



# The Hole Argument without the notion of isomorphism

Joanna Luc<sup>1</sup>

Received: 20 March 2023 / Accepted: 18 January 2024 / Published online: 15 March 2024  
© The Author(s) 2024

## Abstract

In this paper, I argue that the Hole Argument can be formulated without using the notion of isomorphism, and for this reason it is not threatened by the criticism of Halvorson and Manchak (Br J Philos Sci, 2022. <https://doi.org/10.1086/719193>). Following Earman and Norton (Br J Philos Sci 38, pp. 515–525, 1987), I divide the Hole Argument into two steps: the proof of the Gauge Theorem and the transition from the Gauge Theorem to the conclusion of radical indeterminism. In the analysis of the first step, I argue that the Gauge Theorem does not rely on the notion of isomorphism but on the notion of the diffeomorphism-invariance of the equations of local spacetime theories; however, for this approach to work, the definition of local spacetime theories needs certain amendments with respect to Earman and Norton’s formulation. In the analysis of the second step, I postulate that we should use the notion of radical indeterminism instead of indeterminism simpliciter and that we should not decide in advance what kind of maps are to be used in comparing models. Instead, we can tentatively choose some kind of maps for this purpose and check whether a given choice leads to radical indeterminism involving empirically indistinguishable models. In this way, the use of the notion of isomorphism is also avoided in the second step of the Hole Argument. A general picture is that physical equivalence can be established by means of an iterative procedure in which we examine various candidate classes of maps, and, depending on the outcomes, we need to broaden or narrow these classes. The Hole Argument can be viewed as a particular instance of this procedure.

**Keywords** Hole Argument · Symmetry · Isomorphism · Form-invariance

## 1 Introduction

The Hole Argument is one of the most important arguments in philosophical debates about the significance of symmetries and about the (in)determinism of physical theo-

---

✉ Joanna Luc  
joanna.luc.poczta@gmail.com

<sup>1</sup> Institute of Philosophy, Jagiellonian University, Grodzka 52, 31-044 Kraków, Poland

ries. Different versions of the argument differ in terms of, among other things, what conclusion the argument is supposed to support. Einstein, who first formulated the Hole Argument, regarded it as an argument against the general covariance of the laws of physics (but later abandoned it). The Hole Argument was reintroduced by Earman and Norton (1987), this time as an argument against spacetime substantivalism. However, later some authors argued that certain kinds of substantivalism are not prone to this argument (see, e.g., Brighouse, 1994; Hofer, 1996; Pooley, 2006). In this paper, I will consider the Hole Argument as an argument for the thesis that diffeomorphism-related models (cf. Definition 4 in Sect. 2) of local spacetime theories represent the same physically possible world without presupposing anything about the relation of this thesis to spacetime substantivalism.

The aim of this paper is to analyse the recent challenge to the Hole Argument made by Halvorson and Manchak (2022), who suggest that it crucially relies on the notion of isomorphism and that among various possible readings of this notion in the context of General Relativity (GR), none is suitable to underpin the Hole Argument. I will argue that the Hole Argument can be expressed without ever mentioning the notion of isomorphism, so whatever problems are related to this notion, they are not relevant in assessing this argument. This is, I will claim, because of its dynamical nature: the mathematical fact that underpins the Hole Argument is not that certain GR models are isomorphic but that the dynamical equations of GR do not change their form under diffeomorphisms.

The paper is organised as follows. In Sect. 2, certain terminological issues concerning isomorphic, isometric and diffeomorphic models will be clarified. In Sect. 3, I will review in detail the Hole Argument as formulated by Earman and Norton. It is divided into two steps: the proof of the Gauge Theorem (step 1, Sect. 3.1) and the transition from the Gauge Theorem to the conclusion of radical indeterminism (step 2, Sect. 3.2). In Sect. 4, I will sketch out Halvorson and Manchak's criticism of this argument. I will focus on the issue of where exactly in their view the notion of isomorphism is needed in the Hole Argument. In Sect. 5, I will present the version of the Hole Argument that does not use the notion of isomorphism at any point. This requires certain precisifications of this argument with respect to Earman and Norton's version. First, in the formulation of the Gauge Theorem (i.e., step 1) we must restrict their notion of local spacetime theories because the definition in Earman and Norton's paper is too broad. Second, instead of a fully developed definition of indeterminism, we should rather use (in step 2) a partial definition of radical indeterminism that is relativised to the choice of the kind of maps that serve as the standard of comparison of models (which at the same time is a tentative proposal for the standard of their physical equivalence). Then, by means of an iterative procedure, we can find what kind of maps should be the mentioned standard, taking into account whether this particular choice leads to radical indeterminism and what differences between models are empirically detectable. In Sect. 6, some further details of the paper by Halvorson and Manchak will be discussed in light of the approach presented in Sect. 5. Additionally, I will discuss in this section an earlier objection to the Hole Argument due to Weatherall (2018). Finally, Sect. 7 summarises the main points of this paper.

## 2 Isomorphic models, isometric models and diffeomorphic models

The aim of this section is to clarify what could be meant by saying that two models are isomorphic, isometric or diffeomorphic. Let us start with the notion of isomorphism. I will define it in an abstract way:

**Definition 1** Two models,  $\mathcal{M}$  and  $\mathcal{M}'$ , are isomorphic iff they have the same mathematical structure.

This definition on its own does not provide us with a practical criterion for deciding whether two given models are isomorphic. For this aim, we also need to specify the mathematical structure of  $\mathcal{M}$  and  $\mathcal{M}'$ , which might be disputable. Sometimes, this issue is approached in the reverse order: that is, for a given class of models, one can specify which of them are isomorphic and which are not, and then define the mathematical structure of the models in this class as whatever is left invariant by these isomorphisms. However, one cannot have it both ways, that is, one cannot employ this strategy and simultaneously use Definition 1, as this would be a vicious circle.

Let us now turn to the remaining two notions, which are more concrete. One can define the meaning of the expression “models  $\mathcal{M}$  and  $\mathcal{M}'$  are isometric” in the following two ways<sup>1</sup>:

**Definition 2** (*Isometric models, narrow sense*) Two models,  $\mathcal{M} = \langle M, g_{ab}, O_2, \dots, O_n \rangle$  and  $\mathcal{M}' = \langle M', g'_{ab}, O'_2, \dots, O'_n \rangle$ , are isometric in the narrow sense iff there is a diffeomorphism  $\phi : M \rightarrow M'$  such that  $g'_{ab} = \phi_* g_{ab} = g_{ab}$ .<sup>2</sup>

**Definition 3** (*Isometric models, broad sense*) Two models,  $\mathcal{M} = \langle M, g_{ab}, O_2, \dots, O_n \rangle$  and  $\mathcal{M}' = \langle M', g'_{ab}, O'_2, \dots, O'_n \rangle$ , are isometric in the broad sense iff there is a diffeomorphism  $\phi : M \rightarrow M'$  such that  $g'_{ab} = \phi_* g_{ab}$ .

