Robustness and dark matter observation

Abstract

Current cosmological observations place little constraints on the nature of dark matter, allowing the development of a large number of viable models and various methods for probing their properties. At first glance, this variety of models and methods provides ideal grounds for the employment of robustness arguments in dark matter research. The aim of this article is to examine the extent to which such arguments can be used to overcome various methodological and theoretical challenges. The main conclusion is that while robustness arguments have a limited scope in the context of dark matter research, they can still be used for increasing the scientists' confidence about the properties of specific models.

1 Introduction

According to the current received view in cosmology, the ΛCDM model, more than 95% of the observable universe is ‘dark’, consisting primarily of dark energy (≈68%) and dark matter (≈27%). However, despite the fact that the systematic research of dark matter has a history of almost fifty years, its exact nature still remains elusive mainly due to the severe underdetermination of viable dark matter models by the available evidence. If it exists, dark matter could be *anything* – a single type of particle or a collection of particles in a dark sector, excitations of a superfluid Bose-Einstein condensate, neutron stars, and even primordial black holes – insofar as it satisfies a minimum set of constraints based on current cosmological observations.

Martens (2022) aptly describes this minimum set of constraints as the ‘thin common core concept of dark matter’: assuming general relativity, dark matter is a massive field that contributes ≈ 27% to the current total cosmic mass-energy budget, it primarily interacts with baryonic matter via the gravitational force, and if it is a particle, its mass is expected to be between $10^{-3} - 10^{7}$ eV. If one is willing to give up the standard gravitational laws of physics, the thin common core becomes even thinner: dark matter is some sort of ‘stuff’ that contributes ≈ 27% to the total cosmic mass-energy budget or acts as if it does so, and is responsible for the observation of various ‘dark phenomena’ related to structure formation, clusters and galaxies.[1] Hence, depending on whether one wishes to maintain general relativity at

[1]The term ‘dark phenomena’ is borrowed from Martens & Lehmkuhl (2020a) and refers to the various astrophysical phenomena that either contradict the gravitational laws of general relativity or require the postulation of additional ‘invisible’ dark matter that causes the formation of some large scale cosmological structures due to its grav-

low accelerations or is willing to introduce a new modified theory of gravity, the 'thin common core concept of dark matter' embodies the minimum set of constraints that every candidate model of dark matter needs to satisfy in order to be compatible with current cosmological evidence. The thin common core concept of dark matter is therefore 'common' precisely because it is shared by all possible models for the nature of dark matter, and it is 'thin' because it places very little constraints on the exact nature of dark matter. At the same time, it also reflects the progress in the field of dark matter research in that the enrichment of this common core concept via the derivation of model-independent properties of dark matter automatically leads to a stricter set of constraints, thus leading to the reduction of viable candidate models.

As expected, the fact that the nature of dark matter is underdetermined by the available evidence has naturally led to the development of a large number of diverse, but nonetheless viable, models for dark matter. This proliferation of models has in turn led to the development of a variety of methods for the possible observation of dark matter and the probing of its properties. At a first glance, the variety of models and methods of observation in dark matter research seems to provide a fruitful ground for robustness arguments. In its most general formulation, robustness can be defined as 'the state in which a hypothesis is supported by evidence from multiple techniques with independent background assumptions' (Stegenga, 2009). Hence, given that different models for dark matter embody, in one way or another, different hypotheses about the nature of dark matter, the fact that there are different methods

itational pull. Examples of such phenomena are the mass discrepancies in the Coma cluster and the flat rotation curves of nearby galaxies. For a nice review of the observational evidence for dark matter and dark energy based on dark phenomena see Jacquart (2021).

for probing the properties of dark matter can be used for the validation or the invalidation of these hypotheses.

The central aim of the present article is to examine the extent to which robustness arguments from the variability of experiments can be used within the scientific context of dark matter research in order to overcome various methodological and theoretical challenges. The main conclusion is that while these arguments have a limited scope, they can still be used in dark matter research for increasing the scientists' confidence about the constraints on the parameter space of certain model scenarios. In particular, it will be argued that robustness arguments cannot be used for establishing which model (or models) actually captures the true particle nature of dark matter, but they can – and should – be used to increase the reliability of the results concerning the sub-models of model scenarios that can be probed by more than one method.

The present article falls within the recently rising literature in the philosophy of dark matter (e.g. Vanderburgh 2003, 2014; Kosso 2013; Massimi 2018; Weisberg et al. 2018; de Swart 2020; Martens & Lehmkuhl 2020a, 2020b; Smeenk 2020; De Baerdemaeker 2021, Jacquart 2021, and De Baerdemaeker & Boyd 2021) and its broader aim is to explore an existing methodological challenge in cosmology regarding the proliferation of viable models for the nature of dark matter, by embedding it in the current philosophical literature on robustness. Although the main conclusion partly concerns the limits of robustness arguments in dark matter research, the present study nonetheless highlights the need of a common ground of reference in dark matter research for the integration of results from different methods, and should not be seen as introducing any kind of pessimism or scepticism about dark matter research. If anything, it showcases how a collective and rigorous conceptual and methodological analysis of the scientific practice in dark matter research by philosophers of science and scientists could eventually shed light on various methodological conundrums in this field.

These observations will become clearer by elucidating the existing methods of dark matter observation and providing a possible strategy to evaluate their epistemic strength. In what follows, a brief discussion of the literature on robustness will be presented in order to identify the type of robustness analysis that is most relevant for the purposes of this article (Section 2). The discussion will follow with a presentation of the viable models of particle dark matter and the possible methods for probing their properties according to the existing scientific literature (Section 3). In Section 4, an analysis of the three key epistemic virtues of Informativeness, Model Sensitivity and Reliability of the different methods of observation will be given, in order to support the main argument about the limits of robustness analysis from the variability of experiments which will follow in Section 5.

2 Robustness arguments from the variability of experiments

Robustness arguments are widely used from philosophers of science in various forms.[2] For the purposes of this article, *robustness from the variability of experiments* will be understood mainly in terms of what Woodward (2006, pp.233-5) calls 'measurement

[2]See for instance Franklin (1989), Weisberg (2006), Woodward (2006), Kuorikoski, Lehtinen & Marchionni (2010), Parker (2011), Odenbaugh & Alexandrova (2011), Lloyd (2015), Eronen (2015), Basso (2017), Lisciandra (2017), and Schupbach (2018).

robustness' and what Franklin (1989) calls the 'variability of experiments'. The rationale behind measurement robustness is that if different measurement procedures of a physical quantity that are in some sense independent of each other produce nearly the same result, then the result is said to be *robust* and can be used as grounds for increasing our confidence that the quantity has been measured accurately. This is because it is very unlikely that each procedure is subject to exactly the same kinds of error that would give rise to the same result, and so we have good reasons to believe that the result of the different measurements is reliable.

