum algorithms development, as described by Chuang (2020), addresses real-world problems and can be categorized into three main types: (1) search, (2) simulation, and (3) factoring.

Search algorithms aim to find a specific (set of) target state(s). They are derived from Grover (1996)'s algorithm and amplitude amplification, as formalized by (Bennett et al., 1997). These models led to adiabatic computation (Farhi et al., 2000) and the fixed-point adiabatic evolution algorithm (Dalzell et al., 2017). **Simulation algorithms** aim to approximate the evolution and dynamics of quantum system Hamiltonians, $e^{-i\hat{\mathcal{H}}t}$. Originally, these algorithms consisted of universal quantum simulators (Feynman, 2018; Lloyd, 1996). More recently, this has been generalized to quantum random walks (Childs, 2009) and linear combinations of unitaries (Berry et al., 2015). **Factoring algorithms**, in their most general form, aim to find a unitary's eigenphases. Shor (1994)'s algorithm applied period-finding to integer factorization, rendering RSA cryptography obsolete. This was further generalized to quantum phase estimation (Cleve et al., 1998) and solving linear systems of equations (Harrow et al., 2009). Although some recent algorithms, e.g. variational quantum eigensolvers, exist outside these three pillars, their speedup is questionable.

Prior to the QSVT, there was little to no known relation between these distinct quantum algorithm types. All algorithms achieved quantum speedup, but the means of achieving that speedup varied by algorithm type. Each algorithm type could be distilled, more or less, to a distinct subroutine: factoring algorithms use a quantum phase-estimation subroutine, simulation algorithms use a quantum walk subroutine, and search algorithms use a phase-estimation subroutine. However, we expect a *central* source of quantum speed-up, meaning there should be some way to *unify* all these algorithms. This unification, the so-called "grand unification," was achieved recently by the QSVT.

1.3.2 The Quantum Singular Value Transformation

The QSVT algorithm's key contribution was unifying and generalizing prior work on **qubitization** (Low and Chuang, 2019) and **Quantum Signal Processing (QSP)** (Low et al., 2016) – enabling polynomial transformations of (rectangular) quantum sub-systems' singular values. Following Gilyén et al. (2019) and Martyn et al. (2021), I will provide brief descriptions of qubitization, the QSP algorithm, and the QSVT algorithm.

1.3.2.1 Qubitization

Qubitization was proposed to robustly approximate the time-evolution operator $e^{-i\hat{\mathcal{H}}t}$, where $\hat{\mathcal{H}} = \Pi\hat{U}\Pi$ is a symmetric projected unitary encoded Hermitian operator. In place of $\hat{U}$, qubitization finds a unitary $\hat{W}$ that is robust to repeated application leakage, enabling higher orders of $\hat{\mathcal{H}}$. Each eigenvector of $\hat{\mathcal{H}}$ splits into *two* eigenvectors of $\hat{W}$, resulting in 2D subspaces isomorphic to a qubit – hence "qubitization". Specfically, assuming $\hat{\mathcal{H}}$ has eigenvector $|\lambda\rangle$ with corresponding eigenvalue λ, qubitization finds $\hat{W}$ with eigenvectors $|\lambda^{\pm}\rangle$ and corresponding eigenvalues $e^{\pm i \arccos(\lambda)}$, where $|\lambda\rangle$ is a superposition of $|\lambda^{-}\rangle$ and $|\lambda^{+}\rangle$. As demonstrated in the following section, QSP can apply polynomial $\mathscr{P}$ to the spectrum of $\hat{W}$, producing a projected unitary encoding, $\hat{U}'$, of $\mathscr{P}(\hat{\mathcal{H}})$.

1.3.2.2 The QSP Algorithm

Inspired by nuclear magnetic resonance pulse sequences for increasing image contrast (Wimperis, 1994; Minch, 1998; Vandersypen and Chuang, 2005; Wolfowicz and Morton, 2016), QSP demonstrated that the gate sequence[7],

$$\hat{U}_{\vec{\phi}}(\theta) = e^{i\phi_0\hat{Z}}\, e^{i\theta\hat{X}}\, e^{i\phi_1\hat{Z}}\, e^{i\theta\hat{X}}\, e^{i\phi_2\hat{Z}} \ldots e^{i\theta\hat{X}}\, e^{i\phi_d\hat{Z}} = e^{i\phi_0\hat{Z}} \prod_{k=1}^{d} e^{i\theta\hat{X}}\, e^{i\phi_k\hat{Z}}, \tag{1.19}$$

with unknown θ and tuneable $\vec{\phi} = (\phi_0, \phi_1, \phi_2, ..., \phi_d)$, can generate a rich set of quantum unitary operations, $\hat{U}_{\vec{\phi}}(\theta)$. For each $\hat{U}_{\vec{\phi}}(\theta)$, sweeping over $\theta \in [-\pi, \pi]$ results in a function of $|0\rangle \rightarrow |0\rangle$ transition probabilities,

$$P_{\vec{\phi}}(\theta) = |\langle 0|\, \hat{U}_{\vec{\phi}}(\theta)\, |0\rangle\,|^2. \tag{1.20}$$

Modifying pulse sequence values of $\vec{\phi}$ realizes different transition probability functions $P_{\vec{\phi}}$. Furthermore, letting $\theta = -2\cos^{-1}(a)$ with $a \in [-1, 1]$, then

$$e^{i\theta\hat{X}} = \hat{W}(a) = \begin{bmatrix} a & i\sqrt{1-a^2} \\ i\sqrt{1-a^2} & a \end{bmatrix}. \tag{1.21}$$

[7] We use standard notation: $\hat{X}$ is Pauli-X, $\hat{Y}$ is Pauli-Y, $\hat{Z}$ is Pauli-Z, and $\hat{I}$ is identity.

Plugging this back into Equation (1.19), QSP proves the existence of a set of (calculable) QSP angles $\vec{\varphi}$, such that

$$\hat{U}_{\vec{\varphi}}(a) = e^{i\varphi_0 \hat{Z}} \prod_{k=1}^{d} \hat{W}(a)\, e^{i\varphi_k \hat{Z}} = \begin{bmatrix} \mathscr{P}(a) & i\mathscr{Q}(a)\sqrt{1-a^2} \\ i\mathscr{Q}(a)\sqrt{1-a^2} & \mathscr{P}^*(a) \end{bmatrix} \tag{1.22}$$

for any polynomials $\mathscr{P}(a)$ and $\mathscr{Q}(a)$, subject to a few minor constraints[8]. Thus, QSP systematically transforms quantum states according to nearly arbitrary polynomial functions of degree d, using $O(d)$ elementary unitary quantum operations. This result can be extended to multi-qubit systems, using block-encodings of square matrices. However, we will jump straight to the QSVT, since it generalizes to rectangular block-matrices.

1.3.2.3 The QSVT Algorithm

The QSVT uses QSP sequences to polynomially transform singular values of a (possibly rectangular) matrix $\mathbf{A}$[9] block encoded into unitary matrix $\hat{U}$,

$$\hat{U} = \begin{bmatrix} \mathbf{A} & . \\ . & . \end{bmatrix}. \tag{1.23}$$

Every $m \times n$ matrix $\mathbf{A}$ can be decomposed by the **Singular Value Decomposition (SVD)**, defined as,

$$\mathbf{A} = \hat{W}_\Sigma \mathbf{\Sigma} \hat{V}_\Sigma^\dagger, \tag{1.24}$$

where $\hat{W}_\Sigma$ and $\hat{V}_\Sigma$ are unitary matrices and $\mathbf{\Sigma}$ is a (possibly rectangular) diagonal matrix containing non-negative, real singular values $\{\sigma_k\}$. The column vectors of $\hat{W}_\Sigma$ and $\hat{V}_\Sigma$, denoted $|w_k\rangle$ and $|v_k\rangle$, form orthonormal bases. The so-called left singular vectors $|w_k\rangle$ span the left singular vector space, while right singular vectors $|v_k\rangle$ span the right singular vector space. Thus, $\mathbf{A}$ is alternatively expressed as

$$\mathbf{A} = \sum_k \sigma_k \left|w_k\right\rangle \left\langle v_k\right|. \tag{1.25}$$

Utilizing Equation (1.23)'s block-encoding, define projectors

$$\hat{\Pi} := \sum_k \left|v_k\right\rangle \left\langle v_k\right| \tag{1.26}$$

$$\hat{\Pi}' := \sum_k \left|w_k\right\rangle \left\langle w_k\right| \tag{1.27}$$

locating $\mathbf{A}$ in $\hat{U}$,

$$\mathbf{A} = \hat{\Pi}' \hat{U} \hat{\Pi}. \tag{1.28}$$

In other words, $\hat{\Pi}$ selects the rows and $\hat{\Pi}'$ selects the columns of $\hat{U}$ encoding $\mathbf{A}$. As in qubitization, $\hat{\Pi}$ and $\hat{\Pi}'$ identify two distinct Hilbert subspaces relevant to $\mathbf{A}$, namely the left and right singular

[8] A given $\vec{\varphi} = (\varphi_0, \varphi_1, \varphi_2, ..., \varphi_d)$ can only achieve functions of the form $\deg(\mathscr{P}) \leq d$, $\deg(\mathscr{Q}) \leq d-1$, $\mathscr{P}$ with parity $d \mod 2$, $\mathscr{Q}$ with parity $(d-1) \mod 2$, and $|\mathscr{P}|^2 + (1-a^2)|\mathscr{Q}|^2 = 1$

[9] Boldface notation is used to distinguish non-unitary, rectangular matrices.

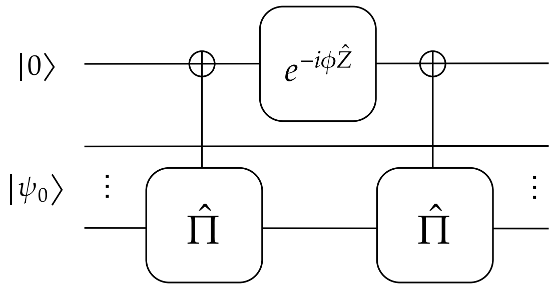

Figure 1.2: Implementation of the $C_{\hat{\Pi}}NOT$ gate.

vector spaces. Every singular value σ_k defines a plane by $\{|v_k\rangle, |v_k^\perp\rangle\}$ and another by $\{|w_k\rangle, |w_k^\perp\rangle\}$. The key QSVT insight is that $\hat{U}$ and $\hat{U}^\dagger$ perform rotations between these two planes as:

$$\hat{U} : \{|v_k\rangle, |v_k^\perp\rangle\} \to \{|w_k\rangle, |w_k^\perp\rangle\} \tag{1.29}$$

$$\hat{U}^\dagger : \{|w_k\rangle, |w_k^\perp\rangle\} \to \{|v_k\rangle, |v_k^\perp\rangle\}. \tag{1.30}$$

Thus, interleaving $\hat{U}$ and $\hat{U}^\dagger$ with controlled phase-shift operations enables polynomial operations on $\mathbf{A}$'s singular values, σ_k. Specifically, for odd polynomial[10] $\mathscr{P}(x)$, there exists a $\vec{\phi} = \{\phi_1, ..., \phi_d\}$, with odd d, such that

$$\hat{U}_{\vec{\phi}} = \hat{\Pi}'_{\phi_1}\hat{U} \prod_{k=1}^{(d-1)/2} \hat{\Pi}_{\phi_{2k}}\hat{U}^\dagger\hat{\Pi}'_{\phi_{2k+1}}\hat{U} = \begin{bmatrix} \mathscr{P}(\mathbf{A}) & \cdot \\ \cdot & \cdot \end{bmatrix}. \tag{1.31}$$

$\mathscr{P}(\mathbf{A})$ is the polynomial transform of $\mathbf{A}$'s singular values,

$$\mathscr{P}(\mathbf{A}) = \sum_k \mathscr{P}(\sigma_k)\,|w_k\rangle\langle v_k| = \hat{W}_\Sigma \mathscr{P}(\mathbf{\Sigma})\hat{V}_\Sigma^\dagger, \tag{1.32}$$

and $\hat{\Pi}_\phi$ and $\hat{\Pi}'_\phi$ are projector controlled phase-shift operations,

$$\hat{\Pi}_\phi := e^{i2\phi\hat{\Pi}} \tag{1.33}$$

$$\hat{\Pi}'_\phi := e^{i2\phi\hat{\Pi}'}, \tag{1.34}$$

imparting phase $e^{i2\phi}$ to the subspace determined by $\hat{\Pi}$ or $\hat{\Pi}'$, respectively. As illustrated in Figure 1.2, $\hat{\Pi}_\phi$ (and similarly $\hat{\Pi}'_\phi$) can be implemented in a quantum circuit with two projector controlled-NOT gates (aka $C_{\hat{\Pi}}NOT$) and a single-qubit Z-rotation by angle ϕ ($e^{-i\phi\hat{Z}}$).

1.3.3 Search via the QSVT

As explained by Gilyén et al. (2019) and Martyn et al. (2021), QSVT can be applied to the problem of search. While Grover (1996)'s algorithm was the first proposed efficient quantum search algorithm, it proved a limiting case of the oblivious fixed-point amplitude amplification (OFPAA) search algorithm (Dalzell et al., 2017). OFPAA aims to find winner state $|m\rangle \in \{|1\rangle, ..., |N\rangle\}$, using unitary oracle

$$\hat{U}_m |j\rangle = (-1)^{\delta_{j,m}} |j\rangle = \begin{cases} |j\rangle, & |j\rangle \neq |m\rangle \\ -|j\rangle, & |j\rangle = |m\rangle \end{cases}, \tag{1.35}$$

[10] A similar expression exists for even polynomials, with even d.

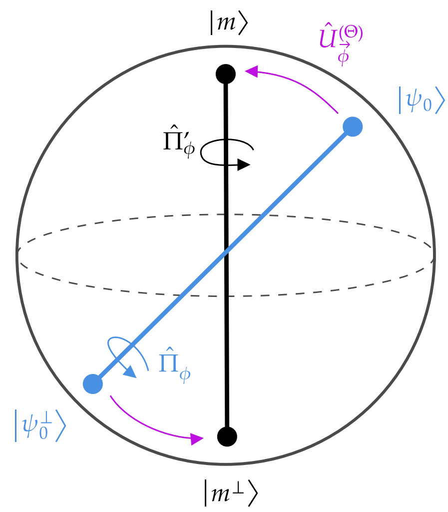

Figure 1.3: Qubitization maps search into two concentric Bloch spheres. QSVT generates circuit $\hat{U}^{(\Theta)}_{\vec{\phi}}$, which maps initial state $|\psi_0\rangle$ to desired state $|m\rangle$. [Inspired by Martyn et al. (2021)'s Fig.2]

which applies a phase-flip to $|m\rangle$. The QSVT OFPAA circuit is initialized to the uniform superposition state,

$$|\psi_0\rangle = \hat{H}^{\otimes N}|0\rangle^{\otimes N} = \frac{1}{\sqrt{N}}\sum_{j=1}^{N}|j\rangle = \frac{1}{\sqrt{N}}\sum_{j\neq m}|j\rangle + \frac{1}{\sqrt{N}}|m\rangle\,, \tag{1.36}$$

which is nearly orthogonal to $|m\rangle$. For arbitrary circuit $\hat{C}$, denote the overlap between states $\hat{C}|\psi_0\rangle$ and $|m\rangle$ as

$$a = \langle m|\hat{C}|\psi_0\rangle\,. \tag{1.37}$$

For the circuit initialization $\hat{C} = \hat{I}$, $a = \frac{1}{\sqrt{N}}$. Thus, OFPAA aims to find the circuit $\hat{C} = \hat{U}_{\vec{\phi}}$ that amplifies $|m\rangle$'s amplitude, by mapping $|\psi_0\rangle$ to $|m\rangle$,

$$|\langle m|\hat{U}_{\vec{\phi}}|\psi_0\rangle|^2 = |a|^2 \to 1. \tag{1.38}$$

Qubitization reduces a potentially multi-qubit search problem to a simple 2D Hilbert subspace by expressing Equation (1.36)'s superposition of non-winner states as a single, normalized state perpendicular to $|m\rangle$,

$$|m^{\perp}\rangle = \frac{1}{\mathcal{N}}\big(\hat{I} - |m\rangle\langle m|\big)|\psi_0\rangle\,, \tag{1.39}$$

where $\mathcal{N}$ is a normalization factor. Thus, defining

$$|\psi_0\rangle = a|m\rangle + \sqrt{1-a^2}\,|m^{\perp}\rangle\,, \tag{1.40}$$

and orthogonal state

$$|\psi_0^{\perp}\rangle = \sqrt{1-a^2}\,|m\rangle - a\,|m^{\perp}\rangle\,. \tag{1.41}$$

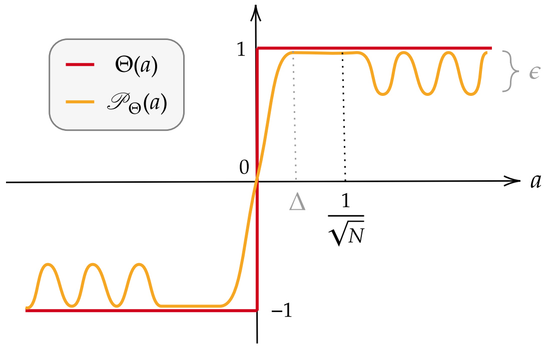

Figure 1.4: The sign function $\Theta(a)$ is approximated by the QSVT as odd-polynomial $\mathscr{P}_\Theta(a)$, mapping the block-encoding's eigenvalue $|a|^2 \to 1$, for search. [Inspired by Martyn et al. (2021)'s Fig.6]

As illustrated in Figure 1.3, the spaces spanned by $\{|\psi_0\rangle, |\psi_0^\perp\rangle\}$ and $\{|m\rangle, |m^\perp\rangle\}$ form two concentric Bloch spheres. Furthermore, the ψ_0-basis to m-basis unitary mapping is defined as

$$\hat{U} = a\Big(|m\rangle\langle\psi_0| - |m^\perp\rangle\langle\psi_0^\perp|\Big) + \sqrt{1-a^2}\Big(|m\rangle\langle\psi_0^\perp| + |m^\perp\rangle\langle\psi_0|\Big), \tag{1.42}$$

expressed in matrix form[11] as,

$$\hat{U} = \begin{array}{c c} & \begin{array}{cc} |\psi_0\rangle & |\psi_0^\perp\rangle \end{array} \\ \begin{array}{c} |m\rangle \\ |m^\perp\rangle \end{array} & \begin{bmatrix} a & \sqrt{1-a^2} \\ \sqrt{1-a^2} & -a \end{bmatrix} \end{array}. \tag{1.43}$$

$\hat{U}$ block-encodes matrix

$$\mathbf{A} = \hat{\Pi}^{(m)}\hat{U}\hat{\Pi}^{(\psi_0)} = |m\rangle\langle m|\hat{U}|\psi_0\rangle\langle\psi_0| = a\,|m\rangle\langle\psi_0|, \tag{1.44}$$

with left singular vector $|m\rangle$, right singular vector $|\psi_0\rangle$, singular value a, and projectors $\hat{\Pi}^{(m)} = |m\rangle\langle m|$, $\hat{\Pi}^{(\psi_0)} = |\psi_0\rangle\langle\psi_0|$. $\hat{\Pi}_\phi^{(m)}$ and $\hat{\Pi}_\phi^{(\psi_0)}$ are constructed from gates: $e^{-i\phi\hat{Z}}$, $C_{\hat{\Pi}^{(m)}}NOT$, and $C_{\hat{\Pi}^{(\psi_0)}}NOT$. Since $|\psi_0\rangle$ is known, $C_{\hat{\Pi}^{(\psi_0)}}NOT$ is easily constructed using the circuit in Figure 1.2. Meanwhile, in the same circuit, $C_{\hat{\Pi}^{(m)}}NOT$ utilizes a controlled-$\hat{U}_m$ operator, with the control sandwiched between Hadamard gates.

