

Quantum systems other than the Universe

David Wallace*

October 12, 2024

Abstract

How should we interpret physical theories, and especially quantum theory, if we drop the assumption that we should treat it as an exact description of the whole Universe? I expound and develop the claim that physics is about the study of autonomous, but not necessarily isolated, dynamical systems, and that when applied to quantum mechanics this entails that in general we should take quantum systems as having mixed states and non-unitary dynamics. I argue that nonetheless unitary dynamics continues to have a special place in physics, via the empirically-well-supported reductionist principles that non-unitarity is to be explained by restriction to a subsystem of a larger unitary system and that microscopic physics is governed by unitary and largely known dynamics. I contrast this position with the ‘Open Systems View’ advocated recently by Michael Cuffaro and Stephan Hartmann.

1 Introduction: Physics as the study of autonomous systems

Quantum theory is often described as our most fundamental theory of physics, and as such, the task of understanding what quantum theory tells us about the world is often taken to be the task of understanding what the world would be like if quantum theory in fact described it completely. To understand quantum theory in this way commits us (here I follow (Wallace 2022a)) to two assumptions: the *Fundamentality Assumption*, that the theory is *exactly true*, and the *Cosmological Assumption*, that the theory describes *the whole of reality*.¹ Of course, no-one literally thinks that any particular quantum theory — or indeed any theory we currently possess — actually satisfies either assumption, but

*Department of History and Philosophy of Science / Department of Philosophy, University of Pittsburgh; david.wallace@pitt.edu

¹For evidence that the two assumptions are widespread, see (Wallace 2022a, section 1). There are exceptions: notably, the conjunction of the two assumptions gives us something close to Ruetsche’s (2011) notion of a ‘pristine interpretation’ of a quantum theory, and the alternative that she proposes is close to what I propose here. See also Williams (2019), who proposes something similar in the context of effective field theories, though focussed more on rejection of fundamentality than of cosmology.

still: much, perhaps most, interpretative work on quantum theory and indeed on theories of physics in general takes them as a starting point.

At the risk of cliché (cf Quine 1951), we might call these two *dogmas* of interpretation — dogmas that, once made explicit, are very hard to argue for. A moment’s observation of physics shows that hardly ever do physicists seek to model the whole universe, and hardly ever do they seek to model every last degree of freedom of even a finite system. What actually happens in physics, in almost all cases, is that a physical theory is used to model some subsystem of the universe, and to model that subsystem only on certain scales. (Even in general relativity, where our models are called ‘spacetimes’ and it is only natural to think of them as modelling entire Universes, in practice the vast majority of the time we are actually using them to model isolated subsystems of a larger Universe. Even when applied to cosmology, our models usually describe patches of a larger cosmos, and normally track only large-scale features of the geometry and matter distribution.)

Of course, it remains logically possible that our understanding of physics modelling practice is somehow derivative on the interpretation of theories under the Fundamentality and Cosmological assumptions. But there seems little positive reason so to interpret them: what we actually have evidence for is the efficacy of physics in modelling scale-dependent features of subsystems, and it seems far more natural to look for a direct understanding of physical theories — and, in particular, quantum theories — as they are so used.

But *which* subsystems are the proper concern of physics? A very natural answer — one I (too quickly) endorsed in Wallace (2022a) — is that physics models *isolated* systems, systems that do not interact with any other system. But this has multiple problems. For one thing, it is not as straightforward as it might seem to say carefully what ‘isolated’ actually means here — the temptation is to say something like ‘evolving the same way as it would if it were alone in the Universe’, but that is scarcely a counterfactual that we can assess empirically. For another, insofar as we can understand ‘isolation’ it seems to apply to systems that are dynamically isolated from one another in some (usually spatial) way, and does not naturally apply when we are modelling some degrees of freedom of a system and neglecting other, perhaps smaller-scale, degrees of freedom of that same system.

We could go formal. Quantum mechanics — at least as it is normally formulated — is a theory with unitary dynamics, where a system evolves under the Schrödinger equation with respect to some fixed Hamiltonian. We could *define* an isolated quantum system as a system that could be so modelled.

One problem with this definition is that unitarity does not coincide with isolation in the physical sense. Sometimes an environment interacts with a system not by introducing non-unitary dynamics but by adjusting — *renormalizing* — the Hamiltonian of the system. But there is a much more serious problem, which also afflicts any attempt to make precise a more physical notion of isolation. The problem is that ‘isolated system’ under any reasonable disambiguation is inadequate to capture all the systems that physics (including quantum physics) in fact succeeds in modelling. The excited atom that decays

through the emission of radiation; the pollen grain undergoing Brownian motion; the friction-bound system that eventually slows to a stop... all these are paradigmatically important, yet paradigmatically *non*-isolated, systems. Many such cases, indeed, are paradigmatically important *quantum* systems, but they are not isolated in either the substantive sense that they can be physically understood without consideration of their environment, or in the formal sense that their dynamics can be reduced to the Schrödinger evolution of pure states.

Let me suggest a better answer to the question of physics' proper concern. What physics models is *autonomous* systems, and a system is autonomous if it can be effectively modeled *as* a system with a state space and a dynamics, such that many initial states of the system are possible and that the dynamics tells us what the future states are as a (possibly stochastic) function of the present and past states. In particular, a system is an autonomous *quantum* system if it can be effectively modelled as a system with a Hilbert space and a dynamics for (not necessarily pure) quantum states on that Hilbert space.

All the systems I mention above are autonomous in this sense — of course various things must hold of the environment for the system to be autonomous, and broad features of the environment (the temperature of the fluid in which the pollen grain moves, for instance) determine the dynamical parameters, but nonetheless it is possible to state, and test, a dynamical equation for the system's degrees of freedom, without regard for the *details* of the state of the environment.

Autonomy is neither trivial nor automatic: most subsystems of a complex interacting system will *not* be autonomous. (For instance, consider three bodies of similar masses interacting gravitationally: the three bodies form a complex interacting system and no two bodies can be modelled effectively, even stochastically, without tracking the third.) Nor need science be concerned only with autonomous systems. A non-autonomous system can be thought of as requiring not only initial-time data but a constant stream of data describing the influence of the environment, data which cannot itself be modelled (at least not as part of studying this system) but is simply an input. The human body, or an organ of the body, can profitably studied this way: we want to know not what it does in isolation but what it does as a function of the way the environment impinges upon it over time.

But for the most part, *physics* is concerned with autonomous systems. The primary subject matter of physics is dynamics, and non-autonomous systems have no dynamics: their behavior cannot be analyzed dynamically except as a shadow on the wall, cast by the dynamics of a larger system.