Woodward's characterisation of measurement robustness is closely related to what Franklin (1989, Ch.6) describes as the *variability of experiments* in his discussion of possible strategies to justify the reliability of experimental results in high-energy physics. For Franklin, a potential agreement between the experimental results of two or more different experimental methods automatically increases our confidence in the reliability of the results (and consequently of the methods themselves), since 'it would be a preposterous coincidence if the same patterns were produced by two totally different kinds of physical systems'.[3] In a similar spirit to Woodward and Franklin, Stegenga (2009) provides a wider definition of robustness as 'the state in which a hypothesis is supported by evidence from multiple techniques with independent background assumptions' (*ibid.*, p.651). Again, the main idea behind Stegenga's definition is that hypotheses are better corroborated by results coming from multiple and independent experimental methods compared to hypotheses that are supported

[3]Hacking (1983), as quoted by Franklin (1989, p.166). See also Franklin and Howson (1984) for a Bayesian argument showing how the variability of experiments increases the confidence of the scientists to the results in a higher degree compared to the repetition of the same experiment.

by the results of a single experimental method.

Given the above definitions of robustness, and keeping in mind the experimental practice in dark matter research which is largely about constraining the parameter space of dark matter models, what we shall call *robustness from the variability of experiments* will be understood in the following way:

> Compared to hypotheses supported by only one experimental method, a hypothesis H_{DM} concerning the nature of dark matter is more robust and more reliable – and thus enjoys a higher degree of confidence from the scientific community – if it is supported by the results of two or more experimental methods. Insofar as different and independent methods provide concordant results, the agreement of results also increases the reliability of these methods.

A hypothesis H_{DM} concerning the nature of dark matter can be either about the value of a quantity of a physical property of dark matter, e.g. the cosmic relic density of dark matter, or – as it is more often in dark matter research – about a constraint (or set of constraints) on the parameter space of a dark matter model regarding, for instance, the mass of a weakly interacting dark matter particle, the couplings of mediator particles to dark matter and so on. On the face of it, the situation in dark matter research in which one finds a number of diverse and competing models and different methods of constraining the parameter spaces of those models, provides an ideal set-up for the employment of robustness arguments. The question that arises – and which is the main focus of this article – is the following: 'How exactly can we take advantage of the fact that there are different methods available for probing the properties of dark matter in order to increase our understanding and confidence about the properties of dark matter?'. Prima facie, it seems that there are two possible ways of employing robustness arguments in dark matter research each one corresponding to the following two aims:

[1] To reduce the number of viable models in order to decide which of the existing model, or models, actually captures the true nature of dark matter

[2] To increase our confidence about the properties of a specific model of dark matter

The first aim is ultimately the most important. At the end of the day, what we are interested in is to figure out which of the available candidate models of dark matter actually captures the nature of dark matter. Hence, given that each model incorporates different hypotheses of dark matter, those hypotheses that will ultimately be supported by the results of more than one method, will become more robust compared to those hypotheses that are supported only by one method – or even worse by no method at all.

The second aim is more modest. Instead of comparing hypotheses from different models of dark matter, the objective is to employ the results of as many methods as possible in order to increase our confidence about the allowed parameter space of a particular model. To put it simply, this would amount for instance in constraining the mass of a WIMP particle, not only by using the results from direct searches, but also by taking into consideration the results from indirect searches.

In what follows, it will be argued that robustness arguments from the variability of experiments unfortunately do not succeed in offering a possible strategy for the accomplishment of the first – and most important – aim, but they do, nonetheless, succeed in fulfilling the second aim. That is, robustness arguments from the variability of experiments are unable to provide stronger grounds for deciding between competing models of dark matter, but at the same time they can – and should – be employed for the fulfilment of the second aim stated above, namely, for increasing the reliability of results concerning the allowed parameter space of specific models via complementarity and compatibility studies.

In order to support this argument, we shall first discuss the current situation in dark matter research by presenting some of the competing candidate models of dark matter and the possible methods for probing their parameter space.

3 Many models, four methods

As already mentioned, the fact that the nature of dark matter is underdetermined by the available cosmological evidence has naturally led to the development of a large number of diverse models regarding the nature of dark matter. Some of these models concern the particle nature of dark matter, while others describe the large scale structure of dark matter in terms of Massive Compact Halo Objects (MACHO's) which may or may not compose of baryonic matter (e.g. primordial black holes and neutron stars). For the purposes of this article, we shall restrict ourselves in models regarding the particle nature of dark matter for reasons of consistency and uniformity, as these models are already abundant enough to be used for robustness arguments, and are easier to compare with each other.

Arguably, the most popular candidate for the particle nature of dark matter comes from the models of dark matter as Weakly Interacting Massive Particles (WIMPs)

(Roszkowski et al. 2018), followed, perhaps, by models of dark matter as axions (Duffy & Van Bibber, 2009) and sterile neutrinos (Boyarski et al. 2019). Roughly speaking, 'WIMP' is an umbrella term referring to a large class of models including particles that interact with ordinary baryonic dark matter via gravity and a non-vanishing force which is either weaker or at least as weak as the weak nuclear force. Axions are hypothetical elementary particles that were initially introduced as a possible solution to the strong CP problem in particle physics, and finally, sterile neutrinos are similar to standard model neutrinos, except that they have right-handed chirality, and do not interact with baryonic matter via any of the fundamental interactions apart from gravity.

However, these three models do not exhaust, by any means, the available options for the nature of dark matter particles. As already mentioned in the introduction, assuming standard gravitational laws, the minimum set of constraints as stated in the thin common core concept of dark matter also allows for the possibility that dark matter particles can be made of collisionless particles, strictly self-interacting particles, or even by an entire hidden dark sector including a number of fundamental dark particles.[4] Given that each one of these models comes in various versions

[4]Just to provide a sense of how long the list of models of dark matter particles is, some further proposed dark matter candidates not mentioned above are: gravitinos, axinos, light-scalar dark matter, dark matter from little Higgs models, wimpzillas, Q-balls, mirror particles, CHArged Massive Particles(CHAMPs), Strongly Interacting Massive Particles (SIMPs), D-matter, cryptons, superweakly interacting dark matter, brane world dark matter, heavy fourth generation neutrinos etc. For a generic review of dark matter research containing useful information on various dark matter candidates see Bertone et al. (2005). For a more detailed review of dark matter candidates see

(especially in the case of WIMPS), for the remaining of this paper we shall refer to these broad classes of models as the possible *model scenarios* of dark matter, whereas the various specific examples of a particular model scenario – e.g. the LSP and LKP examples in WIMPs – will be referred to as the *sub-models* of a model scenario.[5]

As one might expect, the abundance and diversity of viable model scenarios for dark matter has naturally led to the development of a variety of methods for the possible observation of dark matter, since each model scenario is built on a number of different assumptions and requires different experimental setups to be tested. Observation should be understood here in the broad scientific meaning as an act of obtaining information about the properties of one or more physical entities, via any kind of interaction which involves the communication of information from the target of the observation to the observer. This definition closely follows Shapere's (1982) understanding of observation as a subspecies of interaction between two physical systems where information is transmitted from one system to another. In a sense, observation is thus a binary directional relation between the target system(s) being observed – i.e. the physical entities about which the scientific community is gathering information – and an observer. The observer can be either the scientist or any other scientific device which can be used for the collection of information about the properties of a physical entity, which will eventually be conveyed to the scientist via a series of data.

