

Experiment and the Pursuit of Ugly Models

Martin King*

January 13, 2025
DRAFT COPY DO NOT CITE

Abstract

Scientists do not merely choose to accept fully formed theories, they also have to decide which models to work on before they are fully developed and tested. Since decisive empirical evidence in favour of a model will not yet have been gathered, other criteria must play determining roles. I examine the case of modern high-energy physics where the experimental context that once favoured the pursuit of beautiful, simple, and general theories now favours the pursuit of models that are ad hoc, narrow in scope, and complex; in short, ugly models. The lack of new discoveries since the Higgs boson, together with the unlikeliness of a new higher energy collider, has left searches for new physics conceptually and empirically wide open. Physicists must make use of the experiment at hand while also creatively exploring alternatives that have not yet been explored. This encourages the pursuit of models that have at least one of two key features: i) they take radically novel approaches, or ii) are easily testable. I present three models, neutralino dark matter, the relaxation, and repulsive gravity, and show that even if they do exhibit traditional epistemic virtues, they are nonetheless pursuitworthy. I argue that experimental context strongly determines pursuitworthiness and I lay out the conditions under which experiment encourages the pursuit of ugly models.

1 Introduction

Philosophers of science have long tried to identify the characteristics of good scientific theories—the theories that one should rationally opt for among rivals. This includes epistemic criteria, such as internal consistency and the scope of the theory's fit with empirical data. But philosophers have also recognized that some important criteria are perhaps not strictly epistemic, but cognitively beneficial for those working with the theories, such as the theory's simplicity. Of course, scientists do not simply choose to accept theories after they are fully formed, they also have to decide which to work on before they are theoretically developed and empirically tested. A growing literature on the pursuitworthiness

*Munich Center for Mathematical Philosophy, Ludwig Maximilian University of Munich

of theories has emerged stemming from early work by Laudan (1977) who distinguished between the context of acceptance and the context of pursuit. These decisions of pursuit are inevitably made on the basis of considerations that are more pragmatic and heuristic and, I will argue, are strongly influenced by the experimental context.

In order to show this, I will examine the case of contemporary high-energy physics (HEP). The changing experimental context of HEP in recent years make clear the influence on criteria of pursuit. For decades, the Standard Model (SM) has guided much of the experimental and theoretical research in HEP and since the Higgs boson was discovered in 2012, the SM is in many senses complete. Physicists have long been devising alternative theories and higher-energy completions of the SM, but no significant evidence in favour of any of these models has been uncovered at the LHC. Models that predicted new particles that would be clearly visible in the data as sharp resonances above the SM background simply did not manifest. When it comes to new physics models, the low-hanging fruit is gone. The models that remain in the available parameter space and are the subject of most research are by-and-large complex, difficult to calculate, ad hoc, have small empirical scope, are not unifying, not very plausible, etc. In short, the models on the table are *ugly*—they do not exhibit traditional theoretical virtues. What I will argue is that they are nonetheless pursuitworthy since they exhibit a different set of criteria that are important in our current experimental context.

I will highlight two important experimental conditions in the state of HEP: that there has been no evidence in favour of new physics with our current experiments; and second that there is no new higher-energy experiment forthcoming. These conditions lead to two guides for pursuit in HEP: i) radical novelty; and ii) immediate testability. The increasing importance of novelty stems from a sense that physicists have already exhausted the avenues that were most appealing in terms of epistemic virtues—radically different approaches are needed. Immediate testability enters because with no new higher-energy collider on the horizon, there is no possibility of simply turning on a machine and seeing evidence of a new physics theory Beyond the Standard Model (BSM). There is limited empirical progress to be expected in working on theories outside of achievable energies. In order to maximize chances of discovering new physics, physicists have had to get very clever: they have had to re-evaluate long-cherished guiding principles and the very foundations of quantum field theory (QFT); and they have had to get very creative with model-building, experimentation, and data analysis. This points to more than a decline in standards and rather indicates that a different set of criteria are playing a stronger role in determining pursuit.

The paper will begin by discussing epistemic virtues and the shift to pursuitworthiness in Sec. 2. Then in Sec. 3, I look at three ugly models that are pursuitworthy, but do not exhibit the traditional virtues of theories. Finally, I distil two desiderata (novelty and testability) from the actual cases that are difficult to square with extant accounts of pursuit in Sec. 3.2. Detailed descriptive accounts, such as (Franklin, 1993) and decision-theoretic models (Nyrup, 2015) do not stress or capture the experimental context-dependence that I wish

to highlight. As such, the aim of the paper is not to present an account of pursuitworthiness. In fact, a general account may not be possible given the context-dependence I argue for, in agreement with arguments by Shaw (2022). Ultimately, what I hope to show is that the context in which research takes place strongly determines considerations of scientific pursuit—it is the way particle physics is now that leads us to novelty and testability as dominant criteria.

2 Epistemic Virtues and Pursuitworthiness

The debate is often traced back to Kuhn (1974), who outlined the most important criteria for theory choice—accuracy, consistency, simplicity, scope, and fruitfulness¹. He included these not because they are exhaustive, but “they are individually important and collectively sufficient to indicate what is at stake” (p. 357). He sought to reassure people of the objectivity that was possible within the framework developed in (Kuhn, 1962), but noted that how these criteria are assessed is still a subjective matter. Two scientists may agree that simplicity is important, but they may take simplicity to mean different things and they may disagree on how much it ought to weigh against other criteria like accuracy or scope.

An important contribution to the debate that I wish to pick up on was made by Douglas. Douglas (2009, 2013) proposes a finer-grained account of values in science in part to show that they are not necessarily in tension with each other. She does this by making two important distinctions: that between minimal criteria and desiderata, and that between values as applied to theories themselves and as applied to theories in relation to evidence. There are two such minimal criteria for a scientific theory: internal consistency and empirical adequacy; internal consistency is a property of the theory in itself and empirical adequacy is a property of the theory in relation to evidence. She separates other cognitive values into pragmatic (which apply to the theory in itself) and epistemic (which apply to the theory in relation to evidence). The latter two are mere desiderata on her account and include such values as simplicity and breadth of scope. Thus, it is pragmatically beneficial if a theory is simple in itself, but it is epistemically valuable if the theory is simple in relation to evidence. What will be important for our discussion is that even so-called minimal epistemic criteria are not relevant for pursuit.

