Quantitative Parsimony: Probably for the Better

Lina Jansson and Jonathan Tallant[*]

Abstract

Our aim in this paper is to offer a new justification for preferring theories that are more quantitatively parsimonious than their rivals. We discuss cases where it seems clear that those involved opted for more quantitatively parsimonious theories. We extend previous work on quantitative parsimony by offering an independent probabilistic justification for preferring the more quantitatively parsimonious theories in particular episodes of theory choice. Our strategy allows us to avoid worries that other considerations, such as pragmatic factors of computational tractability, etc., could be the driving ones in the historical cases under consideration.

1. Introduction

A series of recent papers have defended the notion that in addition to considerations of qualitative parsimony (minimizing the types of entities postulated) there are episodes of theory choice where a principle of quantitative parsimony (minimizing the number of entities postulated) is plausible.[1] Our aim in this paper is to offer a new justification for preferring theories that are more quantitatively parsimonious than their rivals. In doing so we will discuss cases where it seems clear that those involved opted for more quantitatively parsimonious theories.

However, our *justification* for quantitative parsimony is not an inductive one from these cases. Instead, we extend previous work on quantitative parsimony by offering an independent probabilistic justification for preferring the more quantitatively

---

[*]Authors are listed in alphabetical order and are equal contributors.
[1] See Nolan ([1997]), Baker ([2003]), Huemer ([2009]), Tallant ([2013]) and Baron & Tallant ([ms]).

parsimonious theories in particular episodes of theory choice. This strategy allows us to avoid worries that other considerations, such as pragmatic factors of computational tractability, etc., could be the driving ones in the historical cases under consideration. Since our justification is independent of the specific cases we can show that an epistemic justification for a preference for the quantitatively more parsimonious alternative *can* be given in these cases (whether or not we take that to historically have been the main factor in theory development).

Nolan ([1997]) presents the most frequently discussed example, which concerns the postulation of a spin-½ particle in order to account for the (perceived) missing spin (and features of energy and momentum) in beta decay, and it is here we will begin. This case comes with many historical complications so we will focus on Baker's ([2003], pp. 246-7) overview of the central issue that simplifies the case to deal only with explaining the missing spin. While this does leave out other considerations, we think that this is not a problem for our view since, to repeat, the justification for preferring quantitatively parsimonious theories in particular episodes of theory choice is not one that relies on the specifics of these cases. Moreover, since it is a frequently discussed case, it is useful to make use of it in order to show how our account differs from, for example, Baker's. Here is the case.

> If we focus for the moment on explaining the missing spin, then the following series of alternative neutrino hypotheses can be straightforwardly constructed:
>
> H_1 1 neutrino with a spin of ½ is emitted in each case of Beta decay
> H_2 2 neutrinos, each with a spin of ¼ are emitted in each case of Beta decay
> H_3 3 neutrinos, each with a spin of 1/6 are emitted in each case of Beta decay
>
> and, more generally, for any positive integer n,
>
> H_n n neutrinos, each with spin of 1/2n are emitted in each case of Beta decay
>
> Each of these hypotheses adequately explains the observation of a missing ½-spin following Beta decay. Yet the obvious default hypothesis, both intuitively and from the point of view of actual scientific practice, is that exactly 1 neutrino is emitted in each case.

This case *seems* to be one where considerations of quantitative parsimony could be in play. However, there are complications that lead us to think that the analysis of the case requires review. We will return to the case later in the paper.

Nolan also proposes a second example of quantitative parsimony, much less discussed, that revolves around Avogadro's law. To set up the case, we (following Nolan) must

make explicit three assumptions made by Avogadro. The first assumption is that gases are made up of tiny particles, and that it is the joining or separating of the particles that constitutes the process of chemical reactions (see Nolan [1997] p. 335). The second assumption, '[…] was the Gay-Laussac law of combining volumes, which stated that volumes of gases at equivalent temperatures and pressures combined in fixed ratios, and furthermore that these ratios were in low whole numbers (so that e.g. one volume of oxygen combined with two of hydrogen to produce water, one of nitrogen combines with three of hydrogen to produce ammonia, and so on)' (Nolan, [1997], p. 336). The third and final assumption made by Avogadro was that a given volume of gas, at a given temperature and pressure, would contain the same number of gas-molecules (Nolan, [1997], p. 336). Since the ratio of volumes of hydrogen to oxygen in the production of water is 2:1 Avogadro concludes that '[…] water results from the union of each molecule of oxygen with two molecules of hydrogen' (Avogadro (English translation), [1923], p. 30).[2]

We can then present the problem case:

> Avogadro thought it reasonable to suppose that, for example, since two volumes of hydrogen combine with one volume of oxygen to produce water, there is twice as much hydrogen as oxygen in water. Furthermore, if one volume of oxygen was reacted with two of hydrogen, the natural thing to expect would be that one volume of water be produced (since there are twice as many hydrogen as oxygen in water, there would be as many water molecules as there were oxygen, and half as many as there were hydrogen). However, the experimental result was different: combining two volumes of hydrogen and one of oxygen produced two volumes of steam. Similarly with ammonia. Since three volumes of hydrogen were needed to react with all the nitrogen in one volume of nitrogen, we would think that ammonia was made of one molecule of nitrogen and three of hydrogen. But the reaction of three volumes of hydrogen and one of nitrogen produced two volumes of ammonia, not one as one would expect (for one would expect that there be only as many ammonia molecules as nitrogen molecules). (Nolan, [1997], p. 336)

In order to explain the experimental result, Avogadro drew a distinction between atoms and molecules.[3] The explanation then offered by Avogadro was that each molecule of oxygen and hydrogen gas is made up of two atoms of oxygen and hydrogen, respectively; that each molecule of steam is made up of two hydrogen atoms and one oxygen atom.[4]

---

[2] In Avogadro's usage 'molecules' can be either atoms or molecules (in the modern usage of the terms). In the original French it reads '[…] l'eau résulte de l'union de chaque molécule d'oxigène avec deux molécules d'hydrogène' (Avogadro, [1811], p. 60).

[3] The terms 'atom' and 'molecule' are anachronistic but correspond roughly to the distinction drawn between *molecule élémentaire* and *molécule intégrante* (and *constituante*) (see for example Avogadro, [1811], p. 60). We adopt the modern vernacular in what follows for ease of presentation to the contemporary reader.

[4] Similarly for the ammonia case.

As Nolan notes, this hypothesis was not the only one open to Avogadro. It would be perfectly consistent to adopt a view whereby the numbers of atoms involved are *multiples* of the minimum number required, that Avogadro had identified. Thus, Avogadro could have supposed that each molecule of oxygen and hydrogen is made up of four atoms of oxygen and hydrogen (respectively), and that each molecule of water is made up of four hydrogen atoms and two oxygen atoms.