My main goal in this paper is to develop what an account of quantum systems looks like given that it should be a theory of *autonomous* systems, but not necessarily isolated systems — still less systems that aim to model the Universe. The starting point in this account (developed in section 2) is the observation that *formal* features of quantum theory, when combined with appropriate choices of system and dynamics — imply a robust notion of autonomy that is equivalent neither to unitary dynamics nor to the physical notion of isolation. We can obtain that notion by (i) writing a formally-exact equation for the evolution of the reduced state of any subsystem of a unitarily-evolving quantum system,

and (ii) identifying physical assumptions under which that formally-exact equation reduces (in some appropriate idealization) to an autonomous dynamical equation for the subsystem alone. This analysis shows that in some sense it is inconsistent to suppose that quantum systems always and only obey unitary dynamics: the existence of large systems with unitary dynamics often entails the existence of smaller autonomous subsystems with non-unitary dynamics.

This observation, especially once we reject the Fundamentality and Cosmological Assumptions, might suggest that unitarity is a ladder that can be kicked away, and that we should replace it entirely with a generalized, non-unitary (perhaps also non-Markovian) notion of quantum dynamics. I do not think this is true. Quantum physics may not be about exactly-true models of the whole Universe, but it does provide very powerful (essentially inductive) evidence for a thesis of *quantum reductionism*: that non-unitarity can and should always be understood as the restriction to a subsystem of larger-scale unitary dynamics, and that larger-scale, phenomenological quantum physics is ultimately explicable microscopically in terms of specific, largely known, unitary quantum theories: non-relativistic particle mechanics and various quantum field theories. I develop this position in section 3, where I suggest four assumptions of quantum reductionism which in a way constitute a naturalistic replacement for the Fundamentality and Cosmological assumptions.

I am not the first to develop a view of quantum systems along these general lines, and indeed this paper is substantially inspired by recent pioneering work by Cuffaro and Hartmann (2023). Their *Open Systems View* rejects the primacy of unitary dynamics and aims at analyzing open quantum systems, not just isolated ones, and my second goal in this paper is to engage with their view and see how it relates to mine. In section 4 I argue that their view tacitly requires the notion of autonomy which I have described above, and that once this is recognized, the main difference between our views concerns the status of quantum reductionism: at least on one interpretation, the Open Systems View is committed to the rejection of most or all of my four assumptions of quantum reductionism. As I go on to argue, in large part I think this feature of their view needs to be rejected, for a mixture of reasons: partly that it makes the very notion of ‘interaction’ mysterious in a way which inhibits physics, partly because of the evidential support for reductionism in physics practice. But in section 5 I identify an important — if partial — exception to my arguments, concerned with how *irreversibility* enters into the dynamics of autonomous systems, and relate the differences between my view and the Open Systems View to unresolved issues in the philosophy of statistical mechanics.

The physics I discuss is fairly standard and for the most part I do not provide original references. An up-to-date standard monograph on open quantum systems is (Breuer and Petruccione 2007).

2 Open quantum systems: deriving autonomy from unitarity

Suppose that some unitarily evolving quantum system with Hilbert space \mathcal{H} and Hamiltonian \hat{H} has a tensor-product decomposition into subsystems S and E , $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$, and a concomitant decomposition of its Hamiltonian,

$$\hat{H} = \hat{H}_S \otimes \hat{\mathbf{1}} + \hat{\mathbf{1}} \otimes \hat{H}_E + \sum_n \hat{C}_n \otimes \hat{D}_n. \quad (1)$$

If the third term in this expression vanishes, \mathcal{H}_S and \mathcal{H}_E will define closed systems with Hamiltonians \hat{H}_S, \hat{H}_E respectively. Even in this case, of course, the embedding of these systems in a larger system means that their joint state might be entangled, so that the appropriate notion of *state* for each subsystem is a mixed rather than pure state. (This already suffices to demonstrate that once we drop the Cosmological Assumption, the principle that a quantum system must be in a pure state becomes untenable; cf (Wallace 2023b) for further development of this observation in the context of quantum statistical mechanics.)

However, the vanishing of the interaction term is not the most general circumstance in which a subsystem of a larger quantum system is autonomous. To see this,² let S_E be a fixed (in general mixed) state of \mathcal{H}_E , which we will require to be invariant under the self-Hamiltonian \hat{H}_E of \mathcal{H}_E , $[S_E, \hat{H}_E] = 0$. Then define the following superoperator \mathbf{J} on the joint density operators of $\mathcal{H}_S \otimes \mathcal{H}_E$:

$$\mathbf{J}\rho = \text{Tr}_E \rho \otimes S_E \quad (2)$$

where Tr_E is the partial trace over \mathcal{H}_E , i. e. $\text{Tr}_E \rho$ is the reduced density operator of \mathcal{H}_S . Note that \mathbf{J} is a projection, $\mathbf{J}^2 = \mathbf{J}$, projecting onto the subspace of states of form $\rho = \rho_S \otimes S_E$. (Also note that here and after, I (i) use **boldface** to label superoperators that act on the space of Hilbert-space operators; (ii) drop the ‘hats’ on mixed quantum states, to avoid cluttering the notation.)

Following Nakajima (1958) and Zwanzig (1961) we can now extract an integro-differential equation for ρ_S , the *Nakajima-Zwanzig equation*:

$$\begin{aligned} \frac{d}{dt} \rho_S(t) &= -i [\hat{H}_S, \rho_S(t)] + \text{Tr}_E \int_0^t d\tau \mathbf{L}_H \mathbf{V}(\tau) (\mathbf{1} - \mathbf{J}) \mathbf{L}_H (\rho_S(t - \tau) \otimes S_E) \\ &+ \text{Tr}_E \mathbf{L}_H \mathbf{V}(t) (\rho(0) - \rho_S(0) \otimes S_E) \end{aligned} \quad (3)$$

where:

²The analysis I provide here is one of two main routes in quantum statistical mechanics to derive autonomous dynamics via scale separation. The other — the quantum version of the so-called BBGKY hierarchy — seeks to develop not an autonomous dynamics of subsystems but an autonomous dynamics of N -point correlation functions for some fixed (usually small) N . See Wallace (2016, sections 2-4,10) for a conceptual discussion of this latter approach; see also (Calzetta and Hu 2008, ch.1) for a comparison.

- \widehat{H}_R is the *renormalized system Hamiltonian*,

$$\widehat{H}_R = \widehat{H}_S + \sum_n \widehat{C}_n \text{Tr}(\widehat{D}_n S_E). \quad (4)$$

- \mathbf{L}_H is the superoperator representing the adjoint action of the Hamiltonian,

$$\mathbf{L}_H \rho = -i [\widehat{H}, \rho]. \quad (5)$$

- $\mathbf{V}(t)$ is a modified propagator,

$$\mathbf{V}(t) = e^{(\mathbf{1}-\mathbf{J})\mathbf{L}_H t}. \quad (6)$$

Equation (3) is exact, and so by itself tells us nothing about whether subsystem S has autonomous dynamics. However, in many cases of physical interest there is *separation of scales*, so that the timescale T_E on which the degrees of freedom of E evolve — and, in particular, effectively equilibrate — is much faster than the timescale T_S on which the degrees of freedom of S evolve and interact with E . When this occurs (something that needs to be argued for in any specific case), then if we take S_E to be the appropriate equilibrium state (i.e. $S_E \propto \exp(-\beta \widehat{H}_E)$ for some inverse temperature β),

1. Since effective equilibration of a system means that the matrix elements of those operators describing collective degrees of freedom — such as those describing coupling with other systems — approach their equilibrium values, for $t \gg T_E$ we ought to be able to approximate ρ_0 by $\rho_S \otimes S_E$ in the third term, so that it approximately vanishes.
2. Since the second term tracks the degree to which previous states of S leave a memory trace in E , and since such traces ought to be effectively unraveled by equilibration, the kernel in that second term ought to fall off on timescales $\sim T_E$, so that for $t \gg T_E$ we can replace the upper limit t of integration with some fixed large (but sub-recurrent) time T .