Feng (2010).

[5]The Lightest Supersymmetric Particle (LSP – also known as the neutralino) and the Lightest Kaluza-Klein Particle (LKP) are probably the two most well-studied sub-models for WIMPs, coming from supersymmetry and theories of extra dimensions respectively.

According to current scientific practice, dark matter can be observed via four different methods based on the physical phenomena on which they rely: (a) via cosmological observations (b) directly (c) indirectly and (d) in collider searches. As a precaution, it should be noted that regarding methods (b), (c) and (d), the choice of labelling them as direct, indirect and collider based, reflects the relevant terminology in current scientific literature which clearly distinguishes between these three different methods based on their methodology. However, while the methodology and the fundamental principles of each method are – as we shall see – clearly distinct, whether these methods are indeed direct or indirect, and the interesting question of whether collider searches provide a possible direct or indirect way of observing dark matter, is a separate and very interesting issue that lies outside the scope of this article, but nonetheless deserves to be studied further on its own merit. Finally, let us also note that while the observation of dark matter based on cosmological observations is often considered as providing strong evidence for its existence, the remaining three methods have not provided any significant positive results so far, other than various set of constraints on the parameter space of some models. Below is a brief description of the four possible methods of dark matter observation.

Cosmological Observations: Cosmological observations of dark matter can be divided in two categories: (i) precision measurements of cosmological observables and (ii) observations based on gravitational effects of dark matter on large scale structures. Precision measurements are typically related to the non-gravitational effects of dark matter on large-scale structures and the thermal history of the universe. These effects are captured by various cosmological observables related to the Cosmic Microwave Background (CMB) such as its spectral distortions, polarization, and temperature anisotropies. Other cosmological observables providing information about the nature of dark matter are related to data from distance measurements of Type Ia Supernovae (SN Ia) and Baryon Acoustic Oscillations (BAO). The combination of datasets from CMB observations with BAO and SN Ia measurements provides a precise estimate of the average matter density in the Universe, as well as a tight constraint on the mass of the dark matter particle (Aghanim et al. 2018). Cosmological observations of dark matter via its gravitational effects are typically related – as the name suggests – to the purely gravitational effects of dark matter captured, for instance, by measurements of mass discrepancies in galaxy clusters, galaxy rotation curves, and gravitational lensing.[6]

Direct Searches: Direct detection methods are earth-based experiments based on the interaction of dark matter particles with ordinary baryonic matter. The basic idea behind direct searches is that if the galaxy is full of dark matter particles (e.g WIMPS, axions etc.) that interact weakly with baryonic matter, then a significant amount of them will travel through the Earth, enabling us to search for the interaction of these particles with standard model particles by recording the recoil energy of nuclei as dark matter particles scatter off them. Since the interaction of dark matter particles with the nuclei is expected, by definition, to be extremely weak, direct search experiments take place in ultra-sensitive low-background experiments that are often placed well below the Earth's surface in order to block out spurious particles.[7]

Indirect Searches: Indirect searches for dark matter are based on the astronomical

[6]For more detailed reviews on cosmological observations of dark matter see Lukovic et al. (2014) and Gluscevic et al. (2019). For a review of dark matter gravitational lensing see Massey et al. (2010).

[7]See Schumann (2019) for a review of direct searches on dark matter.

observation of standard model particles that are most likely to be the products of the decay or annihilation of dark matter in the Universe. These searches are based on the assumption that the final states of a dark matter annihilation/decay are either standard model particles of any kind (insofar as they are kinematically accessible), or unknown particles which then decay to standard model particles. Current experiments for the indirect detection of dark matter are mainly focused on the detection of three different products: (i) gamma-rays, (ii) neutrinos and (iii) cosmic rays. These three types of radiation are used for the indirect detection of dark matter for a number of theoretical and practical reasons, such as the fact that the mass scale of WIMPs in the most promising models implies that a large fraction of the generated emission from dark matter annihilation/decay ends up in gamma-ray energies.[8]

Collider searches: The main idea behind collider searches of dark matter is that high-energy collisions of standard model particles, such as the ones taking place at the Large Hadron Collider (LHC) at CERN, can lead to the direct pair-production of dark matter particles. Moreover, if we are lucky enough and the dark matter particle mass is comparable to the electroweak scale, the LHC is also expected to be able to produce large quantities of dark matter particles via the decays of heavier states that are instantaneously created in high-energy proton-proton collisions. Collider searches are thus based on the possible interactions of dark matter particles with standard model quarks and gluons which, in either case, lead to a missing energy (or missing momentum) signature in the final states due to the lack of interaction between dark matter and the material of the detectors, as well as to the detection of unexpected particle products. There are five main processes in which dark matter pair-production could occur at the LHC: (i) mono-jet (ii) mono-V (iii) mono-Higgs (iv) dark matter

[8]See Gaskins (2016) for a detailed review of indirect searches of dark matter.

with top quarks and (v) invisible Higgs decays. In mono-jet processes the dark matter particles are produced in association with one or more QCD jets, in mono-V they are produced in association with a vector boson and so on.[9]

The four different methods for dark matter observation described above and the relevant underlying physical phenomena are summarised in Fig.1, with precision measurements of cosmological observables and dark matter observation via its gravitational effects illustrated separately for reasons of clarity.[10] Each one of these methods can be used to probe and constrain various properties of dark matter, and hence, they are, in principle, valuable tools in the hands of scientists for performing various cross-checks about the properties of dark matter. Ultimately, the question we are interested in exploring is whether this plurality of methods for probing dark matter can be exploited for the use of robustness arguments from the variability of experiments in order to [1] reduce the number of viable models for the nature of dark matter , and [2] increase our confidence about the possible properties of dark matter as captured by specific model scenarios for the particle nature of dark matter. As already mentioned in the introduction, robustness arguments cannot be employed for fulfilling aim [1], but can – and should – be employed for fulfilling aim [2]. To see why, it is helpful to first examine some of the epistemic virtues of the various methods for dark matter observation in terms of their *informativeness*, *model sensitivity*, and *reliability*.