The expression “models  $\mathcal{M}$  and  $\mathcal{M}'$  are diffeomorphic” is also used in two different senses:

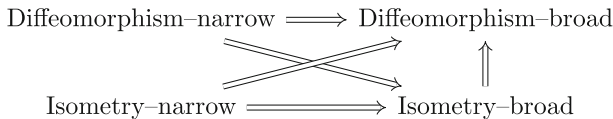
**Definition 4** (*Diffeomorphic models, narrow sense*) Two models,  $\mathcal{M} = \langle M, g_{ab}, O_2, \dots, O_n \rangle$  and  $\mathcal{M}' = \langle M', g'_{ab}, O'_2, \dots, O'_n \rangle$ , are diffeomorphic in the narrow sense iff there is a diffeomorphism  $\phi : M \rightarrow M'$  such that  $g'_{ab} = \phi_* g_{ab}$  and  $O'_i = \phi_* O_i$  for  $i = 2, \dots, n$ .

**Definition 5** (*Diffeomorphic models, broad sense*) Two models,  $\mathcal{M} = \langle M, g_{ab}, O_2, \dots, O_n \rangle$  and  $\mathcal{M}' = \langle M', g'_{ab}, O'_2, \dots, O'_n \rangle$ , are diffeomorphic in the broad sense iff there is a diffeomorphism  $\phi : M \rightarrow M'$ .

Both senses of isometric models and diffeomorphic models can be found in the literature (e.g., Definition 2 is used in Hawking & Ellis, 1973, p. 43 and Wald, 1984, p. 438; Definition 3 is used in Butterfield, 1989, pp. 5–6; Halvorson & Manchak, 2022,

<sup>1</sup> Similar definitions are in Menon and Read (2024, p. 8): their  $\text{Isometry}_1$  corresponds to my isometry in the narrow sense, whereas their  $\text{Isometry}_2$  corresponds to my isometry in the broad sense. The difference is that they consider only the case with  $M' = M$ .

<sup>2</sup> We denote by  $\phi_* O$  the pushforward of  $O$  by  $\phi$ .



**Fig. 1** The relations between Definitions 2–5 (“ $A \Rightarrow B$ ” means “being related by  $A$  implies being related by  $B$ ”). Notice that being related by a diffeomorphism in the broad sense is implied by all other relations, whereas being related by an isometry in the narrow sense implies all other relations for models consisting only of a differential manifold and a metric (but not in general)

p. 5; Hawking & Ellis, 1973, p. 56; Malament, 2012, p. 85; Roberts, 2020, p. 253 and Weatherall, 2018, p. 335; Definition 4 is used in Earman & Norton, 1987, p. 520; and Definition 5 is used in Halvorson & Manchak, 2022, p. 5).<sup>3</sup>

The relations between these notions are depicted in Fig. 1. If two models are isometric in the narrow sense, then they are also isometric in the broad sense but not the other way around; and if two models are diffeomorphic in the narrow sense, then they are also diffeomorphic in the broad sense but not the other way around. If two models consist only of a differential manifold and a metric (or they include some other objects but these objects are uniquely determined by the metric), then their being isometric in the broad sense is equivalent to their being diffeomorphic in the narrow sense. However, if our models have a different form (i.e., they include other objects besides the metric that are not uniquely determined by the metric, such as matter fields), then their being isometric in the broad sense does not guarantee that they are diffeomorphic in the narrow sense; for the same reason, their being isometric in the narrow sense also does not guarantee that they are diffeomorphic in the narrow sense. Therefore, I would prefer to use the notion of diffeomorphic models (in the sense of Definition 4) instead of isometric models because we are going to investigate the whole class of theories, called local spacetime theories, the models of which can have various forms.<sup>4</sup> From now on, “diffeomorphic models” (or “diffeomorphism-related models”) without further specification will be understood in the narrow sense of Definition 4.

### 3 Earman and Norton’s Gauge Theorem and their version of the Hole Argument

In their formulation of the Hole Argument, Earman and Norton (1987) use a very general concept of local spacetime theories. They define such theories by means of their models. A model of a local spacetime theory is an  $(n + 1)$ -tuple  $\mathcal{M} = \langle M; O_1, \dots, O_{k-1}; O_k, \dots, O_n \rangle$ , where  $M$  is a differential manifold and  $O_i$  for  $i = 1, \dots, n$  are fields of geometric objects defined everywhere on  $M$ . For some

<sup>3</sup> It must be noted that some of these authors speak about isometries or diffeomorphisms rather than about models being isometric or diffeomorphic, but I take these two ways of speaking to be naturally inter-translatable.

<sup>4</sup> Butterfield (1989, p. 6) calls models that are diffeomorphic in the narrow sense “isomorphic”. However, this is not suitable for our purposes, as I would like to use the word “isomorphic” in the abstract sense of “having the same mathematical structure” (cf. Definition 1). It is a part of the debate what exactly this structure is and what transformations preserve it, so we should not presuppose at the stage of the choice of terminology what are the answers to these questions.

$k < n$ , the last  $n - k + 1$  objects satisfy a further requirement of being tensor fields<sup>5</sup>; the equations of the theory consist of the vanishing of these tensor fields, that is,  $O_k = 0, \dots, O_n = 0$ . I will call these vanishing tensor fields the “LHSs of the dynamical equations” of the theory (so I will assume that whenever we talk about some dynamical equations, all their terms are moved to the left-hand side so that they consist of equating the combination of these terms to zero).

An example of a local spacetime theory is GR, the models of which can be represented as  $\mathcal{M} = \langle M; g_{ab}, O_2, \dots, O_{k-1}; G_{ab} - T_{ab}, O_{k+1}, \dots, O_n \rangle$ , where  $g_{ab}$  is a metric tensor (field),  $O_2, \dots, O_{k-1}$  are matter fields,  $T_{ab}$  is a stress energy tensor (field) of these matter fields and  $G_{ab}$  is the Einstein tensor (field), expressed in terms of  $g_{ab}$  and its derivatives. The equations of this theory are Einstein’s equations,  $G_{ab} - T_{ab} = 0$  (plus the equations for matter fields,  $O_{k+1} = 0, \dots, O_n = 0$ ).

It is often the case in the literature that instead of such full models, only “impoverished” models of the form  $\langle M, g_{ab} \rangle$  are considered. This can be justified as follows: matter fields are relevant for constraining the metric  $g_{ab}$  only via the stress energy tensor  $T_{ab}$  because it is this quantity that enters Einstein’s equations, so it is enough to consider  $\langle M, g_{ab}, T_{ab} \rangle$ ; but when we have only  $g_{ab}$  and  $T_{ab}$  in our models, we can regard  $g_{ab}$  as fixed and  $T_{ab}$  as determined uniquely by  $g_{ab}$  together with Einstein’s equations, and therefore we can simplify our models even further to  $\langle M, g_{ab} \rangle$ . For models of this form, there is no difference between being diffeomorphic in the narrow sense (Definition 4) and being isometric in the broad sense (Definition 3). However, there are serious conceptual problems with such a simplification. One of them is that GR understood in this way loses its dynamical aspect: the metric is not constrained by any dynamical equations, and the set of kinematically possible models is identical with the set of dynamically possible models. Einstein’s equations appear only as a definition(!) of the stress-energy tensor in terms of the metric. We also lose the idea, expressed in a famous quote by Wheeler, that in GR “matter tells space how to curve and curved space tells matter how to move”. There is no influence of matter on the metric in the “ $\langle M, g_{ab} \rangle$ ” view of GR. While I do not deny that this simplification might be helpful and valid for some purposes, using it in foundational investigations might be misleading, and this concerns the Hole Argument in particular. Therefore, I propose to stick with full models.<sup>6</sup> It is worth noting that even in the vacuum case we cannot disregard  $T_{ab}$ : such-and-such metrics are vacuum solutions of GR *because* they are solutions to Einstein’s equations with  $T_{ab} = 0$ , so the stress-energy tensor has a role to play even in determining the class of vacuum solutions.

