



Nested modalities in astrophysical modeling

Elena Castellani¹ · Giulia Schettino^{2,3}

Received: 28 February 2022 / Accepted: 9 January 2023 / Published online: 21 February 2023
© The Author(s) 2023

Abstract

In the context of astrophysical modeling at the solar system scale, we investigate the modalities implied by taking into account different levels of detail at which phenomena can be considered. In particular, by framing the analysis in terms of the how-possibly/how-actually distinction, we address the debated question as to whether the degree of plausibility is tightly linked to the degree of detail. On the grounds of concrete examples, we argue that, also in the astrophysical context examined, this is not necessarily the case.

Keywords Astrophysical modeling practices · Modal modeling · How-possibly/how-actually distinction · Post-Newtonian models · Metric theories of gravity

1 Introduction

Astrophysics, intended in a broad sense as including both the physics of the solar system and the physics at large and very large scales (i.e., cosmology),¹ notoriously deals with phenomena and processes taking place in extreme conditions, hardly reproducible in a laboratory. Given also the broad range of physical scales and the complexity of the systems considered, the adoption of models and computer simulations becomes indispensable in this field, thus offering exemplary case studies for the role of models in scientific practices.

¹This is a common use of the term though not the only one. See for example Anderl (2016), where astrophysics is characterised as the physics operating at “intermediate” scales, that is between the physics of the solar system and the cosmology of the entire universe

This article belongs to the Topical Collection: *EPSA21: Selected papers from the biennial conference in Turin*

Guest Editors: A.C. Love, C. Marchionni, M. Redei, J. Williamson

✉ Elena Castellani
elena.castellani@unifi.it

In fact, the functions and meaning of modeling and simulations in astrophysics have been objects of extensive philosophical study, especially in the last decade (e.g., Vanderburgh, 2003, 2014; Anderl, 2018; Massimi, 2018; Smeenk & Gallagher, 2020; Gueguen, 2020; Jacquart, 2020, 2021). Here, we take a slightly different direction with respect to this literature, by considering the issue of astrophysical modeling in the framework of the current debate on *modal modeling* in science. In fact, in recent reflections on modeling practices in science, there has been a renewal of interest in the modal aspects implied. In particular, some recent papers (Verrault-Julien, 2019; Sjölin Wirling & Grüne-Yanoff, 2021, 2022; Grüne-Yanoff & Verrault-Julien, 2021) are devoted to tackle the epistemology of modal modeling with a special attention to the meaning and role of the *how-possibly/how-actually distinction*.

Here, we propose to contribute to the analysis of the how-possibly/how-actually distinction by extending it to the context of astrophysics, not typically considered in the modal modeling literature. In particular, the question we are interested in is how the involved modality is related to the level of abstraction or generality at which astrophysical phenomena are considered. On this aim, we have chosen to focus on concrete cases at the solar system scale, including the modeling practices involved in a specific experiment – that is, the case of the test of relativistic theories that will be performed by the radio science experiment of the joint ESA/JAXA BepiColombo mission to Mercury.

The paper is organized as follows. In Section 2, we specify the conceptual framework adopted, that is, how we will use such notions as “model”, “simulation”, “data set” and the how-possibly/how-actually distinction. Section 3 provides an outlook on the distinctions which can be drawn between classes of theories as well as classes of models at the solar system scale. On this background, Section 4 examines the specific case of the relativity experiment of the BepiColombo mission. Finally, in Section 5 we engage with the current philosophical literature on the how-possibly/how-actually distinction and discuss, in this regard, the import of the analysis of the modalities implied in modeling gravity within the solar system, provided in the previous two sections.

2 Conceptual framework

Over time, and especially in the last decades, a huge amount of literature on models has been proposed, debating questions regarding their nature, functions and epistemic import.² As clearly shown by this literature, “models” are meant in different senses, depending on the context in which they are applied and on their intended use. Here, following (Weisberg, 2013), we will adopt the notion of a model in the sense of an *interpreted structure* used for studying physical phenomena, properties or evolution in a given domain.

More precisely, we will focus on astrophysical *theoretical models*, and consider their relations with *data sets* by means of *simulations*. For the aim of this paper,

²See the useful overview provided by Frigg and Hartmann (2020) and references therein.

“theoretical models” are intended in the sense of non-concrete interpreted structures, used and tested to explore a space of physical possibilities characterized in terms of parameters which may take different values according to given purposes (e.g., Datteri & Schiaffonati, 2019). “Data sets” are the results of the procedures (synthesising, filtering, correcting or smoothing) by means of which are processed and elaborated the “raw” astronomical data, collected by using a telescope or a space mission or whatever other observing methodology.³ In this sense, data sets function as the “observational” basis to be taken into account, providing the so-called “observed observables”.

“Simulations”, in the case we consider, come into play by mediating between the theoretical models and the data sets. In some more detail, the models examined are related to the data sets through the computer simulations which are used to obtain numerical results from the model equations.⁴ In substance, once a theoretical model is chosen or properly built, by simulating the corresponding dynamics it is possible to generate the so-called “simulated observables”, which are the data that would be recorded if the model was, in fact, the actual description of the phenomena under study. In other words, the usual procedure consists in building a suitable theoretical model, running a computer simulation on its basis and then comparing the output of the simulation with the available data sets. At this point, simulated and “observed” observables can be directly compared by different sort of fitting algorithms, in order to check the validity of the theoretical model (e.g., Lari et al., 2021).

By definition, theoretical models are possible models, that is, models of possible state of affairs. Is there a way of being more precise about the degree of possibility these models represent? In this respect, a helpful conceptual tool turns out to be the how-possibly/how-actually distinction mentioned in the previous section. Note that this distinction – at the center of a lively debate, especially flourished (in its current form) in the “New Mechanism” literature⁵ – is usually considered in regard to the explanatory role of models, accordingly taking the form of a distinction between how-actually and how-possibly model explanations. Here, we will focus on the contrast between how-possibly and how-actually descriptions in general, leaving on the

³Here we follow the common scientific usage of the term (see Kelleher & Tierney, 2018, Chap. 2). In fact, there is some ambiguity in the use of the term in the literature, especially in the philosophical one, where they are often identified with data models (see, for example, the discussion in Bokulich, 2014; Bokulich & Parker, 2021; Antoniou, 2021).

⁴Here, we take the term “simulation” in the narrow sense of running a computer process and, following (Datteri & Schiaffonati, 2019), we adopt their working definition according to which a (computer) system is said to simulate a theoretical model if it can be characterised in terms of parameters whose values depend on one another according to the regularities mentioned in the theoretical model. Of course, how to define a “simulation system” and under which conditions it can be said to effectively simulate a target system (or a theoretical model of the target system) is not such a simple issue and the different approaches debated in the literature on scientific modeling sensibly depend on the context considered (physics, climate science, economics, social science, ..). See, e.g., the detailed investigation on the relation between models and simulations proposed by Winsberg (2018). A recent philosophical discussion of the role of computer simulations in astrophysics is provided by Jacquart (2020).