While Avogadro considered the possibility of there being gases with different compositions (such as four atoms comprising one molecule of the gas, etc.) he did not pay serious attention to such an option in the case of the production of steam (or the production of ammonia, nitrous oxide, and nitric oxide).[5]

> […] [W]e suppose, namely, that the constituent molecules of any simple gas whatever […] are not formed of a solitary elementary molecule, but are made up of a certain number of these molecules […] and further, that when molecules of another substance unite with the former to form a compound molecule, the integral molecule which should result splits up into two or more parts (or integral molecules) composed of half, quarter, &c., the number of elementary molecules going to form the constituent molecule of the first substance, combined with half, quarter, &c., the number of constituent molecules of the second substance that ought to enter into combination with one constituent molecule of the first substance […] ; so that the number of integral molecules of the compound becomes double, quadruple, &c., what it would have been if there had been no splitting-up, and exactly what is necessary to satisfy the volume of the resulting gas.
> On reviewing the various compound gases most generally known, I only find examples of duplication of the volume relatively to the volume of that one of the constituents which combines with one or more volumes of the other. We have already seen this for water. […] Thus in all these cases there must be a division of the molecule into two; but it is possible that in other cases the division might be into four, eight, &c. (Avogadro (English translation), [1923], p. 32)[6]

---

[5] Partington ([1964], pp. 213-7) contains a discussion of the cases that Avogadro discussed and goes on to address his later applications of the same reasoning.

[6]Here is the original text in French.

> […] c'est de supposer que les molécules constituantes d'un gaz simple quelconque […] ne sont pas formées d'une seule molécule élémentaire, mais résultent d'un certain nombre de ces molécules […], et que lorsque des molécules d'une autre substance doivent se joindre à celles-là pour former des molécules composées, la molécule intégrante qui devroit en résulter se partage en deux ou plusieurs parties ou molécules intégrantes composées de la moitié, du quart, etc. du nombre de molécules élémentaires dont étoit formée la molécule constituante de la première substance, combinée avec la moitié, le quart, etc. du nombre des molécules constituantes de l'autre substance, qui devroit se combiner avec la molécule totale […]; ensorte que le nombre des molécules intégrantes du composé devienne double, quadruple, etc., de ce qu'il devroit être sans ce partage, et tel qu'il le faut pour satisfaire au volume du gaz qui en résulte. En parcourant les différens composés gazeux plus connus, je ne trouve que des exemples de redoublement de volume relativement au volume de celui des composans, qui s'adjoint une ou plusieurs fois son volume de l'autre : on l'a déjà vu pour l'eau. […] Ainsi, dans tous les cas il doit y avoir partage des molécules en deux; mais il est

Here he claims that for all known cases where the volume of the produced gas doubles from what we would expect, the molecules involved in the production have to be composed of two atoms.[7] While he allows that for other, unknown gases, it might be the case that molecules of those gases are composed of four, or eight, etc. atoms, the quantitatively parsimonious alternative is being treated as a default hypothesis when the volume of the produced gas is double of the expected one.

Finally, we need to introduce a third case, also less discussed than the neutrino case. This is the case to which we turn now, namely, the postulation of Neptune. The failure of Newtonian physics together with the known facts about the solar system to accurately capture the motion of Uranus led to the postulation of a planet beyond the orbit of Uranus. The claims of priority and merit of the work of Adams and Le Verrier—that led to the identification of Neptune by Galle in 1846—have been much discussed.[8]

For our purposes, the role that quantitative parsimony played in the postulation of *one* trans-Uranian planet is of particular interest. Gould ([1850], pp. 29-30) describes how Le Verrier rules out several hypotheses based on their incompatibility with the data known at the time. For example, Le Verrier discarded the hypothesis that a comet could have caused the disturbances in the motion. He also ruled out an intra-Uranian planet. No such planet could have accounted for the disturbances without also disturbing the motion of Saturn to a discernible degree. The postulated planet would therefore have to be a trans-Uranian one, but yet could not be too remote. If its orbit was at too great a distance from Uranus, then its mass would have to be so large that it would again be expected to produce detectable disturbances in the motion of Saturn.

---

possible que dans d'autre cas le partage se fasse en quatre, en huit, etc. (Avogadro, [1811], pp. 60-1)

[7] What Avogadro explicitly asserts in a footnote is that '[…] the integral molecule of water will be composed of a half-molecule of oxygen with one molecule, or what is the same thing, two half-molecules of hydrogen' (Avogadro (English translation), [1923], p. 32). However, if we do not take this together with the claim that the division of the molecule is '[…] exactly what is necessary to satisfy the volume of the resulting gas' to imply that water is made up of two hydrogen atoms and one oxygen atom then we raise trouble for the overall understanding of Avogadro's paper. To see why, let us assume for the moment that we should not take the half-molecules discussed to be hydrogen and oxygen atoms, but rather just half of the number of atoms in a hydrogen or oxygen molecule. The project is to determine the relative masses of atoms (elementary molecules) and the relative proportions in which they enter into compounds. If it was enough to satisfy the 2:1 ratio of hydrogen to oxygen in water that water consisted of one hydrogen molecule and one half-molecule of oxygen then H_6 combining with O_2 to form H_6O would satisfy the requirement. However, now we could not determine the relative masses of hydrogen and oxygen atoms in the way suggested for determining the ratio of masses in the first section. To rule this out hydrogen molecules and oxygen molecules have to have the same number of atoms. Since Avogadro does not mention this constraint we take it to be a reasonable reading that he does not feel the need to postulate this constraint. This makes sense if he is assuming that the two half-molecules of hydrogen that make up water are two elementary molecules (atoms) of one molecule of hydrogen and that the half-molecule of oxygen that makes up water is one of two elementary molecules (atoms) of one oxygen molecule. We will only consider alternative hypotheses that keep the number of atoms in hydrogen and oxygen molecules the same.

[8] See for example Gould ([1850]), Grant ([1852], chapter XII, appendix III), and Hanson ([1962]).

It was known at the time that more than one body *could* provide a suitable explanation. Hanson ([1962], p. 361) reports that Hansen speculated about aberrations in the motion of Uranus being due to more than one body as early as 1829, in a letter to Bouvard. Although Hansen denies ever having claimed that the observations *required* the existence of more than one unknown planet, he allows that more than one body *could* be involved in producing the disturbances.[9] Thus, throughout the discussion of these scenarios it is assumed that, in spite of the fact that more than one unknown body *could* be posited, we are looking for *one* body to account for the disturbances of the motion of Uranus. When Le Verrier tries to account for the anomalies in the motion of Mercury in the same way that he tackled those of Uranus he explicitly considers postulating not one body but rather several asteroids.[10] However, this scenario is considered only to solve the difficulties of the failed observation of a single body. Here the quantitatively parsimonious hypothesis is being treated as the default.