Another popular simplification, however, is not harmful: the LHSs of the equations do not need to be included explicitly in the model as additional objects, as Earman and Norton did. Instead, one can just say that a theory has models of the form  $\mathcal{M} = \langle M; O_1, \dots, O_{k-1} \rangle$  that satisfy equations  $O_k = 0, \dots, O_n = 0$ . As I will often refer

<sup>5</sup> Tensor fields are smooth assignments of tensors to each point of  $M$ . One often uses the abbreviation “tensors” for tensor fields, which is strictly speaking not correct but should not lead to any confusions as from the context it is clear that what is meant is the assignment of tensors to the points of the manifold and not tensors themselves.

<sup>6</sup> The same stance is advocated by Menon and Read (2024, p. 5, footnote 7). I thank these authors for the discussion about this issue.

to the details of the paper by Earman and Norton (1987), I will stick with their convention, but nothing hinges on this choice.

GR is not the only example of a local spacetime theory. Many other theories can be formulated in this way: Special Relativity, all relativistic field theories and even non-relativistic spacetime theories (for Special Relativity see, e.g., Pooley, 2017, pp. 120–121; for Newtonian Gravity see, e.g., Weatherall, 2017, pp. 17–21).

Given these background notions, in Sects. 3.1 and 3.2 I will formulate the Hole Argument in detail. Following Earman and Norton, I will divide it into two steps: the Gauge Theorem and its application to obtain the conclusion of radical indeterminism.

### 3.1 Step 1: the Gauge Theorem

The first step in the formulation of the Hole Argument is the following theorem by Earman and Norton (1987, p. 520):

**Theorem 1** (Gauge Theorem) *If  $\mathcal{M} = \langle M; O_1, \dots, O_n \rangle$  is a model of a local spacetime theory  $T$  and  $\phi : M \rightarrow M$  is a diffeomorphism, then the carried along tuple  $\mathcal{M}' = \langle M, \phi_* O_1, \dots, \phi_* O_n \rangle$  is also a model of  $T$ .*

Models related by  $\phi$  are diffeomorphic in the sense of Definition 4. I will present the entire proof of this theorem because its details will be important for my assessment of the objections against it. Let us assume that  $\mathcal{M}$  is a model of  $T$ . Recall that in the models of our theory  $T$ , the last  $n - k + 1$  geometrical objects are the LHSs of the dynamical equations of  $T$ . Therefore,  $\mathcal{M}$  needs to satisfy the following equations:

$$O_i = 0 \text{ for } i = k, \dots, n. \quad (1)$$

These equations can be written in the coordinate form as<sup>7</sup>

$$(O_i)^m = 0 \text{ for } i = k, \dots, n. \quad (2)$$

To show that the transformed tensors  $\phi_* O_i$  are also equal to zero, Earman and Norton use the following trick. The new model  $\mathcal{M}'$  has been obtained by using the diffeomorphism  $\phi$  to transform geometric objects on  $M$ , leaving the coordinates intact. However, we can also carry along the coordinate system  $\{x^m\}$  in which Eq. (1) have been written, obtaining

$$\{x^{m'}\} = \{\phi_* x^m\}. \quad (3)$$

This expression means that  $x^{m'}$ , which is an  $m$ th coordinate in the new coordinate frame, is equal to the  $m$ th coordinate in the old coordinate frame transformed by  $\phi_*$ . Earman and Norton (1987, p. 520) observe that performing both active and passive transformations on a geometrical object at the same time does not change it, so that the

<sup>7</sup> I use the original notation of Earman and Norton's paper, but it must be noted that it is simplified at this point, as  $O_i$  can have more than one index. Therefore, in general, equation (2) should be  $(O_i)^{m_1 \dots m_q} = 0$  for  $i = k, \dots, n$ .

components of any object in the old coordinate frame are the same as the components of an actively transformed object in the new coordinate frame. In symbols,

$$(O_i)^m = (\phi_* O_i)^{m'} \text{ for } i = k, \dots, n. \tag{4}$$

By combining Eqs. (2) and (4), we get

$$(\phi_* O_i)^{m'} = 0 \text{ for } i = k, \dots, n. \tag{5}$$

The next crucial observation is that the last  $n - k + 1$  objects in  $\mathcal{M}$  are tensors. This is important because tensors have a special property: if their components vanish in one coordinate frame, then they vanish in any coordinate frame. Therefore, from (5) it follows that

$$(\phi_* O_i)^m = 0 \text{ for } i = k, \dots, n \tag{6}$$

because  $(\phi_* O_i)^m$  are just  $(\phi_* O_i)^{m'}$  expressed in the old coordinate frame  $\{x^m\}$ . In the coordinate-free formulation, the equations are

$$\phi_* O_i = 0 \text{ for } i = k, \dots, n. \tag{7}$$

However, this is precisely what we wanted to establish—namely, that  $\mathcal{M}'$  satisfies the same dynamical equations as  $\mathcal{M}$ , so it is a model of the same theory  $T$ . At least, this is what Earman and Norton conclude. I think that this conclusion is a bit too quick, and more needs to be said about what it means that two equations are the same and what conditions a local spacetime theory must satisfy for this conclusion to be valid (see section 5.1).

### 3.2 Step 2: from the Gauge Theorem to radical indeterminism

Consider any local spacetime theory  $T$  and one of its models  $\mathcal{M}_{hole} = \langle M, O_1, \dots, O_n \rangle$ . Choose a bounded region  $H \subset M$  and define the hole diffeomorphism  $\phi : M \rightarrow M$  as follows: it acts as the identity outside of  $H$  but differs from the identity within  $H$  (it needs to smoothly become the identity near the boundary of  $H$ ). Now, consider a new model  $\mathcal{M}'_{hole} = \langle M, \phi_* O_1, \dots, \phi_* O_n \rangle$ . By the Gauge Theorem, it is also a model of  $T$ . As the chosen region  $H$  could be arbitrarily small, the pair of models  $\mathcal{M}$  and  $\mathcal{M}'$  witnesses the (radical) indeterminism of  $T$ . In Earman and Norton’s words (1987, p. 516):

In developing the dilemma, we shall see that the equations of these theories are simply not sufficiently strong to determine uniquely all the spatio-temporal properties to which the substantialist is committed. The type of indeterminism involved will be a very radical one indeed. Given some neighbourhood of spacetime we shall see that these theories cannot uniquely determine the fields within the neighbourhood from even the most exhaustive prescription of the

fields outside of it. This is true no matter how small the neighbourhood. We have christened this behaviour ‘radical local indeterminism’.