Recall that the QSVT generates a gate sequence that applies an arbitrary odd-polynomial $\mathscr{P}$ to $\mathbf{A}$'s singular value, a. With a initialized to positive value $\frac{1}{\sqrt{N}}$ and the search goal of mapping $|a|^2 \to 1$, Martyn et al. (2021) show that the sign function (illustrated in Figure 1.4),

$$\Theta(a) = \begin{cases} -1 & a < 0 \\ 0 & a = 0 \\ 1 & a > 0 \end{cases}, \tag{1.45}$$

is an apt odd-function. $\Theta(a)$ can be approximated as odd-polynomial $\mathscr{P}_\Theta(a)$ by the QSVT gate sequence $\hat{U}_{\vec{\phi}}^{(\Theta)}$.

[11] The matrix labels indicate that the columns act on $|\psi_0\rangle$, $|\psi_0^\perp\rangle$ and the rows act on $|m\rangle$, $|m^\perp\rangle$.

1.3.4 The "Grand Unification" of Quantum Algorithms

Prior to the QSVT algorithm discovery, as described in Section 1.3.1, the field of quantum algorithms was disunited – with few, disjointed means of achieving quantum speedup. As described in Section 1.3.2.3 and illustrated by Section 1.3.3's search example, the QSVT applies arbitrary polynomials to the singular values of a (possibly rectangular) matrix block-encoded into a unitary matrix. Furthermore, Gilyén et al. (2019) and Martyn et al. (2021) demonstrate that algorithms for search, simulation, factoring, and non-provably efficient domains (i.e. quantum machine learning, the quantum OR lemma, fractional query implementation, and Gibbs state preparation) can all be mapped into the QSVT framework. Thus, the QSVT provides a general parameterization of quantum algorithms, with tunable phases $\vec{\phi}$ and block-encoding $\hat{U}(\mathbf{A})$, that can be swept over to achieve the full spectrum of known quantum algorithms. In this sense, it is a "grand unification" of quantum algorithms.

1.4 Philosophical Implications of the QSVT

Having presented existing philosophical quantum speedup theories as well as the QSVT's workings and significance, I now make three key arguments: (1) the QSVT algorithm should play a central role in future philosophical quantum speedup discourse, (2) among existing philsophical theories, QSVT's speedup is closest related to that of BDV, and (3) all algorithms that fit within the QSVT framework can be mapped to BDV, generalizing BDV to all known efficient quantum algorithms.

1.4.1 Putting the QSVT Center Stage

As explained in Section 1.2, there is no current consensus on quantum speedup (assuming such a speedup does exist). Each speedup theory is derived primarily from a specific class of quantum algorithms. The QPT is based on oracle algorithms, SIV is derived from measurement-based quantum computation, and BDV is based on hidden-subgroup algorithms. Ultimately, *the source(s) of quantum speedup should be able to explain the success of all efficient quantum algorithms.* This, however, does not promote the coexistence of all three theories discussed in Section 1.2, since they are not consistent with one another. QPT postulates quantum speedup by increased computation, directly contradicting SIV and BDV's argument that speedup comes from reduced computation. SIV attributes speedup to entanglement, while BDV attributes it to the nature of quantum logic. Thus, while multiple sources of quantum speedup may exist, current philosophical speedup theories do not accurately encapsulate and distinguish those different sources. Work remains in unifying the source(s) of quantum speedup.

What tool could better uncover and unify sources of quantum speedup than an algorithm which parameterizes essentially all known (efficient) quantum algorithms? Granted the QSVT's grand unifying nature, as discussed in Section 1.3.4, future philosophical work on quantum speedup *must* conform to the QSVT framework. Furthermore, rather than a high-level philosophical theory, the QSVT is an algorithm in itself, possessing immense potential explanatory power. Beyond enabling a deeper understanding of quantum speedup, close analysis of the QSVT could enable systematic discovery of novel efficient quantum algorithms. Beyond philosophical theories, the QSVT could enable a physical or mathematical theory of quantum speedup. In the remaining sections, I present an initial theory for QSVT's speedup.

1.4.2 Quantum Speedup According to the QSVT

Quantum computer scientists have already begun high-level exploration of QSVT's source of speedup. When motivating the QSVT, Martyn et al. (2021) emphasize that a key discrepancy between classical and quantum computation is the highly non-unitary nature of classical computation, yet unitary nature of quantum computation. Fredkin and Toffoli (1982) attempted to reconcile this by demonstrating (with a small overhead in space and time) reversible classical Boolean functions, simulable by quantum circuits. However, prominent quantum algorithms – such as Grover's and Hamiltonian simulation – do not employ a reversible Boolean function embedding. Shor's algorithm requires a reversible embedding to instantiate the circuit input's modular exponentiation, but the core speedup emerges from clever use of the quantum Fourier transform (with no known classical analog) for factoring. Thus, Martyn et al. conclude that quantum speedup does *not* result from quantum implementation of *reversible* classical logic, but instead from the *non-unitary* dynamical behavior of quantum *subsystems*. QSVT leverages this to realize irreversible and non-linear functions.

QSVT's speedup argument is nearly identical to that of BDV [emphasis my own]:

> A quantum algorithm works by exploiting the non-Boolean logic represented by the projective geometry of Hilbert space to *encode* a global property of a *function* (such as a period, or a disjunctive property) as a *subspace in Hilbert space*, which can be *efficiently distinguished* from alternative subspaces, corresponding to alternative global properties, by a measurement (or sequence of measurements) that identifies the target subspace as the subspace containing the final state produced by the algorithm. (Bub, 2010)

Both BDV and the QSVT attribute speedup to computation performed in non-unitary quantum Hilbert subspaces. Thus, among current philosophical quantum speedup theories, QSVT supports BDV.

1.4.3 Extending BDV to Search

Despite similar speedup intuition, the QSVT instantiates all types of quantum algorithms, whereas BDV appears challenging to generalize beyond hidden-subgroup type algorithms (as discussed in Section 1.2.3.3). However, I will now demonstrate that the QSVT OFPAA algorithm, presented in Section 1.3.3, can be mapped into the BDV framework (analogously to Deutsch's algorithm, as presented in Section 1.2.3.2).

As described in Section 1.3.2.3, the QSVT parameterizes quantum algorithms as block-encoded matrices and manipulates their orthogonal subspaces. The unitary matrix $\hat{U}$, block-encoding matrix $\mathbf{A}$, performs rotations between two distinct sets of orthogonal singular vector subspaces: $\{|v_k\rangle, |v_k^{\perp}\rangle\}$ and $\{|w_k\rangle, |w_k^{\perp}\rangle\}$. For search, qubitization mapped the algorithm into the singular vector subspaces spanned by $\{|m\rangle, |m^{\perp}\rangle\}$ and $\{|\psi_0\rangle, |\psi_0^{\perp}\rangle\}$. Although Figure 1.3, in light of qubitization, visualizes the search subspaces as concentric Bloch-spheres, these subspaces can alternatively be visualized as planes, analogously to those of Deutsch's algorithm in Figure 1.1. Mathematically, QSVT search deals with the two quantum logically defined planes,

$$\mathcal{P}_m = \hat{P}_{|m\rangle} \vee \hat{P}_{|m^{\perp}\rangle} = \hat{\Pi}^{(m)} \vee \hat{P}_{|m^{\perp}\rangle} \tag{1.46}$$

$$\mathcal{P}_{\psi_0} = \hat{P}_{|\psi_0\rangle} \vee \hat{P}_{|\psi_0^{\perp}\rangle} = \hat{\Pi}^{(\psi_0)} \vee \hat{P}_{|\psi_0^{\perp}\rangle}, \tag{1.47}$$

with intersection ray

$$\vec{\mathcal{I}}_{m,\psi_0} = \mathcal{P}_m \wedge \mathcal{P}_{\psi_0}. \tag{1.48}$$

Recall that Deutsch's algorithm made $\mathcal{P}_c$ and $\mathcal{P}_b$ as orthogonal as possible (with overlap solely in $\vec{\mathcal{I}}_{b,c}$), so as to distinguish constant from balanced functions. In OFPAA, the goal is opposite – i.e. maximize overlap between $\mathcal{P}_m$ and $\mathcal{P}_{\psi_0}$ in order to map known initial state $|\psi_0\rangle$ onto unknown desired state $|m\rangle$. The QSVT applies $\mathscr{P}_\Theta$ to $|m\rangle$ and $|\psi_0\rangle$'s corresponding singular value a, mapping $|m\rangle$'s measurement probability to 1,

$$P_m = \left| \langle m| \, \hat{U}_{\vec{\phi}} \, |\psi_0\rangle \right|^2 = \left| \mathscr{P}_\Theta \left(a = \frac{1}{\sqrt{N}} \right) \right|^2 \to 1. \tag{1.49}$$

Geometrically, the QSVT efficiently rotates $\mathcal{P}_{\psi_0}$ onto $\mathcal{P}_m$, increasing the conjunction to intersection plane $\mathcal{I}_{m,\psi_0} = \mathcal{P}_{\psi_0} \wedge \mathcal{P}_m$ and maximizing probability of collapse to $|m\rangle$.

1.4.4 Generalizing BDV via the QSVT

The previous section explicitly extended BDV to the realm of search. I now, more generally, argue that any algorithm within the QSVT framework – including all know efficient quantum algorithms – can similarly be mapped to BDV.

At the beginning of Section 1.3.2.3, I noted that the QSVT generalizes prior work on qubitization and QSP. Qubitization block-encodes the algorithm matrix, $\mathbf{A}$, into a larger unitary matrix, $\hat{U}$, with sets of orthogonal singular vector subspaces: $\{|v_k\rangle, |v_k^\perp\rangle\}$ and $\{|w_k\rangle, |w_k^\perp\rangle\}$. The quantum logical disjunction of these singular vectors' projection operators constitute sets of mostly orthogonal planes,

$$\mathcal{P}_v^{(k)} = \hat{P}_{|v_k\rangle} \vee \hat{P}_{|v_k^\perp\rangle} = \hat{\Pi} \vee \hat{P}_{|v_k^\perp\rangle} \tag{1.50}$$

$$\mathcal{P}_w^{(k)} = \hat{P}_{|w_k\rangle} \vee \hat{P}_{|w_k^\perp\rangle} = \hat{\Pi}' \vee \hat{P}_{|w_k^\perp\rangle}, \tag{1.51}$$

with intersections $\mathcal{I}_{v,w}^{(k)} = \mathcal{P}_v^{(k)} \wedge \mathcal{P}_w^{(k)}$. QSVT generalizes the QSP to calculate a gate sequence, $\hat{U}_{\vec{\phi}}$, that applies arbitrary polynomial, $\mathscr{P}(a_k)$, to $\mathbf{A}$'s singular values. Within each set of planes, the singular value a_k governs the $|v_k\rangle \to |w_k\rangle$ transition probability as,

$$P_k = |\, \langle w_k| \, \hat{U}_{\vec{\phi}} \, |v_k\rangle \,|^2 = |\mathscr{P}(a_k)|^2. \tag{1.52}$$

Geometrically, QSVT gate sequence $\hat{U}_{\vec{\phi}}$ rotates the planes within each set – modifying their respective conjunction to maximize the desired outcome's amplitude prior to system measurement. Therefore, while the Hilbert subspace structure of efficient quantum algorithms may not always be as obvious as that of hidden-subgroup algorithms, the QSVT demonstrates that all efficient quantum algorithms can be mapped to BDV.

Not all QSVT parameterizations result in efficient quantum algorithms, but all efficient quantum algorithms are parameterized by the QSVT. Thus, only certain computations (i.e. $\mathscr{P}(a)$) in certain Hilbert subspaces (i.e. $\mathbf{A}$'s singular vector spaces) exhibit quantum speedup. Although BDV and the QSVT have uncovered a probable source of quantum speedup, they do not explain *why* it is the source of speedup. Why does computation in certain non-unitary subspaces of unitary quantum Hilbert spaces outperform computation in inherently non-unitary classical spaces? Bub claims that quantum disjunctive logic enables richer computation than classical computers. However, this is far from a QSVT block-encoding and polynomial selection criterion for efficient quantum algorithms. A deeper understanding of the QSVT speedup could lead to such a criterion, enabling systematic search for efficient quantum algorithms.

1.5 Conclusion

In this work, I presented existing theories for quantum speedup (the QPT, SIV, and BDV), described a novel breakthrough in quantum algorithms (the QSVT), explained how the QSVT serves as a grand unification of quantum algorithms, argued that future philosophical work on quantum speedup must incorporate the QSVT, demonstrated a shared source of quantum speedup in the QSVT and BDV, and used the QSVT to extend BDV to algorithms beyond Bub's original proposal. Granted that quantum speedup has not been rigorously proven, the QSVT offers a unique avenue for exploring why quantum computers appear faster than classical computers for certain computational tasks. A deeper philosophical understanding of quantum speedup via the QSVT could enable systematic discovery of novel efficient quantum algorithms – a holy grail for quantum computation.

Bibliography

Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Barends, R., Biswas, R., Boixo, S., Brandao, F. G., Buell, D. A., et al. (2019). Quantum Supremacy Using a Programmable Superconducting Processor. *Nature*, 574(7779):505–510.

Bennett, C. H., Bernstein, E., Brassard, G., and Vazirani, U. (1997). Strengths and Weaknesses of Quantum Computing. *SIAM Journal on Computing*, 26(5):1510–1523.

Bernstein, E. and Vazirani, U. (1997). Quantum Complexity Theory. *SIAM Journal on Computing*, 26(5):1411–1473.

Berry, D. W., Childs, A. M., Cleve, R., Kothari, R., and Somma, R. D. (2015). Simulating Hamiltonian Dynamics with a Truncated Taylor Series. *Physical Review Letters*, 114(9).

Bub, J. (2010). Quantum computation: Where does the speed-up come from. *Philosophy of Quantum Information and Entanglement*, pages 231–246.

Childs, A. M. (2009). Universal Computation by Quantum Walk. *Physical Review Letters*, 102(18).

Chuang, I. (2020). Grand Unification of Quantum Algorithms. *University of Waterloo Institute for Quantum Computing Colloquium.*

Cleve, R., Ekert, A., Macchiavello, C., and Mosca, M. (1998). Quantum Algorithms Revisited. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):339–354.

Dalzell, A. M., Yoder, T. J., and Chuang, I. L. (2017). Fixed-Point Adiabatic Quantum Search. *Physical Review A*, 95(1):012311.

Deutsch, D. (1985). Quantum Theory, the Church–Turing Principle and the Universal Quantum Computer. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 400(1818):97–117.

Deutsch, D. (1998). *The Fabric of Reality*. Penguin UK.

Deutsch, D. and Jozsa, R. (1992). Rapid Solution of Problems by Quantum Computation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 439(1907):553–558.

Devoret, M. H. and Schoelkopf, R. J. (2013). Superconducting Circuits for Quantum Information: An Outlook. *Science*, 339(6124):1169–1174.

Duwell, A. (2007). The Many-Worlds Interpretation and Quantum Computation. *Philosophy of Science*, 74(5):1007–1018.