2. Three Desiderata

We have the cases. Now we need some desiderata. In this section we note three desiderata that applications of a principle of quantitative parsimony ought to satisfy. The first desideratum concerns the scope of the set of situations in which the principle may be applied. The challenge turns on the worry that, as typically formulated, a principle of quantitative parsimony may apply to *too many* cases. We must find a way of limiting the application of the principle. The second desideratum concerns whether or not the principle is robust enough to withstand an attack from those who would worry about the overall size of our ontology. The third and final desideratum asks that any justification for preferring quantitatively parsimonious hypotheses should cover all three of the cases, described above. We illustrate each point in turn and treat the satisfaction of these desiderata as requirements on any successful attempt to justify appeals to quantitative parsimony. We concede that the satisfaction of these desiderata by a single principle is only *prima facie* desirable and that there may ultimately be reasons to be given as to why more than one principle must be involved. Nonetheless, in the absence of said reasons being provided, we assume, here, that it would be best if a single principle could be provided that does satisfy these desiderata.

2.1 Limiting
First, many statements of the principle of quantitative parsimony are formulated in such a way that they entail that we should, ceteris paribus, try to reduce the *overall size* of our

---

[9] As printed in Gould ([1850], p. 12), Hansen's letter reads 'Ich kann möglicher Weise geschrieben haben, dass *vielleicht* die bis dahin in der Bewegung des Uranus nicht erklärten Abweichungen von der Theorie nicht von einem, sondern von mehreren auf ihn einwirkenden, unbekannten Planeten herrührten […]'.
[10] 'Mais se pourrait-il qu'un tel astre existât sans avoir jamais été aperçu? Assurément il serait doué d'un très-vif éclat : doit-on croire qu'en raison de sa faible élongation il se fût toujours perdu dans la lumière diffuse du Soleil? Comment admettre qu'on n'eût point été frappé de sa vive lumière durant quelqu'une des éclipses totales de Soleil ? D'où vient qu'on ne l'ait jamais découvert passant sur le disque de cet astre? Toutes les difficultés disparaîtraient en admettant, au lieu d'une seule planète, l'existence d'une série de corpuscules circulant entre Mercure et le Soleil.' (Le Verrier, [1859], p. 382)

ontology.[11] Indeed, this was how *we* glossed the principle in the previous section when we described quantitative parsimony as the practice of 'minimizing the number of entities postulated'. Some considerations of parsimony seem to fit this injunction. For example, in the Uranus case, the preference for postulating one unobserved celestial body, rather than several celestial bodies, seems to minimize the (total) number of entities postulated.

But while it also seems intuitively plausible to prefer the hypothesis that minimizes the number of neutrinos postulated *in order to account for the missing spin* (and energy and momentum), it is not at all clear that discussion of this case supports a *general* principle that requires us to try to minimize the number of entities in our overall ontology.

Imagine (contrary to fact) that the scientific community agreed that the universe will end in a big crunch. An injunction to prefer a theory that postulates fewer entities in the universe would now require us to prefer a theory that states that the universe will end sooner rather than later, since such a theory would (assuming that the theories are otherwise similar and that the rate of beta decay occurrences in them is the same) postulate a lower total number of neutrinos. We call this the 'Early-Big-Crunch' hypothesis. We take it that, at least intuitively, Early-Big-Crunch is *not* a theory that considerations of parsimony should favour on the grounds that it reduces the number of neutrinos postulated to exist.[12, 13]

Here is Wallace ([2012], p. 105) offering the same conclusion from a different evidential base.

> Generally in physics, we try to keep our number of postulates as low as possible. But we're not usually that bothered about *how much* there is in the Universe of any given entity we postulate. For instance, we don't tend to assume that cosmological theories are a priori more or less likely to be true according to how many galaxies they postulate.

Thus, we submit, considerations of scope should allow us to apply considerations of quantitative parsimony only to particular cases. Call this concern the 'LIMITING' concern.

2.2 Robustness

---

[11] See for example Baker ([2013]) and Huemer ([2009], p. 216).

[12] Further, there is no evidence that we can find of Early-Big-Crunch-like hypotheses being preferred within scientific debate. We will return to the question of why and how this might be justified later on in the paper in section 2.4.

[13] A referee reported a different intuition here, that it is equally plausible to describe this as a case where parsimony considerations favor Early-Big-Crunch, but these are outweighed by other theoretical considerations. We are open to alternative explanations, of course, but would want to see the details of such an explanation. Even upon review, we were of the view that this is a case where quantitative parsimony shouldn't get a grip given that all else is equal.

It seems (to us, and others) that we should be able to apply considerations of parsimony, independently of the overall size of our ontology. As Nolan ([1997], p. 340) notes, the preference for quantitatively parsimonious theories does not seem to hinge on the size of the ontology that we are already committed to. Whether or not we are mathematical Platonists seems to have little bearing on how we treat the beta decay case. Yet, if we were concerned only with the overall number of entities postulated, we would expect the overall number of entities in play to be highly relevant. Thus, we think, the justification for preferring quantitative parsimony must remain robust, independently of the overall number of objects that populate our ontology. We call this concern the 'ROBUSTNESS' concern.

Earlier discussion of (what we call) ROBUSTNESS is flawed. Nolan suggests that the concern here should be, not just with how many entities there are, but how many entities of a given kind that there are (Nolan, [1997], p. 340). We think this a step in the right direction, but this modification is not strong enough. Consider the neutrino case again. Even if we make the assumption that the universe is infinite in extent and that there are infinitely many beta decays *in total,* it still seems to us that considerations of parsimony militate in favour of preferring H_1 over $H_{n>1}$, despite the fact that H_1 would not, in this context, lower the number of beta decays in the universe, when contrasted against $H_{n>1}$. We do not see that an infinite universe should undermine our preference for parsimony when dealing with an explanation of beta decay. Thus, we think, Nolan's defence against (what we call) the ROBUSTNESS concern isn't successful.

2.3 Breadth

Baker ([2003]), when attempting to justify a preference for quantitatively parsimonious hypotheses, restricts his attention to cases that he calls 'additive' (meaning, in his terms, that they involve the postulation of qualitatively identical objects to collectively explain some phenomenon by simple summing of their contributions). Baker then relies on showing that, in cases such as that of H_1-H_n, the less quantitatively parsimonious hypotheses run up against what we describe as a dilemma.

Horn 1: the less quantitatively parsimonious hypotheses provide a worse basis for an answer to a question that they themselves make it necessary to answer and that the parsimonious hypothesis provides the basis for a ready explanation of. In the neutrino case, the question that the less parsimonious hypotheses makes it necessary that we answer is why we have not observed spin in fractions other than ½. The more parsimonious hypothesis (H_1) provides the basis for a ready explanation: there are no particles with such a spin; the more complex hypothesis does not, on its own, provide the basis for such an answer.