However, the conclusion of radical indeterminism is avoided if one assumes that although  $\mathcal{M}_{hole}$  and  $\mathcal{M}'_{hole}$  differ mathematically, they do not differ physically, that is, if one assumes that diffeomorphic models of the same local spacetime theory are physically equivalent. Earman and Norton (1987, p. 522) call the principle that “diffeomorphic models [of a given local spacetime theory] represent the same physical situation” Leibniz Equivalence. They assume that any substantialist approach to spacetime must reject this principle, and therefore view the Hole Argument as undermining spacetime substantialism. However, subsequent developments in the literature suggest that the connection between substantialism and Leibniz Equivalence is more subtle. The position called “sophisticated substantialism” (Pooley, 2006) or “dynamic structural realism” (Stachel, 2014) combines the assumption of the reality of spacetime points with Leibniz Equivalence by claiming that spacetime points lack primitive identity. As declared earlier, I will not discuss the issue of substantialism here and focus solely on Leibniz Equivalence,<sup>8</sup>

#### 4 Halvorson and Manchak’s criticism of the Hole Argument

Halvorson and Manchak (2022, p. 2) take the Hole Argument to have three main ingredients: the assumption of substantialism, certain mathematical facts and the conclusion of “pernicious indeterminism” (called by Earman and Norton “radical indeterminism”). They observe that most of the existing literature focused on clarifying the first and the last ingredient—that is, what substantialism exactly is and what it means that a theory is (in)deterministic. In contrast, Halvorson and Manchak want to focus on the second ingredient—that is, what mathematical facts are underlying the Hole Argument. Their argumentative strategy is as follows (2022, p. 3):

It seems that there are two mathematical claims that might be relevant here. The first claim is that there are distinct but isomorphic models. But that fact is not strong enough to support the rest of the argument. The second claim is that there

<sup>8</sup> It should be mentioned that Roberts (2020) recently challenged Earman and Norton’s view on the relationship between Leibniz Equivalence and the Hole Argument. He distinguishes two versions of Leibniz Equivalence, weak and strong. Weak Leibniz Equivalence states that “isomorphic mathematical structures can each be used with equal accuracy to represent a given physical situation (though not necessarily at once)”. Roberts observes that the weak version is irrelevant to the Hole Argument. The fact that  $\mathcal{M}_{hole}$  in one context can represent the same physical situation as  $\mathcal{M}'_{hole}$  in a different context is irrelevant for blocking radical indeterminism because the latter arises from the assumption that  $\mathcal{M}_{hole}$  and  $\mathcal{M}'_{hole}$  represent different physical situations in the same context, which is fully consistent with Weak Leibniz Equivalence. Strong Leibniz Equivalence states that “isomorphic mathematical structures can all be used with equal accuracy to represent a given physical situation, at once”. The strong version is able to block radical indeterminism but has counterexamples. However, all of Roberts’s counterexamples are such that at least one of the models does not represent an entire possible world, so they do not undermine the validity of Strong Leibniz Equivalence for models representing entire possible worlds. Here, I will understand Leibniz Equivalence (formulated as in the main text, i.e., with “diffeomorphic” in place of “isomorphic” and “represent” in place of “can be used to represent”) as a principle about *models representing entire possible worlds* in which case, I think, it is valid and not threatened by Roberts’s argumentation. I hope to analyse this issue at length elsewhere.



are isomorphisms that only move elements inside a hole. But that claim, as we show, is false.

This leads them to conclude that the Hole Argument is wrong because one cannot find appropriate mathematical facts that could serve as the second ingredient.

Halvorson and Manchak's analysis is based on the assumption that the notion of isomorphism plays a central role in the Hole Argument (2022, p. 10).<sup>9</sup> In their reading of Earman and Norton, the notion of isomorphism is needed both to establish the Gauge Theorem (which I called "step 1"; see Sect. 3.1) and to make the conclusion about indeterminism (which I called "step 2"; see Sect. 3.2).

Concerning step 1, according to Halvorson and Manchak (2022, pp. 13–14), the justification of the Gauge Theorem relies on the notion of isomorphism:  $\mathcal{M}$  and  $\mathcal{M}'$  are regarded as models of the same theory because they are regarded as isomorphic. Concerning step 2, it requires the precisification of the meaning of indeterminism. According to Halvorson and Manchak (2022, p. 12), the notion of indeterminism can be spelt out (using my notation, not theirs) in the following way:

**Definition 6** (*Indeterminism according to Halvorson and Manchak*) A theory  $T$  is indeterministic iff there are two models of  $T$ ,  $\mathcal{M} = \langle M; O_1, \dots, O_n \rangle$  and  $\mathcal{M}' = \langle M'; O'_1, \dots, O'_n \rangle$ , such that there is a proper open subset  $O$  of  $M$  (where  $M \setminus O$  contains some initial segment of  $M$ ), and an isomorphism  $\phi : M \rightarrow M'$  that changes things in  $O$  but not outside  $O$ .

This formulation is imprecise because we need to specify what it means that  $\phi$  changes (or does not change) things in  $O$ . For this purpose, we need the standard of comparison for the pair of models  $\mathcal{M}$  and  $\mathcal{M}'$ —that is, another map that tells us which points in the base manifold of  $\mathcal{M}'$  should be regarded as counterparts of which points in the base manifold of  $\mathcal{M}$ . Let us call this map  $\psi : M \rightarrow M'$ . Halvorson and Manchak (2022, p. 13) advocate that this map should be an isomorphism. Then, the definition of indeterminism becomes as follows:

**Definition 7** (*Indeterminism according to Halvorson and Manchak, more precise version with isomorphism as the standard of comparison*) A theory  $T$  is indeterministic iff there are two models of  $T$ ,  $\mathcal{M} = \langle M; O_1, \dots, O_n \rangle$  and  $\mathcal{M}' = \langle M'; O'_1, \dots, O'_n \rangle$ , such that there is a proper open subset  $O$  of  $M$  (where  $M \setminus O$  contains some initial segment of  $M$ ), and isomorphisms  $\psi : M \rightarrow M'$  and  $\phi : M \rightarrow M'$  such that  $\phi|_{M \setminus O} = \psi|_{M \setminus O}$  but  $\phi|_O \neq \psi|_O$ .

Notice that there are two isomorphisms invoked by Definition 7: the first isomorphism,  $\psi$ , is the standard of comparison, whereas the second isomorphism,  $\phi$ , is a map that "generates" indeterminism (i.e., changes things in  $O$  but not outside  $O$ , where "changing" and "not changing" are understood relative to  $\psi$ ). Another option they consider is that  $\psi$  should be the identity map  $\mathbb{1}$  on  $M$  (Halvorson & Manchak, 2022, p. 13), in which case the definition becomes<sup>10</sup>:

<sup>9</sup> I assume that they use the word "isomorphism" in the sense of Definition 1.

<sup>10</sup> Cf. Weatherall (2018), who disregards the choice of the identity map as the standard of comparison and endorses the choice of isomorphisms in this role.

**Definition 8** (*Indeterminism according to Halvorson and Manchak, more precise version with the identity map as the standard of comparison*) A theory  $T$  is indeterministic iff there are two models of  $T$ ,  $\mathcal{M} = \langle M; O_1, \dots, O_n \rangle$  and  $\mathcal{M}' = \langle M; O'_1, \dots, O'_n \rangle$ , such that there is a proper open subset  $O$  of  $M$  (where  $M \setminus O$  contains some initial segment of  $M$ ), and an isomorphism  $\phi : M \rightarrow M$  such that  $\phi|_{M \setminus O} = \mathbb{1}|_{M \setminus O}$  but  $\phi|_O \neq \mathbb{1}|_O$ .

Observe that in contrast to Definition 7, Definition 8 can be formulated only for models with the same base manifold (i.e.,  $M' = M$ ).