⁵For the use of the how-possibly/how-actually distinction in this literature, the seminal papers are Machamer et al. (2000) and Craver (2006). See on this, for example, Glennan (2017), Sect. 3.5

background the explanatory side of the issue (which is, however, important).⁶ How to characterize a how-possibly *vs.* a how-actually description will be examined in Section 5, by taking into account the current discussion. In particular, we will refer to the distinction as discussed in Bokulich (2014), where a special attention is paid to the different levels of abstraction at which the phenomenon under study can be framed.

3 Modeling gravity in the solar system

As a matter of fact, the dynamics of physical systems in astrophysical contexts is led by gravity. Nowadays, the most acknowledged description of gravity is provided by the theory of general relativity (GR), which has reached an impressive predictive success over the decades. Accordingly, most of adopted modeling practices are framed in the context of the experimental testing of GR, a central topic in the actual debate in astrophysics. Although GR has passed a large number of experimental tests, observations over the last decades have pointed out some shortcomings of the theory both at the infrared (i.e., galactic) and ultraviolet (i.e., quantum) scales, thus highlighting that GR could not be the final theory for gravitational interaction (see, for example, Capozziello & de Laurentis, 2011, for a review on this topic).⁷

The issues raised by such recent observational results can be addressed differently at different astrophysical scales. Here, we have chosen to focus on the solar system scale. Notwithstanding the growing interest in modeling and testing gravity in cosmological contexts, the solar system remains a very powerful laboratory for investigating gravitational theories. In fact, there are various advantages in testing gravity at the solar system scale, such as, first of all, the relative proximity of the phenomena under study. From a theoretical point of view, the dynamics within the solar system – that is, in the weak field and slow motion regime – can be more easily handled with respect to cosmological scales. Indeed, solving the equations of motion in their “weak-field form” is typically less demanding. From an experimental point of view, gravitational tests in the solar system have been carried out for a long time. This means that our knowledge of its dynamics is significantly deeper than the knowledge of the dynamics at strong-field regimes or in cosmological contexts. Moreover, the present and near-future on-ground and space-based technologies provide very accurate direct measurements, often allowing for a more straightforward analysis than in cosmological cases.

This section is devoted to highlight the relevant theoretical background for discussing the how-possibly/how-actually distinction in the framework of modeling gravity at the solar system scale. First of all, let us point out a very general distinction which can be drawn between *metric theories of gravity*, such as GR, and *non-metric*

⁶A very detailed, recent discussion of the epistemic value of how-possibly explanations is provided in Grüne-Yanoff and Verrault-Julien (2021).

⁷In this paper, we will consider only “classical” gravitation (i.e., issues at quantum scales will not be taken into account).

theories of gravity.⁸ A general description of metric vs. non-metric theories will be provided in Section 3.1; as for now, let us just underline that while a metric theory satisfies a number of basic physical principles – first of all Einstein’s *equivalence principle* (EEP) – this is not the case for non-metric theories.

As will be discussed in some detail in Section 3.1, observational evidences in the case of the solar system strongly support a metric description of gravity. In general, in the framework of a metric theory, the solar system dynamics can be described by an approximate solution, known as *post-Newtonian* (PN) *approximation*, corresponding to the limit of slow moving particles under the effect of weak gravitational fields. In particular, within the PN limit, any theory of gravity can be formalized in an approximated parameterized form, the *parameterized post-Newtonian* (PPN) *approximation*: this means that any significant dynamical effect due to gravitation can be described by means of a specific PN parameter, the value of which depends on the metric theory chosen. The different metric theories of gravity, all conveniently approximated in terms of their PPN formulation, form the second level of distinctions which we take into account for the aim of the paper. Section 3.2 is devoted to describe in some detail this second level.

Finally, a further, third level of distinctions can be obtained by considering the theoretical models built from the metric theories in their PPN approximation. More precisely, once a metric theory in its PN limit is chosen, it is possible to build a theoretical model – call it a “PN model” – accounting for the dynamics in terms of a specific choice of the values of the theory’s PN parameters.⁹ Thus, we find a number of competing theoretical models, each one identified by its own set of PN parameters. Note that, at this level, the predictions of each model can be compared with the available data sets by means of different kind of fits. This third level will be examined in detail in Section 3.3.

Summing up, three different levels of distinctions can be envisaged, depending on the degree of generality at which we are considering competing descriptions of gravity. The first, more general level, regards the distinction between metric and non-metric theories of gravity; a second, less general level is generated by the different metric theories of gravity; a third, even less general level corresponds to the different PN models, each one belonging to the PPN approximation of a specific metric theory (see Fig. 1).

⁸In fact, over the years a number of competing theories of gravity have been proposed in order to modify, to various extent, the standard formulation of GR. Note that a motivation for an increasing interest in such a kind of research (initially mainly stimulated by theoretical/mathematical driving developments), is undoubtedly due to the recent observational drawbacks of GR.

⁹What is intended, here, by “theoretical models” has been specified in Section 2. In fact, there is a perduring debate about how to characterize models with respect to theories in an astrophysical context and which of the two are the appropriate “units” to consider (a recent discussion can be found in Jacquart, 2021): for the purposes of this paper, we will refer to “theories” as the set of principles, assumptions and equations in terms of which the properties and dynamics of the astrophysical systems are described, and to “models” as the parameterization of a given theory in order to allow the comparison with data sets.

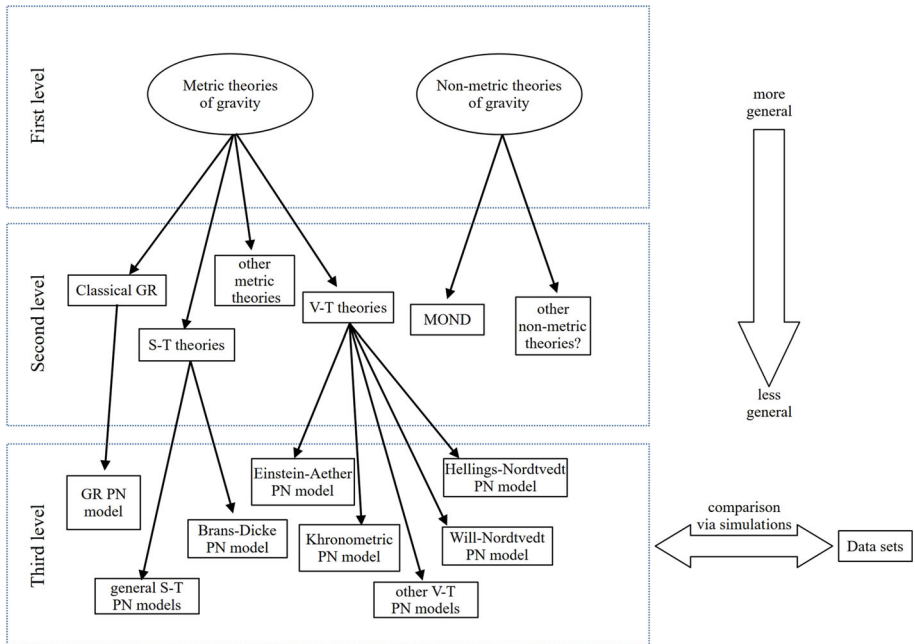


Fig. 1 Sketch of the three levels of distinctions which can be individuated when dealing with gravity in the case of the solar system