Horn 2: the less quantitatively parsimonious hypotheses meet the explanatory challenge just described only by postulating a new law. For instance, to the more complex hypothesis we could add an *additional* law dictating for example that neutrinos are only

ever emitted in pairs and so on, which now makes the non-parsimonious alternatives less syntactically simple by increasing the number of postulates.[14]

Baker's ([2003]) diagnosis fits his description of the beta decay case well (as he shows). But, as it stands, it does not extend to other, similar cases such as Avogadro's hypothesis case; for that reason, we do not think it sufficiently general. Whether we account for the production of 2 volumes of water resulting from the combination of 2 volumes of hydrogen and 1 volume of oxygen by H_2 and O_2 creating H_2O or by H_4 and O_4 creating H_4O_2 *seems* to involve considerations similar to those raised by the beta decay case.[15] But, as we shall demonstrate, appearances are deceptive. We begin by working through the appearances.

2.3.1 A limited success for Baker

There are, as we have noted, a number of hypotheses that Avogadro could have formed:

AH_1: Water results from the combination of H_2 and O_2

AH_2: Water results from the combination of H_4 and O_4

AH_n: Water results from the combination of H_{2n} and O_{2n}[16]

To illustrate the appearance of similarity with the neutrino case, we must now demonstrate that either: AH_2 provides a worse basis for an answer to a question that it makes salient, and that AH_1 provides a ready basis for; or that AH_2 meets this explanatory challenge only by postulating a new law, etc., which makes a theory with AH_2 less syntactically simple than one with AH_1, by increasing the number of postulates.

If we postulate AH_2, we raise the question: what prevented more than 2 volumes of water being produced? Given the weak background assumption that water contains twice as much hydrogen as oxygen, the presence of 2 volumes of H_4 and 1 volume of O_4 gives one the raw material to make, at most, 4 volumes of water. In contrast, given AH_1, there *could not* be more than 2 volumes of water created since there was only 1 volume of oxygen where each oxygen molecule contained only enough oxygen atoms for 2 molecules of water.[17]

---

[14] Baker ([2013]) sometimes restricts the notion of syntactic simplicity to foundational postulates. We will not follow that practice here for reasons that will become clear in section 3.

[15] To have 2 volumes of H and 1 volume of O combining to create H_2O is ruled out (given the background assumptions) since this would only create 1 volume of H_2O (contrary to the observation of 2 volumes of water being produced).

[16] In what follows we focus upon AH_2.

[17] Of course, we could ask, compatible with our background assumptions, why it is that 2 volumes of H_2 and 1 volume of O_2 did not create 1 volume of H_4O_2 instead of 2 volumes of H_2O. However, we can similarly ask why it is that 2 volumes of H_4 and 1 volume of O_4 did not create 1 volume of H_8O_4 instead of 2 volumes of H_4O_2. So, here the two hypotheses are on a par.

We can easily concede that *some* explanation that incorporates AH_2 could be given. But we think it's clear that such an explanation, in the form of a law or additional hypothesis, *would have* to be given—it has not, as yet. If we consider adding a law then, following Baker, the resulting total explanatory hypothesis would be less syntactically simple than AH_1. Syntactic simplicity in the sense that Baker is using it here has to do with elegance and is viewed as the number and complexity of hypotheses required. It is fairly intuitive that adding additional laws would lower the theoretical elegance of the resulting theory. To support this judgement we could compare the laws needed with AH_1 as the basis for the explanation to the laws needed with AH_2 as the explanatory basis. Here the only difference is the extra law needed when AH_2 rather than AH_1 is the base hypothesis. It is less clear-cut how to judge comparative elegance when we are dealing with an 'additional' hypothesis. The reason for this is that we are adding a hypothesis to AH_2 rather than to AH_1. We no longer have a shared category (like the laws) that we are merely adding complexity to. Nonetheless, we take Baker's point—that the alternative package of hypotheses offers a less elegant explanation—to intuitively be right.[18]

2.3.2 Rejecting Baker's analysis

So far, this story seems like it fits Baker's diagnosis. Ultimately, however, the diagnosis does not go far enough. Baker himself explicitly restricts the application of his diagnosis to cases where the entities involved are, as he puts it, qualitatively identical and their contribution to the phenomenon under consideration is additive. To illustrate: for each qualitatively identical neutrino we can say what the spin of it is and the total effect to be explained is then obtained by summing a number of such contributions.[19]

There is a good reason for this restriction. If we did not include the restriction to additive cases, then Baker's account should be expected to apply more widely. But, as Baker himself notes, it is not clear that it functions in the Neptune case. In brief, the worry for Baker's account in the Neptune case is that postulating two bodies, rather than one, does not raise any unanswered questions where the more parsimonious hypothesis provides a good basis for an answer and the less parsimonious hypothesis fails to do so. In the neutrino case, our background theory and assumptions together with H_2 do not have the resources to suggest an answer the question of why we have not observed ¼ values of spin, but the same background theory and assumptions together with H_1 do. In the Neptune case, the very same background theories (of Newtonian mechanics and gravitation) allow us to explain the deviation of Uranus in both cases and there is no question raised by the two bodies hypothesis that the background theories and assumptions do not have the resources to suggest an answer to.[20]

---

[18] We take it to be a virtue of our view that we will later vindicate Baker's overall verdict without directly relying on a judgment of comparative elegance.

[19] For this to make sense it matters that we are trying to explain only the missing spin-1/2 that is needed in order to make it possible for the decay to conserve angular momentum.

[20] This is somewhat of a simplification. There are difficulties of application and computation. However, such considerations are better suited to giving a pragmatic defence of the simplicity considerations at hand and they do not fit with Baker's overall project.

The only non-ad-hoc explanation of why to not try to apply Baker's reasoning to the Neptune case, that Baker gives, is that the Neptune case is not additive.[21] So, without the restriction to additive cases, Baker's justification for preferring quantitatively parsimonious theories *should* be expected to extend to the Neptune case. But, as explained, it does not apply. And that means that Baker *requires* the restriction to *additive cases only*.

But, we think, we cannot keep the restriction as strictly stated and deal with the case involving Avogadro's hypothesis. In the case of Avogadro's hypothesis, there is no *type* of entity that contributes to the explanandum in this additive way. Rather, we have two different kinds of entities (the different hydrogen and oxygen molecules) that jointly contribute to the effect. For instance, in AH_1 we have H_2 and O_2 combining. H_2 and O_2 are not qualitatively identical. (In contrast, each of the neutrinos posited in any of H_1 and H_n *are* qualitatively identical with every other neutrino posited by that hypothesis).