To sum up, the notion of isomorphism appears two or three times in my reconstruction of Halvorson and Manchak's interpretation of Earman and Norton's Hole Argument: once in the justification of the Gauge Theorem and once or twice in the definition of indeterminism (Definitions 7 or 8)—as a transformation between models that witnesses indeterminism (i.e.,  $\phi$ ) and as the standard of comparison (i.e.,  $\psi$ ), unless it is chosen to be the identity map. Having set the dispute in this way, Halvorson and Manchak argue that no candidate for the notion of isomorphism of GR models is suitable to play all these roles at the same time. For the sake of brevity, I will omit the details of their argumentation because they are not important for my subsequent analysis.

## 5 The two steps of the Hole Argument without the notion of isomorphism

In this section, I will argue that Earman and Norton's Gauge Theorem and the Hole Argument (or at least their slightly modified versions) can be formulated without using the notion of isomorphism. If I am right, then the criticism of Halvorson and Manchak does not have any bearing on the validity of this argument. In particular, the issue of what is the proper notion of isomorphism for GR models is irrelevant for the assessment of the Hole Argument understood in this way.

### 5.1 The notion of isomorphism is not needed for step 1 of the Hole Argument

In Sect. 3.1, I presented a detailed proof of Earman and Norton's Gauge Theorem. Halvorson and Manchak suggest that the notion of isomorphism is indispensable in the proof of the Gauge Theorem, but one can see that in Sect. 3.1 this notion was not used even once. The claim that "if  $\mathcal{M}$  is a model of a theory  $T$ , then  $\mathcal{M}'$  is also a model of  $T$ " is justified by Earman and Norton not by appealing to  $\mathcal{M}$  and  $\mathcal{M}'$  being isomorphic but by the fact that they satisfy the same dynamical equations.<sup>11</sup>

<sup>11</sup> Menon and Read (2024) argue independently for this claim, which they phrase in a different but arguably equivalent way: "the real physical content of general covariance is contained in the fact that solutionhood is preserved under diffeomorphisms, not merely that some diffeomorphisms are isometries" (Menon & Read, 2024, p. 5, footnote 7). Landsman (2023) also understands the Hole Argument dynamically. He relates it to the theorem of Choquet-Bruhat and Geroch (1969) about the existence and uniqueness of globally hyperbolic solutions to the Einstein field equations. As I use the concept of radical indeterminism instead of indeterminism defined as the existence of distinct solutions with the same initial data on a Cauchy surface (see Sect. 5.2), my approach also applies to solutions that are not globally hyperbolic.

The assumption that the last  $n - k + 1$  objects<sup>12</sup> (i.e.,  $O_k, \dots, O_n$ ) are tensor fields is used in the transition from (5) to (6), but the property of tensor fields that is exploited here is not that diffeomorphisms are their isomorphisms but that if their components vanish in one coordinate frame, then they vanish in any coordinate frame (which is independent of any claims about which transformations are isomorphisms of tensor fields). Therefore, the crucial notion in the Gauge Theorem is that of the dynamical equations of a theory and not the isomorphisms of its models.

However, there is a significant subtlety here, which will ultimately lead us to a modification of the notion of local spacetime theory. Even granting the above explanations, one could have the following worry. The Gauge Theorem relies on the assumption that Eqs. (1) and (7) are the same equations. However, it has not been made explicit which criterion of the identity of equations is used here. One candidate would be the isomorphism of the tensor fields that appear on the LHS of these equations. Under this approach, Eqs. (1) and (7) are said to be the same because  $O_i$  and  $\phi_* O_i$  are isomorphic (for  $i = k, \dots, n$ ). However, this approach is not available to us, as it smuggles the notion of isomorphism through the back door.

Another candidate for the identity criterion for equations uses the notion of the (functional) form of equations. Let us begin with a simple example. The equations of the Newtonian mechanics of  $n$  pointlike particles with gravity as the only force are as follows:

$$m_i \frac{d^2}{dt^2} \vec{x}_i(t) - \sum_{j \neq i} Gm_i m_j \frac{\vec{x}_i(t) - \vec{x}_j(t)}{|\vec{x}_i(t) - \vec{x}_j(t)|^3} = 0, \tag{8}$$

where  $i, j = 1, \dots, n$ .

It is well known that these equations are invariant under the symmetries of the Galilean group. I will present the idea of the form of equations using the simplest of these symmetries—namely, translations in space. Translations in space amount to a change of the spatial variable  $\vec{x} \mapsto \vec{x}' = \vec{x} + \vec{a}$ , where  $\vec{a}$  is some fixed spatial vector. This transformation acts on the first term in (8) as follows:

$$m_i \frac{d^2}{dt^2} (\vec{x}'_i(t)) = m_i \frac{d^2}{dt^2} (\vec{x}_i(t) + \vec{a}) = m_i \frac{d^2}{dt^2} \vec{x}_i(t) + m_i \frac{d^2}{dt^2} \vec{a} = m_i \frac{d^2}{dt^2} \vec{x}_i(t), \tag{9}$$

whereas on the second term in (8) it acts as follows:

$$\begin{aligned} \sum_{j \neq i} Gm_i m_j \frac{\vec{x}'_i(t) - \vec{x}'_j(t)}{|\vec{x}'_i(t) - \vec{x}'_j(t)|^3} &= \sum_{j \neq i} Gm_i m_j \frac{(\vec{x}_i(t) - \vec{a}) - (\vec{x}_j(t) - \vec{a})}{|(\vec{x}_i(t) - \vec{a}) - (\vec{x}_j(t) - \vec{a})|^3} \\ &= \sum_{j \neq i} Gm_i m_j \frac{\vec{x}_i(t) - \vec{x}_j(t)}{|\vec{x}_i(t) - \vec{x}_j(t)|^3}. \end{aligned} \tag{10}$$

From (9) and (10), it follows that whenever (8) is satisfied, the following “primed” equation is also satisfied:

<sup>12</sup> See the beginning of Sect. 3 for the meaning of the division of  $\{O_1, \dots, O_n\}$  into two classes.

$$m_i \frac{d^2}{dt^2} \vec{x}'_i(t) - \sum_{j \neq i} G m_i m_j \frac{\vec{x}'_i(t) - \vec{x}'_j(t)}{|\vec{x}'_i(t) - \vec{x}'_j(t)|^3} = 0. \quad (11)$$

The crucial observation here is that Eqs. (8) and (11) have the same form: (11) is just (8) with  $\vec{x}_i$  replaced by  $\vec{x}'_i$  and  $\vec{x}_j$  replaced by  $\vec{x}'_j$ . This is often taken as the defining feature of a symmetry: the claim that spatial translations are symmetries of the Newtonian mechanics of  $n$  pointlike particles can be understood as the claim that spatial translations do not change the form of the equations of this theory (and analogously for other symmetries; see, e.g., Brading & Castellani, 2007, pp. 1342–1343; Giulini, 2007, p. 108; Saunders, 2003, p. 299; cf. Pooley, 2017, pp. 114–120). This is sometimes called the “covariance” or “form-invariance” of the dynamical equations under symmetries (although these terms are also used with other meanings).<sup>13</sup>

One should distinguish here between two seemingly similar claims, one trivial and the other nontrivial. The trivial claim is that if we replace in the LHS of (8) all occurrences of  $\vec{x}$  by  $\vec{x}'$ , then we will obtain the LHS of (11), so such a replacement transforms the LHS of (8) into an expression of the same form. This trivial claim holds no matter what the relation is between  $\vec{x}$  and  $\vec{x}'$  (and even if they are entirely unrelated). In contrast, the nontrivial claim is that whenever equation (8) holds for  $\vec{x}$ , the equation of the same form holds for  $\vec{x}' = \vec{x} + \vec{a}$ . This nontrivial claim is true only for very special choices of  $\vec{x}'$ .<sup>14</sup> For example, if we had chosen  $\vec{x}' = \vec{x} + \vec{a}(t)$ , then we would get an additional term  $m_i \frac{d^2}{dt^2} \vec{a}(t)$  in (9), so that equation (11) would follow from (8) only if  $\frac{d^2}{dt^2} \vec{a}(t) = 0$  (i.e.,  $\vec{a} = \vec{a}_1 t + \vec{a}_2$ ), which is a restrictive constraint on  $\vec{a}(t)$ . For other choices of  $\vec{a}(t)$ , equation (11) would simply not follow from (8).