The difference between acceptance and pursuit is not simply a matter of a lower threshold for the same criteria. If it were, models that have not yet been fully developed would always lose out in comparison with established theories. One would then always be pursuing the best developed and empirically established theories, which misses the point of pursuit and misses an important scientific practice. In order to rationally opt to work on some models and

¹Kuhn’s list and arguments sparked a long history of debate about the status these criteria and other criteria, opening what can be described as a new demarcation problem (Douglas, 2009; Lacey, 1999; Laudan, 1984; Longino, 1995; McMullin, 1985, see for example), but that is not the avenue this paper will pursue.

hypotheses before there is conclusive evidence and before all the conceptual problems have been straightened out, one must have a different set of criteria on which this choice can be made.

Pursuitworthiness, as an explicit concept, goes back to (Laudan, 1977) and (McMullin, 1976) and significantly developed by (Achinstein, 1993) and (Franklin, 1993). The idea is to shift the focus from the justification of acceptance to the justification of pursuit, where decisions are not necessarily about truth or realism, but can focus on short timescale decisions. For (Laudan, 1977), pursuit can include refining a hypothesis² empirically, developing it theoretically by solving conceptual problems, and so on. And importantly, one can be justified in working on theories with serious conceptual or empirical problems, if, in line with his account of scientific progress, the theories have resources to potentially solve otherwise unsolved problems. This idea is echoed by Franklin (2016) who says that “a hypothesis can act as a stimulus for further work even if one were skeptical of both the hypothesis and the evidence supporting it” (p. 82) and that “belief in the truth of a hypothesis or in the experimental results is also not a requirement of further theoretical work,” (Franklin, 1993, p. 106). It is not the case that one should pursue the theory that is best supported by evidence, has the largest scope, or even restrict oneself to theories that are consistent. These are the precisely the points at which considerations of pursuit pull apart from those of acceptance. The “basic criteria that any scientific work must meet” (Douglas, 2009, p. 92) are not yet met in the context of pursuit. Decisions to pursue are less constrained by these basic epistemic considerations, which allows many more features of the theory and the context to play leading roles in the decisions. This also brings with it a worrying openness in descriptions of pursuitworthiness assessments. Let us look at a few accounts of how decisions of pursuitworthiness are made.

(McKaughan, 2008) argues that Pierce’s account of abduction and IBE was really about pursuitworthiness. He articulates Pierce’s view as follows: “if we estimate that testing the hypothesis will be easy, of potential interest, and informative, then we should give it a high priority” (McKaughan, 2008, p. 457). This is a very reasonable view, but one could easily imagine that these conditions are not all satisfied, or are satisfied by different models to different degrees. This alone is of little help to those hoping for a normative account of pursuit. As another example, Franklin (2016) takes a detailed historical look at how these decisions were made to pursue empirically and theoretically dubious models (in this case that of a fifth fundamental force). He sums up his analysis as follows:

I suggested several reasons why a scientist might choose to further investigate, or to pursue, a hypothesis. These included the interest

²Franklin often speaks of pursuing hypotheses, others speak of theories or of models. I take the line between model and theory to be more-or-less continuous, but one could reserve the term theory for a well-developed and broadly empirically confirmed model. Thus, theories are the kinds of things that would be subject of theory-choice decisions and models the subject of decisions of pursuit. One could of course pursue just a hypothesis, which may not yet have much mathematical structure or conceptual content and may only be about the value of some observable.

and importance of the hypothesis; its plausibility, based on existing evidence, on its resemblance to other successful theories, or on its mathematical properties; the fact that it fit in with an ongoing research program; and its ease of test, in which I include the conceptual simplicity of the test, which differs from the technical experimental details of the test, which might be quite complex; and whether or not the experiment can be performed with either existing apparatus or with small modifications of it, or with a relatively modest investment in a new apparatus.

(Franklin, 2016, p. 80)

This is a rather long and permissive disjunction that can hardly be wrong. There are surely many reasons that a scientist could consider when deciding to pursue a given hypothesis and I believe that Franklin is quite correct in giving much place to the experimental side of the decision. Although, it is almost certainly true, the above quote is descriptive and less helpful in application than one might hope. Other accounts have provided more detailed breakdowns of how decisions of pursuit are, or ought to be, made.

One such is in given in (Nyrup, 2015), where the problem to be solved is put into a decision-theoretic framework. He says “Scientists need to justify which hypotheses are worth investigating in order to prioritize their resources. Justifying pursuit is, essentially, a decision-theoretic problem of how to optimize the epistemic output of science” (Nyrup, 2015, p. 753). The idea is to model the decision to pursue a hypothesis as a weighted sum of factors based on the expected epistemic value (EEV) of pursuing that hypothesis. One can then use the formalism of probability calculus to obtain rigorous results. In Fig. 1, one sees the calculation of the EEV and various contributions are modelled in, such as the value gained from pursuing the hypothesis even if the hypothesis is false. Decision theoretic models in general can be informative in that separating out relevant factors and examining the role different weights would play can tell us a lot about how kinds of decisions can be affected by the information that goes into them. For any given decision, calculating with real values of course will not be possible, so it is of limited applicability. The messy details about how one comes to the number that feed into the calculation are precisely what most are interested in when discussing pursuitworthiness. For example, I am most interested in the influence of the ease of test and the novelty of the hypothesis, but there is no place for any such values, they must simply be folded in to the estimated value of pursuing the hypothesis.