Moreover, the effect is not achieved in the process *simply* by adding the contributions of the qualitatively identical entities. It's true that in the neutrino cases, to account for the explanandum, we need simply add together the spins of the various neutrinos posited by the hypothesis under consideration. But, in the Avogadro case, we must make substantive assumptions about how that production worked in order to account for the production of 2 instead of 1 or 4 volumes of water. As a consequence, we cannot keep the restriction to additive cases only *and* give a satisfactory account of the Avogadro case.

So, we concede to Baker that he intends to restrict the cases to which his justification of a preference for quantitatively parsimonious theories should apply. But, if we keep that restriction, then (contra Baker) his justification for preferring quantitatively parsimonious theories will not extend to the Avogadro case. We think this surprising and limiting. Since intuitive judgements along the lines of Baker's account seem to apply more generally, we would think it best—all things considered—if our justification of appeals to quantitative parsimony could be applied more generally.

Thus, we suggest: any justification of appeals to quantitative parsimony should apply to more than just one case—ideally, it should apply to all of the cases discussed here. We call this criterion 'BREADTH'.

2.4 The proposal

Our suggestion is that (roughly) we should relativize principles of parsimony to directly competing explanations of the same explanandum. This allows us to address the challenges above and to show what the various cases mentioned have in common.

---

[21] See Baker ([2003], pp. 257-8).

The idea of directly competing explanations requires some clarification. But in the hopes of illustrating what is meant, we think that remarks made above are worth repeating: it seems intuitively plausible to prefer the hypothesis that minimizes the number of neutrinos postulated *in order to account for the missing spin* (and energy and momentum); it is not at all clear that discussion of these cases supports a *general* principle that requires us to try to minimize the number of entities in our overall ontology. We think that this insight is important. Relative to the explanation of some phenomena, we should try to minimize the number of entities posited.

What we have in mind in general, here, are explanations that share the same broad theoretical framework, but that postulate different specific hypotheses to account for some particular explanandum. In the neutrino case we assumed a shared theoretical framework of conservation of energy, momentum, angular momentum, etc. The hypotheses involved are therefore in direct competition with each other in a way that they would not be if we were also allowing theoretical and background assumptions to vary.

Our guiding principle is QP:

> QP: First, assume a framework of theoretical and background knowledge. Second, locate directly competing hypotheses, compatible with that framework, that allow for the explanation of some explanandum. Third, prefer, ceteris paribus, the hypothesis that minimizes the number of entities that the hypothesis involves in the explanation.

In the next sections, we will clarify the ceteris paribus qualification and make the very rough QP more precise.

Notice, that by relativizing the principle to explanations of some given explanandum, we will satisfy LIMITING. Our concern, with LIMITING, was that the principle of parsimony ought to apply only to specific cases, and not lead us to favour hypotheses like Early-Big-Crunch. With QP we are minimizing our ontological commitments *only* relative to a specific explanandum and so we are forced to consider the minimization with respect to a specific case of some sort. This explains what it is that is so unintuitive about favouring Early-Big-Crunch. In order to favour Early-Big-Crunch under QP, it would have to be the case that the number of beta decays is somehow crucially involved in a hypothesis about the end of the universe, and that this hypothesis, complete with beta decays, was explaining some explanandum. The fact that, given our background knowledge, it is extremely *implausible* that beta decay is implicated in such an explanation makes it easy to see why it is unreasonable to favour the Early-Big-Crunch on such grounds.

Our proposal also seems to satisfy ROBUSTNESS. QP ignores the question of the *total number* of entities in existence and asks us to compare *specific explanatory hypotheses*, and the number of entities that they posit. In the beta decay case, for instance, we can agree that

there are infinitely many instances of beta decay quite generally and still have grounds for preferring H_1 to competitors: the phenomena in need of explanation (the apparent missing spin 1/2) is explained by a scientific explanatory hypothesis that posits only *one* neutrino be emitted in each particular instance of beta decay. It remains for us to show that a version of QP can satisfy BREADTH. We must also explain and motivate QP.

3. Probabilistically Additive Hypotheses and a (sort of) Bayesian Account: The Limpid Rationale Relativized and Reconsidered

Our key claim is that it is not simply that the less parsimonious hypotheses raise questions that are harder to answer with these hypotheses than with their more parsimonious alternatives; rather, the key claim is that in order to account equally well for the data, the less parsimonious alternatives will, in these cases, turn out to have a lower prior probability than the parsimonious alternative (given the shared background knowledge and theories). In the cases above we are presented with hypotheses that *seem* to all account for the data equally well. We will show that this is not the case. By modifying the scenarios so that the competing hypotheses do account equally well for the data (given the background assumptions) in the minimal sense of having the same likelihood the less parsimonious alternative ends up with a lower prior. By analysing the cases in probabilistic terms in this way, we can show that all three cases discussed in this paper can fit under the same analysis.[22]

To make the comparison precise, we will make use of the idea of probabilistically additive *hypotheses*. This is an extension of what Sober ([1981], p. 145) refers to as Quine's "limpid" account of parsimony. That is, in general, removing existence claims increases the probability of a hypothesis, since a conjunction cannot be more probable than its conjuncts.

In the cases under consideration, we will treat H_1 as a hypothesis that is relatively quantitatively parsimonious and entails the relevant evidence E. For reasons that will become clear, we will treat H_2 as a hypothesis that is less parsimonious than H_1 *but that does not* entail E. We will assume that H_1 is at least on a par with H_2 when it comes to any part of our total evidence that is not part of E. That is, we take H_2 to at most have the same prior as H_1 (relative to our background knowledge). We will treat H_3 as a hypothesis that is equivalent to the conjunction of H_2 with some other hypothesis (H_4, H_5, etc.) that, collectively, entail the evidence.

To generate our extension we will, thus, not focus on the postulation of existence claims. Rather, we will focus upon the relationship between hypotheses that obtains when the prior probably of hypothesis H_1 is not lower than the prior probability of H_2. We know

---

[22] Moreover, this analysis also lends itself to an easy extension to the cases of parsimony that Sober ([1994]) considers. The cases discussed here just form a natural group with shared properties that we can use to motivate applying a principle of quantitative parsimony to them.

that the probability of a hypothesis H_3—obtained by taking the conjunction of H_4 and H_2—will typically be lower than that of H_1 (assuming that H_4 is not trivial). A conjunction cannot be more probable than its conjuncts. We will exploit this fact.

To extend the terminology introduced by Baker ([2003]), we will call hypotheses related as H_1 and H_3 *probabilistically additive* hypotheses. The terminology is apt since H_3 is composed of one hypothesis not more probable than H_1, *together with* an extra, added, hypothesis (H_4 or H_5, etc.).