Equation (3) then simplifies to

$$\frac{d}{dt} \rho_S(t) = -i [\widehat{H}_R, \rho_S(t)] + \text{Tr}_E \int_0^T d\tau \mathbf{K}(\tau) (\rho_S(t-\tau) \otimes S_E) \quad (7)$$

where $\mathbf{K}(\tau) = \mathbf{J} \mathbf{L}_H \mathbf{V}(\tau) (\mathbf{1} - \mathbf{J}) \mathbf{L}_H$. This is an autonomous, time-translation invariant, evolution equation for ρ_S . It is, however, *non-Markovian*: the system has memory, and its future evolution depends on its past and not just its present state.

Systems which are *sufficiently* non-Markovian are often autonomous only by courtesy. The interactions with the system's environment leave traces in the system's past state, and those traces can be used to reconstruct the environment state. (It is plausible that one could write down a wildly non-Markovian

“autonomous” dynamics for one body in the three-body problem.) Autonomy requires that the kernel $\mathbf{K}(\tau)$ falls off quickly compared to the timescales on which the system evolves, so that only its *recent* past is dynamically relevant. We have already assumed this, however, and so equation (7) indeed describes an autonomous dynamics.

In many cases we can further assume that the memory timescale is short enough that $\rho_S(t)$ is effectively static on that timescale, in which case the second term can be approximated as

$$\mathrm{Tr}_E \int_0^T d\tau \mathbf{K}(\tau) \rho_S(t - \tau) \otimes S_E \simeq \mathrm{Tr}_E \int_0^T d\tau \mathbf{K}(\tau) (\rho_S(t) \otimes S_E). \quad (8)$$

If we then define $\mathbf{\Lambda}$ by

$$\mathbf{\Lambda} \rho_S = \mathrm{Tr}_E \int_0^T d\tau \mathbf{K}(\tau) (\rho_S \otimes S_E) \quad (9)$$

we have extracted a *Markovian*, but in general non-unitary, dynamical equation for ρ_S :

$$\frac{d}{dt} \rho_S(t) = -i [\widehat{H}_R, \rho_S(t)] + \mathbf{\Lambda} \rho_S(t). \quad (10)$$

Fairly general results from the theory of open quantum systems (see, e.g., Nielsen and Chuang 2000) tell us that in this case, $\mathbf{\Lambda}$ must be expressible in Lindblad form: that is, there must exist operators \widehat{L}_i such that

$$\mathbf{\Lambda} \rho = \sum_i \left(\widehat{L}_i \rho \widehat{L}_i^\dagger + \frac{1}{2} \left(\widehat{L}_i \widehat{L}_i^\dagger \rho + \rho \widehat{L}_i \widehat{L}_i^\dagger \right) \right). \quad (11)$$

Finally, in some cases the $\mathbf{\Lambda}$ term in (7) can be neglected entirely — for instance, it might be that the energies required to excite the environment out of its ground state are too large compared to the characteristic system energies. In this case the system’s evolution remains unitary; however, it is governed by the renormalized Hamiltonian \widehat{H}_R , not by the ‘bare’ self-Hamiltonian \widehat{H}_S , and so the influence of the environment is not entirely absent.

To sum up, we have a hierarchy of possibilities for a quantum system. At the bottom of the hierarchy, the system might be a ‘system’ only by courtesy, inextricably dynamically interwoven with its environment and studyable only in conjunction with it. Alternately, it might be autonomous, governed by a dynamics which does not depend on the fine details of the environment (in the case we have considered, it depends only on the temperature) and which, if it has memory effects, has those effects only on timescales fairly short compared to its dynamical timescale. Above that in the hierarchy, it might be Markovian, with the memory effects idealizable into an instantaneous dynamics. And if those dissipative effects are entirely absent or can be idealized as such, the system is unitary.

Note that (as anticipated in section 1) ‘isolated’ does not appear in this hierarchy, and indeed is not straightforwardly definable in formal terms. It

cannot be identified with ‘unitary’: as we have seen, even if dissipative effects of other systems can be neglected, other systems’ influence can appear as a renormalization. And this is common in physics: the famous renormalizations of quantum field theory have exactly this form, for instance, where the effects of high-energy physics manifest at lower energies not through any violation of unitarity, but through the renormalization of the parameters in the unitary dynamics. (See (Boyanovsky 2015) for an explicit presentation of quantum-field-theoretic renormalization in this form.)

3 Four dogmas of quantum reductionism

We have seen that it is not self-consistent to suppose that quantum systems invariably or even usually have unitary dynamics. Once we realize that quantum physics, like physics generally, is concerned with autonomous subsystems of the universe and not simply with the universe as a whole or those of its subsystems that can be treated as entirely isolated, we have to recognize that quantum systems generally must be allowed to have mixed states and that quantum dynamics need not be Markovian and, even when Markovian, need not be unitary.

Once we recognize this, we could in principle adopt ‘quantum pluralism’: the belief that there is nothing particularly special about unitary dynamics (or pure states) and that the right way to do quantum mechanics is simply to discover, system by system, what that system’s dynamics actually are, free of any prejudice that they must be unitary. Quantum pluralism is a fundamentally dis-unified quantum metaphysics, very reminiscent of Nancy Cartwright’s “dappled world” metaphysics, in which we abandon any pretense for universality of any dynamical laws and work system-by-system, and quite closely related to Cuffaro and Hartmann’s ‘Open Systems View’ (to which I will return in section 4).

But the vast majority of practicing physicists do not practice quantum pluralism, and I think they are correct not to do so. Physics practice is overwhelmingly committed to what I call *quantum reductionism*, a position that is encapsulated by four principles — four dogmas of quantum reductionism, if you like. The principles can be understood both methodologically (this is how, as practicing physicists, we *should* explore the physics of quantum systems) and metaphysically (this is, in fact, the *truth* about quantum systems). I should be clear that I use ‘dogma’ playfully here: in my view the four ‘dogmas’ of quantum reductionism are fully justified methodologically and probably largely correct metaphysically, and indeed I offer them as a replacement for the two dogmas we have already encountered (the Fundamentality and Cosmological assumptions).