3.1 First level: Metric vs. non-metric theories of gravity

Assuming that, from a mathematical viewpoint, spacetime should be a four-dimensional differentiable manifold and the equations of gravity together with the mathematical entities in them should be expressed in a covariant form, non-metric theories are not ruled out, in principle, from the list of viable theories of gravity (see Will, 2018, p. 12-16). What distinguishes the two classes of metric and non-metric theories is whether EEP holds (metric theories) or not (non-metric theories). In some more detail, EEP is usually defined as the conjunction of the following three assumptions: (i) the weak equivalence principle (WEP), stating the equivalence of gravitational and inertial mass of test particles, is valid; (ii) the outcome of any local non-gravitational experiment is independent of the velocity of the freely falling apparatus; (iii) the outcome of any non-gravitational experiment is independent of where and when in the universe it is performed.¹⁰ Consequently, it can be shown that gravitation must be a curved-spacetime phenomenon in the following precise sense:

- spacetime is endowed with a metric;
- the world lines of test particles are geodesics of that metric;
- in local freely falling frames, the non-gravitational laws of physics are those of special relativity.

¹⁰See Will (2018), p. 16-17 for a detailed discussion.

At present, there are strong experimental evidences in favour of the validity of EEP, leading to a widespread consensus in favour of metric theories of gravity.¹¹ Note that the equivalence principle is often enunciated in a “stronger” version of the Einstein’s formulation, usually known as *strong equivalence principle* (SEP). In substance, SEP generalizes WEP by including also the case of self-gravitating bodies (not only test particles) and gravitational (not only non-gravitational) experiments.¹² It is worth underlining that, while EEP is fulfilled by any metric theory of gravity, SEP is strictly met only by GR.

Nevertheless, nothing prohibits EEP (or SEP) to break at some level, allowing for a non-metric theory of gravity. The best known example are the so-called Modified Newtonian Dynamics (MOND) theories, which have gained an increasing interest since the first formulation by Milgrom (1983). The main motivation for such theories is the phenomenology of galactic dynamics, usually addressed in the current cosmological paradigm – the Λ Cold Dark Matter (Λ CDM) model – by resorting to the assumption that most of the matter content in the universe is made of *dark matter* (see, e.g., Ferreira, 2019, for a review on the Λ CDM model). MOND theories, on the contrary, address the issue by modifying Newton’s dynamics. More precisely, the break of the Newton’s law in the MOND regime is expressed in terms of the so-called *Milgrom’s law*: introducing a fundamental scale of acceleration a_0 (as provided by empirical evidences), Newton’s equation of motion, $a = GM/r^2$, holds for accelerations larger than a_0 , while in the “MOND regime”, that is, for accelerations smaller than a_0 , the equation of motion is modified.¹³ As a consequence, a break of the equivalence principle is expected in MOND theories. This can be understood in the following way: on the one hand, SEP can be reformulated by asserting that the

¹¹Concerning the first assumption (i), since the famous Eötvös experiment in 1885, WEP has been tested over the years with very high accuracy: the current upper limit on WEP has been set by the MICROSCOPE mission at the level of 10^{-15} (see, e.g., Touboul et al., 2020). The second assumption of EEP (ii) implies the validity of special relativity, that is, the so-called *Local Lorentz Invariance* (LLI). Many experiments have been devised in order to check for possible violations of LLI, leading to very tight experimental constraints on the validity of LLI (see, e.g., Mattingly, 2005, for a general review on this topic). Finally, testing the third assumption (iii) of EEP implies testing Local Position Invariance (LPI), which in fact refers both to spatial and temporal invariance. The tests for spatial LPI consist in gravitational redshift experiments, based on precise atomic clocks measurements, and they typically assume that WEP and LLI are valid. Spatial LPI is tested with less accuracy than WEP and LLI: current best bounds are around one part per million (see, for example, Leefer et al., 2013). The tests for temporal LPI consist, instead, in checking for possible time variation of non-gravitational universal constants, such as the fine structure constant, the weak interaction constant and the electron-to-proton mass ratio. If LPI is violated, the coupling between possible external fields and matter should evolve in time in a non-metric way, causing a variation of some universal constants. Also in this case, the experimental constraints turn out to be very tight (see Will, 2018, p. 34).

¹²In some more detail, SEP can be enunciated as follows: (i) WEP is valid for test bodies and for self-gravitating body; (ii) the outcome of any local, gravitational or non-gravitational, experiment is independent of the velocity of the freely falling apparatus; (iii) the outcome of any local, gravitational or non-gravitational, experiment is independent of where and when in the universe it is performed (see Will, 2018, p. 170). Alternatively, SEP can be rephrased by stating that the outcome of an experiment performed in a sufficiently small freely falling laboratory over a sufficiently short time is indistinguishable from the outcome of the same experiment performed in an inertial frame in empty space.

¹³Reviews of the MOND paradigm can be found, e.g., in Sanders (1990); Famaey and McGaugh (2012); Khoury (2015).

internal dynamics of a system is the same independently of any external constant field in which the system is embedded; on the other hand, the dynamics predicted by MOND depends on the relative magnitude between the total acceleration acting on the system (thus, including also external fields) and the scale acceleration a_0 . This “external field effect”, characterizing MOND theories, implies the break of SEP.¹⁴ Such a violation of SEP should lead to observational effects. Thus, also at the solar system scale, experimental tests of MOND can be conceived (see Section 3.3).

3.2 Second level: Metric theories in PPN approximation

In the framework of a metric description of gravity, a number of alternative theories with respect to GR have been proposed over time. In general, such theories are not aimed at replacing GR, but rather at modifying or extending it in regimes where GR shows its main drawbacks – whence the use of referring to these theories as *modified* or *extended theories of gravity*. This means that competing metric theories of gravity are not expressly designed for solar system tests of gravity, where GR is extremely successful. Nevertheless, any modification or extension of GR should eventually be detectable and tested in some way also in the solar system, where current and near-future experiments are significantly more accurate than at cosmological scales.

The main metric theories of gravity can be grouped as follows (cfr. Fig. 1):¹⁵

- *Scalar-tensor (S-T) theories*: while in Newton’s theory gravity is determined by means of a scalar field and in GR by the metric tensor $g_{\mu\nu}$, in this case gravity is determined both by a metric tensor and a scalar field ϕ , so that the metric can be put in the form $\bar{g}_{\mu\nu} \equiv A^2(\phi)g_{\mu\nu}$ (for a comprehensive introduction see, for example, Fujii & Maeda, 2003). In addition, a characteristic coupling function, $\omega(\phi)$, is introduced. Different formulations can be given depending on the behaviour of the scalar field and the coupling function. The first attempt in this sense was proposed by Brans and Dicke (1961): Brans–Dicke theory assumes that the coupling function is, in fact, a coupling constant, ω_{BD} , such that the larger is the value of ω_{BD} , the smaller is the effect of the scalar field (resorting to GR in the limit of $\omega_{BD} \rightarrow \infty$).
- *$f(R)$ theories*: in this case, the idea is to substitute the Ricci scalar curvature R with a suitable function, $f(R)$, chosen in such a way that at cosmological scales the universe would experience an accelerated expansion, without the need to resort to a cosmological constant or dark energy. This family of modifications of GR was first proposed in Buchdahl (1970). It can be shown that eventually $f(R)$ theories are equivalent to S-T theories (see, for example, Jain & Khoury, 2010).
- *Vector-Tensor (V-T) theories*: in this case gravity is determined both by the metric tensor and by a dynamical four-vector field u^μ . This type of modifications is

¹⁴ See, e.g., Famaey and McGaugh (2012), p. 54–56.