The main claim of this section is that all three cases discussed above are cases where the hypotheses (or alternatives based on the hypotheses) can be understood as being additive in *this* sense and that, as a consequence, the simpler hypothesis in each case (H_1) is to be preferred.

In our rough statement of QP we captured sensitivity-to-evidence in explanatory terms and by restricting the case to directly competing hypotheses. We need this assumption since we do not assume that we can make judgements about the priors involved, or about the extra hypotheses needed, in the absence of such considerations. Moreover, it is only when the theory and background assumptions *do* give us reasons to think that the hypotheses are related in this probabilistically additive way that the defence we give here will have force. This makes our defence a limited one. But Sober is likely right that we should not expect a completely general defence of any principle of parsimony. We would certainly be surprised if it turned out that hypotheses that included more entities were always considered as having a lower prior than those with fewer entities.

This leaves us with a wrinkle that requires ironing out. Above, we suggested that it was a defect of Baker's view that he is not able to account for parsimony being a virtue in the cases involving Avogadro's reasoning and the postulation of Neptune. If appeals to quantitative parsimony are only *ever* justified in particular cases, then what is the benefit of our proposal being able to justify appeals to parsimony in more cases? Here is where we part ways from Sober.

While we cannot give a justification for parsimony without taking into account background conditions involved in particular cases, this does not prevent a general account of when considerations of quantitative parsimony have force. A general account can be given as long as the role that the background conditions play in these cases is relevantly similar. Thus, we think, what we provide here is a defence of appeals to quantitative parsimony in all cases where, when background conditions are taken into consideration, we are considering directly competing hypotheses related as H_1 and H_2. This, we think, is a significant advance on Baker's account which seems—at best—to function only in a relatively small number of cases and that leaves the justification of the principle in those cases directly hostage to other principles of simplicity such as elegance.

3.1.   Neutrinos and beta decay

Let us start by showing how this idea plays out in the case of the postulation of neutrinos to account for the missing spin in the case of beta decay. Take H_1 to be the hypothesis that there is one spin-½ particle emitted in beta decay and H_2 to be the hypothesis that there are two spin-¼ particles emitted. As Baker ([2003]) points out, if there were spin-¼ particles emitted, given our background knowledge, we would expect to see them produced singly in some interactions (barring some reason for restricting interactions to the production of two of them). Moreover, given our background assumptions, we expect two spin-¼ particles to generally be detectable individually and not merely in pairs.[23]

Though this is largely following Baker ([2003]) there is a subtlety in the way that we have described the case that will be of importance. Baker considers the additional explanation to be simply why there is no observation of spin-¼ particles *in general* and notes that H_1 does not explain this on its own either, but is rather compatible with an easy (or easier) explanation for this phenomenon than H_2.

In our description of the case the claim is directly concerned with the neutrinos that are supposed to exist. The question is *not* one of explaining why, in general, we have not seen spin-¼ particles of any kind. It is rather why we have not observed any of the neutrinos emitted in the case of beta decay display spin ¼ in interactions individually. On H_2, given that H_2 postulates the existence of just such particles, this is puzzling. This is not similarly puzzling with the spin-½ particles in H_1. Even though H_1 does not entail that there could be *no* spin-¼ particles of any kind, H_1 does entail that the particles emitted in beta decay are not spin-¼ particles. Hence, it is not puzzling that we have not seen the particles *in beta decay* interact individually to display spin-¼.

Our background theory and knowledge leads us to expect that, if in beta decay there were fractions of spin as postulated by H_2, then it should be possible to observe them singly. We need to add an additional hypothesis in order for H_2 to account for our evidence (including the absence of such an observation). Here we have a few different hypotheses to choose from, but two of the most readily available ones are as below.[24]

> H_4^*: There are interactions that would make it possible to observe the spin-¼ particles, but we have not yet performed the experiments to allow us to do so.

> H_5^*: There are no interactions that would make it possible to observe the spin-¼ particles, since a law forbids them from being emitted or interacting other than as pairs.

---

[23] Although, of course, they do not need to be easy to detect.
[24] This list is not exhaustive.

So, let us be explicit and add to the observations of missing spin ½ in beta decay that constitute our body of evidence, the fact that there have been no observations of spin ¼ *for the particles involved in beta decay*. Let E stand for this enlarged body of evidence, K for our background knowledge, and T for our background theories.

Notice that H_2 (together with the background theory and knowledge) does not entail E (although H_2 together with H_4^* or H_5^* does). Let us assume for now that our background knowledge and theory does not favour there being two new particles emitted over there being one new particle emitted, so that $Pr(H_1 | K \& T)$ is not lower than $Pr(H_2 | K \& T)$. For simplicity, say that $Pr(H_1 | K \& T) = Pr(H_2 | K \& T)$.[25] Given this, we will find that in a direct comparison the evidence favours H_1 over H_2. After all, H_1 together with the background knowledge and theory entails the evidence, E (where, remember, E includes the failure to observe ¼ values of spin for the neutrinos in beta decay singly), but H_2 does not.[26] By Bayes' theorem we have that $Pr(H_1 | E \& K \& T) = Pr(H_1 | K \& T) Pr(E | H_1 \& K \& T)/Pr(E | K \& T)$ and that $Pr(H_2 | E \& K \& T) = Pr(H_2 | K \& T) Pr(E | H_2 \& K \& T)/Pr(E | K \& T)$. Since by stipulation the only term that differs between the two cases is the likelihood and since this is lower for H_2 than for H_1, the evidence will favour H_1 over H_2.[27] This draws on similar reasoning to Huemer's ([2009]) support of a likelihood defence of parsimony. However, such a defence falters when the likelihood is the same and, as is shown below, such a defence captures only part of what makes considerations of quantitative parsimony seem reasonable in the scenarios we discuss.[28]

We noted above that H_2 together with H_4^* or H_5^* (and, as always, T and K) will entail the evidence. Let H_3^* be $H_2 \& H_4^*$ and H_3^{**} be $H_2 \& H_5^*$. In terms of likelihoods, H_3^* and H_3^{**} will be on a par with H_1. A likelihood defence of parsimony will not yet tell us why we should prefer H_1 over H_3^* and H_3^{**} and, thus, prefer H_1 over H_2. As before, we will assume that $Pr(H_1 | K \& T)$ is not lower than $Pr(H_2 | K \& T)$. Given this we have a case where H_1 and H_3^* as well as H_1 and H_3^{**} are related as probabilistically additive hypotheses. Now the prior of H_3^* and H_3^{**} will be lower than that of H_1 (since the

[25] The reasoning will hold a fortiori if $Pr(H_1 | K \& T) > Pr(H_2 | K \& T)$.

[26] We are assuming here that our background knowledge and theory precludes skeptical scenarios like being systematically mistaken about our observations, measurement instruments, etc.