The first principle is that *non-unitarity implies interaction*. A unitarily evolving quantum system may or may not require us to consider its interactions with other systems in order to explain its dynamics (recall renormalization), but if a system is non-unitarily evolving it is always doing so because of its interactions with other systems — or at least, denying this is an absolute last resort. More specifically: the presence of memory terms always indicates that some

additional and unmodelled degrees of freedom are interacting with the system; even in the absence of memory terms, non-unitarity means that the system is becoming entangled with some other system. Methodologically, this principle advises us to look first for interaction-based explanations of non-unitarity, and post intrinsic non-unitarity only as a last resort; metaphysically, it says that emphin fact non-unitarity always has an external cause.

The second principle is that *interaction is derived from dilation*. By a *dilation* of an autonomous quantum system, I mean the inclusion of that system in a larger autonomous quantum system, such that the dynamics of the original system can be derived from the dynamics of the larger system by the means sketched in section 2. Methodologically, this principle advises us to look for a dilation-based account of interaction; metaphysically, it says that such an account exists even if we haven't found it yet.

Section 2 made only minor use of the unitarity of the larger system. The superoperator \mathbf{L}_H defined the dynamics of that system, and I took $\mathbf{L}_H\rho = -i[\widehat{H}, \rho]$, but the derivation of equation (3) would go through largely unchanged if \mathbf{L}_H were itself a non-unitary dynamics. (The main difference is that the first term in the expression would become non-unitary.) But in practice this is not how dilations are normally done in physics, and this is the third principle: that *the dilations that explain interaction are to larger unitary systems*. More precisely: they are to systems that can be idealized as unitary at least for the purposes of explaining the subsystem's dynamics: perhaps their own dynamics, modelled at a sufficiently high level of precision, requires non-unitary terms, but those terms make no significant contribution to the subsystem's autonomous dynamics. (If they did, then by the third principle we did not dilate far enough — but the third principle commits us not just to some in-principle possibility of dilating unitary but to the practical strategy of modelling interaction that way.) Again, methodologically speaking this is advice to the modeller; metaphysically speaking it is a substantive claim about the detailed origin of interaction terms.

At a purely formal level, the dilational assumptions in the second and third principle are almost empty: Stinespring's (1955) dilation theorem (see, e.g., (Nielsen and Chuang 2000) for a modern presentation) says that (under weak additional assumptions; cf (Cuffaro and Hartmann 2023, section 2.2.2)) any well-behaved linear dynamics for a quantum system is the restriction to that system of *some* unitary dynamics on *some* subsystem. The second and third assumptions get their bite from the requirement that the dilation needs to be not just to a fictional system but to some concretely specified system whose dynamics can themselves actually be observed and measured.

Indeed, in physics practice it is ultimately insufficient merely to posit the dynamics of that larger system, even if those posits are observationally confirmed. The fourth principle is that the dynamics of any autonomous quantum system should ultimately be derived, via restriction to subsystems and/or the averaging over microscopic degrees of freedom (insofar as these are distinct), and under appropriate restrictions on a system's state (e.g., fixing temperature or a regime of symmetry breaking), from a short list of "microscopic" — and invariably unitary — quantum theories. The main entries on the list are:

1. Non-relativistic quantum mechanics: The nonrelativistic physics of nuclei and electrons, interacting via Coulomb coupling, occasionally also by Newtonian gravitational coupling, and by coupling to background electric, magnetic, and gravitational fields.
2. Relativistic quantum mechanics: The relativistic generalization of (1), under the self-consistency assumption that interparticle interactions are gentle enough not to lead to particle creation.
3. Quantum optics, the hybrid theory of the quantized electromagnetic field coupled to nonrelativistic nuclei and electrons.
4. Quantum electrodynamics, describing relativistic electrons and nuclei (the latter treated as pointlike) interacting with the quantized electromagnetic field.
5. The Standard Model of Particle Physics, along with its various effective-field-theory approximations (which in fact include (1)-(4).)
6. Low-energy quantum gravity, the effective field theory of gravitation coupled to quantum matter. (Strictly speaking this is the only item we need on the list: (5), and hence (1)-(4), are derivable from it.)

Methodologically, the fourth principle advises us to look to theories on this short list when constructing and/or justifying phenomenological models; metaphysically, it is the reductionist thesis that in fact the systems we encounter are ultimately governed by the theories on the list.

The fourth principle is not inviolate, and is routinely violated in places where ‘new physics’ is expected or at least not ruled out: examples include dark matter, beyond-standard-model particle physics, and the observationally underconstrained mechanisms of CP violation and neutrino mass. But the positing of new physics (like fundamental non-unitarity) is normally a last resort, and this is especially true in terrestrial, low-energy situations — positing a new particle to explain why some high-temperature superconductor works, for instance, would require extremely powerful evidence and extensive argument as to why that particle had not shown up in other contexts.

Why think these four principles in fact constitute standard practice in physics? A really systematic demonstration would require quantitative methods well beyond the scope of this paper, but I will motivate it through two examples. The first and most important is non-equilibrium quantum statistical mechanics, where pretty much any derivation I know adheres to them. The name of the game is to derive equations — complete with quantitatively determined parameters — for collective degrees of freedom such as the spatially-averaged density of a fluid or a particle interacting with an environment. Invariably the derivation begins with a larger system whose dynamics are taken to be unitary; often those dynamics are taken as determined by the specific forms found in microphysical theories non-relativistic quantum mechanics or quantum optics (or, for more exotic applications, relativistic field theory), and where they are not, we are

given at least qualitative reasons to think their general form is consistent with those microphysical theories (or else the theory is specifically described as a toy model, intended to illustrate some general principle rather than describe some specific system). The distinction between ‘phenomenological’ and ‘micro-based’ equations, central in statistical mechanics, is exactly the distinction between equations posited simply to explain the data and equations derived from known microphysics — both have their place, but ultimately we aim to see why the former is a consequence of the latter.

The second is the empirical search for fundamental non-unitary (as occurs in dynamical collapse theories like the GRW theory, continuous state localization, or the Diosi-Penrose model; see, e. g., (Bassi, Dorato, and Ulbricht 2023) for an up-to-date review of extant collapse theories and their empirical status). Here it is essential (the first principle) to rule out the possibility that the non-unitarity is the consequence of interaction with an environment: only if this occurs do we have a genuine violation of quantum mechanics and not just emergent non-unitarity. But (the second principle) it is constitutive of ‘interaction with an environment’ that the interaction can actually be modelled as the consequence of dynamical coupling with a second system. And (the third and fourth principles) evidence of fundamental non-unitarity does not require the experimenter to rule out *every conceivable* dynamical interaction with another system (that would be in practice impossible) but only to rule out those dynamical interactions that our microphysical theories give us reason to expect to be present.

(Of course, fundamental non-unitarity is new physics, and once we are considering new physics, other possibilities appear: perhaps the apparent non-unitarity is after all a consequence of entanglement with a hitherto-unsuspected new particle or field. No single experiment would suffice to falsify the whole quantum framework: the point is that one way or another such a discovery would point to new physics, but it would do so only because of the methodological commitment to attempt first to explain the observations in terms of known — and unitary — microphysical theories.)