Duwell, A. (2018). How to Make Orthogonal Positions Parallel: Revisiting the Quantum Parallelism Thesis. *Physical Perspectives on Computation, Computational Perspectives on Physics*, pages 83–102.

Ekert, A. and Bouwmeester, D. (2000). *The Physics of Quantum Information: Quantum Cryptography, Quantum Teleportation, Quantum Computation*. Springer.

Farhi, E., Goldstone, J., Gutmann, S., and Sipser, M. (2000). Quantum Computation by Adiabatic Evolution. *arXiv preprint quant-ph/0001106*.

Feynman, R. P. (2018). *Simulating Physics with Computers*. CRC Press.

Fredkin, E. and Toffoli, T. (1982). Conservative Logic. *International Journal of Theoretical Physics*, 21(3):219–253.

Gilyén, A., Su, Y., Low, G. H., and Wiebe, N. (2019). Quantum Singular Value Transformation and beyond: Exponential Improvements for Quantum Matrix Arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 193–204. Association for Computing Machinery.

Grover, L. K. (1996). A Fast Quantum Mechanical Algorithm for Database Search. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing*, STOC '96, page 212–219, New York, NY, USA. Association for Computing Machinery.

Harrow, A. W., Hassidim, A., and Lloyd, S. (2009). Quantum Algorithm for Linear Systems of Equations. *Physical Review Letters*, 103(15):150502.

Lloyd, S. (1996). Universal Quantum Simulators. *Science*, pages 1073–1078.

Low, G. H. and Chuang, I. L. (2019). Hamiltonian Simulation by Qubitization. *Quantum*, 3.

Low, G. H., Yoder, T. J., and Chuang, I. L. (2016). Methodology of Resonant Equiangular Composite Quantum Gates. *Phys. Rev. X*, 6.

Martyn, J. M., Rossi, Z. M., Tan, A. K., and Chuang, I. L. (2021). Grand Unification of Quantum Algorithms. *PRX Quantum*, 2.

Minch, M. J. (1998). Spin Choreography: Basic Steps in High Resolution NMR (by Ray Freeman). *Journal of Chemical Education*, 75(2):155.

Preskill, J. (2018). Quantum Computing in the NISQ Era and Beyond. *Quantum*, 2:79.

Shor, P. W. (1994). Algorithms for Quantum Computation: Discrete Logarithms and Factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134. IEEE.

Simon, D. R. (1997). On the Power of Quantum Computation. *SIAM Journal on Computing*, 26(5):1474–1483.

Steane, A. M. (2003). A Quantum Computer Only Needs One Universe. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 34(3):469–478.

Vandersypen, L. M. K. and Chuang, I. L. (2005). NMR Techniques for Quantum Control and Computation. *Reviews of Modern Physics*, 76:1037–1069.

Von Neumann, J. (1932). *Mathematische Grundlagen der Quantenmechanik*, volume 38. Springer-Verlag.

Wimperis, S. (1994). Broadband, Narrowband, and Passband Composite Pulses for Use in Advanced NMR Experiments. *Journal of Magnetic Resonance, Series A*, 109(2):221–231.

Wolfowicz, G. and Morton, J. J. (2016). *Pulse Techniques for Quantum Information Processing*, pages 1515–1528. John Wiley & Sons, Ltd.

The asymmetry in time of the world about us is one of the most fundamental aspects of our experience, and it is necessary to inquire about the origin of this asymmetry.

— Davies (1977)

2

A Low-Entropy Big Bang is Not Needed to Fry an Egg

Contents

2.1 Introduction

Statistical mechanics aims to explain the macroscopic predictions of thermodynamics via a system's underlying microscopic dynamics. Central to the phenomenological laws of thermodynamics is the irreversible process, which introduces a temporal-asymmetry, generally exhibited via entropy increase. We experience this temporal-asymmetry in our everyday lives – e.g. eggs fry in pans, ice-cubes melt in warm water, and milk dissolves in coffee, but not vice-versa. However, system micro-states, whether governed by classical Newtonian mechanics or quantum theory, are described by temporally-symmetric laws. How, then, can statistical mechanics predict thermodynamic macro-phenomena? Clearly, there must be some sort of temporal-asymmetry, but its emergence is contentious – with significant epistemic and metaphysical implications.

In this work, I will introduce a common argument for statistical mechanical temporal-asymmetry, known as the Past Hypothesis – which relies on constrained universal thermodynamic conditions at the Big Bang. However, I argue that this theory is not well-justified, facing challenges ranging from ill-defined entropy to questionable cosmological evidence. Rather than requiring a low-entropy Big Bang to model the frying of an egg, I will assume David Parker's stance that a 'local' hypothesis is necessary. Thus, I conclude by championing the merits of the Branching Hypothesis – which accounts for statistical mechanical temporal-asymmetry via temporary separation of quasi-isolated systems from a main environment.

2.2 Thermodynamics, Statistical Mechanics, and Reversibility

In this section, I concisely review necessary background from thermodynamics and statistical mechanics, enabling discussion of reversibility and time-asymmetry.

2.2.1 The Laws of Thermodynamics

Thermodynamics is a "phenomenological science", with parameters over macroscopic, measurable variables – such as pressure, volume, and temperature. A treatment of classical thermodynamics begins with two empirically tested and adequate laws (Callender, 2021):

The **First Law** of thermodynamics expresses conservation of energy. For thermally isolated systems, work (W) transferred to the system's surroundings is compensated by internal energy (U) loss, as $dW = -dU$. For non-isolated systems, this is extended to $dQ = dU + dW$, where dQ is the reversible heat differential. This conservation of energy has no temporal-asymmetry and enables the First Law to eliminate perpetual mobiles of the 1st kind – those which produce work without energy. However, it does not rule out perpetual mobiles of the 2nd kind – those which spontaneously convert thermal energy into work – necessitating the Second Law of thermodynamics.

The **Second Law** of thermodynamics, as described by Kelvin in the context of steam engines, stated that there exists no thermodynamic process which solely transforms heat (extracted from a source at uniform temperature) completely into work. Similarly, Clausius described the Second Law as a thermodynamic process' inability to solely extract heat from a colder reservoir and deliver it to a hotter reservoir. Note that the Second Law applies only to *isolated* systems. For a reversible, quasi-static transformation from arbitrary state fixed state O to state A, Clausius defined **thermodynamic entropy** as

$$S(A) = \int_O^A \frac{dQ}{T}, \tag{2.1}$$

with temperature T. Given the quasi-static nature of this process, O and A must both be equilibrium states. Using this notion of entropy, the Second Law claims that, for equilibrium states B and C,

$$\Delta S = S(C) - S(B) \geq \int_C^B \frac{dQ}{T}, \tag{2.2}$$

where equality is achieved only for a reversible process. In other words, entropy cannot decrease during spontaneous evolution of a thermally closed system and attains its maximum value at equilibrium.

2.2.2 Statistical Mechanics

Statistical mechanics aims to resolve the thermodynamic macroscopic system state (macro-state) with its underlying microscopic state (micro-state). In the late 1800s, Boltzmann and Gibbs developed two independent statistical mechanical frameworks.

Suppose we have a 3D box filled with N gas particles. In **Boltzmannian statistical mechanics**, the system micro-state is governed by the 3D position ($\vec{q_i}$) and momentum ($\vec{p_i}$) vectors of every particle (i) in the box. The full system micro-state is mathematically represented as vector $\vec{X} = (\vec{p}_1, ..., \vec{p}_N, \vec{q}_1, ..., \vec{q}_N)$ in the 6N-dimensional 'Γ-space' phase-space. Since many micro-states can exhibit the same macroscopic properties, the micro- to macro-state mapping is surjective. Thus, each macro-state, M, consists of multiple Γ-space points and has phase-space volume $|\Gamma_M|$. The system's instantaneous **Boltzmann entropy** is defined as,

$$S_B(M(X)) = k_B \log(|\Gamma_M|), \tag{2.3}$$

where $M(X)$ is the macro-state corresponding to the instantaneous system micro-state and k_B is Boltzmann's constant. Granted combinatorial coarse-graining arguments, the vast majority of Γ-space is dominated by the equilibrium macro-state's volume. Therefore, *entropy is maximized at equilibrium.*

Furthermore, granted assumptions of ergodicity (or otherwise unstructured motion around phase-space), any state initially in a small-volume, non-equilibrium macro-state will rapidly move into the large-volume, equilibrium macro-state. Thus, the *entropy of non-equilibrium states is likely*[1] *to increase in time.* Once in equilibrium, there is a small, non-zero probability that the system will move back out of equilibrium. In this way, entropy increase is not deterministic, but a probabilistic process or tendency.

In place of a singular micro-state traversing Γ-space, **Gibbsian statistical mechanics** defines a probability distribution, ρ, over all possible micro-states. This probability distribution encodes system likelihood of producing a given micro-state and defines the **Gibbs entropy** as

$$S_G(\rho) = -k_B \int_{\Omega_E} \rho(\Gamma) \, \log(\rho(\Gamma)) \, d\Gamma, \tag{2.4}$$

[1] There is no guarantee that the system will immediately move into equilibrium.

where Ω_E is the set of all possible system micro-states. Since equilibrium is defined by constancy of observed system macro-states, the equilibrium distribution is stationary. Finally, the Gibbsian max entropy principle requires ρ maximize S_G – under system constraints such as constant energy and/or particle number – enabling derivation of canonical equilibrium distributions.

2.2.3 Reversibility

Thermodynamics' exact time-asymmetry source is disputed. While most claim it arises in the Second Law, some – including Uffink (2001) and Brown and Uffink (2001) – argue it is an altogether separate assumption (i.e. the "Minus First Law"). Overall, however, spontaneous movement from non-equilibrium to equilibrium, or lower to higher entropy, is assumed and accepted. Furthermore, irreversible processes lie at the heart of thermodynamics. Thus, I take for granted the existence of thermodynamic temporal-asymmetry.

Statistical mechanics, on the other hand, seems incapable of accounting for this temporal-asymmetry. Traditional statistical mechanics is microscopically governed by reversible Newtonian mechanics – i.e. the system's energy function, Hamiltonian $H(\vec{q}_1, ..., \vec{q}_N, \vec{p}_1, ..., \vec{p}_N)$. The system micro-state evolves according to Hamilton's equations of motion,

$$\frac{d\vec{q}_i}{dt} = \frac{\partial H}{\partial \vec{p}_i} \ , \ \frac{d\vec{p}_i}{dt} = -\frac{\partial H}{\partial \vec{q}_i}, \tag{2.5}$$

begging the question of how temporal-asymmetry arises in statistical mechanics?[2]

Two well-established arguments – the **reversibility objections** – challenge statistical mechanical ability to adequately describe thermodynamics, due to lack of temporal-asymmetry. The **Loschmidt paradox** states that, for time-reversible statistical mechanics, any forward-process compatible with the laws must have a corresponding, compatible backwards-process. Thus, entropy-decreasing processes can be no less natural or statistically common than entropy-increasing ones (Loschmidt, 1867). This implies that an ice-cube in warm water is as likely to melt as a portion of the warm water is to freeze into an ice-cube. **Zermello's recurrence** further establishes that any classical system confined to a finite phase-space region will invariably return (arbitrarily close) to its initial conditions (Zermelo, 1896). This implies that if a fried egg survived long enough, it would un-fry.

2.3 Explanations for and Implications of Entropy Increase

In light of these troubling findings, I will present Albert (2001)'s reconciliation of the reversibility arguments with Boltzmannian and Gibbsian statistical mechanics – known as the Turning Argument. I will then demonstrate how the resulting Past Hypothesis leads to worries of "Boltzmann brains" and insufficient justification of a low-entropy past.

[2]Although modern treatment of microscopic physics necessitates quantum theory, it is also time-reversible (governed by unitary operations) and will not affect this work's discussion. For simplicity, I will focus on the classical case, but note that Wallace (2011) believes subtle differences between classical and quantum statistical mechanics can have implications for the Past Hypothesis.

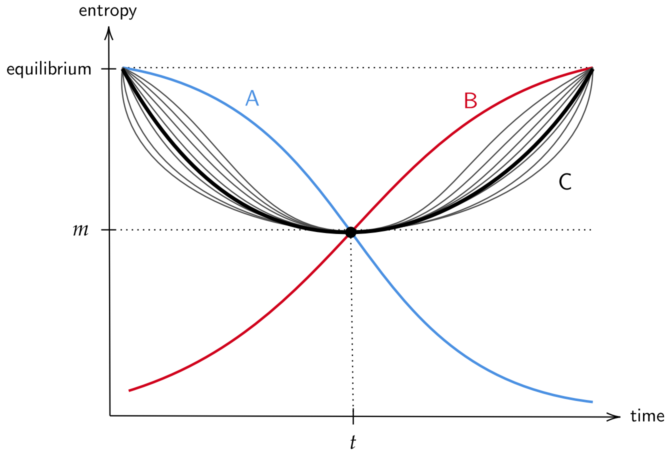

Figure 2.1: If a system is at non-maximal entropy m at time t, most physically-possible trajectories will curve in both temporal directions towards equilibrium (C), not away (A,B). [Inspired by (Albert, 2001) Fig.4.1]

2.3.1 The Turning Argument

Albert claims that the **Turning Argument (TA)** reconciles statistical mechanics and the reversibility objections. As discussed in Section 2.2.2, Boltzmannian and Gibbsian micro-states tend towards the equilibrium macro-state. According to Albert, this entails that "the *overwhelming majority* of the trajectories passing through any particular non-maximal-entropy macrocondition *increase their entropies toward the future*" and, in light of the Loschmidt paradox, "the *past*!" Therefore, every non-maximal-entropy microcondition is a 'turning point' for physically possible trajectories passing through it. While some trajectories may continue decreasing into the future or past, each entropy decrease causes significantly more trajectories to 'turn' and increase in entropy (illustrated in Figure 2.1). Therefore, continual entropy decrease becomes exponentially harder over time.

In Albert's words, "only an unimaginably tiny minority of the physically possible trajectories" will follow a continuously increasing or decreasing trajectory (like Figure 2.1's A and B trajectories). This may seem alarming. Suppose I drop an ice-cube into warm water at time $t = 0$. Consistent with the TA, I expect entropy increase into the future, producing a half-melted ice-cube at at $t = 5$ and a fully-melted ice-cube equilibrium state at $t = 10$. However, looking retrospectively from $t = 5$, Albert's TA implies a similar entropy increase, meaning the ice should be fully-melted at $t = 0$ – a clear contradiction to the starting supposition and real-world experience!

2.3.2 The Past Hypothesis

Albert argues that, by constraining the universe's initial state, the **Past Hypothesis (PH)** resolves this contradiction, aligning statistical mechanical predictions with real-world experience. His argument is rooted in an example consisting of a 'pinballish' ice-machine and several warm water glasses (depicted in Figure 2.2).

Assume, at $t = 0$, the glasses are filled with half-melted ice-cubes. By the TA, at $t = 5$ (future), all ice should be fully melted. Similarly, by the TA, entropy increases from $t = 0$ to $t = -5$, meaning all ice was also fully melted at $t = -5$ (past). However, suppose our real-world memory tells us that, at $t = -5$, the ice-cubes were fully un-melted in the glasses. To

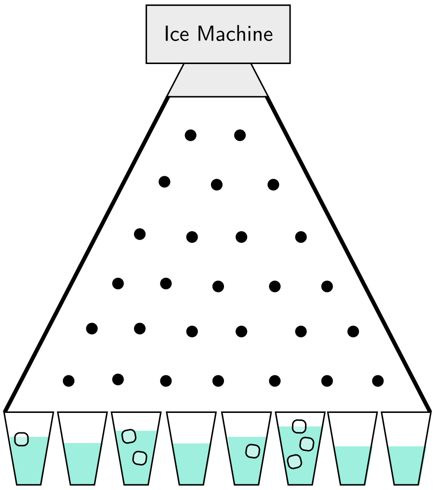

Figure 2.2: Albert's pin-ball ice machine. [Inspired by (Albert, 2001) Fig.4.2]

remedy the TA and match the real-world memory, Albert suggests we instead begin our entropy analysis from the observed $t = -5$ macro-state. Clearly this will lead to accurate predictions for $t \geq -5$, but problematic ones for $t < -5$ (since TA predicts the glasses ice will be half-melted at $t = -10$ and fully-melted for $t \leq -15$).

Thus, Albert once again suggests moving our starting point back, this time to $t = -15$. In this case, however, the outcome is different. At $t = -15$, fully-frozen ice-cubes are observed atop the pinballish device, ready to fall into the glasses. The overall system macrocondition has a slightly lower temperature than the macroconditions at $t = -5$ or $t = 0$ (energy will be gained as ice falls). Furthermore, this starting condition does not guarantee that the ice-cubes will fall into the same configuration, producing overwhelmingly large probability of a different macro-state (each macro-state having relatively little probability). However, starting at $t = -15$ ensures that macroscopic system properties (such as overall energy and how melted the ice-cubes are) will be conserved across most possible $t = -5$ and $t = 0$ macro-state outcomes. Albert argues this constitutes "a fully satisfactory probabilistic theory of the history of this system beginning [15] minutes ago". However, this pinball-ice theory fails anytime $t \leq -15$. To ensure the theory is correct at *any* previous time, Albert concludes "all such posits are bound to fail – unless they concern nothing less than the entirety of the universe at nothing later than its beginning". In conclusion, Albert posits that the universe started in an *extremely* low-entropy macro-state, preventing universal entropy decrease following the Big Bang.