[27] Notice that by holding fixed background knowledge and theory across H_1 and H_2 we are going some way to ruling out the use of our proposal in cases other than tie-breaking cases where all else is equal. This, we think, helps us preserve the claim we made in QP that we do not take our proposal to extend beyond ceteris paribus cases.

[28] We also flag, here, that this seems to be a typical case of where were use parsimony considerations, rather than in cases where evidence tells directly against a particular hypothesis, as it seems to tell against H_2 when H_2 is considered in isolation. A Bayesian defence purely in terms of likelihood principles runs up against the additional challenge that when we apply parsimony considerations we are typically dealing with explanations of known evidence rather than predictions. This brings with it familiar difficulties of how to treat old evidence. Our account in terms competing explanations cashed out in terms of probabilistically additive hypotheses goes some way to alleviate this challenge. We are now dealing with competing explanations of E. This motivates the demand that the competing explanations need to be on equal footing at least when it comes to the entailment of E given the background knowledge and theory. This provides a non-ad hoc reason for not treating E as part of the background knowledge.

probability of a conjunction is generally lower than probability of either conjunct).[29] This gives us reason to favour H_1 over H_3^* and H_3^{**} and so again to favour H_1 over H_2.

This reasoning is particularly nice since it shows that, for example, the stronger reason we have to think that H_5^* holds, the weaker this preference is. That is, if we had some particular strong and independent reason for thinking that there are no interactions that would make it possible to observe the spin-¼ particles singly, since a law forbids them from being emitted or interacting other than as pairs (H_5^*), then that would leave H_3^{**} only very slightly less probable than H_1. This, we think, is the right result. It is not a result that we can see how to recover given Baker's account.

Finally, notice that all of the reasoning above takes place given K and T. It only holds given our background knowledge and theory. Our background theory and knowledge allows us to expand the evidence to include what we would expect to have seen and to compare the priors of H_1 and H_2. Moreover, the reasoning is defeasible (as it should be). It seems plausible that our theory and background knowledge did not favour a many particle hypothesis over a single particle hypothesis, but it could have done so.[30] If this had been the case then the same reasoning that we have just given here shows that it has to be the case that the increased support for H_2 outweighs either the lack of entailment of the evidence or the decrease in probability incurred by moving to H_3^* or H_3^{**}.

3.2.   Avogadro's hypothesis

The reasoning in section 3.1 applies also to the case of Avogadro's hypothesis. Here both the hypothesis (AH_1) that 2 volumes of H_2 and 1 volume of O_2 create water with a 2-1 ratio of hydrogen to oxygen and the hypothesis (AH_2) that 2 volumes of H_4 and 1 volume of O_4 create water with a 2-1 ratio of hydrogen to oxygen are compatible with the evidence of 2 volumes of water (with a presumed 2-1 ratio of hydrogen to oxygen) being created. Given the background assumptions, theory, and AH_1, we find that the possible atomic compositions of water that would respect a 2-1 ratio are H_2O and H_4O_2. We would plausibly view these options as equally probable given the background knowledge and theory. So, the production of 2 volumes of water and 1 volume of water are equiprobable, given the background theory, assumptions, and hypothesis AH_1. However, on hypothesis AH_2 (and the same background assumptions and theory) there are 3 options for the production of water with a 2-1 ratio. We could get 4 volumes of H_2O, 2 volumes of H_4O_2, or 1 volume of H_8O_4. Given the background assumptions and theory these are also plausibly equiprobable.[31] The observation gives us that 2 volumes

[29] When it is clear from the context we will drop the reference to the background theory and knowledge.

[30] Baker ([2003], p. 250) discusses such a case. His focus is, however, on whether inductive reasoning can explain the entire preference for the parsimonious hypothesis. We agree that this is not straightforwardly the case.

[31] Here it is again important that our defence only holds under the assumption of shared background knowledge and theory. The claim is *only* that this reasoning is plausible given that background knowledge and theory. We do not rely on the claim that a principle of indifference is *generally* defensible. We do not offer an account of how

of steam was produced, so the likelihood term for AH_2 is lower than that for AH_1. Even if we regard them as having equal *prior* probability (given the background assumptions and theory), AH_1 should be preferred to AH_2. As in the earlier case, even though AH_1 and AH_2 are both *compatible* with the evidence they are not, as presented, on a par when it comes to the likelihood of the evidence when we take into account background theory and assumptions. However, we can amend the case to make this so.

Let us now move to such a case where we ensure the entailment of the evidence. So, let us now say that AH_1^* is AH_1 conjoined with the principle that the volume is not minimized in the relevant interactions.[32] Now AH_1^* (together with the theory and the background assumptions) entails that 2 volumes of H_2O will be produced. However, when we try to do the same for hypothesis AH_2 we find that we have to add yet another hypothesis. Simply ruling out minimization of volume (taking us to AH_2^*) is not enough; we also have to rule out maximization of volume (let us call this new hypothesis AH_3^*).[33] Under the assumption that the prior of AH_1 is not lower than that of AH_2 we get that the prior probability of AH_1^* is not lower than that of AH_2^*. We now know that AH_3^* and AH_1^* are related as probabilistically additive hypotheses, so the prior of AH_3^* will be lower than that of AH_1^*. Again we have reason to prefer AH_1 over AH_2.

Now we have a way of justifying the application of a principle of quantitative parsimony in this case. By taking into account the background theory and assumptions we have argued that a likelihood defence can favour AH_1 over AH_2. On its own this has not yet convincingly shown that we have an epistemic reason for a principle of quantitative parsimony that allows us to prefer AH_1 over AH_2. After all, it is easy to modify the description of the case so that the competing hypotheses are on a par and a likelihood defence does not apply. However, when we do so we end up with a new hypothesis of which AH_2 is part that has as lower prior than the competing one of which AH_1 is a part (relative to the background theory and assumptions). Yet again, then, we have reason to favour AH_1 over AH_2.

We can now see that given our background theory and assumptions we have robust reasons to prefer AH_1 over AH_2. When the likelihood defence applies it favours AH_1 and when it does not AH_1 is favoured by considering the priors of the new competing hypotheses. Moreover, it is the intuitively non-parsimonious nature of AH_2 that is the source of the trouble. In this case a principle of quantitative parsimony is on solid ground in favouring AH_1.

3.3. Postulation of Neptune

---

these probability judgements are made. It is, however, a substantive assumption of our account that they can be made.

[32] We could have claimed that volume is maximised, but that would rule out AH_2 for simply being incompatible with the evidence.

[33] Ruling out minimization allows us to block the production of 1 volume of H_8O_4 and ruling out maximization allows us to block the production of 4 volumes of H_2O.