[9]For a detailed review on LHC dark matter searches see Kahlhoefer (2017).

[10]It should be noted however, that this rather helpful depiction of the different methods of dark matter observation is not exhaustive of all dark matter research, especially with respect to the level of the phenomena that are responsible for each type of observation.

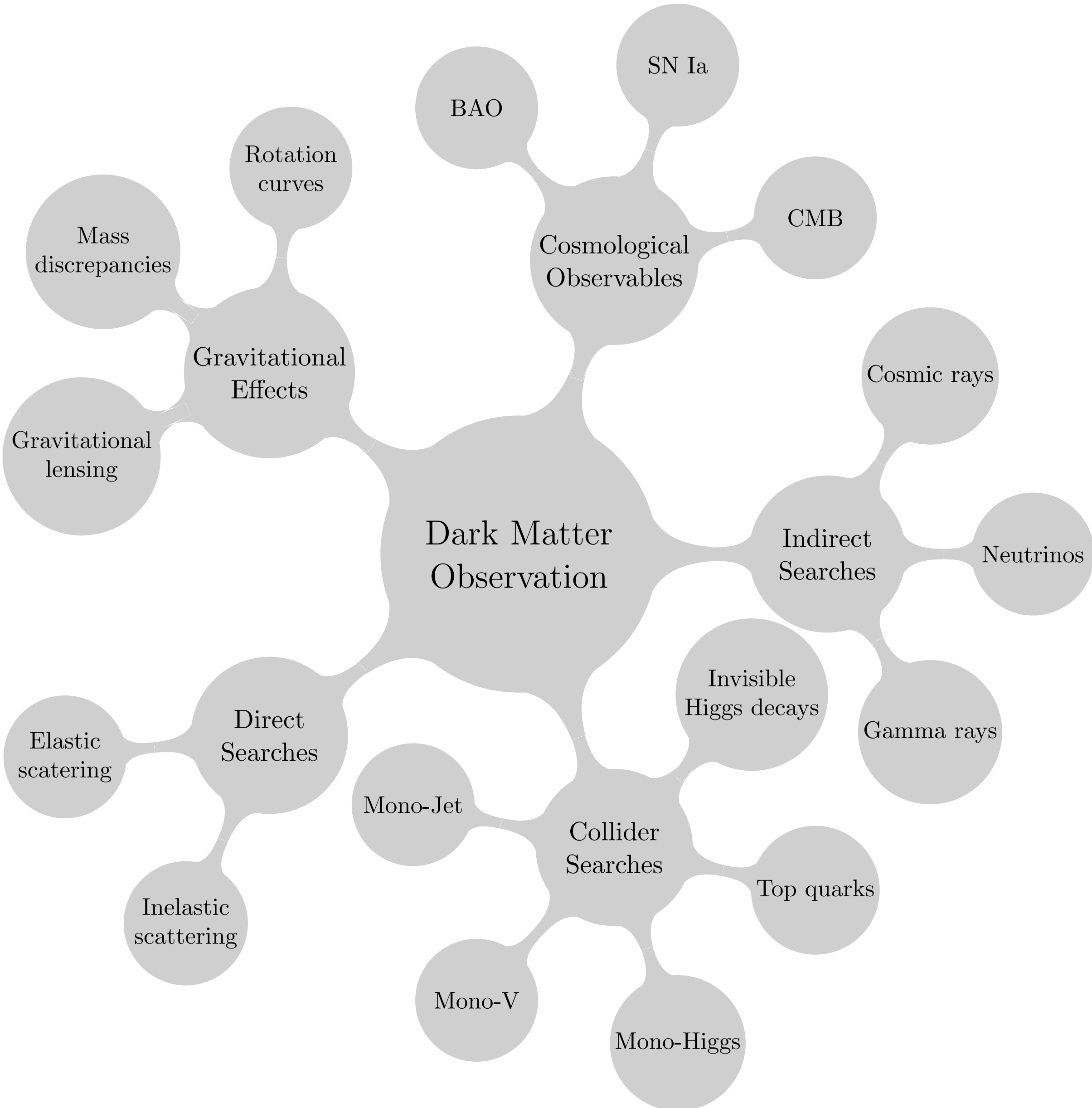

Figure 1: Dark matter observation: The first level of the diagram shows the possible methods of dark matter observation (i) gravitational effects and precision measurements of cosmological observables, (ii) direct searches (iii) indirect searches and (iv) collider searches. The second level of the diagram shows the various phenomena responsible for each type of observation.

4 Informativeness, Model Sensitivity, and Reliability

A necessary condition for employing robustness arguments from the variability of experiments in order to ensure the reliability of results from the various methods of dark matter observation, is that these methods are actually probing and constraining the same quantities. Siska De Baerdemaeker highlights this point in her discussion of the implications of methodological pluralism in dark matter research by noting that 'a crucial condition for measurement robustness is that the *same parameter or quantity* is being measured by the different experiments' (2021, p.140, emphasis added). She then suggests that the common core within the different experiments is provided by the fixed definition of the target of the observation. In her own words, 'the definition of the target system remains fixed under the employment of different methods. It provides, a common core that might underlie multiple methods attempting to probe the same target. Without this agreement on the common core, it is not obvious that methods that detect different phenomena are still probing the same target system and that measurement robustness arguments therefore apply' (*ibid.*).

While it is true that all methods discussed above share the common goal of probing the physics of dark matter and the results they provide do indeed concern the properties of dark matter as De Beardemaeker notes, what I aim to show is that the thin definition of dark matter alone does not suffice to ensure that the agreement of results between different methods can be achieved. In order to establish that the results of two different methods are reliable via robustness arguments, we first need to ensure that the extracted information concerns the properties of *the same dark* *matter models* and relies on the *same assumptions*. However, as will be shown, this is rarely the case in dark matter research. The concepts of *informativeness* and *model sensitivity* will help us clarify this point.

Informativeness. The informativeness of a method concerns its ability to provide information on a number of different *properties* of dark matter either by providing specific values for these properties or – as it is often the case in dark matter research – by constraining the parameter space of a model. The properties of dark matter for which one can derive information from a particular method could be the cosmic relic density of dark matter, the mass of a dark matter particle, the self-interaction cross section, the cross-section between dark matter and baryonic matter and so on. Assessing the informativeness of a method is crucial for the employment of robustness arguments since – as already mentioned – such an analysis requires that the involved methods provide information about the same property. However, in practice, the nature and the amount of information about the properties of dark matter that can be inferred from an observational method is typically determined by a number of factors which make the comparison of the information from different methods a much more complicated – if not impossible – process.