3 Hunting for New Physics

The 27km long LHC at CERN began operation in earnest in 2010, colliding protons at an unprecedented 7 TeV. On July 1, 2012, the two large experiments at the LHC, ATLAS and CMS, announced the discovery of a Higgs-like boson

$$\begin{aligned}
EEV(p(H)) = & EV(a(H) \& H) \cdot Pr(a(H) \mid H \& p(H)) \cdot Pr(H) \\
& + EV(a(H) \& \sim H) \cdot Pr(a(H) \mid \sim H \& p(H)) \cdot Pr(\sim H) \\
& + EV(r(H) \& H) \cdot Pr(r(H) \mid H \& p(H)) \cdot Pr(H) \\
& + EV(r(H) \& \sim H) \cdot Pr(r(H) \mid \sim H \& p(H)) \cdot Pr(\sim H).
\end{aligned}$$

Figure 1: A decision theoretic model for pursuit of scientific hypotheses, where $EEV(p(H))$ is the expected epistemic value of pursuing H , based on the probabilities of accepting $a(H)$, rejecting $r(H)$, or remaining agnostic $\neg a(H) \wedge \neg r(H)$, conditionalised on the truth of H (Nyrupe, 2015).

at 125 GeV. Early on, there was still room for this particle to exhibit non-SM-like properties, but as more analyses were conducted, it looked more and more like the ‘vanilla’ Higgs of the SM, and the door began closing on new physics. The LHC, now operating at 14 TeV during Run 3, will soon be undergoing a major upgrade to higher luminosity³, which should be completed around 2028. While this will allow an order of magnitude more data than the nominal LHC runs, expectations of new physics are low. The stated physics goals of the upgrade do include two particular avenues for discovering new physics: stops, and neutralinos and charginos (see (Schmidt, 2016) for more details). While the additional data will help increase the mass range that can be probed, the space is already heavily constrained. ATLAS has managed a 95% exclusion limit on chargino-neutralino pairs up to around 850 GeV and stop pair production is excluded up to around 1 TeV (Aad et al., 2023; Williams, 2024) and the High-Luminosity LHC (HL-LHC) will push these beyond 1.5 TeV (Rühr, 2016).

There are plans for the 90km long Future Circular Collider (FCC) at CERN that could achieve energies up to 100 TeV, but it is still in the proposal phase. A feasibility study will be completed in 2025 and, if all goes well, the FCC high-energy phase could be operational in the mid-2070s. There are two large sources of trepidation: whether or not this is the best use of the 15 billion Francs this would cost; and whether or not this is the best target of the physics community’s efforts. These too are important matters of pursuit, but far outside the scope of this paper, whose target is the features of particular models. Arguments that helped make the case for the LHC, in particular naturalness arguments for the need to discover the mechanism of electroweak symmetry breaking near the electroweak scale and for new physics at low energies (in particular supersymmetry), are now regarded with suspicion. They are precisely the guides and assumptions being reevaluated in face of the lack of new discoveries. There are no widely accepted arguments that new physics should be discovered below the achievable 100 TeV—the SM could be a perturbatively accurate description of nature up to the Planck scale (10^{18} GeV).

Part of an approach to find new physics elsewhere has involved shifting focus

³Luminosity is the ratio of the number detected events per time per area and roughly corresponds to the amount of data the accelerator can generate.

from the high-energy to the low-energy frontier (Jaeckel and Ringwald, 2010), emphasising the prospects of discoveries with small experimental searches for axions (and axion-like particles). Another emphasis has been on finding long-lived particles the large detectors at the LHC are not sensitive to by building a small detector further down the tunnel. This experiment (FASER) has provide constraints on dark photons (Boyd et al., 2023a), and observed for the first time neutrinos produced at collider experiments (Abreu et al., 2023). There are however, no hints of new physics here either.

The lack of new physics discoveries has changed the field of particle physics in a number of significant ways. One is that many physicists are increasingly turning to model-independent methods and to models that would not be pursued in a different experimental context. To put it in a rather pessimistic way: physicists are scraping the bottom of the barrel for models, trying to get the most they can out of the experiments they have on hand. However, one never knows where a breakthrough will come from or when a small discrepancy will ultimately lead to a crucial test of a new theory. Hunting for new physics requires a multi-pronged approach involving among other things precision measurements, model testing, and pure theory. Physicists are working on uglier models in general and they should be, given the current open-ended context, but of course not all of them. One may describe the situation by saying that there are shifts in the division of cognitive labour. Where once there may have been an overwhelming number of theorists working on string theory and supersymmetry, there are now fewer, and there are more working on the remaining LHC-accessible models, even though they are ugly. This is a reasonable response to the situation.

3.1 Ugly Models

In this section, I present three ugly models of BSM physics that are nonetheless pursuitworthy: neutralino dark matter; the relaxion; and the aptly-named repulsive gravity. I choose three to highlight different problems, namely ad hocness, consistency, and implausibility, while drawing attention to the two features that make them most pursuitworthy regardless, novelty and testability.

Neutralino Dark Matter

The first model is the singlino-like neutralino model of dark matter, which is a next-to-minimal supersymmetric model (Ellwanger et al., 2010). The model represents a class of solutions to a problem of the SM, in this case the dark matter problem.⁴ The kind of solution it represents is weakly interacting massive particle (or WIMP) dark matter, where some new matter field is introduced, often in a new larger symmetry, which is light and stable, so that other particles can decay into it and then it will stick around and float about the universe. A particle is a DM candidate if it is neutral and interacts at most only weakly with

⁴The basic problem is that it seems that visible matter along with known physics falls drastically short of being able to reproduce basic observable elements of the cosmos. One can find more about this from a philosophical point of view in (Boyd et al., 2023b).

other matter, so that we would not have seen it in our precision experiments to date, but has mass and so will have the gravitational influence on matter that we seem to see in astronomical observations. If one wants to study a model of dark matter, WIMP solutions are a perfectly reasonable avenue to pursue.

Perhaps the best way to understand some of the relevant parts of the model is to see how the model was arrived at. The basic WIMP idea is to add a new particle to account for missing mass in the universe and we might think that SUSY is attractive for a variety of reasons. We might try to search for minimal supersymmetry (MSSM), which has four neutralinos, the lightest of which (\tilde{N}_1^0) is a DM candidate. The model is theoretically capable of supporting predictions and is even LHC-accessible, so it can be empirically investigated. Further, as a supersymmetric model it could possibly solve many problems of the SM in one fell swoop, potentially unifying the coupling constants of the fundamental forces at high energies and providing a solution to the Higgs naturalness problem as well. But the model starts to get ugly when we modify and constrain it on the basis of it not having shown up in our experiments.