Finally, let us consider the case of the postulation of Neptune to account for the aberrations in the motion of Uranus. This case too will follow the structure above. Given our background knowledge and theory, the non-parsimonious hypothesis either has a lower likelihood on the evidence or a lower prior than its parsimonious rival.

Let us call the postulation of one new celestial body to account for the deviation in the motion hypothesis UH_1. We also know that two new celestial bodies could have accounted for the deviation and let us call this hypothesis UH_2. As before, let us assume that our background knowledge and theory gives us no reason to assign UH_2 a prior higher than that of UH_1. For the sake of simplicity, let us say that they are given equal priors. Now, given our background knowledge and theory, the likelihood of UH_1 on the evidence is higher than that of UH_2. In this case, neither hypothesis entails the evidence without further specification of the masses and the orbits involved. Our focus here, however, is that part of our evidence is that the aberration *can* be accounted for by the presence of a single body. We also know that it *could* be accounted for by two or more bodies. But in order for two or more bodies to account for this motion, we would have to restrict their orbits and masses with respect to one another. Our background knowledge and theory does *not* give us reason to think that they typically are so restricted and this makes the additional assumption required a costly one.

In a move that is now familiar, we could consider a more specific hypothesis UH_3^* that adds to UH_2 the requirement that the motion and masses of the two bodies are orchestrated as to mimic the periodic perturbation that could be produced by the presence of a single mass (UH_4^*). Now, however, UH_1 and UH_3^* are related as probabilistically additive hypotheses. Yet again we reason to prefer UH_1 over UH_2.

4.   Conclusion

All of the cases that we have considered share a similar structure. The problem for the non-parsimonious hypotheses, as we see it, is that in order to form packages of hypotheses that entail the evidence, we typically have to add extra, costly, assumptions. The more parsimonious hypotheses do not come with this cost and are, as such and in the range of cases described, to be preferred to their competitors. Our approach has the advantage that it can deal with probabilistically additive hypotheses in general and not merely additive cases in Baker's ([2003]) sense.

We have taken considerations of quantitative parsimony to come into play at the level of directly competing explanations of the same explanandum. This means that we have not provided an argument that the parsimonious hypothesis will generally be preferable (on epistemic grounds) to the disjunction of the non-parsimonious competitors. This strikes us as correct. We would not want to claim, for instance, that perfectly generally we have

epistemic reason to prefer a parsimonious hypothesis to the disjunction of less parsimonious ones.[34]

Finally, our account shows how, in these generally characterized cases, the likelihood defence translates to a defence in terms of priors and how we can move between the two. Although Sober is not focused on quantitative parsimony, the approach here diffuses the seemingly large difference between Sober's two cases of parsimony where one is motivated by considering priors and the other by considering likelihoods. These cases motivate Sober ([1994], p. 141) to claim that '[…] the legitimacy of parsimony stands or falls […] on subject matter specific […] considerations' and to reject general logical and mathematical defences of parsimony. We have shown how a mathematical and not merely local defence can be given even while accepting Sober's ([1994], p. 152) point that '[…] whether one hypothesis (H_1) provides a better explanation of the observations (O) than another hypothesis (H_2) does […] depends on further auxiliary assumptions A'.

Acknowledgements

Lina Jansson
Philosophy Programme
14 Nanyang Drive
Nanyang Technological University
Singapore 637 332
idajansson@ntu.edu.sg

Jonathan Tallant
Department of Philosophy
Room C45 Humanities Building
University Park
Nottingham
NG7 2RD
UK
jonathan.tallant@nottingham.ac.uk

---

[34] Whether or not the parsimonious hypothesis is to be preferred (on epistemic grounds) to a disjunction of non-parsimonious ones will depend on how high the likelihood of the disjunction of the non-parsimonious hypotheses is on the evidence, whether the extra hypotheses needed in the non-parsimonious cases are compatible or incompatible, how independently likely those extra hypotheses are, etc. The case where we have the strongest reason to prefer the parsimonious hypothesis even over a disjunction of non-parsimonious ones will be when the non-parsimonious cases involve the postulation of very costly additional hypotheses (such as additional laws) that are incompatible with one another and where the likelihood of the individual non-parsimonious hypotheses on the evidence is low without these additional hypotheses. Several of the cases that we have discussed follow this pattern.

References

Avogadro, A. [1811]. 'Essai d'une manière de déterminer les masses relatives des molécules élémentaires des corps, et les proportions selon lesquelles elles entrent dans ces combinaisons', *Journal de Physique, de Chemie et d'Histoire Naturelle*, LXXIII, 58-76

Avogadro, A. [1922]. 'Essay on a manner of determining the relative masses of the elementary molecules of bodies, and the proportions in which they enter into these compounds', *Alembic Club Reprints, No. 4 : Foundations of the Molecular Theory comprising papers and extracts by John Dalton, Joseph-Louis Gay-Lussac and Amedeo Avogadro (1808-1811)*, Oliver and Boyd: Edinburgh

Baker, A. [2003]. 'Quantitative Parsimony and Explanatory Power', *British Journal for the Philosophy of Science*, 54, 245-259

Baker, A. [2013], 'Simplicity', The Stanford Encylopedia of Philosophy (Fall 2013 Edition), Edward N. Zalta (ed.)

Baron, S. and Tallant, J. [MS]. 'Do not revise Ockham's razor without necessity'

Gould, B. [1850], *Report on the History of the Discovery of Neptune*, Washington: Smithsonian Institute

Grant, R. [1852], *History of Physical Astronomy, from the earliest ages to the middle of the nineteenth century*, London: Henry G Bohn

Hanson, N R. [1962], 'Leverrier: The Zenith and Nadir of Newtonian Mechanics', *Isis*, 53(3), 359-378

Huemer, M. [2009]. 'When is Parsimony a Virtue?', *Philosophical Quarterly*, 59, 216-236

Le Verrier, U. [1859], Lettre de M. Le Verrier à M. Faye sur la théorie de Mercure et sur le mouvement du périhélie de cette planète. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, 49, 379-383

Nolan, D. [1997]. 'Quantitative Parsimony', *British Journal for the Philosophy of Science*, 48, 329-443

Partington, J. R. [1964]. *A History of Chemistry*, volume 4, Macmillan, London; St. Martin's Press, New York

Sober, E. [1981], 'The Principle of Parsimony', *British Journal for the Philosophy of Science*, 32, 145-156

Sober, E. [1994]. 'Let's razor Ockham's razor', chapter 7 in Sober, E. *From a Biological Point of View*, Cambridge University Press

Tallant, J. [2013]. 'Quantitative Parsimony and the Metaphysics of Time: Motivating Presentism', *Philosophy and Phenomenological Research*, 87, 688-705

Wallace, D. [2012]. *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation*, Oxford University Press