2.3.3 Boltzmann Brains

Let's take a step back. Suppose the current universal macro-state, P, has a uniform distribution, D_P, over all possible corresponding phase-space micro-states, M_P. By the TA alone (without postulating a PH), this setup is only compatible with our future expectations. Past memories and artifacts appear to falsify TA predictions of entropy increase towards the past. However, all this evidence (e.g. my mental state in conjuring a memory or a photograph of my great grandmother) is fully described by P. As illustrated in Figure 2.1, the majority of trajectories leading into P

come from higher-entropy states and are in the process of turning. A trajectory from the initial low-entropy state – supported by our memories and artifacts – is *so* unlikely that Albert admits such a memory or artifact probably "formed, spontaneously, as a matter of pure chance".

Generally speaking, everything we believe to have previously experienced – all our thoughts and memories – are far more likely byproducts of a random, instantaneous collection of molecules in a high-entropy universal macro-state. As described by Albrecht and Sorbo (2004), if the universe spent eternity in heat death, eventually a rare thermal fluctuation would spawn a **Boltzmann brain** – indistinguishable from an intelligent observer – which perceives the 'present' and all our 'past memories'. Such a Boltzmann brain fluctuation, is far more likely than the fluctuation necessary to produce a low-entropy system at the scale of the observable universe. To reconcile this, Albert updates the PH – requiring, not only a low-entropy initial macro-state, but also a restricted phase-space distribution that ensures the universe evolves to its current macro-state. While this phase-space distribution prevents a Boltzmann brain catastrophe within the PH, it does not actually justify the PH. Therefore,

> the flip-side of the insight of Boltzmann and Gibbs is that there can be nothing at all about the present macrocondition of the world which can possibly count as evidence that the world's entropy has ever previously been lower. (Albert, 2001)

2.4 Further Challenges for the Past Hypothesis

In light of this Boltzmann brain *reductio ad absurdum*, the PH does not appear justified in its claims about past entropy. In the following section, I will raise further potential concerns with the PH, challenging its: (1) notion of universal entropy, (2) universal scope of entropy increase, (3) desired cosmological evidence, and (4) explanatory power.

2.4.1 Is Universal Entropy Well-Defined?

The PH relies on a definition of the universe's entropy. Assuming a finite universe[3], Boltzmannian or Gibbsian statistical mechanics can define universal entropy via a course-graining or probability distribution over Γ-space's description of every atom in the universe. However, many philosophers believe universal *thermodynamic* entropy is not well-defined. For example, Planck (1897) argues it is simply undefined. The integral of Equation (2.1) defines thermodynamic entropy between two different system states. However, if system A represents the state of the whole universe, where could it possibly absorb heat from? Even if the universe could be treated as an arbitrary adiabatic isolated system, Uffink (2001) is skeptical that such an irreversible process could be closed by a reversible process to form a cycle (as needed to define entropy between initial and final states). Historically, thermodynamics was proposed to explain the macroscopic workings of steam engines. While, admittedly, thermodynamics has proven successful for systems far more complex than steam engines, extrapolating to a universal scale seems unwarranted.

[3]In the infinite case, Reichenbach (1956) argues statistical mechanical probability and entropy cannot be defined.

2.4.2 Entropy Beyond the Observable Universe?

Even if universal entropy was well-defined, why must entropy increase at the universal scale? Feynman (1965) argues that,

> from the hypothesis that the world is a fluctuation, all of the predictions are that if we look at a part of the world we have never seen before, we will find it mixed up, and not like the piece we just looked at. If our order were due to a fluctuation, we would not expect order anywhere but where we have just noticed it. We therefore conclude that the universe is not a fluctuation, and that the order is a memory of conditions when things started.

In other words, Feynman believes that, because the observable universe possesses the same low-entropy dynamics as we experience, the whole universe must be low-entropy – with the PH as the only possible justification.

This is a bold extrapolation. Granted that we can only measure a small portion of the universe (the 'observable' universe), we cannot dismiss the possibility of future observations revealing high-entropy dynamics beyond the current observable universe. In fact, Price (2004) claims that we should not expect the low-entropy region of the universe "to be any more extensive in space than we already know it to be" – meaning we may inhabit a low-entropy pocket of the universe. Even if this low-entropy pocket was unlikely, a low-entropy environment is crucial for the survival of organisms like ourselves. Therefore, were such a pocket to exist, it would be unsurprising that we would inhabit it. In conclusion, the PH's universal macro-state is underdetermined, posing a challenge for realist commitment to the theory.

2.4.3 Cosmological Evidence?

Like many philosphers, such as Carroll (2020), Albert is eager to dismiss the Boltzmann brain argument. However, as discussed in Section 2.3.3, the additional restrictions placed on the universe's initial phase-space were *ad hoc*. Thus, the PH appears solely justified in preventing skeptic catastrophe. This, however, is unsatisfying – leading many to seek justification for the PH in the form of cosmological evidence.

In fact, Price (2004) urges astronomers to determine cosmological conditions – during and prior to the Big Bang – for the sake of justifying the PH. However, this proposal expresses a fundamental lack of scientific objectivity. Furthermore, any cosmological observation made regarding the Big Bang faces significant verifiability challenges. In addition to typical astronomy hardships that come with indirectly observing distant objects and phenomena (via theory-laden apparatuses), claims deduced regarding the early universe cannot and never will be directly verifiable by human experience. Thus, cosmological evidence alone appears unsubstantial in justifying the PH. Earman (2006) endorses this view when describing several contentious, modern cosmological theories and findings.

2.4.4 Parker's Argument for a 'Local' Hypothesis

Although there exist further PH objections, I conclude with Parker (2005)'s presentation of, arguably, the most intuitive, yet rigorous criticism. Parker attacks the PH's explanation for everyday thermodynamic experiences:

> The informality of the argument aside, it appears implausible that the mere stipulation of a nonequilibrium state of the universe somewhere in the distant past could justify my memory of an unmelted ice-cube ten minutes ago, somehow make it altogether improbable that the ice cube formed as a spontaneous fluctuation ten minutes ago, and testify to the veracity of any records to that effect.

He conjectures that the PH can only be a useful theory if conditionalizing on the present state of a glass-ice system and the universe, as well as the universe's starting condition, produces a high probability of un-melted ice 10 minutes ago. Mathematically, the PH must satisfy the inequality

$$\mathbb{P}(\, U \mid H \,\&\, B \,\&\, M\,) > \frac{1}{C}, \tag{2.6}$$

where B is the phase-space region containing all micro-states compatible with the universe's initial low-entropy macro-state; U is the region compatible with an unmelted ice-cube 10 minutes ago (the PH); H is the region compatible with a currently half-melted ice-cube; M is the region compatible with the macro-state of the rest of the universe (excluding the ice and water)[4]; and C is a positive constant[5].

Through simple manipulation of conditional probabilities, paired with insight about the PH and phase-space, Parker shows that Equation (2.6) reduces to

$$\mathbb{P}(\, B \,|\, U \,\&\, M\,) \gg \mathbb{P}(\, B \,|\, H \,\&\, M\,). \tag{2.7}$$

Equation (2.7) posits that a low-entropy Big Bang is more likely if the ice-cube was un-melted 10 minutes ago than if it is currently half-melted. However, Albert's TA claims that nothing in the present universe could justify the universe being further from equilibrium than it is now. Therefore, the PH does not satisfy Equation (2.7), nor (2.6).

Parker suggests the PH lacks explanatory power because Albert's simplified pinballish ice-machine example does not generalize to a universal scale. Specifically, universal time magnitudes are larger and universal subsystems interact (unlike the glasses of water). Furthermore, the transition from unmelted to melted ice-cubes in the pinballish ice-machine example (a 'local' PH) is likely, while the transition from the Big Bang macro-state to our present universe's macro-state (Albert's 'global' PH) is not. In fact, Parker demonstrates that the pinball ice-machine example satisfies the original inequality of Equation (2.6). In this case, M fixes the system of interest's history, such that H and U can be discarded, resulting in

$$\frac{\mathbb{P}(\, B \,\&\, M\,)}{\mathbb{P}(\, B \,\&\, M\,)} = 1 > \frac{1}{C}. \tag{2.8}$$

Thus, Parker argues that we should not conditionalize on initial conditions of the universe to explain irreversible processes. Instead, a *local* theory is needed.

2.5 A 'Local' Hypothesis

I conclude with a discussion of Reichenbach (1956) and Davies (1977)'s alternative, local theory of statistical mechanical temporal-asymmetry, known as the **Branching Hypothesis (BH)**. I defend the BH against Albert's objection, Boltzmann brains, and other challenges faced by the PH. With this, I argue that the BH provides a justified explanation for both past and future entropy.

[4]Note: $M \cap H$ is the phase-space region compatible with the entire, current universe.
[5]Parker argues for $C \geq 2$.

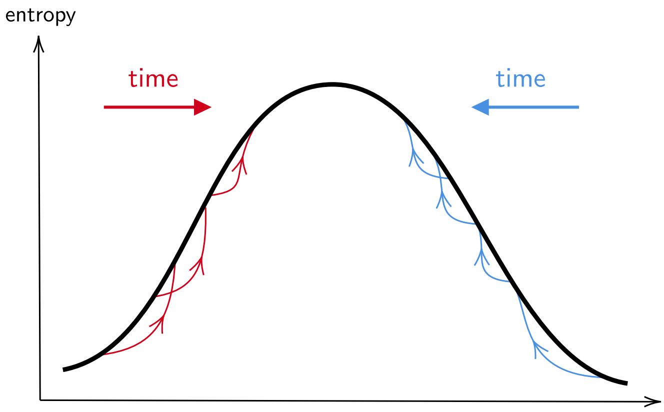

Figure 2.3: Branch systems (red and blue curves) temporarily separate from the main environment (primary black curve), defining time in the direction of entropy increase. [Inspired by (Reichenbach, 1956) Fig.21]

2.5.1 The Branching Hypothesis

The BH argues that the key to an arrow-of-time is **isolation** – or lack thereof it. As described in Section 2.2.1, the Second Law of thermodynamics assumes all involved systems are isolated. In practice, however, this is an extremely restricting assumption. Davies notes that: (1) It is impossible to experimentally produce perfectly isolated systems – at best, we can hope for quasi-isolated systems with relaxation times short enough to prevent noticeable disturbance by the environment. (2) Even if a perfectly isolated system was possible, the system could not be isolated for all time. Thus, Reichenbach argues that low-entropy systems should be modeled as "subsystems of comprehensive systems". The comprehensive system's total entropy increases, while the subsystem of interest is put into a state of *relatively* low entropy[6]. For example, frying an egg requires heat flow from the stove-top and energy dissipation by the chef's muscles. These quasi-isolated systems that separate from the main, higher-entropy environment are known as **branch systems**. For example, in the previous ice-cube in warm water example, the ice-water system only came into existence once the ice-cube was submerged in the water and began melting. This is a quasi-closed system because the melting can be explained to high-accuracy[7] without reference to anything beyond the ice and water.

Reichenbach's key insight is relating the evolution of branch systems *ensembles* to the mixing processes of gas molecule ensembles. Through manipulations of probability lattices[8], he demonstrates that ensembles of branch systems, unlike temporal ensembles of singular isolated systems (as referenced in the TA), defy Loschmidt's paradox. Thus, when a branch system breaks off from the main environment, Reichenbach shows that "the probability that a low-entropy state is followed by a state of high entropy is greater than the probability that the same low-entropy state is preceded by a state of high entropy."

As depicted in Figure 2.3, branch systems usually only separate from the main environment for a limited time. For example, someone will eventually drink a glass of melted ice-water, or it will

[6]Here, "relative" is with respect to other configurations of the subsystem, not the full universe.

[7]A more accurate explanation would require incorporating more of the environment – e.g. simulating air surrounding the glass that will cool.

[8]For the sake of brevity, I do not explain these derivations in this work. The interested reader can refer to (Reichenbach, 1956) Chapter 14.

evaporate. Thus, a primary curve represents the main environment's entropy[9], with small branch systems constantly breaking off and recombining to explain everyday, observed lower-entropy phenomena. Reichenbach argues that the direction of these quasi-isolated branch systems define the 'positive' direction of time. Increases in the primary curve's entropy support our normal intuition that time moves in the 'forward' direction. However, even when the primary curve's entropy decreases, branch systems still trend towards higher entropy, thus time is defined in the 'backward' direction. Therefore, time is not consistently defined and changes with the environment. Reichenbach notes that such environmental changes from increasing to decreasing entropy, or vice versa, are separated by periods of entropy too high or low to support human life. Thus, in line with our experiences, we could never experience a time direction 'flip'.

2.5.2 Albert's Argument Against the Branching Hypothesis

Before arguing for its merits, I will briefly dispel Albert (2001)'s key concern with the BH. Albert does not give the BH much credit – decrying it as "sheer madness". His only objection to the hypothesis is presented as a series of rhetorical questions:

> How is it (to begin with) that we are to decide at *exactly* what moment it was that the glass of water with ice in it first came into being?...How is it (exactly) that the medium-sized system we decided to focus on was the glass of water with the ice in it and not (say) the *room* in which that glass is currently *located*, which also contains the table on which the glass is currently sitting, and the freezer from which the ice was previously removed, and the person who first got it into his head to do the removing? The uniform probability-distribution over the possible microconditions of the macrocondition of *that* system...[will] differ quite radically...from the one we have just been talking about!

In other words, Albert claims branch systems are ill-specified and changing their scope would result in drastically different micro- and macroconditions.

However, as argued by Reichenbach, if we had a full specification of, say, every particle's position and velocity in a gas chamber, we could deterministically calculate the particles' motion backwards in time and verify that their entropy does, in fact, decrease. However, we cannot obtain the full microcondition of non-trivial systems of interest, necessitating statistical mechanical assumptions and approximations. In this way, I argue that the start, end, and scope of a branch system is no more arbitrary than, say, Boltzmannian coarse-graining or Gibbsian probability distributions.

If anything, the BH more closely mimics practical scientific intuitions than the PH. Reichenbach notes that "by 'isolated' we do not mean complete isolation; it is sufficient if the process within the subsystem represents energy exchanges which are large compared with the interaction with the environment". If a reasonable branch start and end – e.g. when the ice cube is dropped in the glass and when the ice-water is drunk – are chosen, as well as a reasonable system scope – e.g. the glass containing the water and the ice – this will enable decent calculations. In making real-world predictions, physicists do not model the entire universe's phase-space, but narrow their scope to the system of interest. The Big Bang should not be necessary to calculate the entropy of a frying egg. Of course, widening the time interval of or incorporating more of the environment (e.g. modeling air particles hitting off the ice and dispersing back around the room) into a branch system could enable more accurate calculations. However, accuracy gains would be negligible relative to the exponentially increasing problem complexity as system scope widens.

[9]Note that this environment does not need to be the full universe, but could for example be the observable universe. In fact, as in Section 2.4.1, Reichenbach argues that entropy is not even defined at the full scale of the universe.

2.5.3 No Need for a Past Hypothesis

Beyond the previous concern, Albert simply asks "why in God's name bother with all [these branch systems], when the uniform probability-distribution over the possible microconditions compatible with the macrocondition of the world, at the moment when it came into being, will very straightforwardly give us everything we need?" However, in Sections 2.3.3 and 2.4, I presented several clear challenges to justification of the PH: (C1) Boltzmann brains, (C2) ill-defined universal entropy, (C3) unknowable entropy beyond the observable universe, (C4) problems with cosmological evidence, and (C5) need for a more local hypothesis. I now argue that the BH avoids all these issues, warranting explanatory merit.

Albert's TA suggested a low-entropy system is most likely at an entropy minimum (i.e. entropy will increase both forwards and backwards in time). However, backwards entropy increase is inconsistent with our everyday experience and, thus, problematic. Parker (2005) argues that "if a branch system is *formed* in a random low entropy state, it simply did not exist in the 'past' for entropy to increase that way...It is through branch systems that the customary intuitive notion that entropy increases with time is derived". By evading Loschmidt's paradox, the BH: does not necessitate the TA, remains consistent with our everyday observations, and does not imply that the current universe is unlikely or arose from a statistical fluctuation – avoiding Boltzmann brains (C1). Furthermore, the BH – concerned only with a quasi-isolated system's initial conditions and not those of the universe – satisfies Parker's criterion for a 'local' theory (C5). An individual branch systems' scope is far smaller than that of the universe. Even the main environment, from which these subsystems break off, need not be the scope of the full universe (C3). This removes worries about a definition of universal entropy (C2) and eliminates need for cosmological evidence of the universe's distant past (C4).

In a final, futile attempt at justification, Albert argues that the PH deserves law-like status, in light of its successful empirical predictions:

> And so the fact that the universe came into being in an enormously low-entropy macrocondition cannot possibly be the sort of fact that we know, or ever *will* know, in the way we know of straightforward everyday particular *empirical* facts. We know it *differently*, then. Our grounds for believing it turn out to be more like our grounds for believing general theoretical *laws*. Our grounds (that is) are *inductive*; our grounds have to do with the fact that the [PH]...turns out to be enormously helpful in making an enormous variety of particular empirical *predictions*.