What does the form-invariance of equations look like in the case of GR? Consider the Einstein tensor. It is a function of the metric as well as its first and second derivatives—schematically,  $G_{ab}(g_{ab}, \partial g_{ab}, \partial^2 g_{ab})$ . Again, one should distinguish between two seemingly similar claims, one trivial and the other nontrivial. The trivial claim is that if we replace in  $G_{ab}(g_{ab}, \partial g_{ab}, \partial^2 g_{ab})$  all occurrences of  $g_{ab}$  by  $g'_{ab}$ , then we will obtain  $G_{ab}(g'_{ab}, \partial g'_{ab}, \partial^2 g'_{ab})$ , so such a replacement transforms  $G_{ab}(g_{ab}, \partial g_{ab}, \partial^2 g_{ab})$  into an expression of the same form. This trivial claim holds no matter what the relation between  $g_{ab}$  and  $g'_{ab}$  is (and even if they are completely unrelated). In contrast, the nontrivial claim is that  $\phi_* G_{ab}(g_{ab}, \partial g_{ab}, \partial^2 g_{ab}) = G_{ab}(\phi_* g_{ab}, \partial \phi_* g_{ab}, \partial^2 \phi_* g_{ab})$ , as it depends on the details of the functional form of  $G_{ab}$  and on  $\phi$  being a diffeomor-

<sup>13</sup> It might happen that the relation of having the same mathematical structure (i.e., of being isomorphic) and being related by the transformation that does not change the form of the dynamical equations of the theory (i.e., being symmetry-related) coincide extensionally (i.e., two models are related by one of them iff they are related by the other), but (i) this is not always the case (cf. the last paragraph of Sect. 7) and (ii) even if this is the case, the two relations are still conceptually different, despite having the same extension. If we adopt a liberal understanding of isomorphisms from category theory, then there will always be a category of models such that the symmetries of a theory count as isomorphisms (cf. Dewar 2019). However, this is not the reason why symmetries are distinguished, as there might also be other categories involving the same set of models but with a different choice of isomorphisms. The reason why symmetries are distinguished is their connection to the dynamical laws of the theory.

<sup>14</sup> And it is true because of the special features of Eq. (8): in the first term, the addition of constant  $\vec{a}$  is cancelled out by the derivation, whereas in the second term it is cancelled out by taking a difference.

phism. The nontriviality comes from the fact that this does not hold for all functions of the metric and its derivatives but only for special ones, such as  $G_{ab}$ .

More generally, for any local spacetime theory, the form-invariance of  $O_i$ 's (for  $i = k, \dots, n$ ), understood as functions of geometric objects  $O_1, \dots, O_{k-1}$  and their derivatives, under diffeomorphisms means that for any diffeomorphism  $\phi$ ,

$$\begin{aligned} &\phi_* O_i \left( O_1, \dots, O_{k-1}, \right. \\ &\quad \frac{\partial}{\partial x^1} O_1, \dots, \frac{\partial}{\partial x^N} O_1, \dots, \frac{\partial}{\partial x^1} O_{k-1}, \dots, \frac{\partial}{\partial x^N} O_{k-1}, \dots, \\ &\quad \left. \frac{\partial^N}{\partial x^1 \dots \partial x^N} O_1, \dots, \frac{\partial^N}{\partial x^1 \dots \partial x^N} O_{k-1} \right) \\ &= O_i \left( \phi_* O_1, \dots, \phi_* O_{k-1}, \right. \\ &\quad \frac{\partial}{\partial x^1} \phi_* O_1, \dots, \frac{\partial}{\partial x^N} \phi_* O_1, \dots, \frac{\partial}{\partial x^1} \phi_* O_{k-1}, \dots, \frac{\partial}{\partial x^N} \phi_* O_{k-1}, \dots, \\ &\quad \left. \frac{\partial^N}{\partial x^1 \dots \partial x^N} \phi_* O_1, \dots, \frac{\partial^N}{\partial x^1 \dots \partial x^N} \phi_* O_{k-1} \right). \end{aligned} \tag{12}$$

Similarly, the form-invariance of the equations of a local spacetime theory means that whenever

$$\begin{aligned} &O_i \left( O_1, \dots, O_{k-1}, \right. \\ &\quad \frac{\partial}{\partial x^1} O_1, \dots, \frac{\partial}{\partial x^N} O_1, \dots, \frac{\partial}{\partial x^1} O_{k-1}, \dots, \frac{\partial}{\partial x^N} O_{k-1}, \\ &\quad \dots, \\ &\quad \left. \frac{\partial^N}{\partial x^1 \dots \partial x^N} O_1, \dots, \frac{\partial^N}{\partial x^1 \dots \partial x^N} O_{k-1} \right) = 0 \end{aligned} \tag{13}$$

holds for  $i = k, \dots, n$ , then also

$$\begin{aligned} &O_i \left( \phi_* O_1, \dots, \phi_* O_{k-1}, \right. \\ &\quad \frac{\partial}{\partial x^1} \phi_* O_1, \dots, \frac{\partial}{\partial x^N} \phi_* O_1, \dots, \frac{\partial}{\partial x^1} \phi_* O_{k-1}, \dots, \frac{\partial}{\partial x^N} \phi_* O_{k-1}, \\ &\quad \dots, \\ &\quad \left. \frac{\partial^N}{\partial x^1 \dots \partial x^N} \phi_* O_1, \dots, \frac{\partial^N}{\partial x^1 \dots \partial x^N} \phi_* O_{k-1} \right) = 0 \end{aligned} \tag{14}$$

holds for  $i = k, \dots, n$ . This follows from the form-invariance of the LHSs of these equations, given by (12).