This is because even in the cases where two or more different methods are ostensibly providing information about the same property (e.g. constraints on the mass of WIMPs), this information is conditional on a number of factors which vary significantly in each experiment. That is, in order to extrapolate a meaningful result from an experimental process, numerous assumptions need to be implemented both during the construction of the experimental apparatus and the analysis of data. Hence, ensuring that two different methods are probing the same quantity/parameter and provide concordant results requires taking into consideration the effects of these assumptions in the results of the experiment. In the case of dark matter observation, these factors can be grouped into three different categories: (i) the experimental models of the experiment (ii) the extrapolating assumptions and (iii) the model scenarios of dark matter and their sub-models.[11]

Experimental models cover a broad category of models referring to every possible modelling activity that facilitates the construction of an experiment and the completion of a measurement process. For instance, in collider experiments this includes the various competing physical models for calculating the interactions of the produced particles with the different parts of the detector, while in direct searches such models would describe the ionization process of the liquid detectors and the interactions of photons with photomultiplier tubes. Experimental models also cover the required modelling for the simulation of large-scale dark matter formation that is often necessary for making measurements of dark matter properties.[12] As expected, the interpretation of data from an experiment and the further extrapolation of results strongly depends on the adopted experimental models in the various stages of the experiment, since the implementation of different models would give rise to a different set of data. The final effects of the adopted experimental models are often implemented in the results in the form of uncertainties, although it is also possible that a number of different results is derived, based on the selection of a specific

[11]It should be noted that these three sets of assumptions are not always entirely independent of each other. Rather, the above categorisation is a useful conceptual tool in order to illustrate how the extraction of information from an observational method is conditional on a number of assumptions which differ in scope.

[12]See Gueguen (2020) and Smeenk & Gallagher (2020) for some nice discussions on the use of simulation in cosmology.

combination of models.[13]

Extrapolating assumptions are those assumptions needed for carrying out the required calculations for deriving information about the properties of dark matter, *after the acquirement of data from an experiment.* A profound example of a set of extrapolating assumptions comes from the interpretation of experimental results in direct searches for dark matter. The results from direct searches are necessarily extrapolated on the basis of some standard simplified assumptions such as the local density ρ_0 of dark matter, an isothermal profile of dark matter density and a Maxwell-Boltzmann velocity distribution, incorporated in what is known as the Standard Halo Model. In the absence of knowledge about the exact properties of dark matter in the local region, the introduction of these assumptions is essential for carrying out the necessary calculations for the derivation of constraints on various properties of dark matter.

Finally, the information an observational method yields, also depends on the model scenario under consideration and its various sub-models. For instance, indirect searches are based on the fundamental assumption that dark matter is self interacting and its self-annihilation produces standard model particles in the form of gamma rays, neutrinos and cosmic rays. Similarly, current collider experiments are only able to provide constraints based on the assumption that dark matter consists of WIMPs that can be produced in high-energy collisions. However, as already mentioned, the model scenario for WIMPs covers a broad class of specific models and depending on which sub-model is taken into consideration, a method can produce more than one set of results, which means that *the derived constraints on a dark matter candidate*

[13]For more detailed discussions on the impact of experimental models to the final results of an experiment see Mayo (1996, Ch.5), Staley (2020) and Antoniou (2021).

from various experiments are highly model dependent in nature. In other words, depending on which model provides the assumptions needed for the interpretation of data, the experimental data from an observational method often provide different sets of constraints for different models of dark matter.

Model Sensitivity. The fact that the extrapolation of results from a particular method crucially depends on model related assumptions implies that the extracted information from a particular method of observation is most of the time *model specific*. For instance, mass constraints on dark matter particles from collider experiments rely on the assumption that dark matter particles are WIMPs that interact weakly with baryonic matter, since if dark matter consists from purely collisionless particles such constraints cannot be derived from particle collisions. Model sensitivity is the epistemic virtue that concerns the ability of a method to provide model specific information on a range of different model scenarios of dark matter and their sub-models, and it is thus assessed with respect to the number of model scenarios (and their sub-models) for which a method can determine their properties and constrain their parameter space. More model sensitive methods provide information about more dark matter models and vice versa. Hence, although informativeness and model sensitivity are closely related, their difference lies on the fact that the former concerns the number of *properties* of dark matter for which a method yields information (mass, cosmic relic density, cross-sections etc.), and the latter concerns the number of *models* of dark matter for which a method is able to provide information (WIMPs, axions, sterile neutrinos etc.).

Arguably, a good example of a highly model sensitive method comes from the indirect searches of dark matter, since these experiments are able to provide constraints on a number of different model scenarios including self-interacting dark matter (SIDM),

	Cosmological	Direct	Indirect	Collider
Collisionless DM	✓			
SIDM	✓		✓	
WIMPs	✓	✓	✓	✓
Sterile Neutrinos			✓	
Axions		✓	✓	
Hidden/Complex DM		✓	✓	✓
Light Gravitinos				✓

Table 1: Model sensitivity of the various methods of dark matter observation

WIMPs, sterile neutrinos and models of complex dark matter. On the other hand, the cosmological observation of dark matter via precision measurements of CMB observables is considered to be highly model insensitive since the derived cosmic relic density from this method is insensitive to the specific models scenarios of dark matter. In other words, measurements of temperature anisotropies on the CMB provide the current relic density and the stability of dark matter on the cosmological scale, but insofar as it is possible that dark matter is made from more than one component, this information places no model-specific constraints in the relevant models for the particle nature of dark matter. It is only under the additional assumptions that dark matter consists of a weakly interacting massive particle, that CMB measurements are able to provide constraints on the dark matter-proton scattering section (Glucevic et al. 2019), as well as on the mass and the dark matter annihilation rates of WIMPs (Natarajan 2013).

The upshot is that when taking into consideration the informativeness of a particular method of dark matter observation it is important to highlight the degree to which the constraints imposed by its results are tied to specific models. Table 1 illustrates a tentative depiction of the model sensitivity of each method of observation with respect to various model scenarios of dark matter at the present time. The check mark indicates that a method can provide information for at least one parameter of the relevant model, but it should be noted that the situation might well change in the future. For instance, axion-like particles are expected to be searched for in next generation colliders (Bauer 2019) and there is also a possibility of directly detecting collisionless dark matter via its gravitational effects. The table also illustrates the fact that WIMPs can be probed by all four methods of dark matter observations, which partly explains their increasing popularity.

Reliability. The consideration of the informativeness and model sensitivity of each method provides a good way of evaluating the nature of information that can be extracted and the various models that can be constrained by each method. What remains to be seen is how these two virtues relate to the reliability of each method, since ultimately, what is of utmost importance for achieving the necessary progress in dark matter research is whether the extracted information from a method is reliable and can be used to enrich the common core concept of dark matter. In principle, the reliability of a method concerns its ability to consistently provide robust and infallible results that accurately describe the physical world. In practice, ensuring that a method provides reliable results amounts to ensuring (i) that the material part of the relevant experimental equipment has limited and controllable flaws, and (ii) that the underlying theory and assumptions that are necessary for the derivation of results are correct. In the case of dark matter research, a perfectly reliable method would thus be one whose results consistently have zero deviation from the actual values of the physical properties of dark matter precisely because the material equipment works properly and the underlying theory and assumptions are correct. The crucial question however, is how we can ensure that this is the case, and it is here that robustness arguments become relevant.