¹⁵For a general discussion, see Will (2018), ch. 5 and references therein.

motivated by the idea of exploring possibilities for a violation of Lorentz invariance in gravity, thus allowing for preferred frame effects (a detailed description is provided in Jacobson & Mattingly, 2001). V-T theories can be distinguished in *constrained* and *unconstrained V-T theories*. Constrained V-T theories assume that the dynamical field is constrained, as in the Einstein–Aether theory, where u^μ is constrained to be time-like with unit norm (see, e.g., Eling et al., 2004) or in the khronometric theory, where the vector field is required to be hypersurface orthogonal (see, e.g., Blas et al., 2010). Otherwise, the dynamical field can be unconstrained, as in Will–Nordtvedt theory (see Will & Nordtvedt, 1972) and in Hellings–Nordtvedt theory (see Hellings & Nordtvedt, 1973).

More recently, mainly inspired by results in quantum physics and cosmology, further alternative proposals have found a fair consensus in the scientific community.¹⁶ Examples are *Tensor-Vector-Scalar (TeVeS) theories*¹⁷ and *massive gravity theories*.¹⁸

As already mentioned, in the framework of a metric theory, the solar system dynamics can be suitably described by means of the PN approximation, corresponding to the limit of slow moving particles under the effect of weak gravitational fields.¹⁹ In fact, within the solar system, gravitational fields are weak enough to consider relativistic effects as “corrections” of Newtonian gravity. In other words, Newtonian acceleration represents the zero-order term and relativistic corrections are added to the equation of motion as higher-order terms. As a consequence, any metric theory can be expressed by expanding the general spacetime metric about the Minkowski metric as a sum of PN adimensional gravitational potentials of varying degrees of smallness, each potential being a functional of matter variables. The result is that each metric theory is described by its PN metric, and the only way in which one PN metric can differ from another is in the values of the coefficients that multiply each term in the metric.²⁰

3.3 Third level: PN models and data models

The comparison of metric theories with each other in their PN limit underpins what we have indicated as the second level of distinctions. The third, less abstract level emerges when taking into account the theoretical models based on these metric theo-

¹⁶These are indicated as “other metric theories” in Fig. 1.

¹⁷TeVeS theories, characterized by three different gravitational fields (the metric, a dynamical four-vector field and a dynamical scalar field) have been devised to provide a fully relativistic theory of gravitation which is also capable to mimic MOND dynamics at regimes where MOND shows its best success, such as at galactic scales (see, e.g., Bekenstein, 2004).

¹⁸Massive gravity theories have been devised on the attempt to ascribe a mass to the gravitational field, usually referred as “graviton” in this context (see, e.g., Hinterbichler, 2012, for a review).

¹⁹In fact, as of now, most of the solar system tests of gravity can be performed in such approximated framework with sufficient accuracy (see Will, 2014).

²⁰For details, see Will (2018), p. 88.

ries, which are each one characterized by its own set of PN parameters. At this more “phenomenological” level, the simulated data based on a specific model can be compared with the available data sets by means of different kind of fits, thus allowing to discriminate between experimentally suitable and experimentally unsuitable models.

Let us enter into some more detail about the construction of the PN models forming this third level. As already mentioned, in the PN limit the expansion of the metric is written in terms of a sum of PN adimensional potentials, each one characterized by a multiplying coefficient whose value depends from the metric theory considered. In particular, in the PPN formalism, dimensionless arbitrary parameters are put in the place of the coefficients of the potentials, where each of these PN parameter describes a specific property of the spacetime metric. To be more precise, the PPN formalism provides for a total of 10 PN parameters.²¹ Hence, any metric theory of gravity predicts a specific set of values for the 10 PN parameters.

In the case of GR, there are only two not null PN parameters: the Eddington parameters γ and β , whose value is expected to be unity. All the other PN additional effects are null in GR, since no additional fields are expected beside the metric field. Hence, the GR PN model is built by setting the values of γ and β equal to 1 and the values of the other 8 PN parameters identically null. Also in the case of S-T theories, only γ and β are expected to be different from zero, but their value can be written as a function of $\omega(\phi)$. In contrast with the case of GR, this function can be different from unity. In the particular case of Brans–Dicke theory, where the coupling function is in fact a constant, the value of β is expected to be unity, as in the GR PN model, while γ is a function of the coupling constant and its value can be different from unity. Finally, in the Brans–Dicke PN model it holds that the other 8 PN parameters are identically null as in GR. In the general S-T PN models, both β and γ are functions of $\omega(\phi)$ (thus, they are different from the case of the GR PN model), while the other 8 PN parameters are identically null also in this case. Conversely, in the case of V-T theories also the PN parameters describing possible preferred frame effects (labeled with α_1, α_2) are expected to be different from zero.²²

For the sake of completeness, it is worth pointing out that, while present data from solar system observations strongly support a metric description of gravity, small violations of EEP could anyway take place below the accuracy level of current tests, thus leaving room for the chance of non-metric theories of gravity. As a consequence, a number of tests of alternative non-metric theories (such as MOND) have been recently proposed also within the solar system, resorting to the weak-field limit of such non-metric theories. In particular, the MOND paradigm could be tested phenomenologically in the solar system and constrained by fitting the available planetary data (see, e.g., Milgrom, 2009; Blanchet & Novak, 2011; Hees et al., 2014).

A different situation holds if we consider SEP instead of EEP. SEP is strictly fulfilled only by GR, while the other metric theories of gravity can violate it in different

²¹A list of the PN parameters, with their meaning and the values assumed in different metric theories, can be found, for example, in Will (2014), p. 31.

²²A detailed derivation of the PN limit of the main metric theories of gravity can be found in Will (2018), ch. 5.

ways. Examples are: the Nordtvedt effect, which implies the violation of WEP for massive bodies (while it still holds for test particles, as expected if EEP is valid), preferred-frame and preferred-location effects, temporal variation of the gravitational constant G . The PPN framework turns out to be particularly suitable for testing such violations in the weak-field regime. Indeed, possible violation of SEP can be mirrored by different values of individual PN parameters or combination of PN parameters. Hence, constraining the values of PN parameters within the solar system is a powerful tool for discriminating between competing theories of gravity.