The above considerations suggest that for the Hole Argument to work, we need a modification of Earman and Norton's notion of local spacetime theory. To its characteristics reviewed in Sect. 3, one should add the following conditions:

- The last  $n - k + 1$  objects (i.e.,  $O_k, \dots, O_n$ ), which are the LHSs of the dynamical equations of the theory, are functions of  $O_1, \dots, O_{k-1}$  and their derivatives.
- $O_k, \dots, O_n$  (as functions of  $O_1, \dots, O_{k-1}$  and their derivatives) are form-invariant under diffeomorphisms.
- Two models  $\mathcal{M}$  and  $\mathcal{M}'$  are models of the same theory iff they involve the same number of geometric objects of each type (e.g., if  $O_i$  is a vector field, then  $O'_i$  is also a vector field) and their dynamical equations have the same form.<sup>15</sup>

In the original characterisation of local spacetime theories by Earman and Norton, nothing is said about  $O_k, \dots, O_n$  other than that they are tensors. However, this is not enough for our purposes. For these tensors to be the LHSs of the equations of a theory, they need to put some constraints on the behaviour of other objects that appear in the models of this theory (i.e.,  $O_1, \dots, O_{k-1}$ ), which is where the first new condition comes from. Once their functional form is identified, one can talk meaningfully about the invariance of this form under diffeomorphisms, which is the subject of the second condition.<sup>16</sup> Finally, all models of the theory should satisfy the same dynamical equations, where “the same” is understood as “having the same form” (which is not an arbitrary reading of “the same” because equations of the same form can be regarded as just the same equation written in a different notation).

One can ask how restrictive our additional conditions on local spacetime theories are. Consider the example that is of the foremost importance in the discussions about the Hole Argument—namely GR. GR is a local spacetime theory not only in Earman and Norton's broader sense reviewed in Sect. 3 but also in our more constrained sense. This is because  $G_{ab}$  is a function of the metric  $g_{ab}$  together with its first and second derivatives, and  $G_{ab}$  and  $T_{ab}$  do not change their form under diffeomorphisms. The latter means that whenever  $g_{ab}$  satisfies Einstein's equations  $G_{ab} - T_{ab} = 0$ , the transformed metric  $g'_{ab}$  satisfies the equations of the same form but with  $g_{ab}$  replaced by  $g'_{ab}$  in all its occurrences (both in  $G_{ab}$  and  $T_{ab}$ ) and  $O_i$  replaced by  $O'_i$  for  $i = 2, \dots, k - 1$  in all their occurrences. The same is true for the diffeomorphic-invariant formulation of Special Relativity (see, e.g., Pooley, 2017, pp. 120–121) because there the LHS of the dynamical equations is the Riemann curvature tensor, which is diffeomorphism-invariant. To tackle this issue more generally, we would need to know whether any

<sup>15</sup> Instead of this condition, Earman and Norton (1987, p. 517) have only the completeness condition, according to which if a local spacetime theory  $T$  has models of a certain form that satisfy certain equations, then any model of this form that satisfies these equations is also a model of  $T$ . However, this does not exclude that  $T$  also has models of different form or satisfying different equations, which I think should be excluded.

<sup>16</sup> One can object that if we assume the diffeomorphism-invariance of the equations at the outset, the intermediary steps in the proof in Sect. 3.1 are redundant, as we could have moved straight from (1) to (7). The validity of this objection depends on what exactly is assumed to be diffeomorphism-invariant: the equations of the theory or the tensors  $O_k, \dots, O_n$ . If the former, then these intermediary steps are indeed not needed, but if the latter, then these steps are indispensable, as from the fact that  $\phi_* O_i$  has the same form as  $O_i$  (for  $i = k, \dots, n$ ) it does not automatically follow that whenever  $O_k = 0, \dots, O_n = 0$  holds,  $\phi_* O_k = 0, \dots, \phi_* O_n = 0$  also holds. In the main text, the latter variant is chosen.

tensor that is a function of some geometric objects and their derivatives is form-invariant under diffeomorphisms. If this is so, then Earman and Norton's definition of local spacetime theories together with my first and third conditions entail the second condition, so it does not add any genuinely new constraint. However, I am not aware of any proof that this is indeed true.

## 5.2 The notion of isomorphism is not needed for step 2 of the Hole Argument

According to Halvorson and Manchak, to define indeterminism precisely, we need to use the notion of isomorphism at least once (Definition 6) and perhaps even twice (Definition 7). If they are right, then step 2 of the Hole Argument relies on the notion of isomorphism; but I will argue that this is not the case.

It seems that to decide whether two models related by a hole diffeomorphism are a witness of indeterminism, we should define precisely the notions of determinism and indeterminism. However, this task is beset with difficulties. A general idea of indeterminism (as formulated for physical theories) is that two models of a theory are a witness of indeterminism iff they agree on some (sufficiently large) region but disagree elsewhere. To make this precise, one should clarify what "agreement" and "disagreement" mean in this context and specify what kind of "sufficiently large regions" should be taken into account. Starting with the second issue, let us call such regions " $S$ -regions". An idea explored in the mathematical physics literature is that  $S$ -regions should be Cauchy surfaces (see, e.g., Choquet-Bruhat & Geroch, 1969; Ringström, 2009; cf. Landsman, 2023), but this option is available only for globally hyperbolic models of GR, which form a proper subclass of all models of GR. This is especially problematic because there are good candidates for witnesses of the indeterminism of GR that are not globally hyperbolic (see, e.g., Doboszewski, 2017, 2019). Another idea is to use a family of notions of (in)determinism, one for each choice of  $S$ -regions (see, e.g., Butterfield, 1989, pp. 7–9 and Doboszewski, 2017, pp. 10–11, who define  $S$ -determinism instead of determinism *simpliciter*).

Fortunately, the family of models generated by hole diffeomorphisms from some given model  $\mathcal{M}$  is so specific that it would count as a witness of indeterminism for all reasonable definitions of this notion. This is because the hole diffeomorphism can be defined no matter how small the "hole".<sup>17</sup> Therefore, no matter how large the  $S$ -regions are, there will always be a hole diffeomorphism that is the identity on some  $S$ -region and is different from the identity elsewhere, which is why this kind of indeterminism has been termed "radical" by Earman and Norton.<sup>18</sup> What is more, we do not even need the full definition of radical indeterminism here—it suffices that we provide a partial definition (which specifies a sufficient condition for radical indeterminism), as long as it is satisfied in all cases of interest.

<sup>17</sup> By the "hole" I do not mean here the region without matter (as was the case in Einstein's original formulation of the Hole Argument) but just the region on which a given hole diffeomorphism is not the identity map.

<sup>18</sup> Observe that among other differences between indeterminism and radical indeterminism, there is this one: an instance of the former is a pair of models, whereas an instance of the latter is a family of models such that one of them is the "original" model and the others are generated from it by applying to it hole diffeomorphisms with various possible choices of the "hole".



Another problem is how to tell whether two models (or their parts) “agree” or “disagree”—that is, whether they are (physically) equivalent. Weatherall (2018) uses the notion of the “standard of sameness” or “standard of comparison”; the latter term is also adopted by Halvorson and Manchak. The standard of comparison is a kind of maps  $K$  that determine which points in a given model should be compared with which points in another model in assessing whether these two models are the same or different.<sup>19</sup> If there is no map of kind  $K$  between two models, then they are surely different; but if there is a map of kind  $K$  between two models, then in comparing these models point by point we should use this map. The problem is that it seems that we should decide in advance what is the standard of comparison before we even start assessing whether a class of models is an instance of radical indeterminism and this choice has significant influence on our verdict. According to Weatherall and Halvorson and Manchak, the notion of isomorphism is crucial here because for each theory, the isomorphisms of its models should serve as the standard of comparison. However, as I will argue, the impression that we need to choose in advance the kind of maps  $K$  that serve as the standard of comparison is wrong. Finding  $K$  is a part of our task of analysing the theory instead of being the starting point of such an analysis, and the Hole Argument has a role to play in finding  $K$  for local spacetime theories. What we can do is to choose  $K$  tentatively and see what the consequences of a given choice are; and if some choice has unfavourable consequences, we should try another one. For this purpose, we can relativise our partial definition of radical indeterminism to the choice of the kind of maps  $K$  that serve as the standard of comparison between models of a given theory.