As already discussed in Section 2, robustness arguments from the variability of experiments are inferences to the best explanation about the agreement of results obtained by two or more different experimental methods. Such concordant results can be used in order to *both increase the confidence of the scientific community that a result is robust and that the methods producing the results are reliable.* The main idea is that the best explanation for the fact that two distinct and independent types of experiments provide concordant results is that the material equipment is working properly and the common underlying theory in these methods is correct, otherwise as Hacking aptly noted 'it would be a preposterous coincidence' if the two methods produced the same patterns as a result of the same kind of errors. The value of robustness arguments from the variability of experiments thus lies on the fact that they provide a rational argument for increasing the confidence of the scientific community both on the reliability of a method, and, consequently, the 'truthfulness' of a given hypothesis. Reliable methods provide reliable results, and the more reliable methods support a given hypothesis, the more robust the hypothesis becomes.

In Section 2 we noted that, prima facie, robustness arguments can be used in dark matter research [1] to reduce the number of viable models for the particle nature of dark matter and [2] to increase the scientists confidence about the properties of a specific model of dark matter. Ideally, the first aim would be achieved if two or more methods support a specific hypothesis about a property of dark matter, e.g. a set of model-independent constraints in the mass of dark matter particles that excludes certain models. Such a hypothesis, if supported by the results of two or more methods, would be significantly more robust compared to any competing hypotheses supported by only one method, and would provide strong grounds to believe that the hypothesis regarding the mass of a dark mater particle is correct. The second and more modest aim, requires that model specific constraints enjoy a higher degree of confidence if they are derived on the basis of two or more methods.

In the next and final section we shall use the concepts of informativeness, model sensitivity and reliability to argue that the use of robustness arguments is not a particularly useful strategy for the accomplishment of the first aim, however, it can be used for achieving the second aim, that is, for deriving more reliable model specific constraints on the viable models of dark matter.

5 The limits and the value of robustness arguments

As already made salient, a necessary condition for the employment of robustness arguments from the variability of experiments is that the results of these methods concern the *same parameter*, i.e. they are concordant. Stegenga (2009) argues that while robustness is a valuable epistemic guide in 'ideal epistemic circumstances', when it comes to real scientific practice it faces important limitations. The main problem with robustness arguments, according to Stegenga, is that, in practice, most of the time multiple and independent experimental techniques provide results that are inconsistent (i.e. one method suggests x and another method suggests $\neg x$) or incongruent (i.e. one method suggests x and another method suggests y), and hence, it is not clear what kind of epistemic support is provided to the relevant hypotheses. While this seems to be the case in dark matter research as well, it will be argued that even in cases where multiple methods ostensibly provide concordant results, the dependence of such results on a number of factors concerning the introduced assumptions in the experiments and the dependence of the experiments on specific dark matter models, makes the task of establishing that the results of different experiments are actually the same extremely difficult. Hence, robustness arguments from the variability of experiments in dark matter research are limited in scope, not only because the results of multiple methods rarely agree, as Stegenga notes, but also because even in cases of a potential agreement on the surface, establishing the concordance of these results is challenging.

The first and most straightforward complication in establishing that different methods provide concordant results comes from the fact that, as shown in our discussion regarding model sensitivity, compared to the huge variety of models, there is relatively little overlap between different methods that are able to probe the same models, let alone the same parameters of these models. To put it simply, for a large number of viable candidate models for dark matter, obtaining results from more than one method is just not possible and thus, robustness arguments cannot be used for the corroboration of hypotheses concerning those models. This is the case, for instance, with sterile neutrinos since, as Table 1 shows, the former can only be probed via direct searches, and the latter can only be probed via indirect searches.

Moreover, even in the overlapping cases where certain models can be probed by two or more methods, most of the time the extracted information concerns fairly small and disjoint areas of the relevant parameter space, making the comparison of results impossible. This is the case, for instance, with WIMPs, which van be better probed in the low mass regions with colliders, whereas models with heavier particles are better constrained by direct searches. This complication reflects Stegenga's (2009) observations on the limits of robustness arguments based on the fact that scientists do not always have multiple techniques to generate common results, and that, often, the results obtained by multiple techniques are incongruent.[14]

[14]Regarding Stegenga's note on inconsistency, a well-known example of incon-

An additional complication however, concerns the fact that even in cases where multiple techniques are ostensibly able to provide concordant results (e.g. constraining the same parameter space for the mass of a dark matter particle), ensuring that these results are indeed in agreement is a very difficult – if not impossible – task. The main source of this difficulty relates to the informativeness and model sensitivity of each method and concerns the fact that the extracted information from the various different methods is almost always model specific and conditional on the experimental models, the extrapolating assumptions and the model dependence of each method. Hence, given that each model comes with its own additional assumptions, constraining, for instance, the mass of WIMPs with method A, is not the same as constraining the mass of axions with method B, despite the fact that both methods provide information about 'the mass of a possible dark matter particle'. To ensure the robustness of a hypothesis H_{DM} concerning a property of dark matter that can be included in the common core concept by taking advantage of the variability of methods, such a hypothesis must be model-independent. This is rarely the case, however. The only model-independent properties of dark matter that are currently available concern the cosmic relic density of dark matter, the upper and lower limits of its particle mass, and the fact that dark matter is cold and dark that are already encompassed in the thin common core concept of dark matter assuming

---

sistent results in dark matter research comes from the controversial result of the DAMA/LIBRA collaboration (Bernabei et al. 2008) claiming evidence of dark matter particles in the galactic halo. Subsequent repetitions of the same experiment by different experimental groups (Xenon100, CDMSII and CoGent) failed to reproduce the same results, decreasing this way the reliability of the initial positive results by the former. For a nice discussion of the reliability and the robustness of the DAMA/LIBRA results and their controversy see Hudson (2009).

standard gravitational laws. The real challenge is to enrich this common core by deriving further properties (e.g. by further constraining the allowed model-independent mass range) however, such properties cannot be derived by using robustness arguments from the variability of experiments due to the model dependence of the results from the available methods.

This means that robustness arguments can only be employed within a range of different sub-models in a specific model scenario, such as the model scenario that dark matter consists of WIMPs. As we have seen in our discussion of informativeness, in order to extrapolate a meaningful result from an experimental process, numerous assumptions need to be implemented during the construction of the experimental setup and the analysis of data. Ensuring that two different methods are probing the same quantity/parameter of the same model in order to employ robustness arguments thus requires taking into consideration the effects of these assumptions in the results of the experiment.