However, his examples of successful predictions – i.e. high likelihood of digging a second boot out of the sand and the prior existence of Napoleon – are similarly (if not more easily and intuitively) predicted by the BH. Additionally, similar to Section 2.5.2, the PH's empirical merits appear irrelevant to practical day-to-day use of statistical mechanics by scientists. The PH relies upon the universe's initial entropy, but, most likely, no scientist will ever definitively claim such knowledge (as argued in Section 2.4.3). This, however, has not barred successful application of statistical mechanics. A scientist using statistical mechanics to model a frying egg does not start calculations from the Big Bang, but instead from the inception of the egg-pan-stove system. A PH proponent could protest – arguing scientists need not explicitly reference the Big Bang, since the PH is already implicitly embedded in the arrow-of-time assumption. However, there exist alternative explanations (i.e. the BH) for the statistical mechanical arrow-of-time.

Let's put aside the extensive arguments for the BH, presented in this work. If all were equal, why choose the PH – which relies on the unlikely occurrence of an extremely low-entropy Big Bang and constraints ensuring its unlikely evolution to the present universe – over the BH – which utilizes temporarily quasi-isolated subsystems that are observed in our daily experiences and match practical scientific intuition? In the words of Reichenbach, "[i]f we wish to find a way of defining a direction of time, it is advisable to study the actual procedure which is used when inferences concerning time direction are made."

Bibliography

Albert, D. Z. (2001). *Time and Chance.* American Association of Physics Teachers.

Albrecht, A. and Sorbo, L. (2004). Can the Universe Afford Inflation? *Physical Review D*, 70(6).

Brown, H. R. and Uffink, J. (2001). The Origins of Time-Asymmetry in Thermodynamics: The Minus First Law. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 32(4):525–538.

Callender, C. (2021). Thermodynamic Asymmetry in Time. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, Summer 2021 edition.

Carroll, S. M. (2020). Why Boltzmann Brains are Bad. In *Current Controversies in Philosophy of Science*, pages 7–20. Routledge.

Davies, P. C. W. (1977). *The Physics of Time Asymmetry.* University of of California Press.

Earman, J. (2006). The "Past Hypothesis": Not Even False. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 37(3):399–430.

Feynman, R. P. (1965). The Feynman Lectures on Physics. *American Journal of Physics*, 33(9):750–752.

Loschmidt, L. (1867). *Sitznngsberichte der Akademie der Wissenschaften in Wien*, 73.

Parker, D. (2005). Thermodynamic Irreversibility: Does the Big Bang Explain What it Purports to Explain? *Philosophy of Science*, 72(5):751–763.

Planck, M. (1897). *Vorlesungen über Thermodynamik, von dr. Max Planck.* Veit & Comp.

Price, H. (2004). On the Origins of the Arrow of Time: Why There is Still a Puzzle About the Low-Entropy Past. *Contemporary Debates in Philosophy of Science*, pages 219–239.

Reichenbach, H. (1956). *The Direction of Time*, volume 65. University of California Press.

Uffink, J. (2001). Bluff Your Way in the Second Law of Thermodynamics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 32(3):305–394.

Wallace, D. (2011). The Logic of the Past Hypothesis. *The Probability Map of the Universe: Essays on David Albert's Time and Chance.*

Zermelo, E. (1896). *Annalen der Physik*, 57.

To my own taste the philosophical position which is, or should be, of most interest to the physicist is the realist one, although again one must be careful to distinguish different answers to the question, realist about what? Is it the entities, the abstract structural relations, the fundamental laws or what?

— Michael Redhead (1999)

3
"Relationships All the Way Down"

Contents

3.1 Introduction

Core to the philosophy of science is how our scientific theories relate to the real world. Namely, do the entities, principles, properties, relations, phenomena, actions, and structures described by scientific theories truly exist? Or are they simply useful instruments for making empirical predictions? Such questions challenge the core pursuit of science, raising skepticism over a 'fundamental working' of the universe that we can ever understand, describe, or model.

In this work, I assess the realist versus anti-realist debate in light of a modern and promising theory of scientific realism: structural realism. Specifically, I argue for the merits of ontic structural realism – positing relationships as ontologically primitive, with properties and objects derivable from relationships. I defend this view against the key 'relations without relata' challenge, arguing for its merits over alternative flavors of realism – i.e. semi- and entity realism. Finally, I discuss the implications of structural realism for modern science.

3.2 Challenges for Traditional Scientific Realism

I begin by introducing the scientific realism debate, elaborating on its historical context and key arguments. I then demonstrate how anti-realist attacks – e.g. underdetermination and pessimistic induction – pushed scientific realists to modern forms of entity, semi-, and structural realism.

3.2.1 Observability

Philosophy of science is largely concerned with observability, which is intertwined with human sensory capability and constitutes an entire philosophical literature. Here, I narrow the scope of '**observable**' to mean anything detectable via the unaided senses – e.g. a dog, sunlight, or warmth of a coffee. Meanwhile, '**unobservable**' refers to anything undetectable in this way – e.g. electrons or the cosmic microwave background. Chakravartty (2007) further categorizes unobservables into detectable and undetectable. **Detectable unobservables** can be observed if the senses are extended through measuring instruments. **Undetectable unobservables** cannot be detected, but are posited for theoretical or explanatory reasons.

3.2.2 Scientific Realism and Anti-Realism

Realists and anti-realists debate the role scientific observables and unobservables play in the real world. **Realism** is typically characterized by the *existence* of unobservables, *independent* of one's beliefs and conceptual schemes (Miller, 2021). Whereas realists typically take a positive epistemic attitude towards both observables and unobservables, anti-realists[1] argue against the description and understanding of unobservables.

[1] Although sometimes attributed to a singular theory, "anti-realism" here refers to a spectrum of theories, opposing different aspects of realism.

3.2.3 The Argument for Scientific Realism

Scientific realism was traditionally proposed and defended via the **miracles argument** (Brown, 1982). According to Putnam (1978), realism “is the only philosophy that doesn’t make the success of science a miracle”. Our best scientific theories – from quantum theory to general relativity – are incredibly successful, enabling accurate predictions and approachable explanations of complex phenomena. Realists argue that truth of these scientific theories is the only plausible explanation of their success. If untrue, such theory success would have to be a miracle. Between accepting scientific theories as true or miraculous, realists argue the obvious choice is truth.

Realists believe the miracles argument is further strengthened by **corroboration**. An unobservable solely measureable via a single experiment is theory-specific and could be overturned or modified by another theory. However, unobservables measurable in multiple independent ways are robust and theory-independent (Eronen, 2015). Thus, realists argue it would be an extraordinary coincidence (miracle) if corroborated unobservables do not exist.

3.2.4 Arguments Against Traditional Scientific Realism

Anti-realists have challenged traditional scientific realism, especially the miracles argument, on several fronts.

In objection to the miracles argument, Van Fraassen et al. (1980) denies the need to explain theory success. He instead proposes the **Darwinian evolution argument**, in which “any scientific theory is born into a life of fierce competition, a jungle red in tooth and claw. Only the successful theories survive”. Thus, scientific theories can be thought of as well-adapted organisms. Darwinians do not question *why* species run from predators, but claim that, if they do not, they will not survive and become irrelevant. Similarly, anti-realists argue that theory success is unsurprising – unsuccessful theories become irrelevant. Therefore, theory success does not reflect correspondence with reality, but a test of survival. Magnus and Callender (2004) further challenge the miracles argument via the **base-rate fallacy**, which dismisses success as a proxy for a scientific theory’s truthfulness. Since we cannot gauge the base-rate of true theories amongst all theories, we cannot calculate the probability of a theory being true versus a “false positive” (i.e. a successful scientific theory wildly different from the truth).

Reiss and Sprenger (2020) challenge corroboration on the feasibility of **objectivity**, which is hard to achieve in practice because of theory-ladenness and incommensurability. Duhem (1991) also argues that theoretical preconceptions can influence observations. Thus, anti-realists challenge corroboration as a metric of theory truth, instead claiming it as a metric of the scientific community’s *belief* in theory truth.

Further anti-realist concerns emerge from theory **underdetermination**, in which distinct scientific theories make successful predictions about the same phenomena. Clearly the realist can commit to only one theory, but this theory may consist of drastically different unobservables to others. In the case of, say, superfluous unobservables, the theorist could commit, via *Ockham’s razor*, to the simplest theory. However, determining theory simplicity or, more generally, theory selection criteria is subjective.

Finally, Chakravartty (2007) claims the strongest anti-realist objection is the **pessimistic induction argument** which challenges full epistemic commitment to scientific theories, based on historical evidence. Considering the history of scientific theories, there is a clear trend of theory overturn as scientific knowledge progresses. In his highly influential work, *The Structure of Scientific Revolutions*, Kuhn (1970) establishes a recurring pattern of transitions between so-called 'normal science' periods and revolutions. During 'normal science' periods, scientific theories remain fairly constant. Meanwhile, during revolutions, e.g. the transition from classical to quantum mechanics, scientific theories are radically rethought. Thus, the historicist anti-realist argues that realist commitment to the reality of scientific theories requires commitment to incompatible entities over time. Furthermore, given the long history of theory overturn, there is high likelihood that our current theories will be overturned. In light of this, how can any current theory fully describe reality?

3.2.5 Modern Proposals for Scientific Realism

To address these anti-realist concerns, especially pessimistic induction, 20$^{\text{th}}$-century realists reformulated their commitment to one of "approximate" truth. While scientific theories are likely to eventually be overturned, aspects will be reflected in the overturning theories – often as limiting cases. Although a given scientific theory has high probability of falsity, modern realists defend *proximity* to the truth. Many argue the goal of science is *converging* scientific theories to truth over time (Hardin and Rosenberg, 1982).

Realists are typically fallibilists, committing only to sufficiently mature and non-ad hoc theories, so as to minimize likelihood of commitment to a fully false theory (Worrall, 1989). Determining which aspects of approximately true theories are worth commitment, known as **selective skepticism**, has led to the development of multiple flavors of realism, notably: entity, semi-, and structural realism. At a high level, **entity realism** (Cartwright and McMullin, 1984; Clarke, 2001) suggests commitment to entities with strong causal underpinnings, e.g. unobservables describing multiple phenomena, but not to theories describing those entities. **Structural realism** (Ladyman, 2020) recommends commitment not to entities, but to structures. However, determining what constitutes a structure, is a key point of contention. Epistemic structural realism denies knowledge acquisition beyond structures, whereas ontic structural realism denies existence beyond structures. Finally, **semirealism** (Chakravartty, 2007) aims to combine the best aspects of entity and structural realism by committing to entities and structures likely to be retained in scientific theories. In light of modern and historical scientific theories, I will argue that ontic structural realism provides the best defense against anti-realism.

3.3 The Fundamental Nature of Relationships

As previously mentioned, the structures of structural realist commitment are widely debated. In this section, I argue that relationships are fundamental structures, while objects and properties are derivable from relationships. Although my personal view is that objects and properties do not exist (they are merely illusory) and unworthy of realist commitment, it is beyond the scope of this work to argue so. Therefore, this argument is open to alternative possible commitments, e.g. where objects and properties are non-fundamental but exist. Finally, (a disclaimer) although I will occasionally reference objects and properties – e.g. electrons and colors – I do not mean to imply that such entities exist, but instead mean to refer to their underlying relational constituents.

3.3.1 Deriving Objects

Drawing inspiration from **bundle theory** (Campbell, 1981; Orilia and Paolini Paoletti, 2022), I argue that bundles of properties comprise observable and detectable unobservable objects[2]. Granted a mind-independent reality, I assume our most reliable means of acquiring information about reality is via our senses – i.e. taste, smell, vision, touch, and hearing. While, in our day-to-day experiences, we learn about and interact with the items referred to as ‘objects’ (e.g. dogs, apples, cars) via these senses, I argue such ‘objects’ are actually inferred from our experience of various *properties* (e.g. shape and color). In this way, ‘objects’ instantiate our experience of properties. For example, the experience of a sheep is fully characterized by observed properties – e.g. seeing its white color and round shape, feeling its fluffy wool, smelling its barn odors, and hearing its “baas”. Furthermore, although distinct in a spatial-temporal sense, other animals are characterized as sheep because they possess the same characteristic properties. Thus, ‘objects’ are derived from bundles of properties.

The scientific realist hopes to extend this argument to unobservable entities. As previously mentioned, knowledge is most reliably gained through personal sensory experiences[3]. Using measuring instruments to aid the senses introduces additional observational uncertainty and risk of theory-laden observations. However, this should not deter the scientist. Continual improvement to measuring devices as well as corroboration can increase trust in acquisition of indirect knowledge. Detectable unobservables, I argue, are well described by their properties. For example, an electron cannot be directly observed, but is characterized by its measured mass, charge, and spin. Measurement of a positive charge ensures an electron was not observed. Although the role of an electron has changed radically over time, from ancient Greek atoms to Bohr’s atomic model to modern quantum field theory, its characterizing properties have remained more consistent.

Let’s now consider undetectable unobservables, or ‘objects’ unobservable with modern scientific capability, but posited for theoretical or explanatory reasons. I argue these undetectable unobservables are comprised of *relational* properties (defined in Section 3.3.2) to other observables or detectable unobservables. For example, properties of the number 2 include being ‘more than 1’ and ‘less than 3’. For a more physical example, consider dark matter, which does not appear to interact with electromagnetic radiation and thus is extremely difficult to detect. This matter is posited to account for 85% of matter in the universe, explaining discrepancies in various astrophysical observations. However, had these discrepancies not been observed, there would be no reason to posit dark matter. In general, scientists should not commit to undetectable unobservables that are not necessary for describing the workings of other observables or unobservables. Thus, undetectable unobservables exist *relationally* to more substantiated aspects of scientific theories. This will be further elaborated in Section 3.4.3’s discussion of detectable and auxiliary properties.

3.3.2 Deriving Properties

Despite the rich metaphysical literature on the nature of properties (Orilia and Paolini Paoletti, 2022), this work will not delve into nuanced classifications and types of properties. Instead

[2]Since this work focuses on the philosophy of science, I will not delve into details of the bundle theoretic debate, e.g. universals versus tropes or identity of indiscernibles. Such discussion can be found in (Van Inwagen and Zimmerman, 1991; Loux, 2001).

[3]The skeptic can argue that our senses might be wrong, e.g. dreams or hallucinations. I will not address such skeptic scenarios, but assume that our senses are reliable.

it focuses on the distinction between *relational* and *non-relational* properties. Distinct from a *relationship* – which is borne between one thing and another (in certain cases to itself) – a **relational property** is the property of bearing a relationship to something (MacBride, 2020). For example, if a bears relationship R to b, then a possesses the relational property of 'bearing R to b' and b possesses the relational property of 'a bearing R to it'. Meanwhile, **non-relational properties** refer to properties something has of itself.[4]. Combining ideas from bundle theory and class nominalism, I will now demonstrate how all properties (and, thus, objects) are derivable from relationships, establishing the fundamental nature of relationships.

Let's begin with non-relational properties, which appear less intuitively connected to relationships. For observables, I use the commonplace non-relational property of color[5] as an example. If everything was green – i.e. we had no distinct perception of the colors red, blue, orange, etc. – we would not identify color as a property. This means we would not attribute to, say, a turtle the property of being green. In this fictional world, a turtle would be identified and categorized with other turtles via properties such as shell hardness and swimming ability. However, relative to other objects or animals (which we have established are simply bundles of properties), there would be no shared or differing property of greenness. This logic applies to unobservables as well. For example, spin is an electron property because some electrons are spin-up while others are spin-down. If all electrons were spin-up, the spin property would be redundant to our electron theory. Therefore, along the lines of **class nominalism** (Rodriguez-Pereyra, 2019), I argue that non-relational properties (of both observables and unobservables) are defined only in *relation* to similar and distinct non-relational properties.

In Section 3.3.1, I leveraged ideas from bundle theory to argue that objects are, from a set theoretic perspective, simply sets of properties. In the turtle example, the `turtle` 'object' is instantiable as a set of properties,

$$\texttt{turtle=}\{\texttt{green, hard shell, swims, ...}\}. \tag{3.1}$$

The last paragraph, leveraging ideas from class nominalism, further established the relational nature of non-relational properties. In the turtle example, the `green` 'property' is instantiable as a set of all the objects we experience as green,

$$\texttt{green = }\{\texttt{turtle, pea, clover, grass, ...}\}. \tag{3.2}$$

However, (3.2) appears to be a simple semantic shift from (3.1) – instead of objects as sets of properties, properties are sets of objects. While it may seem circular, I argue that this semantic shift is critical to enable infinite regress and an entirely relational recursion.

[4] What I refer to as relational (non-relational) properties are often labeled "extrinsic" ("intrinsic") properties in the literature. However, granted concerns over properties which are both intrinsic *and* relational (Marshall and Weatherson, 2018), I refrain from such nomenclature.

[5] Although colors are often classified as relational properties, I mean them in the simple, intrinsic, non-relational, non-reducible, and qualitative sense of Color Primitivist Realism (Maund, 2022). The argument can also be extended to any other non-relational property with multiple determinates – e.g. mass and shape.

To illustrate this, plug the `green` property set (3.2) and others into the `turtle` object set (3.1)[6], resulting in expanded set,

```
turtle={green={pea, clover, grass, ...},
        hard shell = {armadillo, snail, roly-poly, ...},        (3.3)
        swims = {dolphin, frog, fish, ...}, ...}.
```

Recursively expanding this set, by plugging in property sets constituting the objects and object sets constituting the properties, generates an extremely large set of subsets.

Ontologically, the key question is whether the set is finite or infinite. An *infinite* set implies the recursion is infinite, meaning neither objects nor properties need be fundamental[7]. However, a *finite* set necessitates fundamental properties and/or objects as the set's base elements. If we could simply assume there were infinite objects and/or non-relational properties, the sets would be infinite and the challenge would be easily resolved. While such an assumption appears highly non-trivial, I argue it does not matter. Instead, *relational* properties offer a far easier path to infinite regress.