Put aside whether the four principles correctly describe physics practice. Why think that they are correct? The obvious answer is just that the strategy of accounting for the dynamics of quantum systems in this way has been spectacularly successful. We have extremely good evidence that, for instance:

- Where nuclear interactions and electromagnetic radiation can be neglected, systems are accurately described by nonrelativistic quantum mechanics.
- Where nuclear interactions but not electromagnetic radiation can be neglected, systems are accurately described by quantum electrodynamics (and, in the nonrelativistic regime, by quantum optics).
- Where effects at energy scales above 10 TeV can be neglected, systems are accurately described by the standard model.
- Where Planck-scale physics can be neglected, self-gravitating systems (insofar as quantum effects are also present) are accurately described by low energy quantum gravity (cf Wallace 2022b and references therein).

And in each of these cases, the ‘can be neglected’ clause is not empty tautology but can be quantitatively assessed (for instance, in terms of typical energy levels) and normally checked if necessary by evaluating leading-order corrections from a more fundamental description. In short, both the metaphysical and methodological versions of the four dogmas of quantum reductionism are justified by the evidence that we actually do know what the correct microphysical theories of the world are, outside some fairly exotic contexts, and that *in fact* those theories are unitary quantum theories.

Furthermore, this ‘extremely good evidence’ does not simply come from microphysical explorations: it does not, that is, just consist of evidence from particle accelerators and the like combined with an *a priori* commitment to reductionism. It consists in large part of the thousands of successes of the overlapping fields of (equilibrium and non-equilibrium) quantum statistical mechanics, condensed matter physics, and effective field theory in deriving empirically-confirmed claims at the phenomenological level from quantum microphysics. We have not by any means explained *every* quantitatively-known quantum dynamical system in this way, but we have explained a great many: to pick just a few examples, we have quantitatively-confirmed accounts of:

- the linear responses of many forms of matter to external perturbations (through the relation of linear-response-theory coefficients to two-time expectation values and the evaluation of the latter in perturbation theory; see (Zwanzig 1965) for a classic review);
- the master equations governing atomic and nuclear decay of excited states;
- the phenomena of superconductivity and superfluidity in a wide class of solids and fluids via the spontaneous breaking of (respectively) local and global phase symmetry;
- the effective theory of pion scattering (via the identification of the pion with the Nambu-Goldstone boson associated to spontaneous breaking of chiral symmetry in quantum chromodynamics);
- the suppression of quantum interference in mesoscopic systems through decoherence theory.

Reductionist claims of this modest kind are fairly uncontroversial in physics: Anderson, for instance, is often held up by philosophers as an arch-critic of reduction, but his famous (1972) discussion of the subject begins thus:

The reductionist hypothesis may still be a topic for controversy among philosophers, but among the great majority of active scientists I think it is accepted without question. The workings of our minds and bodies, and all the animate or inanimate matter of which we have any detailed knowledge, are assumed to be governed by the same set of fundamental laws, which except under certain extreme conditions we feel we know pretty well.

But Anderson isn't wrong that they are controversial in philosophy of science. And notwithstanding the various quantitative successes of quantum statistical mechanics, both philosophers and physicists have claimed that various paradoxes of principle block its derivations: the measurement problem, the thermodynamic limit, the irreversibility problem of statistical mechanics. (I return to the latter in section 5.) Defending the thesis systematically is far more than I can do here; I will settle for the weaker claims that (a) a methodological commitment to reductionism in this sense has been intimately entangled with quantum theory itself since its inception (and indeed before), and (b) that this methodology has been extremely successful scientifically.

Although my focus here is on quantum theory, modern physics' use of *classical* physics adheres to a modified version of the same four dogmas. Here too there is a plurality of dynamics, with many classical systems obeying stochastic dynamics (like the Langevin equation) and/or dissipative dynamics (like the Navier-Stokes equation) instead of Hamiltonian dynamics (the classical analog of unitary dynamics). But physics is methodologically committed (1) to explaining non-Hamiltonian dynamics by dilation to larger systems, insofar as it can be explained classically at all, and (ii) to model those larger systems via Hamiltonian dynamics. The difference is that in classical physics there is always the possibility of a *quantum* explanation of the form of a classical equation, and there is no simple, short list of classical microphysical theories that suffice to explain all of classical dynamics: unification to that degree requires quantum mechanics. For more on this point, see (Wilson 2006, pp.196-7) and (Wallace 2020a, section 4).

4 The Open Systems View: a comparison

Cuffaro and Hartmann (2023; henceforth CH) have recently expounded and defended what they call the 'open systems view' of quantum mechanics, according to which we should take open, rather than closed, quantum systems as fundamental and treat closed systems as a special case of open systems.³ The Open Systems View clearly has a lot in common with the view I have been developing and advocating; in this section I want to consider how the two are related.

Firstly: what exactly is the Open Systems View? CH give two characterizations, which are somewhat subtly related. The first is a *physical characterization*: an open system is one which is dynamically coupled to its environment, while a closed system is isolated from it. (Open systems are "systems evolving under the influence of their environment".) The second is a *formal characterization*: an open quantum system is a quantum system with dynamics that is

³CH distinguish between (i) the Open Systems View, which is the general position that open systems are prior to closed systems in science, and (ii) the 'general quantum theory of open systems (GT)', which (if I interpret them correctly) they take as an implementation of the Open Systems View in the particular context of quantum mechanics. Since my concern here is exclusively with quantum mechanics, I use 'Open Systems View' to refer to CH's quantum-mechanical position; for ease of comparison with CH's own paper, read 'Open Systems View' as 'GT'.

linear and preserves the consistency of the Born probability rule but need not be unitary; a closed quantum system has unitary dynamics. (CH are mostly interested in Markovian open systems, but I don't think the Markovian restriction is essential to them.)

Thus stated, the Open Systems View seems compatible with the present view, but it contains an important tacit assumption worth making explicit: the 'open systems' CH consider are required to be *autonomous* systems. In general, as we have seen, 'systems evolving under the influence of their environment' might have no self-contained dynamics at all, but CH are specifically concerned with systems that do have well-defined dynamics; they simply want to drop the assumption of unitarity.

Once this is recognized, one of CH's original arguments for the Open Systems View loses its force: they argue that 'there are reasons stemming from physics to believe that the cosmos is the *only* closed system that truly exist', because no system other than (perhaps) the cosmos can be entirely isolated from its environment. But strictly speaking, no system has exactly autonomous dynamics either: autonomy, like isolation, is always an idealization, and manifestly there are many systems studied in quantum theory which are satisfactorily idealized as isolated (and unitary).

This is not, however, a serious problem for CH. What matters in their argument, and what does the bulk of the work in their paper, is the physical argument that there are plentiful examples of autonomous quantum systems that are not isolated, and the formal argument that subsystems of unitary quantum systems can display autonomous but non-unitary dynamics. I endorse both claims in the current paper (indeed, CH's formal argument is closely related to my argument in section 2, though CH are less concerned with the mechanisms that ensure autonomy); thus far, the Open Systems View seems fully compatible with the present view. (And CH provide an invaluable explication and defense of some of the metaphysics of open systems, and its interaction with issues like the quantum measurement problem, most which is applicable here also.)