There has been no evidence of the MSSM at the LHC. The energy scale at which SUSY may dwell can be pushed back by compromising features of the model. For example, a major benefit of the model was its ability to solve the Higgs naturalness problem, but as the SUSY scale gets heavier, the hierarchy problem that was alleviated with loop corrections, begins to grow again and we are left with a ‘little hierarchy problem’ between the electroweak and DM scales. Important for our DM model in particular, the MSSM parameter space combined with DM relic abundance and direct detection data from LUX and XENON1T, is effectively excluded (Abdallah and Khalil, 2016; Akerib et al., 2017; Aprile et al., 2018; Guchait and Roy, 2020). If a preferred model’s parameter space is getting too small, the obvious next step is to expand the parameter space, i.e. move to the less minimal model, viz. next-to-minimal SUSY (NMSSM). This model adds a Higgs singlet to the MSSM’s extra Higgs doublet. This is an important addition as it increases the singlino and neutralino states of the model, changing its phenomenology and requiring new searches. When this fails to turn up any evidence in searches, one can simply move to the next-to-next-to-minimal SUSY model and try again.

This model is very ad hoc. The model is less simple, more computationally involved, requires an even larger number of hypothetical particles, each specified by many new free parameters. Some of these particles, including seven non-SM-like Higgses, are highly constrained by existing data (Hayrapetyan et al., 2024). While many might give SUSY models in general high plausibility, this model resides a tiny, but not excluded, area of the parameter space. If one wants to examine the viability of such a model, one can numerically scan the parameter space to see if it is worth investigating. Elements of the parameter space are sampled and combinations are calculated to see if they remain in the empirically available parameter space. For NMSSM, there were only 2500 possible combinations of values out of 25 million sampled in 2018 (Mou et al., 2018). One only looks at this model because other models have been ruled out and one only looks at these specific exotic regions of the parameter space

because every where else has been excluded.

It is a remote possibility, but that is enough to make it pursuitworthy in the context where not much else is available. Searching for DM candidates is still “one of the most exciting and challenging programs” of HEP (Guchait and Roy, 2020, p. 2) and is treated by CMS and ATLAS as a ‘high-priority’ analysis (Abercrombie et al., 2020). The fact that it can be probed with existing apparatus and the parameter space constrained means that there is something to be learned in performing these analyses. There is also a rather significant payout if any part of the model were to be suggested by the data. Getting a hint of supersymmetry or of a dark matter candidate would be a massive payout for a relatively low cost of continuing to probe accessible parameter space.

Relaxion

The second model represents a class of solutions to the Higgs naturalness or hierarchy problem that involves no new dynamics at the weak scale.⁵ The relaxion is an alternative to the kind of solutions that SUSY provides and which may be more promising as the LHC precision data rules out models up to and beyond the TeV scale where new physics limit the amount of fine tuning the SM requires. The idea is to dynamically relax the electroweak scale via cosmological evolutions. In this model, an axion (more specifically, an axion-like particle) is added that couples to an inflaton (a field of cosmological inflation). This can create a slow rolling periodic potential that ‘naturally’ selects a light Higgs mass when the Higgs vacuum expectation value (vev) becomes non-zero as the relaxion stabilises after it passes a critical point.

The characteristic rolling potential of the relaxion is depicted in Fig. 3.1 taken from the original paper proposing the solution in 2015 (Graham et al., 2015). The Higgs vev is selected when the slope of the linear scanning term balances with that of the periodic potential, so-called dynamical stopping. The authors describe the model as follows:

“During inflation, ϕ will slow-roll, thereby scanning the physical Higgs mass. At some point in the ϕ potential, the quadratic term for the Higgs crosses zero and the Higgs develops a vacuum expectation value. As the Higgs vev grows, the effective heights of the bumps, Λ^4 , in the periodic potential grow. When the bumps are large enough they become barriers which stop the rolling of ϕ shortly after m_h^2 crosses zero. This sets the Higgs mass to be naturally much smaller than the cutoff.”

(Graham et al., 2015, p. 3)

⁵The Higgs naturalness problem is that the Higgs is a scalar field and in Quantum Field Theory one would expect to find the field around the energy level of the cut-off of the theory. But for the SM, this is (as far as we know) at the Planck scale. If the Higgs is light (and in this context 125GeV is light), then an extremely delicate cancellation of quantum contributions to thirty orders of magnitude is required. There is quite some debate about the best way of posing the naturalness problem and I will not go into this here. For an overview on the relevant positions and a philosophical analysis, see (Harlander and Rosaler, 2019).

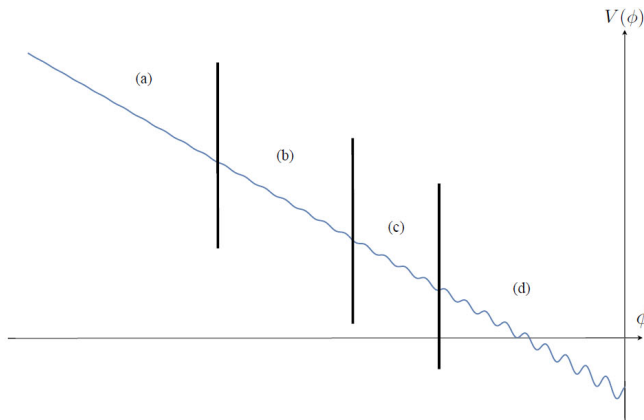


FIG. 2: A close up of the region of ϕ 's potential as the barriers appear. The evolution in these regions are (a) classical rolling dominated, (b) dominated by quantum fluctuations in the steps but classical rolling between steps, (c) classically stable, but quantum fluctuations/tunneling rates shorter than N e-folds, and (d) classically stable, quantum transition rates longer than both N e-folds and 10 Gyr. Again, for clarity, the potential is not to scale.

Figure 2: A depiction of the relaxion model's rolling periodic potential.

This is a radically novel approach to the naturalness problem that intrigued and inspired many physicists to work on variants of the model, as Franklin noted in some cases of pursuit, even though there were conceptual and evidential issues with the model.

Just like the Higgs itself, the model is very ad hoc: the model is introduced to solve a particular problem without any empirical support. But beyond its ad hocness, the model as it was first proposed has a number of empirical and conceptual problems. While the model attempts to solve the naturalness problem, it is not even *technically natural* (in the sense of ('t Hooft, 1980)), since it breaks a gauge symmetry. The axion was first introduced as a potential solution to the strong-CP problem (Peccei and Quinn, 1977; Wilczek, 1978), but here the particle is not a true axion (only axion-like) and in many models actually contributes to the strong-CP problem. In short, a Higgs fine-tuning has been replaced with a CP-fine-tuning problem, but at least it is orders of magnitude smaller. The aspect I would like to highlight is that the model is also inconsistent with cosmological observations and many accepted cosmological models. The mechanism for stopping the relaxion requires a long period of inflationary expansion so that the whole range can be scanned. Rather than the 50 or 60-odd e-foldings⁶ one would expect on most cosmological models, in the minimal relaxion model, one needs 10^{33} or so (Germán et al., 2023; Nelson and Prescod-Weinstein, 2017).