The `turtle` set of (3.1) lists generic non-relational properties, applicable to all turtles. However, describing a specific turtle, e.g. `turtle-Ted`, necessitates relational properties. In fact, I argue `turtle-Ted` bears *infinite* relational properties. To name a few, `turtle-Ted` is '2 inches from turtle-Bob', '1 meter from his food bowl', and '2,000 miles from Paris'. Assuming space is infinite or has infinite resolution, `turtle-Ted` possesses infinite relational properties solely regarding his distance to other spatial points[8]. In addition to `turtle-Ted`, these relational properties apply to infinite spatial points, e.g. on the surface of a 2,000 mile radius sphere surrounding Paris. Therefore, achieving the desired infinite regress simply requires: 1) expanding the `turtle-Ted` object set with these infinite relational properties, 2) expanding those infinite relational property sets with the infinite object sets sharing those properties, and 3) continued infinite recursive expansion of the object and property sets. In this way, relational properties enable infinite regress, or as Stachel et al. (2006) eloquently put it, "relations all the way down".

3.4 Arguments for Ontic Structural Realism

Granted previous arguments, objects and properties are reducible to relationships. **ontic structural realism (OSR)**, typically associated with French and Ladyman (2011) and promoted in this work, holds that relationships are the *only* fundamental metaphysical primitive – properties and objects are secondary.

This may seem shocking and counter-intuitive. How can there be relationships without any 'relata' – i.e. objects or properties to relate? Can this theory withstand pessimistic induction? Even accepting OSR, how could the realist identify which relationships are worthy of commitment? What implications, if any, would this philosophical view have for science? In this final section, drawing inspiration from modern physics and the history of science, I will address each of these questions – defending OSR against key criticisms and bolstering its merits relative to other forms of scientific realism.

[6] To prevent trivial infinite regress by circularity, each expansion of an object/property set must *exclude* that object/property and any earlier objects/properties from which it was expanded.
[7] Section 3.4.1 will address criticisms on the basis of 'relations without relata'.
[8] Similarly, there are infinite relational properties pertaining to metrics such as relative height, weight, speed, etc.

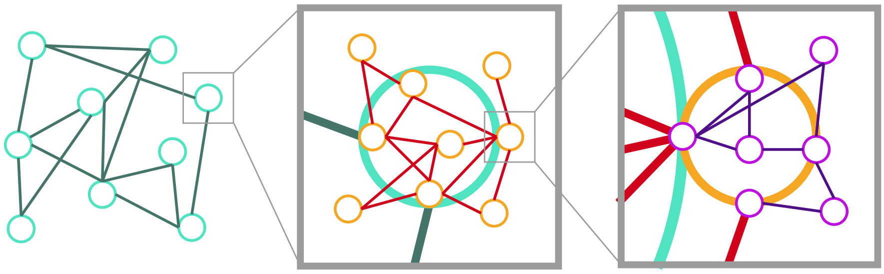

Figure 3.1: Visualization of a relations-based ontology, in which bundles of relationships constitute 'relata', enabling infinite regress. Zooming into the 'relata' of relationships reveals that they are simply further bundles of relationships.

3.4.1 Relations Without Relata?

Among the strongest objections to OSR is the challenge of 'relations without relata' (Psillos, 2001; Busch, 2003; Chakravartty, 2007; Ainsworth, 2010; Briceno and Mumford, 2016). Namely, if neither objects nor properties are fundamental, there is nothing to have relationships between. How can there be relationships without 'relata'?

Chakravartty (2007) challenges the explanatory value of the ontic structural framework, arguing that the relationship-relata dependence is a necessary conceptual dependence. Thus, forms of scientific realism which treat both structures and entities as fundamental, i.e. semirealism, have superior explanatory power. However, in defining semirealism, Chakravartty argues that properties are derivable from relationships and relationships are derivable from properties. Thus, he admits relationships are fundamental, but maintains the fundamental nature of properties largely to avoid 'relations without relata'. I instead argue that 'relations without relata' is a non-concern for OSR, minimizing potential benefits of semirealism over OSR.

While the counter-intuitive nature of OSR arguably stems from difficulties in comprehending the nature of infinite regress, Figure 3.1 illustrates how 'relationships without relata' are not necessarily problematic. Although the vertices appear to be entities, i.e. objects or bundles of properties, zooming into a vertex reveals it is a collection of further vertices and edges. To be concrete, let's return to the turtle example, in which the first blue graph represents relationships between animals. These relationships imbue a turtle with its high-level observable properties – i.e. greenness, shell hardness, swimming ability, etc. Zooming into a turtle vertex, the turtle is comprised of an orange graph of relationships between its organs. For example, one vertex represents the turtle's skin and another, its brain. The organs differ on or share properties such as texture, softness, etc. Zooming into the brain vertex, we find another, this time purple, graph describing relations between different brain constituents – e.g. grey matter, white matter, nerve cells, and blood vessels. Zooming again into any vertex will reveal yet another graph, giving rise to further 'objects' or bundles of 'properties'. Notice how each zoom-in reveals that these are not 'relata' in the typical sense of objects or properties, but *bundles of relationships.*

The key remaining question is whether this regression continues infinitely ("relations all the way down") or stops somewhere relationships connect fundamental entities ("relations part of the way down")? Although Section 3.3.2 presents an abstract argument for the former, one could argue it does not account for our intuitive understanding of fundamental physical 'relata', dating back to the ancient Greek notion of the atom (stemming from the ancient Greek word *atomos* for "uncuttable"). Even with fundamental relationships, it seems that there must also be fundamental entities upon which these relationships act. Physically, "relationships part of the way down" defines physical matter by relationships between fundamental 'atoms'.

Of course, physics has come a long way since the time of ancient Greeks, and I argue that modern quantum theory enables a physical "relationships all the way down" ontology – as desired by OSR. The 20$^{\text{th}}$-century advent of quantum mechanics introduced probabilities into the description of fundamental particles, governed by wavefunctions. Underdetermination of the individuality of quantum particles[9] largely motivated Ladyman (1998)'s transition to OSR. Alternatively, the Ithaca interpretation of quantum mechanics challenges physical 'relata' via Mermin (1998)'s argument that internal correlations, encoded by the state and wavefunction, comprise physical reality. However, I believe that it is the more recent development of quantum field theory (QFT) which most radically shifts our understanding of 'fundamental particle' ontology. Namely, in QFT, particles emerge as manifestations of a *field* – and what is a field, if not relational? Mathematically, a field is defined as an algebraic *structure* consisting of an infinite system of real or complex numbers, such that addition, subtraction, multiplication, and division of any two numbers again yields a number of the system (Dirichlet et al., 1999). Structuralists, such as Shapiro (1997) and Resnik (1997), define a *system* as an instantiation of *structures*, or places that stand in structural relations to one another. Derivatively, fields do not describe objects, but relative positions within structures. For example, in Section 3.3.1's discussion of undetectable observables, the number 2 is defined as less than 3 and more than 1. Attempts at instantiating these structures, e.g. via sets, pose a challenge for ontological commitment, since any instantiation has isomorphic counterparts. For example, Benacerraf (1965) demonstrates that $\{\{\{\emptyset\}\}\}$ and $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ are equally suitable for playing the role of 3. In fact, as an eliminative structuralist, Benacerraf denies existence of abstract mathematical objects with properties other than their place in a relational structure. In conclusion, if we aim to make fundamental particles the 'relata' of physical existence, QFT's mathematical description of particles as fields offers a purely relational ontology.

Rovelli (2018) offers an alternative means to a relations-only ontology. He claims that relationality is ubiquitous in most of modern physics:

> In classical mechanics the velocity of an object has no meaning by itself: it is only defined with respect to another object. The color of a quark in strong-interaction theory has no meaning by itself: only the relative color of two quarks has meaning. In electromagnetism, the potential at a point has no meaning, unless another point is taken as reference; that is, only relative potentials have meanings. In general relativity, the location of something is only defined with respect to the gravitational field, or with respect to other physical entities; and so on. (Rovelli, 2018)

However, he believes that quantum theory "takes this ubiquitous relationalism to a new level". By this, Rovelli refers to Relational Quantum Mechanics (RQM) (Rovelli, 1996), which rejects commitment to the quantum particle and its wavefunction – as originally argued for by Schrödinger and assumed in our previous fundamental particle discussion. Instead, Rovelli argues that, like velocity, the value of any quantum physical system quantity is only meaningful in relation to

[9] Quantum mechanical wavefunctions appear to attribute the same intrinsic and relational properties to distinct particles, violating Leibniz's principle of the identity of indiscernibles (Busch, 2003).

another system. Candiotto (2017) explains that RQM: 1) asserts nonexistence of a perspective-independent description of the universe, 2) argues against the notion of 'objects' as 'entities' possessing intrinsic properties, and 3) proposes an ontology relativizing properties or states of 'objects' to other 'objects'. Thus, RQM instantiates the OSR ontology[10], where physical interactions between systems and instruments are viewed as *primitive relations*. Candiotto argues that, by engendering a new understanding of 'objects', RQM addresses the challenge of 'relations without relata'. Dating back to Aristotle, relationships have been thought of as a properties of objects. However, Candiotto claims "RQM invites us to think of reality not as starting from things, which would be then connected by relationships, but as processes that manifest 'things' as the result of their intertwining". Namely, quantum mechanics describes interactions between *processes* – or transitions between interactions. Thus, reality is comprised of a series of events, not objects. In RQM, relations are not connections between objects, but modalities of processes. They are the structures through which systems interact and communicate. Thus, Candiotto argues that RQM "provides a description of reality that poses relations as real and prior to objects".

Although I began this section by noting substantial philosophical concerns over 'relations without relata', there exist several arguments beyond those presented – drawing largely from modern quantum and general relativity theory – promoting "relationships all the way down" (Schaffer, 2003; Saunders, 2003; French and Krause, 2006; Ladyman et al., 2007). As Rovelli (2018) claims, "rather than letting our philosophical orientation determine our reading of [physics, we should] be ready to let the discoveries of fundamental physics influence our philosophical orientations".

3.4.2 Pessimistic Induction?

Having alleviated the 'relations without relata' concern, commitment to structures alone proves highly effective at resolving the key criticism of scientific realism – pessimistic induction. Entity realism and semirealism argue for the fundamental nature of entities. However, in the history of science, entities have proven the aspect of scientific theories most prone to significant modification. Take, for example, the transition from phlogisten to oxygen or the ever-changing electron. Meanwhile, mathematical structures, governing the relationships of a theory, are generally preserved. As scientific knowledge grows, old scientific theories are still often found correct, but in more restricted settings than those described by the novel, generalized theory. For example, Newtonian mechanics is a limiting case of quantum theory, for macroscopic systems experiencing significant decoherence. In light of pessimistic induction, commitment to relationships rather than entities appears far more robust to theory change. Next, I explore how a realist can determine OSR structures worthy of commitment.

3.4.3 Which Structures are Worth Commitment?

Beyond 'relations without relata', Chakravartty (2007) criticizes OSR and bolsters semirealism in ability to determine aspects of a scientific theory worth commitment.

[10]Because RQM explains structures through which we know the world's *nature*, Candiotto (2017) argues it supports ontic, not epistemic, structural realism. She claims that relationships are the building blocks of reality, constituting observer-independent information. Information is exchanged via physical interactions and interactions are primitive to the structure of matter.

Chakravartty claims that semirealism, with commitment to both relationships and properties, is well-suited to providing a clear distinction between entities that merit realist commitment and those that do not. To do so, he introduces an epistemic distinction between detection and auxiliary properties of unobservables. **Detection properties** are detectable causal properties or, more specifically, properties that are "causally linked to the regular behaviours of our detectors" (i.e. observables and detectable unobservables). Meanwhile, **auxiliary properties** are any other properties attributed to unobservables (i.e. undetectable unobservables). Chakravartty argues simply and practically for realist commitment to detection properties. Thus, scientific investigation must find grounds for discarding auxiliary properties or transforming them into detection properties. Chakravartty further argues that causal detection properties worthy of commitment must provide a minimal interpretation of their governing mathematical equations.

> Anything that exceeds a minimal interpretation, such as interpretations of equations that are wholly unconnected or only indirectly connected to practices of detection, goes beyond what is minimally required to do the work of science: to make predictions, retrodictions, and so on. The excess is auxiliary. (Chakravartty, 2007)

Chakravartty believes that neither entity realism nor structural realism are capable of providing such a clear selection criteria for realist commitment, since neither stance is committed to both entities *and* strutures. He argues that it is the *relationship* of the detection *properties* to the measurement apparatuses which enables such a delineation.

However, in light of our arguments for OSR, I challenge the necessity of properties for Chakravartty's demarcation. Ultimately Chakravartty is concerned with the minimal interpretation of *mathematical equations*. Although equation variables may be attributed to properties or objects, I argue that, alternatively, such variables are non-individualistic and can be replaced by further expressions. Furthermore, the detection properties described by Chakravartty are simply relational properties constituting observables and detectable unobservables, which I established in Section 3.3 are fundamentally relational. In this way, Chakravartty's proposal for a semirealist selection criterion, also seems well-suited to OSR. Namely, in the immense graph of relationships constituting the world, OSR suggests increased realist commitment to relationships closer linked to a scientist's direct sensory experiences. If a scientist can gain detection evidence of an unobservable (comprised of relationships), it moves from auxiliary to detectable and is worthy of commitment.

3.4.4 Implications for Science?

I conclude by discussing OSR's implications for the future of physics, and science more generally. Specifically, I present the most pressing debate in modern physics: quantum gravity. At a high-level, general relativity (GR) and QFT employ incompatible ideas of space and time (Rickles et al., 2006). Namely, GR is *background-independent*, positing a *dynamic* metrical structure and, hence, geometry of spacetime. Meanwhile, QFT is *background-dependent*, employing a *fixed* metric. Among the first serious attempts to resolve this incompatibility were background-dependent, covariant perturbation quantizations – i.e. gravitons, string theory, and 'supergravity' theories. However, problems with these old covariant perturbation methods arose largely from imposed background-dependence. Namely, sustaining various metrics as small perturbations becomes challenging at the Planck scale, with a flat and fixed background. This prompted the development of background-independent, non-perturbative canonical quantization methods – i.e. loop quantum gravity – which quantize the full metric. Interestingly, in loop quantum gravity, spin networks form a basis for quantum states and provide a delocalized structure, encoding *relational* features – i.e. relations between fields. A physicist committed to OSR would clearly be attracted to the relational structures of background-independent theories. In fact, this seems where most of the modern physics community is headed (Rickles et al., 2006). As argued by Rickles (emphasis my own),

> the fundamental ontology of [quantum gravity] is given by *relational structures* rather than individual objects; inasmuch as objects exist at all, they *derive* their properties and individuality from the *relational network* in which they are embedded. (Rickles et al., 2006)

Relationality in GR and QFT has largely informed the growing popularity of OSR. Such OSR is now leading physicists to commit to relational structures and theories of quantum gravity. This serves as a precedent for how OSR can influence theory selection in all domains of science.

Figure 3.2: TURTLES = RELATIONSHIPS

Bibliography

Ainsworth, P. M. (2010). What is Ontic Structural Realism? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 41(1):50–57.

Benacerraf, P. (1965). What Numbers Could Not Be. *The Philosophical Review*, 74(1):47–73.

Briceno, S. and Mumford, S. (2016). Relations All the Way Down? Against Ontic Structural Realism. In Marmodoro, A. and Yates, D., editors, *The Metaphysics of Relations*, pages 198–217. Oxford University Press.

Brown, J. R. (1982). The Miracle of Science. *The Philosophical Quarterly*, 32(128):232–244.

Busch, J. (2003). What Structures Could Not Be. *International Studies in the Philosophy of Science*, 17(3):211–225.

Campbell, K. (1981). The Metaphysic of Abstract Particulars. *Midwest Studies in Philosophy*, 6:477–488.

Candiotto, L. (2017). The Reality of Relations. *Giornale di Metafisica*, 2.

Cartwright, N. and McMullin, E. (1984). *How the Laws of Physics Lie*. American Association of Physics Teachers.

Chakravartty, A. (2007). *A Metaphysics for Scientific Realism: Knowing the Unobservable*. Cambridge University Press.

Clarke, S. (2001). Defensible Territory for Entity Realism. *The British Journal for the Philosophy of Science*, 52(4):701–722.

Dirichlet, P. G. L., Dirichlet, P., and Dedekind, R. (1999). *Lectures on Number Theory*. 16. American Mathematical Society.

Duhem, P. M. M. (1991). *The Aim and Structure of Physical Theory*, volume 13. Princeton University Press.

Eronen, M. I. (2015). Robustness and Reality. *Synthese*, 192(12):3961–3977.

French, S. and Krause, D. (2006). *Identity in Physics: A Historical, Philosophical, and Formal Analysis*. Oxford University Press.

French, S. and Ladyman, J. (2011). *In Defence of Ontic Structural Realism"*, pages 25–42. Springer Netherlands.

Hardin, C. L. and Rosenberg, A. (1982). In Defense of Convergent Realism. *Philosophy of Science*, 49(4):604–615.

Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*. Chicago University of Chicago Press.

Ladyman, J. (1998). What is Structural Realism? *Studies in History and Philosophy of Science Part A*, 29(3):409–424.