Summing up: dealing with gravitational phenomena in the solar system accounts for different distinctions which arrange at different levels, spanning from families of theories to families of models. Such a general arrangement of possibilities is schematically illustrated in Fig. 1. How the highlighted distinctions act in practice can be seen by resorting a specific concrete case. This is the subject of the following section, where we turn to examine, from the viewpoint of the modalities involved, the case of the relativity experiment of the BepiColombo mission to Mercury.

4 Case study: Testing metric theories of gravity with BepiColombo

BepiColombo is an ESA/JAXA space mission for the exploration of the planet Mercury and the inner solar system (e.g., Benkhoff et al., 2013).²³ The spacecraft was launched at the end of 2018 and it is planned for orbit insertion around Mercury at the end of 2025. It is equipped with a competitive suit of instruments to perform different scientific experiments.

One of the main on-board experiments, the Mercury Orbiter Radio science Experiment (MORE), has two major scientific goals (see, e.g., Iess et al., 2021): (i) determining Mercury's gravitational field and rotational state (*gravimetry-rotation experiments*) and (ii) performing a very accurate test of relativistic theories of gravity (*relativity experiment*). Such ambitious scientific goals will be achieved by processing ultra-accurate radio observations. In this section, we will focus on the second goal of MORE, that is, the test of relativistic theories of gravity. There is a remarkable advantage in performing such kind of test by means of a space mission at Mercury. Indeed, Mercury is the best placed planet in the solar system in order to test for gravitational theories, as it is the nearest planet to the Sun and, therefore, the most subject to its gravitational effects. Thanks to the possibility of achieving a very accurate determination of both the heliocentric orbit of Mercury and the mercurycentric orbit of the spacecraft, the MORE relativity experiment will be capable of constraining with very high accuracy the value of the main PN parameters by means of a non-linear least squares fit. The possibility of putting a tighter constraint on the value

²³The mission is named after Giuseppe "Bepi" Colombo (1920–1984), an Italian mathematician and astronomer who first discovered the 3:2 spin-orbit resonance of the planet Mercury and contributed to develop the 'gravity assist' technique to reach the planet.

of the PN parameters will significantly help in discriminating between competing metric theories of gravity.²⁴

At the beginning, that is, in the 1990s, the MORE relativity experiment was devised on the specific aim of testing the validity of GR. Let us give a basic idea of the procedure on which the experiment is based. Assuming GR in its PPN approximation, a PN model accounting for the dynamics of the system is built. This GR PN model describes the dynamics of Mercury (and of the spacecraft around the planet) by considering a Newtonian zero-order gravitational term plus the addition of PN corrections due to GR.²⁵ Such corrections can be written as additional accelerations terms in the equations of motion for Mercury, with each term multiplied by the corresponding PN parameter.²⁶ The first step consists in setting the values of the PN parameters to those predicted by GR (we will call them PN_set_0). Then, given the GR PN model with PN_set_0 , it is possible to run the orbit determination code, ORBIT14 (see Lari et al., 2021): the resulting simulation represents the initial “nominal” solution of the problem. The output of this solution consists in a set of simulated radio observations. This set can be directly compared and fitted against the observed data set. Such a fit is performed by a non-linear least squares fit, in the form of a differential corrections method (e.g., Milani & Gronchi, 2010, ch. 6). The scope of the fit consists in determining the set of PN values that minimizes the difference between simulated and observed data.²⁷ The procedure is then iterated until the best fit of the values of the PN parameters has been obtained.²⁸

Adopting the same procedure as for the test of GR, any metric theory of gravity, written in PPN approximation, could be eventually tested using MORE radio observations. In this case, the orbit determination code will adopt, as the input dynamical

²⁴For the sake of completeness, we point out that the BepiColombo MORE experiment is not the only ongoing effort to constrain the PPN parameters. Indeed, a number of experiments and techniques have been devised, based on very precise measurements both from ground and from space. A well-known example is the case of Lunar Laser Ranging.

²⁵Moreover, the model accounts also for perturbative effects due to the other planets and the main bodies of the solar system (asteroids, etc.).

²⁶The details of the dynamical model adopted can be found in Milani et al. (2002, 2010).

²⁷The difference is defined in terms of the residuals between simulated, i.e., computed, and observed observations (see, e.g., Lari et al., 2021, for details).

²⁸In some detail, the whole iterative procedure can be described as follows (for a deeper description, see, e.g., Schettino & Tommei, 2016; Lari et al., 2021): **Iteration 1**: the nominal simulated observations (obtained by setting the values of the PN parameters to the nominal ones, PN_set_0) are compared with the data set and the fit provides an updated set of values for the PN parameters, PN_set_1 ; **Iteration 2**: we run again the orbit determination code, updating the dynamical model with the new values for the PN parameters, given by PN_set_1 ; the output is an updated set of simulated observations, which are again compared with the data set; the new fit provides a new set of values of PN parameters, PN_set_2 : such new set represents an improved fit of the values of the PN parameters and the residuals between simulated and observed observations should be smaller than at the previous iteration; **Iteration 3**: then PN_set_2 is used as the updated input parameters for the dynamical model to run an updated simulation; the updated simulated observations are again fitted with the data set and an updated fit of the values PN parameters, PN_set_3 , is determined; **Iteration n**: the process continues by iterating the previous steps until the residuals between iteration $(n - 1)$ and iteration n are small enough that the differential correction process has arrived at convergence, that is, the best fit of the values of the PN parameters has been obtained (where by “best fit”, we mean the set of values which minimizes the residuals, i.e., the difference, between simulated and observed observations).

model, a different PN model, based on that specific metric theory, which can predict values of the PN parameters different with respect to the GR PN model. An example could be the case of Brans-Dicke theory, which predicts that the PN parameter γ should be different from unity, while the other PN parameters are expected to assume the same values as in GR. With the MORE relativity experiment it could be possible to perform straight comparison between the two competing theories, GR and Brans-Dicke theory: roughly speaking, if the best fitting value of the parameter γ would turn out to be different from unity, GR should be discarded in favour of any competing metric theory which predicts a different value for that parameter. Note that the current knowledge on γ is set at unity with an accuracy at the level of 2×10^{-5} , as provided by the Cassini spacecraft (see Bertotti et al., 2003). This means that, as of today, any possible departure from GR should be below the 10^{-5} threshold. In the case of the MORE relativity experiment, the constraint on γ is expected to be improved by, at least, one order of magnitude (see, e.g., Schettino & Tommei, 2016; Serra et al., 2018; Schettino et al., 2018; 2020).

In the light of the three levels discussed in Section 3, to conclude, the modeling practice at the basis of the MORE relativity experiment can be understood as follows:

- First level: relativistic metric theories of gravity are assumed to be the actual framework to describe gravitational physics in the solar system; non-metric theories are discarded.
- Second level: despite the impressive accuracy that BepiColombo-MORE radio science observations are expected to achieve, the PN limit is certainly an accurate approximation to describe the gravitational iterations of interest for the experiment. In principle, different relativistic PN theories can be adopted to describe the experiment, depending on the specific metric theory that needs to be tested. The standard approach consists in adopting classical GR theory in its PN limit (see, e.g., Milani et al., 2010, for an extensive discussion), but other attempts are currently under study (see, e.g., Schettino et al., 2020);
- Third level: competing relativistic metric theory of gravity can be tested by means of the MORE relativity experiment. For each theory, the corresponding PN model is built and fitted against the data set. After the differential correction process, the best fit of the values of PN parameters is determined and, accordingly, a theory can be discarded (when found to be inconsistent with the observations) or acknowledged (when representing a possible scenario subject to the level of accuracy provided by MORE).