The following remarks from Goodman et al. (2010) offer a rather illuminating example of how the results of an experiment are conditional to a number of factors. In a paper presenting a set of constraints on dark matter properties from collider experiments, the authors begin their discussion by stating that the interpretation of the results depends on the nature – and hence the adopted sub-model – of the dark matter particle:

> We consider the cases where the DM particle is a scalar [boson] or a fermion; if a scalar, it can be real or complex, and if a fermion, it can be Majorana or Dirac. Each of these cases is considered separately (*ibid.*, p.2).

They then continue by listing the extrapolating assumptions in the experiment in order to yield their results:

> We shall be considering the situation where the WIMP [...] is the only particle in addition to the standard model fields accessible to colliders [...] For simplicity, we assume the WIMP is a singlet under the SM gauge groups, and thus possesses no tree-level couplings to the electroweak gauge bosons. We also neglect couplings with Higgs bosons. (*ibid.*, p.2)

After the presentation of their results Goodman et al. proceed to conclude that the presented constraints on the strength of interactions of WIMPs with hadrons also depend on the mass of the dark matter candidate, as well as the coupling preference of dark matter: if dark matter primarily couples to gluons, the constrains from colliders become significantly tighter (*ibid.*,p.8).

The above remarks by Goodman et al. indicate that the various constraints placed on the interactions between WIMPs and hadrons from collider experiments are conditional on a set of introduced assumptions and are of course model specific to the model scenario of dark matter as weakly interacting massive particles. Given that different methods necessarily involve different assumptions, any comparison between these constraints and the constraints obtained by a different method (e.g. from direct searches) must therefore be made by taking into consideration the effects of these assumptions on the extrapolation of the constraints. This is precisely the aim of compatibility and complementarity studies in dark matter research, however, the severe lack of such studies highlighted by many physicists (e.g. Bauer et al. 2015), indicates the degree of difficulty in achieving this task. The current situation in dark matter research comprises a vast collection of largely unrelated papers placing *model* *specific constraints* on different model scenarios of dark matter, without examining the possible concordance of their results with alternative experiments. It is here that robustness arguments from the variability of experiments become valuable, and can be used via complementarity studies with the aim of combining the results between different methods in order to improve the reliability and accuracy of the currently available model specific constraints on dark matter.

A nice example of such work comes from a study by Cerdeno et al. (2016) who present an improvement on dark matter exclusion limits from direct detection experiments by incorporating the results from indirect searches based on gamma rays from annihilation in the Milky Way halo. Interestingly, the authors begin the abstract of their article by stating that 'When comparing constraints on the Weakly Interacting Massive Particle (WIMP) properties from direct and indirect detection experiments it is crucial that the assumptions made about the dark matter (DM) distribution are realistic and *consistent*' (*ibid.*, emphasis added). They then proceed to calculate a consistent and improved exclusion limit on the dark matter-nucleon scattering cross section by taking into consideration the introduced assumptions about dark matter contribution both in direct and indirect searches. These studies nicely illustrate how robustness arguments from the variability of experiments can be used in dark matter research in the context of complementarity and compatibility studies to improve the reliability of the constraining limits in the parameter space of certain models.

There are however, two important caveats. The first is that such results are always model specific since they are limited to the various sub-models of certain model scenarios, and as such they leave the essential question about which model (or models) best captures the actual nature of dark matter untouched. The second caveat is that such studies are only available for the sub-models of those model scenarios that can be probed by more than one method, and hence, they are of limited use to model scenarios that can be currently probed by only one method, such as sterile neutrinos and self-interacting dark matter. Nevertheless, robustness arguments from the variability of experiments can play a significant role in improving the reliability of results concerning the allowed parameter space of the various competing models, insofar as the different assumptions on which the results of each method are depended are taken into consideration during the comparison of concordant results.

6 Conclusions

In summary, the aim of this article was to examine the extent to which robustness arguments from the variability of experiments can be used within dark matter research in order to overcome various methodological and theoretical challenges, by exploring the concepts of informativeness, model sensitivity and reliability as useful epistemic virtues. Apparently, the most important challenge in dark matter research is to reduce the number of viable models of dark matter by enriching the minimum set of constraints in the thin common core concept in a way that eliminates certain candidate models. The first conclusion is that robustness arguments are of limited use in achieving this aim, due to the the fact that most of the time the extracted information from the various methods of dark matter observation is model specific. The extraction of model-independent limits from a particular method – let alone from two or more methods for the needs of robustness arguments – remains one of the most important challenges in dark mater research.

The second conclusion is that robustness arguments from the variability of experiments can – and should – be used for constraining the parameter space of certain model scenarios via compatibility and complementarity studies. This can be achieved by extracting exclusion limits for the properties of those models scenarios that can be probed by more than one method, by taking into consideration the different assumptions introduced in each method and making sure that they are consistent. The upshot is that while robustness arguments from the variability of experiments cannot be employed for the elimination of possible model scenarios, they are still a valuable tool for increasing the confidence of the scientific community about the reliability of the current methods of dark matter observation and the allowed parameter space of certain models, contributing this way to the achievement of a more modest, but nonetheless, important aim.

References

Aghanim, N., Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, and S. Basak (2018). Planck 2018 results. *VI. Cosmological parameters 1807.*

Antoniou, A. (2021). What is a data model? *European Journal for Philosophy of Science 11*(4), 1–33.

Basso, A. (2017). The appeal to robustness in measurement practice. *Studies in History and Philosophy of Science Part A 65*, 57–66.

Bauer, D., J. Buckley, M. Cahill-Rowley, R. Cotta, A. Drlica-Wagner, J. L. Feng, S. Funk, J. Hewett, D. Hooper, and A. Ismail (2015). Dark matter in the coming decade: Complementary paths to discovery and beyond. *Physics of the Dark Universe 7*, 16–23.

Bauer, M., M. Heiles, M. Neubert, and A. Thamm (2019). Axion-like particles at future colliders. *The European Physical Journal C 79*(1), 74.

Bertone, G., D. Hooper, and J. Silk (2005). Particle dark matter: Evidence, candidates and constraints. *Physics Reports 405*(5-6), 279–390.

Boyarsky, A., M. Drewes, T. Lasserre, S. Mertens, and O. Ruchayskiy (2019). Sterile neutrino Dark Matter.

Cerdeno, D. G., M. Fornasa, A. M. Green, and M. Peiró (2016). How to calculate dark matter direct detection exclusion limits that are consistent with gamma rays from annihilation in the Milky Way halo. *Physical Review D 94*(4), 43516.