Ladyman, J. (2020). Structural Realism. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2020 edition.

Ladyman, J., Ross, D., Collier, J., Spurrett, D., Spurrett, D., Collier, J. G., et al. (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford University Press on Demand.

Loux, M. J. (2001). *Metaphysics: Contemporary Readings*. Psychology Press.

MacBride, F. (2020). Relations. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2020 edition.

Magnus, P. D. and Callender, C. (2004). Realist Ennui and the Base Rate Fallacy. *Philosophy of Science*, 71(3):320–338.

Marshall, D. and Weatherson, B. (2018). Intrinsic vs. Extrinsic Properties. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2018 edition.

Maund, B. (2022). Color. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2022 edition.

Mermin, N. D. (1998). What is Quantum Mechanics Trying to Tell Us? *American Journal of Physics*, 66(9):753–767.

Miller, A. (2021). Realism. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2021 edition.

Orilia, F. and Paolini Paoletti, M. (2022). Properties. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2022 edition.

Psillos, S. (2001). Is Structural Realism Possible? *Philosophy of Science*, 68(S3):S13–S24.

Putnam, H. (1978). Mathematics, Matter and Method. *Philosophy of Science*, 45(1).

Reiss, J. and Sprenger, J. (2020). Scientific Objectivity. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2020 edition.

Resnik, M. D. (1997). *Mathematics as a Science of Patterns*. Oxford University Press.

Rickles, D., French, S., and Saatsi, J. T. (2006). *The Structural Foundations of Quantum Gravity*. Clarendon Press.

Rodriguez-Pereyra, G. (2019). Nominalism in Metaphysics. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2019 edition.

Rovelli, C. (1996). Relational Quantum Mechanics. *International Journal of Theoretical Physics*, 35(8):1637–1678.

Rovelli, C. (2018). Space is Blue and Birds Fly Through It. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2123):20170312.

Saunders, S. (2003). Structural Realism, Again. *Synthese*, 136(1):127–133.

Schaffer, J. (2003). Is There a Fundamental Level? *Noûs*, 37(3):498–517.

Shapiro, S. (1997). *Philosophy of Mathematics: Structure and Ontology*. Oxford University Press.

Stachel, J. et al. (2006). Structure, Individuality, and Quantum Gravity. *The Structural Foundations of Quantum Gravity*, pages 53–82.

Van Fraassen, B. C. et al. (1980). *The Scientific Image*. Oxford University Press.

Van Inwagen, P. and Zimmerman, D. W. (1991). *Metaphysics: the Big Questions*. Wiley.

Worrall, J. (1989). Structural Realism: The Best of Both Worlds? *Dialectica*, 43(1-2):99–124.

Yesterday is history, tomorrow is a mystery, but today is a gift.
That is why it is called the present.

— Master Oogway (*Kung Fu Panda*)

4 The Open Past and Many-Worlds Presentism

Contents

4.1 Introduction

Arguably, one of the most iconic films of my childhood was *Kung Fu Panda*. As a 10-year-old, I was particularly struck by the wise words of turtle sensei Master Oogway:

> Yesterday is history, tomorrow is a mystery, but today is a gift. That is why it is called the present.

I did not realize it then, but this simple statement encapsulates several fundamental aspects of the human temporal experience – namely, the knowledge asymmetry. We generally regard the past ("yesterday") as fixed, known, and unchanging ("history"). Meanwhile, the future ("tomorrow") is uncertain and modifiable (a "mystery"). Finally, the present ("today") is distinguished, vivid, and special relative to other moments in time (a "gift"). We feel control over the present and future in a way we do not about the past. The past appears behind us, to be forgotten, while the future lies ahead, yet to be realized.

However, this common intuition for time – informed by day-to-day experiences – makes several assumptions about the metaphysics and ontology of time. Most modern philosophers and physicists believe that time exists analogously to space, constituting a 4^{th}-dimension in an unchanging Block Universe. Others believe the universe is a growing block, in which only the past and present exist. However, this work argues for past openness and Many-Worlds Presentism – in which only the present exists, as a timeless ensemble of momentary possible worlds. A Many-Worlds Presentist "best match" method is proposed to account for the knowledge asymmetry, providing strong connections to the thermodynamic asymmetry.

4.2 Temporal Asymmetries

While the universe appears symmetric along its spatial dimensions, time appears distinct from space because of several significant temporal asymmetries (Dainton, 2016). I will now define the two temporal asymmetries – knowledge and thermodynamic – considered in this work.

4.2.1 The Knowledge Asymmetry

The **knowledge asymmetry** is defined loosely, encompassing what Callender (2021) refers to as the epistemological, mutability, and psychological arrows of time. The *epistemological arrow* accounts for perceived greater knowledge about the past than the future. For example, I know yesterday's winning lottery number, but not tomorrow's. The *mutability arrow* encompasses perceived '**openness**' or indeterminacy towards the future, which is not perceived towards the past. The future appears mutable, while the past appears fixed for all eternity. Finally, the *psychological arrow* describes our differing attitudes towards the past and future. For example, the 'argument from pain' describes general preference for pain in the past rather than the future. The psychological arrow also accounts for perception of a moving 'now' or the 'flow' of time, in which present events transition from future to past. Overall, the knowledge asymmetry encompasses asymmetries based on human temporal experiences, psychology, and beliefs.

4.2.2 The Thermodynamic Asymmetry

The 19$^{\text{th}}$-century advent of thermodynamics assured physicists that the temporal asymmetry and perceived 'arrow of time' are not just experiential, but physical. Specifically, the **thermodynamic asymmetry**, governed primarily by the 2$^{\text{nd}}$ Law of Thermodynamics[1], accounts for macroscopic irreversible processes and generally observed increases in entropy. For example, eggs fry and ice-cubes melt in hot water, but eggs do not "un-fry" and ice-cubes do not spontaneously form in warm water.

From the Past Hypothesis to local-branch theories, this asymmetry's emergence at microscopic levels is largely debated (Reichenbach, 1956; Albert, 2001; Wallace, 2011; Rovelli, 2022). While I will not delve into such debates, for later reference I will briefly describe statistical mechanical entropy increase. Namely, the ergodic system-state has a proclivity to move into state-space regions occupied by high-entropy macroscopic states (**macro-states**), which possess larger state-space volumes and more compatible microscopic states (**micro-states**).

Finally, the relationship between the thermodynamic and knowledge asymmetry is debated. Although Horwich (1987)'s derived explanatory hierarchy treats them independently, (Rovelli, 2022) believes the thermodynamic asymmetry is the most fundamental temporal asymmetry.

> No phenomenon reveals any detectable difference between the past and the future directions of time, unless it includes a process irreversible in a (general) thermodynamic sense. This is a fact. (Rovelli, 2022)

As will be elaborated in Section 4.4.2, Rovelli argues that past memories and records are thermodynamic mechanisms. By the end of this essay, I hope to establish a strong link between the thermodynamic and knowledge asymmetries.

4.3 Temporal Ontologies

Following from Dainton (2016) and Le Bihan (2020)'s descriptions, I will briefly introduce three established metaphysical ontologies of time: 1) Eternalism, 2) the Growing Block Theory, and 3) Presentism. I emphasize key benefits and drawbacks of the theories, informing later arguments for Many-Worlds Presentism.

4.3.1 Eternalism

In **Eternalism**, also known as the **Block Universe** theory, all temporal existence is equally real (Smart, 1963; Mellor, 1998; Sider et al., 2001). Rather than a 3-dimensional space modulated by the passage of time, the Block Universe treats spacetime as a 4-dimensional unchanging block. The past, present, and future all exist, but the present is not distinguished and time does not 'flow'. There are two key challenges for Eternalism: 1) perceived temporal passage and 2) truthmaking.

Granted a non-dynamic Block Universe, how and why would humans perceive time as 'flowing' and distinct from space? Several philosophers have attempted to argue that this dynamic temporal perception is consistent with an underlying Block Universe. Mellor (2001) cites McTaggartian notions of change to argue that time is the 'causal dimension', enablinging our actions and

[1]The exact source of thermodynamic temporal asymmetry is disputed. While most claim it is the 2$^{\text{nd}}$ Law of Thermodynamics, some (Uffink, 2001; Brown and Uffink, 2001) argue it is a separate "Minus 1$^{\text{st}}$ Law". However, spontaneous movement from non-equilibrium to equilibrium, or lower to higher entropy, is generally accepted.

notions of persistent identity. Meanwhile, leveraging modern psychology, cognitive science, neuroscience, and biology findings, Callender (2017) argues human perception of simultaneity and a common "now" is artificial and illusory. Hartle (2005) uses simplistic **information gathering and utilizing systems (IGUSs)** to demonstrate how, in a Block Universe, evolution would prefer human temporal perception.

Finally, truthmaking acts as a two-sided sword for Eternalism (Le Bihan, 2020). Parts of the Block Universe's spacetime can act as truthmakers for past *and* future statements. While such truthmaking is consistent with a determined and fixed past, it poses a threat of existential future determinism.

4.3.2 The Growing Block Theory

Concerns over this future determination led Broad (1923) to propose the **Growing Block Theory (GBT)**. In line with the knowledge asymmetry, the GBT holds that the past and present exist, but the future does not. "Present" simply labels the most recent time-slice added to reality, meaning there is no significant tensed distinction between the present and past. Unlike the non-dynamic Block Universe, the GBT posits a dynamic, continuously *growing* reality.

The GBT has been challenged on several fronts, especially for its temporal directionality and dynamics. Dainton (2016) argues that the Growing Block's 'arrow of time' does not neccesarily align with the universe's 'arrow of time', as exemplified in a Gold Universe – i.e. a universe in which entropy increases from its conception until a moment of contraction, after which entropy decreases. Beyond temporal orientation, Dainton challenges the dynamics of the Growing Block. Instead of growing, could the universe be seen as shrinking? If time slices can be created or removed at one end of universal reality, what bars their creation or removal at the other end? In this way, Dainton disputes the very nature and existence of time, beyond the distinguished present. However, a major epistemic objection to the GBT is, in fact, the individual's inability to know whether they exist in this objective, distinguished present. Miller (2018) thus promotes a version of the GBT in which objectively present experiences are subjectively distinguishable from identical objectively past experiences. However, this 'distinguishability imperative' requires making past-tensed truthmakers identical to present-tensed truthmakers. Therefore, Miller argues supporters of the GBT should instead support Presentism.

4.3.3 Presentism

Belief in only the present's existence constitutes **Presentism**. There are multiple flavors of Presentism, notably: Solipsistic, Modal, and Dynamic (Dainton, 2016). Before describing these, I should note that a key objection to Presentism is truthmaking. Like Eternalism posed a threat of future determinism, Presentism poses a threat of past indeterminism. However, this work will establish the openness (or indeterminacy) of the past, alleviating this concern.

Solipsistic Presentism is the most radical version of Presentism, arguing that nothing exists beyond *this* present. To account for our beliefs in the past and future, Solopsistic Presentism must rely on a **Boltzmann brain** type skeptic catastrophe – i.e. a highly-unlikely statistical fluctuation in the universe's heat-death forms a brain-structure perceiving my present experience and beliefs of the past and future. However, even this extreme Boltzmann brain *reductio ad absurdum* is incompatible with Solopsistic Presentism, since the low-probability emergence of a Boltzmann brain would require significant time before the present. If reality was reduced to a single present instance of time, no physical models would be valid and time-passage would be illusory. Thus, Solipsistic Presentism appears highly problematic in light of our experiences, intuitions, and theories of physics. I, therefore, will not consider it any further in this work.

Many-Worlds Presentism holds that many presents can and do exist. Namely, reality includes many momentary presents that are not temporally related and do not succeed each other, but constitute brief co-existing, self-contained worlds. At any given time, we can only inhabit one such world, meaning our perception of time amounts to hopping between these distinct, possible worlds. It appears unclear how such a construction would achieve our perceived, continuous experience of reality. Many-Worlds Presentists propose **"best match" methods** – which choose and order the subset of existing worlds that we will inhabit – in attempt to reconstruct our continuous realities and enable our perception of time. A key results of this work is Section 4.5.3's proposal of a "best match" method. Although this proposed method does not provide a *unique* ordering, as desired by De et al. (2019), it draws inspiration from the thermodynamic asymmetry to account for our epistemic beliefs about the past and future – i.e. knowledge asymmetry.

Finally, **Dynamic Presentism** is a dynamic theory of time, in which a succession of instantaneous presents constitute reality. The present comes into existence temporarily, before being annihilated and replaced by a new present. Dainton (2016) discusses several challenges for Dynamic Presentism, which I will not go into here. Instead, Section 4.5.1 presents arguments for the benefits of Many-Worlds over Dynamic Presentism, for modal reasoning and truthmaking.

4.4 The Argument for Past Openness

Barring the possibility of time-travel, we cannot physically access the past or future. Towards the past, humans merely have present physical access to *memories* and *records* – i.e. mental states and subsets of the universe's physical state. Trust in beliefs surrounding these memories and records produces a sense of past knowledge. Why, however, do we believe memories and records provide evidence of a fixed past, but no equivalent evidence exists for the future? Why do we believe we have greater epistemic access to the past than the future? Why are we more confident in past than future beliefs? In the following sections, I will challenge such beliefs, which comprise the knowledge asymmetry. Specifically, I argue that we should *not* be confident in past beliefs and, therefore, that the past must be open – analogous to future openness.

4.4.1 Past Skepticism

The fixed-past intuition arises from the belief that memories and records provide information about past *existence*, acting as truthmakers about past statements. Meanwhile, we do not believe predictions are truthmakers of the future. Thus, the future appears largely uncertain. For example, I feel *entirely* certain that the sun rose yesterday. While I feel *almost entirely* certain that the sun will rise tomorrow, I cannot say I am *entirely* certain. What if the sun spontaneously combusts over night? What if Earth is hit by massive space debris and launched out of orbit? While extremely unlikely, the mere possibility of such events bar me from expressing full certainty in the future, even for something as simple as the sunrise.

This past-future knowledge asymmetry may seem basic and fundamental. However, I argue that it is challengeable via skepticism. For example, why do I feel confident that the sun rose yesterday? I have memories and records – e.g. photos, videos, and news reports – indicating that the sun did rise. However, a skeptic would argue that the memories could be dreamed or hallucinated and the records tampered or fake. I could exist as a Boltzmann brain, meaning there is no actual sun to rise. I could be part of an alien simulation that just started and was imbued with memories of the sunrise, so as not to raise any suspicions. Realistically, either of these realities are highly unlikely. However, nothing physically rules out their possibility. Therefore, in the same way that I cannot be *entirely* certain the sun will rise tomorrow, I cannot be *entirely* certain the sun rose yesterday. Furthermore, while Boltzmann brains or alien simulations may seem unrealistic, we realistically do dream, misremember, and hallucinate past realities – all establishing uncertainty in the past.

4.4.2 Past Non-Robustness

The previous section illustrated how skeptics can challenge the *reliability* of past memories and records. Beyond such skepticism, I argue that past knowledge is not *robust*.

Beginning with memory, Hartle (2005) describes humans as glorified IGUSs. Although we gather high-resolution information about present surroundings, evolutionary pressures have caused us to develop schematic memory. Namely, our memories do not store high-resolution present moments, but instead extrapolate and store important low-resolution information, forming a schema that guides our future actions and behavior. In this process, even someone with perfect memory would forget most present details, over time. In reality, humans do not possess perfect meory – our extrapolation, memory storage, and memory recall mechanisms are all error-prone. We commonly misremember or forget information, no matter how *important* or *recent*. For example, I frequently forget where I put my keys a few minutes ago. At a spelling bee, I may misremember how to spell words that I normally spell easily.

Rovelli (2022) claims that memories are a specific type of record. Via information theory, he argues that all records – e.g. memories, photographs, footprints, and DNA – are "natural mechanisms converting past low entropy into macroscopic information". By Landauer's Principle, record destruction – in which information is converted to heat and lost to the environment – results in entropy increase (Bennett, 2003). Granted the thermodynamic trend towards increased entropy, physics dictates that past information is naturally lost over time. For example, memories are forgotten, photos fade, footprints blow away in the breeze, and DNA degrades. In the same way we are less certain about the distant future, physics ensures we will be less certain about the distant past.

4.4.3 Past Corroboration

Drawing inspiration from the philosophy of science, I will now argue that humans attribute certainty to past *and* future knowledge via a single method: corroboration.

Scientific theories are among the most accurate means of confidently predicting future phenomena. However, philosophers of science debate whether these theories truly reflect reality.

Scientific realism offers a positive epistemic attitude towards our best theories, claiming real-world existence of both observables and unobservables (Chakravartty, 2017). Central to scientific realism is the miracle argument (Brown, 1982), which claims the incredible success of our scientific theories would be a miracle if untrue. Such scientific theory success is primarily quantified and established through **corroboration**. The argument from corroboration holds that the more ways a scientific theory is independently verified, the more miraculous it would be if untrue (Chakravartty, 2017). Thus, we place more confidence in the ability of corroborated scientific theories to predict the future, attributing more certainty to their predictions.

I now argue that corroboration, beyond bolstering certainty in future knowledge, also bolsters certainty in past knowledge. If several independent memories and records exist indicating a past event, it would be miraculous if that past event did not occur[2]. For example, a detective gathers distinct records to increase certainty in a specific theory of the past. The more evidence available to support a belief of the past, the more reason to believe in that past. Therefore, it is unsurprising that we are usually pretty certain in large-scale past events or trends. For example, I am confident I did an undergraduate engineering degree. I have vivid memories of the university, my classes, and my friends. I also have access to various records, such as photos, a graduation diploma, and online records of my transcript. Therefore, while I cannot be *completely* certain, I am pretty confident that I did an engineering degree.