5 Discussion: Nested modalities

To sum up on the basis of the previous analysis: when dealing with gravitational phenomena in the solar system, three levels of distinctions can be envisaged, depending on the degree of generality or abstraction at which competing descriptions of gravity are considered. At a first, very general level, metric and non-metric theories of gravity are distinguished; at a second, less general level, different more specific theories or families of theories can compete, either within the class of metric theories

or within the class of non-metric theories of gravity. At a third, even less general level, distinct models can be built on the basis of the different theories individuated at the second level. At this third, “phenomenological” level, the predictions (simulated data) of each model can be compared with the available data sets by means of different kind of fits, thus allowing to discriminate between experimentally suitable and experimentally unsuitable models.

Now, let us look at such an arrangement of possibilities – the possible descriptions distinguished at the different levels of generality in the study of gravity within the solar system – from the perspective of the philosophical discussion on the how-possibly/how-actually distinction.

First of all, regarding how to characterize the distinction, there is no uniform view in the literature, also given the variety of contexts taken into account over the course of the years.²⁹ As already mentioned, the distinction has been mainly discussed with respect to explanation and, especially, model-based explanation. In this respect, we find different positions in the literature, depending substantially on two points of discussion. On the one side, positions differ on whether a continuum can be envisaged between how-possibly and how-actually explanations: in other words, on whether the difference between how-possibly and how-actually is just epistemic (how-possibly models are conjectures about the actual) or, on the contrary, how-possibly models represent something other than the actual, a sort of “just-so stories”.³⁰ On the other side, from the epistemic viewpoint, different views can be taken with respect to the following question: along which dimension the how-possibly/how-actually distinction is to be measured? More precisely, positions can differ on whether the distinction is assumed to be just a matter of the level of detail at which the description is given, or the relation between the degree of abstraction and the modality involved (how-possibly *vs.* how-actually) is less straightforward than what could appear at first sight.

This second point is specifically the one we want to deal with in this paper. Our analysis is precisely focused on the question as to whether we can establish a direct link between the level of abstraction or generality at which the descriptions are considered and their how-possibly/how-actually characterization. Before turning to consider, from this viewpoint, the import of the distinctions individuated in the case of the solar system, let us enter into more detail by recalling some representative positions to be found in the literature.

According to (Brandon, 1990), to begin with, a how-possibly explanation – “one where one or more of the explanatory conditions are speculatively postulated” – can be moved “along the continuum until finally we count it as a how-actually explanation”, and this passage from how-possibly to how-actually is determined by getting more and more empirical evidence (p. 184). In the same spirit, for (Machamer et al.,

²⁹See for example Bokulich (2014), pp. 322-325, for a discussion of differing views on the how-possibly/how-actually distinction such as those of William Dray in the 1950s in the context of explanations in history, of Robert Brandon in the 1990s for evolutionary mechanisms and, more recently, of Patrick Forber when discussing the role of biological constraints.

³⁰For a detailed discussion of a number of these different positions, see for example Bokulich (2014), Sect. 1, and Glennan (2017), Sect. 3.5.

2000) mechanistic explanations render phenomena intelligible by showing “*how possibly, how plausibly, or how actually things work*” (p. 21). In discussing this intelligibility process by means of examples in neurobiology and molecular biology, the authors introduce a further distinction between “sketches” and “schemata”, where the former are considered abstract, incomplete versions of the latter, and the movement from the former to the latter is by adding missing details (pp. 15–18). One of the authors, Craver, further elaborates on this by arraying mechanistic models along the following two axes (Craver, 2006, 2007):

- (a) The “possibly-plausibly-actually axis”: an explanatory continuum from how-possibly to how-plausibly to how-actually, where at one extreme, *how-possibly models* show how a mechanism “might work” (thus being heuristically useful in “constructing a space of possible mechanisms”), while at the other extreme *how-actually models* show how a mechanism “works”. Between these two extremes, there is a range of *how-plausibly models* that are more or less consistent with the known constraints on the details of the mechanism that in fact produces the phenomenon (Craver, 2006, p. 361).³¹
- (b) The “sketch-schemata axis”: between a speculative sketch, leaving many details out, to an “ideally complete description”, lies a continuum of schemata, that abstract away to a greater or lesser extent from the details (Craver, 2006, p. 360).

Progress in the explanation means movement along both axes (a) and (b) (Craver, 2007, p. 114). This view has raised a number of discussions in the literature. Gervais and Weber (2013), in particular, criticize the debate on mechanistic explanations (referring especially to Craver’s work) for conflating two features of models: plausibility (corresponding to the dimension of Craver’s axis (a)), and richness of information (corresponding to the dimension of Craver’s axis (b)).³² In contrast to plausibility, richness is not necessary for a model to be explanatory, they argue for. In the same vein, (Glennan, 2017) (Sect. 3.5) critically discusses what he considers to be Craver’s tight link between the axis (a) (“possibly-plausibly-actually axis”) and the axis (b) (“sketch-schema-mechanism axis”), that is, between the degree of plausibility and the degree of sketchiness, where by sketchiness one means abstraction. While moving from how possibly to how-actually undoubtedly constitutes scientific progress – he claims – this is not always the case when moving to less “sketchy” models. A sketchier model may be a better one for some purposes, as he shows by means of concrete examples (Glennan, 2017, p. 69).

³¹A constraint, for Craver, is “a finding that either shapes the boundaries of the space of plausible mechanisms or changes the probability distribution over that space”. In short, constraints on the space of possible mechanisms “constitute the relevant evidence for evaluating how-possibly descriptions of mechanisms”, and the progress from how-possibly to how-actually descriptions of a mechanism can thus be conceived “as a process of shaping and constricting the space of plausible mechanisms” (Craver, 2007, pp. 247–248).

³²More in detail: by plausibility, Gervais and Weber (2013) mean “the degree of probability that a model is accurate in the existence of, and distinctions between, the various entities and activities it postulates”, while richness “concerns the degree of detail a model provides in its description” (p. 139).