Looking at the details of CH's view, however, reveals very substantial points of disagreement, which can be summarized as saying that CH appear to reject most or all of the four principles of quantum reductionism.

The first principle is that non-unitarity implies interaction. CH are equivocal about this principle: they note (section 5.3) that non-unitary terms in a system's dynamics 'are most naturally interpreted as due to the system's interaction with an external system', but are willing to countenance non-unitary dynamics for the cosmos as a whole and in the context of dynamical-collapse theories. (As an aside: as we have seen, renormalization means that we should not think of interactions as simply introducing non-unitarity: some part of a system's unitary dynamics might also be explained by its environment.)

CH *seem* to deny the second principle, that interaction is derived from dilation. (And insofar as they do, they effectively deny the third and fourth principles, which presume the second.) They say explicitly (section 2.2) that

[u]nlike [unitary quantum mechanics]. for which the dynamics of

an open system of interest, \mathcal{S} , is obtained via a contraction ... of the total dynamics of $\mathcal{S}+\mathcal{E}$ to the state space of \mathcal{S} ; in [the Open Systems View], the dynamical equations that govern the evolution of \mathcal{S} pertain directly to \mathcal{S} itself. Systems are modelled as genuinely open. . . we do not describe the influence of the environment on \mathcal{S} in terms of an interaction between two systems, but instead represent the environment’s influence on the dynamical equations that we take to govern the evolution of \mathcal{S} .

I quote at length because this seems to encapsulate much of the essence of CH’s view, and also because I find it somewhat difficult to interpret. CH seem to be saying that a dynamical equation describing the evolution of \mathcal{S} in the presence of its environment ‘pertains directly to \mathcal{S} itself’ if it is posited directly, but not if it is derived from contraction of the dynamics of $\mathcal{S}+\mathcal{E}$. But at the formal level, the dynamics of \mathcal{S} pertains to \mathcal{S} simply by virtue of *being* an autonomous dynamics for \mathcal{S} : that is, being a rule that determines the future state of \mathcal{S} from its present and past states, perhaps given some broad-brush information (e.g. temperature) about \mathcal{E} but without regard to \mathcal{E} ’s fine-grained state. And at the physical level, surely what it means for a term in the equation to pertain to $\mathcal{S}+\mathcal{E}$ is that it represents the effect of \mathcal{E} on \mathcal{S} : if, physically, some interaction pertains only to \mathcal{S} , it’s difficult to see how it can represent the effects of something other than \mathcal{S} .

More seriously: if we drop the idea that the effects of an environment on a system can be analyzed in terms of dynamical couplings between the two, it becomes difficult — practically but also conceptually — to distinguish intrinsic non-unitarity from the effects of the environment, and to understand just how it is that the environment influences the system. To say that a system is influenced by the environment is normally to say something, even if schematically, about how that influence occurs and how it might be manipulated, and saying that in quantum theory seems to require us to model the dynamical interaction between the two — which is what CH seem to say we should not do.

In quantum computation, for instance, it is of great importance to minimize the effects of environment on system. Doing so requires us to understand the physics of those effects, so as to manipulate them: we want to know, for instance, whether and how environment-induced non-unitarity depends on environmental temperature, and whether various forms of screening might reduce the environment’s effects. Doing so — at least in the standard practice of physics, but arguably also in principle — requires us to model system plus environment together, so that the induced dynamics of the system can be related to the environment’s macrostate and the form of the interaction between the two. (It matters, of course, that the practice of modelling system-environment interactions in this way is highly successful: physics goes far beyond the phenomenological here.)

Similarly, in testing dynamical-collapse theories, we have seen already that the name of the game is to identify and eliminate environment-induced decoherence, so that any residual non-unitarity must be regarded as intrinsic. If the

effects of the environment are not to be analyzed in the usual way via dynamical interactions, it is opaque — to me at any rate — how even in principle we could obtain evidence for intrinsic non-unitarity.

Finally, there is a question of self-consistency for the Open Systems View thus understood. Suppose \mathcal{S} and \mathcal{T} are both open systems. There is a system $\mathcal{S} + \mathcal{T}$ that comprises both \mathcal{S} and \mathcal{T} ; it, too, presumably is an open system. The dynamics for \mathcal{S} alone, and \mathcal{T} alone, will be derivable from the joint dynamics of \mathcal{S} and \mathcal{T} , and that derivation will reveal the dynamics of \mathcal{S} and \mathcal{T} separately to be, after all, derived from contraction of the joint dynamics of $\mathcal{S} + \mathcal{T}$.

For these reasons, I doubt that it is viable for the Open Systems View to reject the second principle — and, partly because of this, I am unclear as to whether CH intend to do so. If after all we interpret them as *accepting* the second principle, but continue to take them seriously about the inessential role played by unitary dynamics in particular, then we would obtain a coherent version of the Open Systems View which rejects the third principle (that dilations are always to larger, closed systems). In a view like this, we can always dilate a system in order to study its interactions with its environment, but the dilated system in general will itself be an open system — there is no expectation that eventually this process terminates in a closed (i.e., unitarily-evolving) system.

A view of this kind could be combined either with the denial or the acceptance of the fourth principle (that the systems to which we dilate are, outside exotic situations, governed by known dynamical laws. If the fourth principle is denied, we end up with an anti-reductionist position that nonetheless has a consistent methodology for scientifically investigating the dynamical causes of non-unitarity in any given system. If it is accepted, we get a view where the fully dynamics of any given system is a combination of known and unitary microdynamics plus a non-unitary interaction term; the latter term can be analyzed by dilation, but only to a system which in turn has non-unitary interactions, with the whole process continuing without limit.

Both of these views seem coherent. The problem with either is inductive: *in fact* physics, most notably in the (overlapping) realms of statistical mechanics, condensed matter, and effective field theory, has had remarkable success explaining the quantitative dynamics of open quantum systems by dilation to larger *closed* systems, and by assuming that at sufficiently short lengthscales systems are governed by unitary, and largely known, dynamics. This both provides support for the methodological validity of all four principles, and would be difficult to explain without their metaphysical validity.

At this point I should note that CH give several concrete examples of how rejecting the primacy of unitary dynamics supposedly might benefit physics. For reasons of space I will not discuss all of the examples (the case of quantum reference frames is especially interesting and merits further attention), but I will discuss two of their main examples: the increased generality of the Open Systems equations for non-unitary evolution, and the evaporation of black holes.

The first of these is best understood by looking again at equation (11), the Lindblad equation for non-unitary evolution of a density operator. As I noted, any autonomous Markovian dynamics obtained for a system by tracing over the

environment can be put in Lindblad form. But as CH point out, this is not the most general (probability-preserving, Markovian) dynamics that can be written down for a density operator: there exist open-system dynamical equations that cannot be generated from dilation to a larger closed system.