More work has been done on the model in order to alleviate some of these problems. The relaxion's CP problem can be averted in the so-called 'landscape

⁶The amount of time required for growth by a factor of e .

relaxion’ (Nelson and Prescod-Weinstein, 2017), in which the relaxion populates a landscape of universes, there is a noted temperature-dependence of the axion mass, which implies that the QCD contribution to the potential becomes much stronger after the inflationary period when temperatures drop well below the QCD scale. This model is then even more complex and ontologically richer with a landscape of universes, and still involves a finely-tuned cosmological constant. The gigantic e-folding number can be reduced, for example, by replacing the original dynamical stopping mechanism with a thermal stopping from particle production (Hook and Marques-Tavares, 2016), or from Hubble friction (You, 2017), or from double-scanning with additional scalars (Espinosa et al., 2015). A stochastic stopping mechanism with a ‘paper boat’ potential⁷ of the relaxion is proposed in (Wang, 2019). The model has many virtues compared to other relaxion models, but importantly, it is very ad hoc. The Higgs mechanism was charged with being ad hoc as it was only introduced to give the W and Z bosons mass. This relaxion model aims to resolve the problem by introducing a new very specific potential that would explain that other specific potential if it were just so. One ad hoc model is attempting to explain another. Each of the models in the relaxion-class of solutions brings one or more critical issues; resolving certain problems while introducing others and inevitably complicating the picture. The consistency issue with cosmological data is far from resolved, but rather than being a barrier to the pursuitworthiness, it is target and inspiration for future work.

The model’s novelty and testability make it pursuitworthy. Models of axion-like particles “are among the most compelling candidates for physics beyond the Standard Model (SM) of particle physics” (Choi et al., 2021, p. 1). The models represent an interesting and novel class of solutions to an existing problem. The models are also able to be tested and constrained by existing and future experiments. For example, there is sensitivity to the axion-nucleon couplings (via CASPER and proton storage ring) and to the Higgs-axion mixing angle (from AION and AEDGE)(Choi et al., 2021). Of course, there is also cosmological data with which the models are rather poorly fit. The fact that there is conflict with experimental data that may be overcome means that there are targets and resources for continuing research and that progress can be gauged by this contact with experimental data.

Repulsive Gravity

Lastly, I will discuss the very aptly named repulsive gravity.⁸ The hypothesis behind repulsive gravity is that there is a repulsive gravitational force that acts between matter and antimatter and that this could provide enough energy to drive cosmic expansion and explain the matter-antimatter asymmetry of the universe. As such, these models represents classes of possible solutions to the

⁷Another name for the Mexican-hat shaped potential characteristic of the Higgs scalar field.

⁸Repulsive gravity models are loosely related to the fifth force models and experiments discussed by Franklin (2016).

dark energy (DE) problem (Hohmann and Wohlfarth, 2010) and to the baryon asymmetry problem (Noyes and Starson, 1991). Since the discovery of accelerated cosmic expansion in the late 1990s, cosmological models, such as Λ CDM, have had to include a term that represents negative pressure but is otherwise theoretically empty. The old idea of anti-gravity was given a new lease on life. Some of these models assumed that antimatter has negative gravitational charge (Hajdukovic, 2012), in violation of the weak equivalence principle, or that it follows as a prediction of GR if the transformation of matter into antimatter requires not just a transformation of charge conjugation (C) but of parity (P) and time reversal (T) as well (i.e. CPT) (Villata, 2011, 2012).

Some models are more baroque than others. (Hohmann and Wohlfarth, 2010) introduce $N \geq 3$ ‘dark’ copies of the SM, each with its own gravitational metric. The metrics describe attractive gravitation within the groups, but repulsive gravity between them. The resulting matter-antimatter gravitational repulsion could explain the expansion of the universe while the additional matter would constitute the missing dark matter. This model could potentially solve two long-standing problems, but suffers from a number of conceptual and empirical issues. It was nonetheless interesting enough to for many physicists to pursue. It is, like the other models discussed above, very complex, and ontologically rich, proposing to triplicate⁹ all known particles. The model is externally inconsistent with a very widely accepted theory, viz. General Relativity. It had no empirical contact when it was proposed; sufficiently precise tests could not be conducted due to a lack of stable test antiparticles.

This model is quite radical and general expectations for successful testing are were likely very low. Experimental tests require extreme precision and technical apparatus, but tests were recently conducted at CERN where antiprotons were already being produced as a matter of course. The ALPHA experiment at CERN has managed to make the production and capture of anti-hydrogen rather routine. The ELENA ring can deliver 10^6 antiprotons per minute, many of which can be cooled and trapped (Anderson et al., 2023). The effect of repulsive gravity could be tested in a conceptually simple way: trap antiprotons in a container, open the top and bottom, and compare the number of annihilations out of each end, as demonstrated by an early proof-of concept experiment in 2013 following studies in (Gabrielse, 1989). Ordinary matter under the influence of gravity would predict that 80% of the particles would escape through the bottom.

Repulsive gravity was ruled out as evidence indicated that, indeed, antiprotons act just the same as protons under gravity. The experiment was able to conclude the following: “The probability that our data are consistent with the repulsive gravity simulation is so small as to be quantitatively meaningless (less than 10^{-15}). Consequently, we can rule out the existence of repulsive gravity of magnitude $1g$ between the Earth and antimatter. The results are thus far in conformity with the predictions of General Relativity. Our results do not support cosmological models relying on repulsive matter–antimatter gravitation.” (An-

⁹It triplicates, because the $N=2$ case was proven a no-go, so the model proceeded with $N \geq 3$.

derson et al., 2023). The negative result was strongly expected. The experiment was nonetheless conducted, because CERN already had anti-protons available and this hypothesis could be clearly tested with a straightforward experiment. It was pursuitworthy even though few had any hopes of this experiment would successfully overturn GR.