De Baerdemaeker, S. (2021). Method-driven experiments and the search for dark matter. *Philosophy of Science 88*(1), 124–144.

De Baerdemaeker, S. and N. M. Boyd (2020). Jump ship, shift gears, or just keep on chugging: Assessing the responses to tensions between theory and evidence in contemporary cosmology. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 72*, 205–216.

de Swart, J. (2020). Closing in on the Cosmos: Cosmology's Rebirth and the Rise of the Dark Matter Problem. In A. Blum, R. Lalli, and J. Renn (Eds.), *The Renaissance of General Relativity in Context*, pp. 257–284. Cham: Springer.

Duffy, L. D. and K. Van Bibber (2009). Axions as dark matter particles. *New Journal of Physics 11*(10), 105008.

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

Feng, J. L. (2010). Dark matter candidates from particle physics and methods of detection. *Annual Review of Astronomy and Astrophysics 48*, 495–545.

Franklin, A. (1989). *The neglect of experiment*. Cambridge University Press.

Franklin, A. and C. Howson (1984). Why do scientists prefer to vary their experiments? *Studies in History and Philosophy of Science Part A 15*(1), 51–62.

Gaskins, J. M. (2016). A review of indirect searches for particle dark matter. *Contemporary Physics 57*(4), 496–525.

Gluscevic, V., Y. Ali-Haimoud, K. Bechtol, K. K. Boddy, C. Bœhm, J. Chluba, F.-Y. Cyr-Racine, C. Dvorkin, D. Grin, and J. Lesgourgues (2019). Cosmological probes of dark matter interactions: The next decade. *arXiv preprint arXiv:1903.05140*.

Goodman, J., M. Ibe, A. Rajaraman, W. Shepherd, T. M. P. Tait, and H.-B. Yu (2010). Constraints on dark matter from colliders. *Physical Review D 82*(11), 116010.

Gueguen, M. (2020). On Robustness in Cosmological Simulations. *Philosophy of Science 87*(5), 1197–1208.

Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge university press.

Hudson, R. (2009). The methodological strategy of robustness in the context of experimental wimp research. *Foundations of Physics 39*(2), 174–193.

Jacquart, M. (2021a). Dark Matter and Dark Energy. In E. K. Knox and A. Wilson (Eds.), *Routledge Companion to the Philosophy of Physics*, pp. 731–743. New York: Routledge.

Jacquart, M. (2021b). ΛCDM and MOND: A debate about models or theory? *Studies in History and Philosophy of Science Part A 89*, 226–234.

Kahlhoefer, F. (2017). Review of LHC dark matter searches. *International Journal of Modern Physics A 32*(13), 1730006.

Kosso, P. (2013). Evidence of dark matter, and the interpretive role of general relativity. *Studies in History and Philosophy of Science Part B - Studies in History and Philosophy of Modern Physics 44*(2), 143–147.

Kuorikoski, J., A. Lehtinen, and C. Marchionni (2010). Economic modelling as robustness analysis. *The British Journal for the Philosophy of Science 61*(3), 541–567.

Lisciandra, C. (2017). Robustness analysis and tractability in modeling. *European Journal for Philosophy of Science 7*(1), 79–95.

Lloyd, E. A. (2015). Model robustness as a confirmatory virtue: The case of climate science. *Studies in History and Philosophy of Science Part A 49*, 58–68.

Luković, V., P. Cabella, and N. Vittorio (2014). Dark matter in cosmology. *International Journal of Modern Physics A 29*(19), 1443001.

Martens, N. C. M. (2022). Dark Matter Realism. *Foundations of Physics 52*(1), 1–19.

Martens, N. C. M. and D. Lehmkuhl (2020a). Cartography of the space of theories: an interpretational chart for fields that are both (dark) matter and spacetime. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 72*, 217–236.

Martens, N. C. M. and D. Lehmkuhl (2020b). Dark Matter = Modified Gravity? Scrutinising the spacetime − matter distinction through the modified gravity/dark matter lens. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 72*, 237–250.

Massey, R., T. Kitching, and J. Richard (2010). The dark matter of gravitational lensing. *Reports on Progress in Physics 73*(8), 86901.

Massimi, M. (2018). Three problems about multi-scale modelling in cosmology. *Studies in History and Philosophy of Modern Physics 64*, 26–38.

Mayo, D. G. (1996). *Error and the growth of experimental knowledge*. University of Chicago Press.

Natarajan, A. (2013). Bounds on Dark Matter from CMB Observations. In *Sources and Detection of Dark Matter and Dark Energy in the Universe*, pp. 67–71. Springer.

Odenbaugh, J. and A. Alexandrova (2011). Buyer beware: robustness analyses in economics and biology. *Biology and Philosophy 26*(5), 757–771.

Parker, W. S. (2011). When climate models agree: The significance of robust model predictions. *Philosophy of Science 78*(4), 579–600.

Roszkowski, L., E. M. Sessolo, and S. Trojanowski (2018). WIMP dark matter candidates and searches – current status and future prospects. *Reports on Progress in Physics 81*(6), 66201.

Schumann, M. (2019). Direct detection of WIMP dark matter: concepts and status. *Journal of Physics G: Nuclear and Particle Physics 46*(10), 103003.

Schupbach, J. N. (2018). Robustness analysis as explanatory reasoning. *The British Journal for the Philosophy of Science 69*(1), 275–300.

Shapere, D. (1982). The concept of observation in science and philosophy. *Philosophy of science 49*(4), 485–525.

Smeenk, C. (2020). Some reflections on the structure of cosmological knowledge. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 71*, 220–231.

Smeenk, C. and S. C. Gallagher (2020). Validating the Universe in a Box. *Philosophy of Science 87*(5), 1221–1233.

Staley, K. W. (2020). Securing the empirical value of measurement results. *The British Journal for the Philosophy of Science 71*(1), 87–113.

Stegenga, J. (2009). Robustness, discordance, and relevance. *Philosophy of Science 76*(5), 650–661.

Vanderburgh, W. L. (2003). The Dark Matter Double Bind: Astrophysical Aspects of the Evidential Warrant for General Relativity. *Philosophy of Science 70*(4), 812–832.

Vanderburgh, W. L. (2014). On the interpretive role of theories of gravity and 'ugly' solutions to the total evidence for dark matter. *Studies in History and*

Philosophy of Science Part B - Studies in History and Philosophy of Modern Physics 47(1), 62–67.

Weisberg, M. (2006). Robustness analysis. *Philosophy of science 73*(5), 730–742.

Weisberg, M., M. Jacquart, B. Madore, and M. Seidel (2018). The Dark Galaxy Hypothesis. *Philosophy of Science 85*(5), 1204–1215.

Woodward, J. (2006). Some varieties of robustness. *Journal of Economic Methodology 13*(2), 219–240.