But this possibility is not an advantage for the Open Systems View. Ultimately, it is a matter of experiment which dynamical equation applies to a system. If at some point we were to empirically confirm that there are open quantum systems whose evolution cannot be derived from a larger unitary system, there would be no further need to discuss the respective philosophical merits of the Open Systems View and a more conventional reading of quantum mechanics: that ‘more conventional reading’ would have been flat-out disconfirmed. But by the same token, the mere logical possibility of that disconfirmation does not speak in favor of the Open Systems View. Indeed, at least methodologically the reverse is the case: precisely because of the profound significance of a disconfirmation of orthodoxy (as much because of its incompatibility with the Standard Model as because of its nonunitarity *per se*), it is advisable to try hard to model a theory within the standard framework and only to abandon doing so if it proves unviable. (This is just the standard principle that one eliminates the possibility of an explanation within extant physics before positing new physics.)

As for the second example — black hole evaporation — Hawking (1975) proposed that the evaporation process was non-unitary; CH observe that this is an entirely legitimate possibility on the Open Systems View and that (section 1) “at least part of the motivation for wanting to reject it seems to amount to nothing more than that it runs counter to quantum theory and the fundamental dynamics described by it [on a Closed Systems view]”.

At some level this is a semantic dispute: we could take intrinsic (i. e., not environment-induced) non-unitarity to be included in quantum mechanics, or we could take it as a modification of quantum mechanics. (On either approach, *experimental* evidence for intrinsic non-unitarity would require prior elimination of environmental factors, but that does not apply in this theoretical context.) What matters is whether physicists were disinclined to consider Hawking’s proposal due in part to an in-principle rejection of non-unitary dynamics, and it is not at all clear to me that this is the case — of course many, many people have commented on the information-loss paradox, and some of those comments (especially in popular and semi-popular articles and books) have claimed a contradiction with quantum mechanics, but more detailed and technical arguments seem to have carried more weight — notably Banks, Peskin and Susskind’s (1984) argument that non-unitary effects would be magnified to empirically-ruled-out scales by loop corrections in quantum field theory, and more decisively, the evidence for AdS/CFT duality (Maldacena 1998; Witten 1998; cf Wallace (2020b) and references therein for a conceptual discussion), which if correct shows that black hole evaporation is dual to a manifestly unitary process.

In any case, the choice between quantum reductionism and the Open Systems View is largely empirical, and rests on whether to accept my claim that reductionism (in the weakish sense I intend) is indeed underwritten by the successes of physics. Even if this claim is granted, though, it has one interesting

exception, one place where even a reductionist analysis of open quantum systems potentially needs resources not available in unitary quantum mechanics: the problem of the arrow of time in non-equilibrium statistical mechanics.

5 The arrow of time

Unitary dynamics is reversible: the Schrödinger equation can be run backwards in time as easily as forwards, and there is no intrinsic difference between the two. This is not true for non-unitary dynamics: an autonomous but non-unitary dynamics is irreversible, in the sense that two distinct density operators evolving under that dynamics approach one another in trace norm. Physically (and assuming a dilation-based description of the system-environment interaction) this corresponds to information loss: the system gets entangled with the environment; taking the partial trace discards this information. In any case, the future direction in time is picked out as the direction of information loss and of convergence in trace norm.⁴

It is then in principle impossible to find an *exceptionless* derivation of autonomous dynamics for \mathcal{S} from unitary dynamics for $\mathcal{S} + \mathcal{E}$, except in the special case where \mathcal{S} is not just autonomous but itself unitary. And so some initial assumption must have been made in the “derivation” sketched in section 2. It is easy to identify that assumption: I assumed that the third term in the Nakajima-Zwanzig equation (3) vanished, something I argued for on the grounds that the microscopic degrees of freedom of the environment “effectively equilibrated”. Of course, equilibration is itself a time-directed process, and the assumption that a given state equilibrates amounts to the assumption that the state lacks the delicate correlations that would prevent equilibration (see Wallace (2023a) for more on this point).

The conclusion of all this is that deriving autonomous dynamics from unitarity requires an assumption about the initial state of system-plus-environment — a harmless-seeming, intuitively-reasonable assumption, perhaps, but an assumption which holds for *initial* states but on pain of contradiction cannot hold in time-reversed form for *final* states. In most approaches to the foundations of statistical mechanics, it is this initial-state assumption that underpins irreversibility and which must be explained to explain irreversibility: proposed explanations have included an operationalist (and itself inherently time-directed) requirement that systems must always be thought of as prepared by an agent who is unable to prepare them in sufficiently complicated states to prevent effective equilibration (Myrvold 2021), a (time-directed) principle of causal reasoning that legitimates certain no-correlation assumptions for initial but not final states (Maudlin 2007), and a cosmological boundary condition of one kind or another (Penrose 1994; Albert 2000; Wallace 2023a).

There is, however, another tradition in foundations of statistical mechanics:

⁴Note that this distinction is independent of whether the unitary dynamics being considered is time-reversal-invariant in the usual sense: even time-reversal-noninvariant unitary processes are reversible.

the *interventionist* tradition which seeks to explain irreversibility by the effects of an external environment. Applied to our combined system $\mathcal{S} + \mathcal{E}$, an advocate of this tradition would point out that (i) realistically $\mathcal{S} + \mathcal{E}$ is itself not completely isolated, and (ii) adding even a small amount of non-unitarity to the combined dynamics of $\mathcal{S} + \mathcal{E}$ probably suffices to destroy the delicate correlations that would allow any given initial state *not* to proceed effectively to equilibrium with respect to the microscopic degrees of freedom of \mathcal{E} .

The majority view in foundations of statistical mechanics (defended explicitly by, e.g., Sklar (1993, pp.250-254)) is that interventionism is not the right way to explain irreversibility, in part not because the argument is incorrect in itself but because it just pushes the problem of explaining irreversibility back one step, leaving unanswered the *general* question of why systems are so aligned as to produce a consistent direction of irreversibility. I will not attempt to resolve this controversy here. But it is worth observing that environment-based accounts of irreversibility fit very naturally into the Open Systems View — especially into the rather modest version of that view I discussed at the end of section 4, which endorsed the first, second and fourth principles of quantum reductionism while dropping the requirement that autonomous dynamics are always explainable via dilation to a *closed* system. Here at least, the advocate of an Open Systems View has a potential rejoinder to the inductive argument based on the success of closed-system, reductive analyses of autonomous dynamics.

6 Conclusions

If quantum mechanics were to be understood first and foremost as an exactly-true theory of the universe, with its applications to subsystems and to restricted scales understood as derivative or instrumental, then it would be natural to assume that unitarity, and the use of pure rather than mixed states, were fundamental. But there is no reason to do this: the Fundamentality Assumption (that physical theories should be understood as if they were exactly true) and the Cosmological Assumption (that physical theories should be seen as modeling the whole Universe) have no justification in scientific practice or naturalized epistemology. Outside a few explicitly cosmological contexts, our goal in physics (and specifically in quantum physics) is in general to model the dynamics of autonomous, though not necessarily isolated, subsystems of the larger Universe.