3.2 Novelty, Testability, and Pursuit

These models I presented are not cherry-picked as the ugliest models on the market, since those I probably would not have heard of. These are relatively popular models and are representative of the best and more pursuitworthy models particle physicists have to work on. As mentioned, I have called these models ugly because the traditional epistemic virtues that make theories beautiful, elegant, simple, intuitive, understandable, etc. are typically absent or meagre. The models I presented: are ad hoc; have narrow empirical scope; are highly constrained; are fine-tuned; inconsistent with well-established models in other domains; not compatible with data in a broader scope; and compared to other models that have been ruled out are less simple, more ontologically expensive, more difficult to calculate, and not very plausible.

What I mean in saying that these ugly models are pursuitworthy is that they should be worked on—it is reasonable to work on these models given our experimental context. Pursuitworthy models should instantiate the basic conditions of an adequate scientific model like being consistent enough to generate predictions and being mostly empirically adequate within a given domain (even though many pursuitworthy models have not yet had any empirical contact.)

In addition to this, there currently seem to be two pursuitworthy features of the models we have examined:

- (i) the immediate and easy testability of the model or hypothesis
- (ii) the class of solutions represented by the model is radically novel

These are not desiderata that should be maximized: the easiest testable hypotheses are likely trivial, and the most radically novel are probably incredibly unlikely. They are rather features that may constitute independently sufficient reasons for pursuit.

By immediate and easy testability, I mean that a model makes clear predictions in achievable energies, it might already be discoverable in existing or forthcoming data, or makes use of existing facilities (or with relatively small cost-effective amendments), very much along the lines discussed by Franklin (1993). The NMSSM can be constrained with existing LHC data; the relaxion star halos are observable and there are table-top axion experiments that have implications for the model; repulsive gravity was tested with antiprotons already a by-product of the LHC. Some hypotheses are being pursued for almost no other reason than that they can be easily tested, such as looking for magnetic monopoles in the discarded LHC beampipe simply because it was there (Joergensen et al., 2012). No one would have built the LHC in order to generate

a used beampipe to test for an implausible hypothesis, but the cost was low and the potential payoff was high. Contrary to what many have said, plausibility need play little role in pursuit.

By the class of solutions being radically novel, I mean that a model takes a genuinely new approach (rather than merely expanding on or continuing an old approach). This virtue is extolled by the relaxion proposal, and to some degree by the other two models. The NMSSM is a part of an existing class of supersymmetric models, but the relaxion made for such a pursuitworthy model because it in part introduced this new class of solutions. The relaxion makes new use of an existing concept—using an cosmological implications of the axion as a way to solve the Higgs naturalness problem is the kind of clever thinking that excited physicists and can kickstart a small research program. Of course, what is novel is time-relative, and I am not trying to claim that the relaxion is more pursuitworthy than supersymmetry was. Rather, I mean to say that supersymmetry today is no longer a radically novel approach.

Novelty represents a different approach to finding new physics than testability. Testability reflects the need to close the gaps, constrain parameter space, and make use of existing apparatus. Novelty reflects the need to take new approaches and reevaluate theoretical foundations, assumptions, and guides and come at problems from a different angle. Models that are highly novel may not be testable and may nonetheless be pursuitworthy. Much theoretical work may need to be done before there can be any empirical access. A great deal of creativity may be required to envision how to test the predictions of a model. The achievable scope, scale, and precision of experimental physics lags far behind what theorists can dream of. Of course, unexpected empirical access may be opened up at any time, such as what happened with the studies of analogue black holes (Barceló et al., 2005).

Accounts, such as those of Franklin (1993), Laudan (1977), McMullin (1976), and Achinstein (1993), mostly picked up on at least six common criteria of pursuit. The first is often called the i) interest in the hypothesis, which indicates how interesting it would be to the community if the hypothesis were true. Second, there is ii) informativeness, which is how much one would learn (how much progress would be made) by pursuing the hypothesis. Third, there is the iii) plausibility of the hypothesis, which is that it is reasonable or is somewhat empirically supported. Fourth, there is the iv) ease of test, which has to do with how easily or clearly one can obtain a positive or negative result. As Franklin points out, this includes both how easily one can distinguish the consequences of the hypothesis being true from it being false and how cost effective (including time) it would be to implement an actual experiment. Fifth, there is the v) analogousness of the hypothesis to past successes, as this may increase plausibility, or as it may facilitate mode transfer (à la (Nyrup, 2020)). Lastly, a model or hypothesis can be pursuitworthy if it has the conceptual resources for solving problems also known as its vi) heuristic power. Let me go over these in turn in virtue of the case studies presented here.

i) The interest-if-true would be enormous for any BSM model. A statistically significant discovery of a new effect (antigravity) or a new particle (relaxion or

neutralino) would upend the current situation in particle physics, spurring a frenzy of model building, theorising and experimentation. One only needs to look at the frenzy that followed the 2015 diphoton excess at 750 GeV to see how eagerly and enthusiastically physicists would welcome new physics.¹⁰ Because of this, it would be hard to say if any new physics discovery is more interesting than another. ii) The informativeness of the study estimates how much would be learned even if the experiment or study is not successful (otherwise it would get folded into (i) interest). This is realistically quite low for these models, since a well-constrained parameter space only becomes better constrained. In doing theoretical work, one learns about the model, but not necessarily about the world in the case these models do not obtain. iii) The plausibility of the models is also quite low. Previously, searches for BSM physics were given much higher estimates of plausibility, but the remaining ad hoc models in their tiny parameter spaces are not widely considered likely to succeed. Of course, the fact that there is a chance, even a small chance, at a major breakthrough is enough. It is more of a possibility than a plausibility that counts. iv) With respect to ease of test, from what we have seen it almost cannot be stressed enough how strongly it constrains and guides what is pursuitworthy. If something is not easy or it is too expensive (in terms of various resources) then it will not likely be pursued; if something is easy, one may pursue it regardless of how interesting, informative, or plausible it may be. v) The virtue of analogy is only likely to play a role where there is a tighter conceptual or mathematical analogy between an accepted and a novel application of model or idea, because the very idea that past successes will continue to work is something is being given up in pursuit of novel solutions. While, other things being equal this could be a benefit for the pursuit of some studies or models, it was not so relevant in the cases we examined, and rather, taking a new approach seemed to be more of a virtue. vi) Lastly, the problem-solving ability is certainly important insofar as models like the NMSSM come with tools that can be made ready to solve various problems and that classes of solutions inspire variants that may involve more clever thinking to avoid elimination. However, I should again note that the empirical scope of models has become narrower in that pursuing supersymmetry and grand unified theories that may solve many problems at once are being passed over in favour of models that simply solve one for the reason that they are more empirically accessible. Perhaps supersymmetry is a genuine symmetry of nature, but physicists are less inclined to think it will be discovered at low-energies in the forms that HEP phenomenologists have written down.