A critical discussion of the assumption of a tight correlation between the degree of plausibility and the level of detail can be found as well in Bokulich (2014), though from a different perspective. By taking as a case study the geological phenomenon known as “tiger bush” (a characteristic striking periodic banding of vegetation appearing in semi-arid region) and its various possible model explanations, Bokulich shows how alternative models can compete both at the how-actually level and at the how-possibly level, forming a kind of hierarchical branching tree. On the one side, how-actually explanations, i.e., explanations referring in some way to the observable effects of the phenomenon, are shown to be deployed also at a very abstract level. On the other side, within the corresponding class of how-actually models, Bokulich shows how it is possible to identify a split at a second most abstract level between different how-possibly models, each providing a possible further specification of the explanatory mechanism (pp. 331–332). In this case, the how-actually/how-possibly distinction does not merely refer to a more or less detailed description of the phenomenon (p. 334). Bokulich, thus, provides an account of the tiger bush phenomenon that has, in her own words, “the somewhat counterintuitive consequence that one can move from a rather well-confirmed how-actually explanation of tiger bush at a high level of abstraction [...] to a how-possibly model explanation as one tries to fill in some of the further details of that mechanism” (p. 335).

Now, let us go back to the descriptions of gravity at the solar system scale, discussed in Sections 3 and 4. The resulting level structure illustrated in Fig. 1 represents a hierarchical branching tree, moving down from more abstract to more detailed descriptions. The question we want to address, at this point, is whether this corresponds to a parallel movement from a how-possibly to a how-actually level of description.

As we have seen, at a first, very general level, two classes of theories of gravity – metric and non-metric – can be distinguished. Now, on the grounds of the available experimental results in the case of the solar system, there is a general consensus on assuming that gravity is described by a metric theory. Thus, at this first, more abstract level one could assume that the actual description belongs to the class of metric theories of gravity. Once this choice is made, at the second, less abstract level we have at our disposal a number of how-possibly metric theories, including GR. Then, at the third, still more detailed level, within each class of metric theories a number of competing PN models can be provided. In the particular case of testing gravitational theories with BepiColombo, different PN models represent different how-possibly scenarios to be compared with the data set provided by the experiment. Finally, by comparing the simulated data with the data sets, one can progressively restrict the space of possibilities by discarding those theories which are found inconsistent with the observations.³³

³³Analogously, if one wanted to opt for the class of non-metric theories at the first level in contrast to the previous choice, this would mean to assume that the how-actually description is of a non-metric type. Accordingly, at the second level, MOND could be a how-possibly description of gravitation in the solar system. To this aim, a suitable MOND model could be built as a how-possibly model at the third level and, then, it could be checked against the data sets (see, e.g., the tests of MOND proposed in the solar system, cited in Section 3.3).

Thus, in terms of the three levels distinguished when studying gravity at the solar system scale, we can say that the resulting hierarchical structure represents a web of *nested modalities*, rather than a continuum from how-actually to how-possibly descriptions. Indeed, at each level, a given description can be interpreted as one of the how-possibly options along one arm of the branching generated at the higher level of abstraction. At the same time, this very description, in turn, can be interpreted as a how-actually scenario giving rise to a further branching at the subsequent, less abstract level. In other words, one can move from a rather well-confirmed how-actually description at a high level of abstraction to how-possibly models as one tries to fill in some further details. Thus, this shows that we cannot establish, in general, a direct link between the level of abstraction or generality at which the descriptions are considered and their how-possibly/how-actually characterization.

Summing up, we arrive at a very similar conclusion as the one drawn, as seen above, by Gervais and Weber (2013), Bokulich (2014), and Glennan (2017) in completely different contexts: there is not necessarily a strict correspondence between the level of abstraction and the kind of modality implied, or, in other words, the degree of possibility is not necessarily directly linked to the degree of details.

Acknowledgements Elena Castellani acknowledges financial support from MIUR PRIN 2017 (Prot. 2017-ZNWW7F-004). Giulia Schettino acknowledges the Italian Space Agency (ASI), grant agreement no. 2017-40-H.0. The authors are very grateful to the audiences of the *Modeling the Possible* Vienna's online 2021 Workshop, the EPSA21 Conference and the ASTRO2022 ESO online Conference, for helpful, precious feedback. The authors thank the two anonymous reviewers for their very useful comments.

Funding Open access funding provided by Università degli Studi di Firenze within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anderl, S. (2016). Astronomy and astrophysics. In P. Humphreys (Ed.) *The oxford handbook of philosophy of science* (pp. 652–670). Oxford University Press.
- Anderl, S. (2018). Simplicity and simplification in astrophysical modeling. *Philosophy of Science*, 85(5), 819–831.
- Antoniou, A. (2021). What is a data model? an anatomy of data analysis in high energy physics. *European Journal of Philosophy of Science*, 11(101), 1–33.
- Bekenstein, J. D. (2004). Relativistic gravitation theory for the modified Newtonian dynamics paradigm. *Physical Review D*, 70, 083 509.
- Benkhoff, J., Casterena, J. V., Hayakawa, H., & et al. (2013). BepiColombo-comprehensive exploration of mercury: Mission overview and science goals. *Planetary and Space Science*, 58, 2–20.
- Bertotti, B., Iess, L., & Tortora, P. (2003). A test of general relativity using radio links with the cassini spacecraft. *Nature*, 425(6956), 374–376.