Thus understood, quantum mechanics inevitably leads to the conclusion that quantum states need not be pure and that quantum dynamics need not be unitary. Quantum statistical mechanics provides many examples in which a subsystem of a larger unitary system displays autonomous dynamics not because its degrees of freedom are simply uncoupled to those of the rest of the system but because scale separation and effective equilibration of the system's environment leads to autonomous, but non-unitary, dynamics for the subsystem.

This does not, however, imply a simple pluralism of dynamics in which unitary evolution has no special place. Physics is committed, and justifiably so, to four broadly-reductionist principles that can be seen as a naturalized replace-

ment for the unjustified Fundamentality and Cosmological Assumptions, namely (i) that non-unitary dynamics should be explained (except as a last resort) by interaction with an environment; (ii) that ‘interaction with an environment’ should be mathematically modelled through dilation to a larger quantum system; (iii) that for the purposes of modelling an autonomous subsystem it always suffices to dilate to a *unitary* system; (iv) that we should seek to analyze any such system, ultimately, in terms of the known, and unitary, dynamics of microscopic physics: nonrelativistic particle mechanics, QED, the Standard Model, and so forth. (i) is largely justified methodologically (non-unitarity is evidence for *intrinsic* non-unitarity only once interaction with an environment is ruled out). (ii) is constitutive of what it means for a term in the dynamics of a system to represent interaction with other systems. (iii) and (iv) are not true *a priori*, but receive very strong evidential support from the manifold successes of quantum mechanics, especially in statistical mechanics, condensed-matter physics, and effective field theory.

Cuffaro and Hartmann, in developing their Open Systems View, also reject the Fundamentality and Cosmological assumptions, and also center the principle that autonomous quantum systems often evolve non-unitarily. Their view differs from the view here by its rejection of at least some of these four principles of quantum reductionism, though it is to some degree a matter of interpretation just which principles they wish to reject. At least on my reconstruction of their views, it is a semantic matter whether or not they accept the first principle, and a methodological necessity for them to accept the second principle. The status of the fourth principle depends on one’s assessment of the success of reductionist methods (in the moderate sense I intend for ‘reductionist’) in modern quantum mechanics. The third principle is most interesting: its acceptance or rejection is tied to how the problem of irreversibility in statistical mechanics is to be solved.

Acknowledgements

Thanks to Michael Cuffaro and Stephan Hartmann for the stimulating work that inspired this paper, and for helpful correspondence; thanks also to two anonymous reviewers for thoughtful comments on the draft version of the paper.

References

- Albert, D. Z. (2000). *Time and Chance*. Cambridge, MA: Harvard University Press.
- Anderson, P. W. (1972). More is different: Broken symmetry and the nature of the hierarchical structure of science. *Science* 177, 393.
- Banks, T., L. Susskind, and M. E. Peskin (1984). Difficulties for the evolution of pure states into mixed states. *Nuclear Physics B* 244, 125–134.
- Bassi, A., M. Dorato, and H. Ulbricht (2023). Collapse models: A theoretical, experimental and philosophical review. *Entropy* 25, 645.

- Boyanovsky, D. (2015). Effective field theory out of equilibrium: Brownian quantum fields. *New Journal of Physics* 17, 063017.
- Breuer, H.-P. and F. Petruccione (2007). *The Theory of Open Quantum Systems*. Oxford: Oxford University Press.
- Calzetta, E. A. and B.-L. B. Hu (2008). *Nonequilibrium Quantum Field Theory*. Cambridge: Cambridge University Press.
- Cuffaro, M. E. and S. Hartmann (2023). The open systems view. *Philosophy of Physics* 2, 6.
- Hawking, S. (1975). Particle creation by black holes. *Communications in Mathematical Physics* 43, 199.
- Maldacena, J. M. (1998). The large N limit of superconformal field theories and supergravity. *Advances in Theoretical and Mathematical Physics* 2, 231–232.
- Maudlin, T. (2007). What could be objective about probabilities? *Studies in History and Philosophy of Modern Physics* 38, 275–291.
- Myrvold, W. C. (2021). *Beyond Chance and Credence*. Oxford: Oxford University Press.
- Nakajima, S. (1958). On quantum theory of transport phenomena: steady diffusion. *Progress of Theoretical Physics* 20, 948–959.
- Nielsen, M. A. and I. L. Chuang (2000). *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.
- Penrose, R. (1994). On the second law of thermodynamics. *Journal of Statistical Physics* 77, 217–221.
- Quine, W. V. O. (1951). Two dogmas of empiricism. *Philosophical Review* 60, 20–43.
- Ruetsche, L. (2011). *Interpreting Quantum Theories*. Oxford: Oxford University Press.
- Sklar, L. (1993). *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. Cambridge: Cambridge University Press.
- Stinespring, W. (1955). Positive functions on c^* -algebras. *Proceedings of the American Mathematical Society* 6, 211.
- Wallace, D. (2016). Probability and irreversibility in modern statistical mechanics: Classical and quantum. To appear in D. Bedingham, O. Maroney and C. Timpson (eds.), *Quantum Foundations of Statistical Mechanics* (Oxford University Press, forthcoming). Preprint at <https://arxiv.org/abs/2104.11223>.
- Wallace, D. (2020a). Lessons from realistic physics for the metaphysics of quantum theory. *Synthese* 197, 4303–4318.
- Wallace, D. (2020b). Why information loss is paradoxical. In N. Huggett, K. Matsubara, and C. Wuthrich (Eds.), *Beyond Spacetime: the Foundations of Quantum Gravity*. Cambridge: Cambridge University Press.

- Wallace, D. (2022a). Isolated systems and their symmetries, part I: General framework and particle-mechanics examples. *Studies in the History and Philosophy of Science* 92, 239–248.
- Wallace, D. (2022b). Quantum gravity at low energies. *Studies in History and Philosophy of Science* 94, 31–46.
- Wallace, D. (2023a). The logic of the past hypothesis. In B. Loewer, B. Weslake, and E. Winsberg (Eds.), *The Probability Map of the Universe: Essays on David Albert's Time and Chance*, pp. 76–109. Cambridge, MA: Harvard University Press. Originally published online in 2011 (<http://philsci-archive.pitt.edu/8894/>).
- Wallace, D. (2023b). Thermodynamics with and without irreversibility. To appear in Olimpia Lombardi and Cristian Lopez (eds.), *The Arrow of Time: From Local Systems to the Whole Universe* (Cambridge University Press, forthcoming). Preprint at <https://philsci-archive.pitt.edu/22273/>.
- Williams, P. (2019). Scientific realism made effective. *British Journal for the Philosophy of Science* 70, 209–237.
- Wilson, M. (2006). *Wandering Significance: An Essay on Conceptual Behavior*. Oxford: Oxford University Press.
- Witten, E. (1998). Anti de Sitter space and holography. *Advances in Theoretical and Mathematical Physics* 2, 253–291.
- Zwanzig, R. (1961). Memory effects in irreversible thermodynamics. *Physical Review* 124, 983.
- Zwanzig, R. (1965). Time-correlation functions and transport coefficients in statistical mechanics. *Annual Review of Physical Chemistry* 16, 67–102.