What we find in the end is a partial connection between past discussions of pursuit and the actual virtues of pursuit in our current research context. A decision-theoretic model, such as (Nyrup, 2015), also makes a partial accord with this study. The model makes the terms involved in the overall pursuitworthiness assessment explicit. The real drawback is not the failure of the model in any real sense, but the schematic nature of the model and the lack of ability to actual get

¹⁰For more about this response see (Ritson, 2020).

any numbers out that could help understand changes like we have experienced in pursuitworthiness in particle physics. It is nice to envision in a simple calculus in which a term exists for informativeness (the expected epistemic value of pursuing a hypothesis given that that hypothesis is false), but it does not help up determine that value. If one is estimating all of the values of terms in the decision theoretic model, then this is really only one step away from simply estimating the overall pursuitworthiness from the start. The account is fully based on the truth or falsehood of the hypothesis and this may make it difficult to represent something like the novelty of an approach. It is not clear how one would assess the pursuitworthiness of a scientific model difficult, as it is not a proposition and not truth-valued. It also lacks any resources to discuss the context in which the research is being pursued and thus is also unable to explain differences that may arise from distinct research contexts. If it is capable of reflecting this difference it would only be in the values that are put in by hand.

One might be tempted to claim that these models are not really pursuit-worthy and that they should not be pursued, perhaps arguing that this kind of HEP should be abandoned in favour of pure theory or devising ways of testing more beautiful models. When it comes to the normative part of the argument I want to argue that the response in adopting these two desiderata is a rational one in the given experimental context. This discussion necessarily brackets larger issues about whether one should pursue these models and experiments rather than focusing on an entirely different branch of physics. But this is not something an examination of pursuit should have to do. No account of pursuitworthiness I have come across has made a case for why a physicist should work on a particular model rather than take up biology or the flute. Nor do I intend to address whether the best course for discovering new physics lies in a new higher-energy accelerator or from letting deep neural networks loose on CERN data. My aim has been to demonstrate the rationality of pursuing ugly models given our experimental context in HEP.

4 Conclusion

Novelty and testability are pursuitworthy qualities of models in our current situation and they function basically as independent sufficient conditions for pursuit. Approaches based on maximizing theoretical virtues simply have not borne fruit and there are limited new experimental horizons available. It makes a great deal of sense that these two virtues have come to the fore; these are two very reasonable responses to the situation where no new physics has been found and no new higher energy frontier is on the doorstep. It may still be that these ‘ugly’ models still exhibit some of the classic ‘beautiful’ characteristics; all things being equal, a less complex model might be preferable if it was just as testable. But the fact remains that these ugly models do not exhibit theoretical virtues very much at all, and this alone means that we need to pay attention to pursuitworthiness in addition to acceptance. A logic of pursuit could only get us part of way, since, as I have shown, the context in which research taking place

strongly determines in some cases what the dominant features of pursuitworthy models are. It has not always been the case, but in particle physics today, novelty and testability are dominant and they should be. In contexts where there is no forthcoming slew of new data and where old ideas have failed to open the door to new discoveries, pursuing models that are easy to test or that propose radical solutions to classes of existing problems is the reasonable course of action.

References

- Aad, G. et al. (2023). Search for direct pair production of sleptons and charginos decaying to two leptons and neutralinos with mass splittings near the w-boson mass in $\sqrt{s} = 13$ tev pp collisions with the atlas detector. *Journal of High Energy Physics*, 2023(6).
- Abdallah, W. and Khalil, S. (2016). MSSM dark matter in light of higgs and LUX results. *Advances in High Energy Physics*, 2016:1–10.
- Abercrombie, D. et al. (2020). Dark matter benchmark models for early lhc run-2 searches: Report of the atlas/cms dark matter forum. *Physics of the Dark Universe*, 27:100371.
- Abreu, H. et al. (2023). First direct observation of collider neutrinos with faser at the lhc. *Physical Review Letters*, 131(3).
- Achinstein, P. (1993). How to defend a theory without testing it: Niels bohr and the "logic of pursuit". *Midwest Studies in Philosophy*, 18(1):90–120.
- Akerib, D. et al. (2017). Results from a search for dark matter in the complete lux exposure. *Physical Review Letters*, 118(2).
- Anderson, E. et al. (2023). Observation of the effect of gravity on the motion of antimatter. *Nature*, 621:716–722.
- Aprile, E. et al. (2018). Dark matter search results from a one ton-year exposure of xenon1t. *Physical Review Letters*, 121(11).
- Barceló, C., Liberati, S., and Visser, M. (2005). Analogue gravity. *Living Reviews in Relativity*, 8(1).
- Boyd, J. et al. (2023a). First Results from the Search for Dark Photons with the FASER Detector at the LHC. Technical report, CERN, Geneva.
- Boyd, N. M., Baerdemaeker, S. D., Heng, K., and Matarese, V., editors (2023b). *Philosophy of Astrophysics: Stars, Simulations, and the Struggle to Determine What is Out There*. Springer Verlag.
- Choi, K., Im, S. H., and Sub Shin, C. (2021). Recent Progress in the Physics of Axions and Axion-Like Particles. *Ann. Rev. Nucl. Part. Sci.*, 71:225–252.