- Blanchet, L., & Novak, J. (2011). External field effect of modified Newtonian dynamics in the Solar system. *Monthly Notices of the Royal Astronomical Society*, 412(4), 2530–2542.
- Blas, D., Pujolàs, O., & Sibiriyakov, S. (2010). Consistent Extension of Hořava Gravity. *Physical Review Letters*, 104(18), 181302–0.
- Bokulich, A. (2014). How the tiger bush got its stripes: “How possibly” vs. “how actually” model explanation. *The Monist*, 97(3), 321–338.
- Bokulich, A., & Parker, W. (2021). Data models, representation, and adequacy-for-purpose. *European Journal of Philosophy of Science*, 11(1), 1–26.
- Brandon, R. (1990). *Adaptation and environment*. Princeton University Press.
- Brans, C., & Dicke, R. H. (1961). Mach’s principle and a relativistic theory of gravitation. *Physical Review*, 124(3), 925–935.
- Buchdahl, H. A. (1970). Non-linear lagrangians and cosmological theory. *Monthly Notices of the Royal Astronomical Society*, 150, 1–8.
- Capozziello, S., & de Laurentis, M. (2011). Extended theories of gravity. *Physics Reports*, 509(4–5), 167–321.
- Craver, C. F. (2006). When mechanistic models explain. *Synthese*, 153, 355–376.
- Craver, C. (2007). *Explaining the brain mechanisms and the mosaic unity of neuroscience*. Oxford University Press.
- Datteri, E., & Schiaffonati, V. (2019). Robotic simulations, simulations of robotics. *Minds and Machines*, 29, 109–125.
- Eling, C., Jacobson, T., & Mattingly, D. (2004). Einstein-Aether Theory. arXiv e-prints gr-qc/0410001.
- Famaey, B., & McGaugh, S. S. (2012). Modified Newtonian Dynamics (MOND): Observational phenomenology and relativistic extensions. *Living Reviews on Relativity*, 15, 10.
- Ferreira, P. G. (2019). Cosmological tests of gravity. *Annual Review of Astronomy and Astrophysics*, 57, 335–374.
- Frigg, R., & Hartmann, S. (2020). Models in sciences. The Stanford Encyclopedia of Philosophy at <https://plato.stanford.edu/entries/models-science/>.
- Fujii, Y., & Maeda, K. (2003). *The Scalar-Tensor theory of gravitation*. Cambridge University Press.
- Gervais, R., & Weber, E. (2013). Plausibility versus richness in mechanistic models. *Philosophical Psychology*, 26(1), 139–152.
- Glennan, S. (2017). *The new mechanical philosophy*. Oxford University Press.
- Grüne-Yanoff, T., & Verrault-Julien, P. (2021). How-possibly explanations in economics: Anything goes? *Journal of Economic Methodology*, 28(1), 114–123.
- Gueguen, M. (2020). On robustness in cosmological simulations. *Philosophy of Science*, 87(5), 1197–1208.
- Hees, A., Folkner, W.M., Jacobson, R.A., & et al. (2014). Constraints on modified Newtonian dynamics theories from radio tracking data of the Cassini spacecraft. *Physical Review D*, 89(10), 102002.
- Hellings, R. W., & Nordtvedt, K. (1973). Vector-Metric Theory of gravity. *Physical Review D*, 7(12), 3593–3602.
- Hinterbichler, K. (2012). Theoretical aspects of massive gravity. *Reviews of Modern Physics*, 84(2), 671–710.
- Iess, L., Asmar, S. W., Cappuccio, P., & et al. (2021). Gravity, geodesy and fundamental physics with bepiColombo’s MORE investigation. *Space Science Reviews*, 127(1), id.21.
- Jacobson, T., & Mattingly, D. (2001). Gravity with a dynamical preferred frame. *Physical Review D*, 70, 024 028.
- Jacquart, M. (2020). Observations, simulations, and reasoning in astrophysics. *Philosophy of Science*, 87(5), 1209–1220.
- Jacquart, M. (2021). Λ cdm and mond: A debate about models or theory? *Studies in History and Philosophy of Science Part A*, 89, 226–234.
- Jain, B., & Khoury, J. (2010). Cosmological tests of gravity. *Annals of Physics*, 325, 1479–1516.
- Kelleher, J. D., & Tierney, B. (2018). *Data science*. The MIT Press.
- Khoury, J. (2015). An alternative to particle Dark Matter. *Physical Review D*, 91, 024 022.
- Lari, G., Schettino, G., Serra, D., & et al (2021). Orbit determination methods for interplanetary missions: Development and use of the Orbit14 software. *Experimental Astronomy*, forthcoming.
- Leefer, N., Weber, C.T.M., Cingöz, A., & et al. (2013). New limits on variation of the fine-structure constant using atomic dysprosium. *Physical Review Letters*, 111(6), 060801.

- Machamer, P., Lindley, D., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, 67(1), 1–25.
- Massimi, M. (2018). Three problems about multi-scale modelling in cosmology. *Studies in History and Philosophy of Modern Physics*, 64, 26–38.
- Mattingly, D. (2005). Modern tests of lorentz invariance. *Living Reviews in Relativity*, 8(1), 5.
- Milani, A., & Gronchi, G. F. (2010). *Theory of orbit determination*. Cambridge University Press.
- Milani, A., Tommei, G., Vokrouhlický, D., & et al. (2010). Relativistic models for the bepiColombo radioscience experiment. In SA. Klioner, PK. Seidelmann, & MH. Soffel (Eds.) *Relativity in fundamental astronomy: Dynamics, reference frames, and data analysis*, pp 356–365. <https://doi.org/10.1017/S1743921309990652>.
- Milani, A., Vokrouhlický, D., Villani, D., & et al. (2002). Testing General Relativity with the bepiColombo radio science experiment. *Physical Review D*, 66, 082 001.
- Milgrom, M. (1983). A modification of the newtonian dynamics as a possible alternative to the hidden mass hypothesis. *The Astrophysical Journal*, 270, 365–370.
- Milgrom, M. (2009). MOND Effects in the inner Solar system. *Monthly Notices of the Royal Astronomical Society*, 399(1), 474–486.
- Sanders, R. H. (1990). Mass discrepancies in galaxies - Dark Matter and alternatives. *Astronomy and Astrophysics Review*, 2, 1–28.
- Schettino, G., Serra, D., Tommei, G., & et al. (2018). Addressing some critical aspects of the bepiColombo more relativity experiment. *Celestial Mechanics and Dynamical Astronomy*, 130(11), 72.
- Schettino, G., Serra, D., Tommei, G., & et al. (2020). A test of gravitational theories including torsion with the bepiColombo radio science experiment. *Universe*, 6(10), 175.
- Schettino, G., & Tommei, G. (2016). Testing general relativity with the radio science experiment of the bepiColombo mission to mercury. *Universe*, 2(3), 21.
- Serra, D., Di Pierri, V., Schettino, G., & et al. (2018). Test of general relativity during the bepiColombo interplanetary cruise to mercury. *Physical Review D*, 98(6), 064059.
- Sjölin Wirling, Y., & Grüne-Yanoff, T. (2021). The epistemology of modal modeling. *Philosophy Compass*, 16(10), e12775.
- Sjölin Wirling, Y., & Grüne-Yanoff, T. (2022). Epistemic and objective possibility in science. *British Journal for the Philosophy of Science*, forthcoming.
- Smeenk, C., & Gallagher, S. C. (2020). Validating the universe in a box. *Philosophy of Science*, 87(5), 1221–1233.
- Touboul, P., Rodrigues, M., Métris, G., & et al (2020). Microscope mission analysis, requirements and expected performance. [arXiv:2012.06472](https://arxiv.org/abs/2012.06472).
- Vanderburgh, W. (2003). The dark matter double bind: Astrophysical aspects of the evidential warrant for general relativity. *Philosophy of Science*, 70(4), 812–832.
- Vanderburgh, W. (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, 62–67.
- Verrault-Julien, P. (2019). How could models possibly provide how-possibly explanations? *Studies In History and Philosophy of Science Part A*, 73, 22–33.
- Weisberg, M. (2013). *Simulation and similarity: Using models to understand the world*. Oxford University Press.
- Will, C. M. (2014). The confrontation between General Relativity and experiment. *Living Reviews on Relativity*, 17, 4.
- Will, C. M. (2018). *Theory and experiment in gravitational physics*. Cambridge University Press.
- Will, C. M., & Nordvedt, K. L. (1972). Conservation laws and preferred frames in relativistic gravity. i. preferred-frame theories and an extended PPN formalism. *The Astrophysical Journal*, 177, 757.
- Winsberg, E. (2018). *Philosophy and climate science*. Cambridge University Press.

Affiliations

Elena Castellani¹ · Giulia Schettino^{2,3} 

✉ Giulia Schettino
g.schettino@ifac.cnr.it

¹ Department of Humanities and Philosophy (DILEF), University of Florence,
Via della Pergola 60, Florence, 50121, Italy

² Institute of Applied Physics “Nello Carrara” (IFAC), CNR, Via Madonna del Piano 10,
Sesto Fiorentino (FI), 50019, Italy

³ Department of Mathematics, University of Pisa, Largo B. Pontecorvo 5, Pisa, 56127, Italy