However, zooming into smaller-scale past events, there exist fewer memories and records offering evidence of those events. Limited access to records of a past event magnifies the unreliability and non-robustness of those records. For example, while I am pretty certain of my engineering degree, I am less certain about which classes I took. Arguably, the strongest evidence is my transcript (which lists all my classes), but what if it contains a mistake? I took classes with friends, but what if they mix up the classes we took together? Thus, I am less certain about which classes I took during my degree than the fact that I did the degree. Pushing this further, I am almost completely uncertain about what I ate for lunch on my first day of classes. I have no memory of the meal. I potentially have a time-stamped photo of food from that day, but how can I be sure it was my lunch and not dinner? What if the picture is of someone else's food?

Hopefully, by now, it is clear that our knowledge does not 'fix' the past. Past records and memories are unreliable and non-robust. We misremember the past and physics dictates that past evidence will eventually be lost. However, we attribute certainty to past events according to how well they are corroborated by independent records and memories. Corroboration helps minimize the effects of unreliable records and makes past theories more robust. Generally, we are more confident in coarser-grain past events and trends. In all, assigning probabilities to past events is analogous to assigning probabilities to future predictions. This is the argument for past openness. However, as will be discussed in Section 4.5.3, the past need not be *as* open as the future.

4.5 An Argument for Many-Worlds Presentism

Granted Section 4.4's arguments, I will now explore the implications of an epistemically open past for the ontology of time. Specifically, in light of past openness, I will argue for the merits

[2]Of course, such a miracle can occur – e.g. the previously discussed Boltzmann brain or alien simulation skeptic catastrophes.

of Presentism over the GBT and Eternalism. In doing so, I propose a "best match" method which links the knowledge and thermodynamic asymmetries.

4.5.1 Epistemology and Ontology

As mentioned at the beginning of Section 4.4, we cannot access the past (future) physically, only epistemically via physical access to *present* records and memories (predictions). Section 4.4 further establishes past openness, analogous to future openness. Thus, there appears no *epistemological* distinction meriting an *ontological* distinction between the past and the future. Furthermore, as discussed in Sections 4.3.1 and 4.3.2, future openness establishes ontological preference for the future's non-existence (as in the GBT) over the future's existence (as in Eternalism). By the same logic, past openness should establish ontological preference for the past's non-existence. To best account for both past *and* future openness (non-existence), Presentism appears the best ontology of time.

Although there exist several forms of Presentism, as discussed in Section 4.3.3, I will now argue that Many-Worlds Presentism is preferable, on the basis of modal reasoning and truthmaking. Even if I knew my own past, present, and future to exist in a specific way, I could still *imagine* alternative possible realities. Although I would not exist in these alternative realities, why should I deny their existence? In fact, Everettian quantum mechanics suggests a relative-state formulation, with an objectively real universal wavefunction, in which all possible quantum measurement outcomes are physically realized in *some* world (Barrett, 2018; Vaidman, 2021). Furthermore, past and future openness limit certainty in our own past and future realities. Some may argue this uncertainty is merely epistemic – i.e. there exists a singular past and future, whether we know it or not. However, a core tenet of modal realism is the Principle of Plenitude, which states that "absolutely *every* way that a world could possibly be is a way that some world *is*" (Lewis et al., 1986). Many-Worlds Presentism's commitment to the existence of all *possible* 'past' and 'future' worlds enables modal truthmaking about the 'past' and 'future', which Dynamic Presentism does not (De et al., 2019)[3]. In conclusion, Many-Worlds Presentism appears best suited to the modal nature of an open past and future.

4.5.2 Limited Present Epistemic Access

I now hope to alleviate Dainton (2016)'s expressed concerns with Presentism, bolstering past openness and Many-Worlds Presentism as accurate descriptions of reality.

In particular, Dainton worries that existence of the present alone leads to implausible constraints on the past.

> [S]ince facts about the past are constituted by what exists now, the truth about the past is restricted by whatever relevant evidence now exists. If we take "evidence" here to mean traces that would enable human investigators to reconstruct accurately a past occurrence, the consequences of this restriction are dramatic: the bulk of the past would vanish. The claim that only a minuscule percentage of dinosaurs left fossilized remains would be false, since the only dinosaurs that existed would be those that left fossilized remains! (Dainton, 2016)

[3] I use quotes around 'past' and 'future' to emphasize that they are perceived. In practice, they are *present* realities we have inhabited or will inhabit.

This argument, however, appears fallacious. Namely, it makes several strong assumptions, which I argue are: 1) predisposed to rule out Presentism and 2) wrong or unfavorable in light of past openness. To begin, Dainton assumes that there are "*facts* about the past". However, Section 4.4's argument for past openness establishes that there are no "facts" about the past. The non-robustness and unreliability of records ensure that we can never be *fully* certain in a past occurrence. At best, we attribute high likelihood to past occurrences. Arguably more problematic are Dainton's assumptions of "*truth* about the past" and that a past occurrence can be "accurately" reconstructed, by which he commits to the ontological existence of a singular, true past. This in and of itself directly contradicts Presentism's commitment to the non-existence of the past, which Section 4.5.1 argues is favorable in light of past openness and modal truthmaking.

Putting aside these anti-Presentist assumptions, Dainton's argument does not actually appear to reduce the merits of Presentism. Rather, it magnifies the limits of human knowledge and consequences of the status of universal natural laws. By claiming that "the bulk of the past would vanish" based on what "human investigators" could accurately infer from present evidence about the past, Dainton seems to conflate human *epistemic* and *ontic* access to the past. In other words, Dainton assumes that what *we*, as humans, accurately *know* about the past is all that *existed* in past. However, humans have extremely limited epistemic access to the present. Knowing *all* that exists in the present would require epistemic access to the full universal micro-state, which clearly is beyond human capability. Therefore, most past evidence actually lies beyond human reach. For example, while we may know of only a few dinosaur fossils, this does not rule out the possibility of many more buried in unexplored territories. Furthermore, evidence of dinosaurs probably exists beyond simple fossils, i.e. DNA particles floating in the air. Just because we have access to limited present dinosaur records, does not mean we can or should dismiss the possibility that more could exist beyond our epistemic reach. While we cannot be certain that millions of dinosaurs roamed the Earth, we also cannot rule out the possibility that dinosaurs "existed" beyond "those that left fossilized remains" – reiterating Section 4.4's argument for past openness.

Even an omniscient human, with full access to the present universal micro-state, would not necessarily be able to accurately reconstruct the past. Only if the omniscient human knows the *natural laws* of the universe and if those natural laws are *governing* and *deterministic*, would the omniscient human be able to accurately calculate past universal micro-states. However, it is highly optimistic to assume humans will ever know the true natural laws – that is, if the natural laws are even governing (Carroll, 2020). Furthermore, there are many arguments for indeterministic natural laws (Hoefer, 2016), largely based on developments in quantum theory.

If anything, Dainton's argument highlights limited human epistemic access to the present (not just the past and future), which generates more past uncertainty and evidence for past openness. In fact, the present epistemically accessible macro-state generally prohibits full certainty in past events of interest. For example, imagine I have pudding in my mouth, a dirty spoon in my hand, and a half-eaten pudding in front of me. Most likely, I put a spoonful of pudding from the bowl into my mouth. However, unless I can verify that no stranger is currently nearby my house, how can I dismiss the possibility of a stranger running into my house, taking half of my pudding, and putting a spoonful of different pudding into my mouth? While highly unlikely, such a past seems physically possible and would result in my present epistemically limited macro-state. In a less contrived

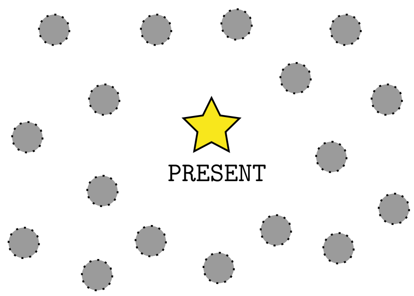

Figure 4.1: An abstract depiction of the Many-Worlds Presentism reality.

example, I can feel that the pudding is cold. However, from this macroscopic knowledge alone, I cannot tell whether the pudding was previously in the kitchen fridge or garage ice-box. In so much as "the proof of the pudding is in the present", my limited epistemic access to the present universal micro-state prevents full certainty even in simple beliefs about the near-past of the pudding.

4.5.3 The Entropic Knowledge "Best Match" Method

I conclude by proposing a Many-Worlds Presentism "best match" method – the Entropic Knowledge Method – which is compatible with past openness, accounts for the knowledge asymmetry, and establishes an interesting link to the thermodynamic asymmetry.

Figure 4.1 abstractly visualizes the Many-Worlds Presentism ontology. Our currently inhabited, instantaneous `PRESENT` world is represented by a star. The star is surrounded by grey circles, representing alternative possible worlds that we do not currently inhabit. Under Many-Worlds Presentism, all possible worlds exist simultaneously with the `PRESENT`. However, we will ever only inhabit a portion of possible worlds. By hopping between different `PRESENT`s, we perceive a passage of time.

To establish a sense of continuity between our inhabited `PRESENT`s, the Entropic Knowledge Method imposes a constraint on our motion between `PRESENT`s. Namely, the `PRESENT` macro-state must be compatible with the previously inhabited `PRESENT` and the next inhabited `PRESENT`'s macro-states, via the natural laws. In other words, the trajectory of inhabited `PRESENT`s must obey the dynamics of the laws of physics. This constraint enables sorting of the large set of possible worlds, which exist under Many-Worlds Presentism, into smaller subsets of potential 'past' and 'future' `PRESENT`s. Furthermore, since the constraint is imposed at the macroscopic (not microscopic) level, many possible worlds' macro-states are compatible with the `PRESENT`'s macro-state – leaving the 'past' and 'future' open[4]. As illustrated in Figure 4.2, the `PRESENT` star is connected to blue circles representing the multiplicity of physically possible 'past' `PRESENT`s and green circles representing the multiplicity of physically possible 'future' `PRESENT`s. Ultimately, however, only some 'past' and 'future' `PRESENT`s are inhabited, defining the red trajectory of inhabited `PRESENT`s.

Perceived 'past' and 'future' openness arise from uncertainty in which 'past' worlds we inhabited and 'future' worlds we will inhabit. Naïvely, this uncertainty could be quantified probabilistically

[4]If the constraint was imposed on the universal micro-state and the governing natural laws were deterministic, only one 'future' and 'past' micro-state would be compatible, which is undesirable in our 'open' ontology.

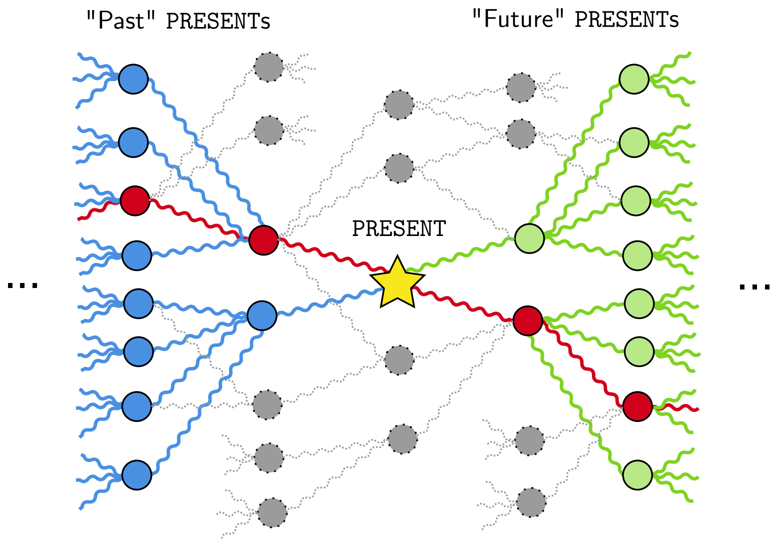

Figure 4.2: An abstract depiction of Many-Worlds Presentism's open 'past' and 'future'. The red connected circles represent our trajectory of inhabited PRESENTs.

via a branch-counting type procedure[5]. Furthermore, each red 'past' and 'future' PRESENT has its own sets of possible 'past' and 'future' PRESENTs. Therefore, once we move into a 'future' PRESENT, our current PRESENT star will become just one of many possible past PRESENTs. Moving further into the 'future' opens up more possible 'past' PRESENTs and reduces our certainty that we previously inhabited the star PRESENT.

As mentioned at the end of Section 4.4.3, although the past is open, this is not to say the past is *as* open as the future. In fact, the knowledge asymmetry claims that we know more about the past than the future. This is reflected in the Entropic Knowledge Method by imposing yet another constraint on the inhabited PRESENTs. Namely, we are more likely to inhabit PRESENTs which possess more possible 'futures' than possible 'pasts'. In other words, we are more likely to inhabit PRESENTs which are more open towards the future than the past. I will not go into metaphysical arguments justifying this preference. However, it should be noted that this tendency towards 'future' openness or uncertainty is analogous to the asymmetric motion of thermodynamic macro-states from lower to higher entropy – i.e. motion towards macro-states occupying larger state-space regions, with more compatible micro-states.

The thermodynamic trend towards higher entropy is statistical – governed by ergodic motion in state-space. While thermodynamic systems generally tend towards higher-entropy micro-states, they occasionally revert to low-entropy micro-states. Similarly, via the Entropic Knowledge Method, we generally move towards PRESENTs with less 'future' than 'past' certainty. However, we should occasionally expect to inhabit a PRESENT with less 'past' than 'future' certainty. For example, when I forget where I put my keys, I have little certainty about the past few minutes, but am very certain that I will spend the next few minutes searching for the keys. In this way, a PRESENT's relative 'future' to 'past' openness acts like a measure of entropy. The Entropic

[5] In light of well-established challenges for branch-counting in Everettian Quantum Mechanics (Saunders, 2021) and given that there may exist infinite possible worlds, a more sophisticated technique – beyond the scope of this work – would need to be developed.

Knowledge Method can, therefore, reconstruct a trajectory of `PRESENT`s by mimicking statistical mechanical motion towards high-entropy states. Our experienced knowledge asymmetry results from inhabiting `PRESENT`s according to the laws of physics and ergodically traversing the 'state-space' of possible `PRESENT`s, generally moving towards higher 'entropy'.

In conclusion, Many-Worlds Presentism and the Entropic Knowledge Method appear well-suited to past openness, account for our temporal experience, and establish a strong symmetry between the knowledge and thermodynamic asymmetries.

Bibliography

Albert, D. Z. (2001). *Time and Chance*. American Association of Physics Teachers.

Barrett, J. (2018). Everett's Relative-State Formulation of Quantum Mechanics. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2018 edition.

Bennett, C. H. (2003). Notes on Landauer's Principle, Reversible Computation, and Maxwell's Demon. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics*, 34(3):501–510.

Broad, C. D. (1923). *Scientific Thought*. Routledge & Kegan Paul.

Brown, H. R. and Uffink, J. (2001). The Origins of Time-Asymmetry in Thermodynamics: The Minus First Law. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 32(4):525–538.

Brown, J. R. (1982). The Miracle of Science. *The Philosophical Quarterly (1950-)*, 32(128):232–244.

Callender, C. (2017). *What Makes Time Special?* Oxford University Press.

Callender, C. (2021). Thermodynamic Asymmetry in Time. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2021 edition.

Carroll, J. W. (2020). Laws of Nature. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2020 edition.

Chakravartty, A. (2017). Scientific Realism. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2017 edition.

Dainton, B. (2016). *Time and Space*. Routledge.

De, M. et al. (2019). Presentism and Modal Realism. *Analytic Philosophy*, 60(3):259–282.

Hartle, J. B. (2005). The Physics of Now. *American Journal of Physics*, 73(2):101–109.

Hoefer, C. (2016). Causal Determinism. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2016 edition.

Horwich, P. (1987). *Asymmetries in Time: Problems in the Philosophy of Science.* MIT Press.

Le Bihan, B. (2020). String Theory, Loop Quantum Gravity, and Eternalism. *European Journal for Philosophy of Science*, 10(2):1–22.

Lewis, D. K. et al. (1986). *On the Plurality of Worlds*, volume 322. Blackwell Oxford.

Mellor, D. H. (1998). *Real Time II.* Routledge.

Mellor, D. H. (2001). The Time of Our Lives. *Royal Institute of Philosophy Supplements*, 48:45–59.

Miller, K. (2018). The New Growing Block Theory vs Presentism. *Inquiry*, 61(3):223–251.

Reichenbach, H. (1956). *The Direction of Time*, volume 65. University of California Press.

Rovelli, C. (2022). Back to Reichenbach.

Saunders, S. (2021). Branch-Counting in the Everett Interpretation of Quantum Mechanics. *Proceedings of the Royal Society A*, 477(2255):2021.0600.

Sider, T. et al. (2001). *Four-Dimensionalism: An Ontology of Persistence and Time.* Oxford University Press on Demand.

Smart, J. J. C. (1963). *Philosophy and Scientific Realism.* Routledge.

Uffink, J. (2001). Bluff Your Way in the Second Law of Thermodynamics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 32(3):305–394.

Vaidman, L. (2021). Many-Worlds Interpretation of Quantum Mechanics. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, Fall 2021 edition.

Wallace, D. (2011). The Logic of the Past Hypothesis. *The Probability Map of the Universe: Essays on David Albert's Time and Chance.*