- Douglas, H. (2009). *Science, Policy, and the Value-Free Ideal*. University of Pittsburgh Press, Pittsburgh, PA.
- Douglas, H. (2013). The value of cognitive values. *Philosophy of Science*, 80(5):796–806.
- Ellwanger, U., Hugonie, C., and Teixeira, A. M. (2010). The next-to-minimal supersymmetric standard model. *Physics Reports*, 496(1–2):1–77.
- Espinosa, J., Grojean, C., Panico, G., Pomarol, A., Pujolàs, O., and Servant, G. (2015). Cosmological higgs-axion interplay for a naturally small electroweak scale. *Physical Review Letters*, 115(25).
- Franklin, A. (1993). Discovery, pursuit, and justification. *Perspectives on Science*, 1(2):252–284.
- Franklin, A. (2016). *The Rise and Fall of the Fifth Force: Discovery, Pursuit, and Justification in Modern Physics*. Imprint: Springer, Cham.
- Gabrielse, G. (1989). Trapped antihydrogen for spectroscopy and gravitation studies: Is it possible? *Hyperfine Interact.*, 44(1-4):349–355.
- Germán, G., Gonzalez Quaglia, R., and Moran Colorado, A. (2023). Model independent bounds for the number of e-folds during the evolution of the universe. *Journal of Cosmology and Astroparticle Physics*, 2023(03):004.
- Graham, P. W., Kaplan, D. E., and Rajendran, S. (2015). Cosmological Relaxation of the Electroweak Scale. *Phys. Rev. Lett.*, 115(22):221801.
- Guchait, M. and Roy, A. (2020). Light Singlino Dark Matter at the LHC. *Phys. Rev. D*, 102(7):075023.
- Hajdukovic, D. S. (2012). Quantum vacuum and virtual gravitational dipoles: the solution to the dark energy problem? *Astrophysics and Space Science*, 339(1):1–5.
- Harlander, R. and Rosaler, J. (2019). Higgs naturalness and renormalized parameters. *Foundations of Physics*, 49(9):879–897.
- Hayrapetyan, A. et al. (2024). Search for heavy neutral Higgs bosons A and H in the $t\bar{t}Z$ channel in proton-proton collisions at 13 TeV.
- Hohmann, M. and Wohlfarth, M. N. R. (2010). Repulsive gravity model for dark energy. *Physical Review D*, 81(10).
- Hook, A. and Marques-Tavares, G. (2016). Relaxation from particle production. *Journal of High Energy Physics*, 2016(12).
- Jaeckel, J. and Ringwald, A. (2010). The low-energy frontier of particle physics. *Annual Review of Nuclear and Particle Science*, 60(1):405–437.

- Joergensen, M. D., De Roeck, A., Hachler, H. P., Hirt, A., Katre, A., Mermod, P., Milstead, D., and Sloan, T. (2012). Searching for magnetic monopoles trapped in accelerator material at the Large Hadron Collider.
- Kuhn, T. S. (1962). *The Structure of Scientific Revolution*. University of Chicago Press, Chicago, IL.
- Kuhn, T. S. (1974). Objectivity, value judgment, and theory choice. In *The Essential Tension: Selected Studies in the Scientific Tradition*. University of Chicago Press, Chicago.
- Lacey, H. (1999). *Is Science Value Free?* Routledge, New York, NY.
- Laudan, L. (1977). *Progress and its Problems: Toward a Theory of Scientific Growth*. University of California Press.
- Laudan, L. (1984). *Science and Values: The Aims of Science and their Role in Scientific Debate*. University of California Press, Berkely, CA.
- Longino, H. (1995). Gender, politics, and the theoretical virtues. *Feminism and Science*, 104(3):383–397.
- McKaughan, D. (2008). From ugly duckling to swan: C. s. peirce, abduction, and the pursuit of scientific theories. *Transactions of the Charles S. Peirce Society*, 44:446–468.
- McMullin, E. (1976). The fertility of theory and the unit for appraisal in science. In Cohen, R. S., Feyerabend, P. K., and Wartofsky, M., editors, *Essays in Memory of Imre Lakatos*, pages 395–432. Reidel.
- McMullin, E. (1985). Galilean idealization. *Studies in History and Philosophy of Science*, 16(3):26.
- Mou, Q., Wu, H., and Zheng, S. (2018). Light neutralino dark matter in the nmssm.
- Nelson, A. and Prescod-Weinstein, C. (2017). Relaxion: A landscape without anthropics. *Physical Review D*, 96(11).
- Noyes, H. P. and Starson, S. (1991). Discrete antigravity.
- Nyrup, R. (2015). How explanatory reasoning justifies pursuit: A peircean view of ibe. *Philosophy of Science*, 82(5):749–760.
- Nyrup, R. (2020). Of water drops and atomic nuclei: Analogies and pursuit worthiness in science. *British Journal for the Philosophy of Science*, 71(3):881–903.
- Peccei, R. D. and Quinn, H. R. (1977). CP conservation in the presence of pseudoparticles. *Phys. Rev. Lett.*, 38:1440–1443.

- Ritson, S. (2020). Probing novelty at the lhc: Heuristic appraisal of disruptive experimentation. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 69:1–11.
- Rühr, F. (2016). Prospects for bsm searches at the high-luminosity lhc with the atlas detector. *Nuclear and Particle Physics Proceedings*, 273-275:625–630. 37th International Conference on High Energy Physics (ICHEP).
- Schmidt, B. (2016). The high-luminosity upgrade of the lhc: Physics and technology challenges for the accelerator and the experiments. *Journal of Physics: Conference Series*, 706(2):022002.
- Shaw, J. (2022). On the very idea of pursuitworthiness. *Studies in History and Philosophy of Science Part A*, 91(C):103–112.
- 't Hooft, G. (1980). *Naturalness, Chiral Symmetry, and Spontaneous Chiral Symmetry Breaking*, pages 135–157. Springer US, Boston, MA.
- Villata, M. (2011). Cpt symmetry and antimatter gravity in general relativity. *Europhysics Letters*, 94(2):20001.
- Villata, M. (2012). “dark energy” in the local void. *Astrophysics and Space Science*, 339(1):7–12.
- Wang, S.-J. (2019). Paper-boat relaxion. *Physical Review D*, 99(9).
- Wilczek, F. (1978). Problem of strong p and t invariance in the presence of instantons. *Phys. Rev. Lett.*, 40:279–282.
- Williams, S. L. (2024). Searches for supersymmetry and additional Higgses. Technical report, CERN, Geneva.
- You, T. (2017). A dynamical weak scale from inflation. *Journal of Cosmology and Astroparticle Physics*, 2017(09):019–019.