What is the meaning of this choice of the standard of comparison between models? I think that it should always be regarded as a tentative proposal (i.e., a candidate) for the relation of physical equivalence. This is because we are not interested here in models of a given theory as abstract mathematical entities but as representations of physical possibilities. If we choose as our standard of comparison a kind of maps  $K$ , then models that do not differ as compared by a map of kind  $K$  would count as the same relative to this choice; but this choice would be inappropriate in the context of physics if the equivalence according to maps of kind  $K$  did not coincide with physical equivalence. Therefore, the standard of comparison of models and the standard of physical equivalence should be thought of as two aspects of the same notion.<sup>20</sup> However, this does not mean that by tentatively choosing a certain kind  $K$ , we thereby settle the issue of which models are physically equivalent. This is only a

---

<sup>19</sup> This is similar to the idea of counterparts in Lewisian modal metaphysics (cf. Lewis, 1968). However, our question also makes sense outside the context of Lewis’s philosophy, as it concerns models rather than possible worlds, and the former do not need to be in one-to-one correspondence to the latter. In fact, once we decide on the standard of comparison, we can regard models that are the same according to it as representing the same possible world.

<sup>20</sup> Gomes and Butterfield (2023) argue that there does not exist a single kind of maps that should be used to compare spacetime points in all situations. They give examples of contexts in which points are compared by means of a map that is not an isomorphism (the definition of the Lie derivative, Noether’s second theorem and limits of spacetimes). However, I think that their observations do not undermine the claim that there exists a single kind of maps that should be used to compare models *when our aim is to establish their physical equivalence*; this is consistent with other kinds of maps being validly used to compare models for other purposes. Therefore, I take it that their claims are not in tension with the suggestion in the main text that for any theory there exists a single kind of maps that is the right standard of comparison (in this sense, related to physical equivalence), and one of our aims in interpreting this theory is to find it.



tentative proposal put forward for investigation, and whether a given choice of  $K$  is accurate depends on the nature of the physical world and not on our conventions.

I would like to propose the following partial definition of radical indeterminism of a theory relative to maps of kind  $K$ :<sup>21,22</sup>

**Definition 9** (*Partial definition of radical indeterminism relative to maps of kind  $K$* )

If a theory  $T$  has a model  $\mathcal{M} = \langle M; O_1, \dots, O_n \rangle$  such that there exists an infinite series of regions of  $M$ ,  $H_0 \supseteq H_1 \supseteq H_2 \supseteq \dots$ , such that (i) for any  $\epsilon > 0$ , there exists  $H_j$  that is included in an open ball with a diameter  $\epsilon$  or smaller,<sup>23</sup> and (ii) for any  $H_j$ , there exists another model of  $T$ ,  $\mathcal{M}' = \langle M; O'_1, \dots, O'_n \rangle$ , with the following properties:

- (1) there is a map between these models,  $\phi : M \rightarrow M$ , such that  $\phi|_{M \setminus H_j}$  is a map of kind  $K$ ,
- (2) but there is no map of kind  $K$  between these models (i.e., on the whole  $M$ ),

then  $T$  is radically indeterministic relative to maps of kind  $K$ .

This definition expresses the following intuition: the radical indeterminism of a theory means that no matter how large is the subset of  $M$  on which we specify the values of physical quantities (i.e., no matter how large is  $M \setminus H_j$ ), this would be not sufficient to determine the values of those quantities on the entire  $M$ . Importantly, this definition is relative to the choice of  $K$ , so the same theory might be radically indeterministic under some choices of  $K$  but not under others.

Earman and Norton's version of the Hole Argument can be conceived as using the above partial definition of radical indeterminism with the full identity as  $K$  (where the difference between the full identity and the standard identity map is that the latter preserves only points of the manifold, whereas the former also preserves all geometric objects at these points). For this choice of  $K$ , the above partial definition becomes:

<sup>21</sup> This definition is to some extent similar to Butterfield's (1989) definition of determinism Dm2 in that it involves two maps, the first of which relates the corresponding regions of two models and the second of which relates the two models taken as wholes. In contrast, it is significantly different from Definition 7 used by Halvorson and Manchak (see Sect. 6).

<sup>22</sup> It has been suggested to me that such a relativisation of the notion of indeterminism to the choice of  $K$  is problematic because indeterminism is supposed to be a metaphysical notion. Although I agree with classifying this notion as metaphysical, I also assume that the only possible way of answering the question of whether our world is (in)deterministic is by examining the question of whether our physical theories are (in)deterministic. As it turns out, the interpretation of these theories is often not straightforward; in particular, from the bare formalism of a theory it does not follow which states should be regarded as physically equivalent and which not, which is of pivotal importance for our question. Therefore, I propose to proceed by examining various candidates for the relation of physical equivalence between the states instead of choosing one of them in advance; the consequences of each choice should help us to decide which of them is the most plausible one. In this approach, the questions of indeterminism and of physical equivalence (which are both metaphysical and not purely formal) are considered jointly and not one before the other.

<sup>23</sup> This condition captures the idea that a hole can be chosen to be arbitrarily small. We define  $\epsilon$  in terms of a metric. Alternatively, we can postulate that the series  $\{H_j\}$  converges to an empty set, in which case we can express condition (i) using only topology, without referring to a metric. Despite the fact that the limit of this series (defined topologically) is an empty set, every element of it is a region, which I take to be by definition non-empty; therefore, for every  $H_j$ , two models differing on  $H_j$ ,  $\mathcal{M}$  and  $\mathcal{M}'$ , will always have different values of geometric object fields at some points of  $M$ .

**Definition 10** (*Partial definition of radical indeterminism relative to the full identity*)

If a theory  $T$  has a model  $\mathcal{M} = \langle M; O_1, \dots, O_n \rangle$  such that there exists an infinite series of regions of  $M$ ,  $H_0 \supseteq H_1 \supseteq H_2 \supseteq \dots$ , such that (i) for any  $\epsilon > 0$ , there exists  $H_j$  that is included in an open ball with a diameter  $\epsilon$  or smaller, and (ii) for any  $H_j$ , there exists another model of  $T$ ,  $\mathcal{M}' = \langle M; O'_1, \dots, O'_n \rangle$ , with the following properties:

- (1) there is a map between these models,  $\phi : M \rightarrow M$ , such that  $\phi|_{M \setminus H_j}$  is the full identity (i.e., for any  $p \in M \setminus H_j$ ,  $\phi(p) = p$  and  $O'_i|_{M \setminus H_j} = \phi_* O_i|_{M \setminus H_j} = O_i|_{M \setminus H_j}$  for  $i = 1, \dots, n$ ),
- (2) but there is no full identity map between these models (i.e., there is no map  $\phi : M \rightarrow M$  such that for any  $p \in M$ ,  $\phi(p) = p$  and  $O'_i = \phi_* O_i = O_i$  for  $i = 1, \dots, n$ ),

then  $T$  is radically indeterministic relative to the full identity.

To see that Definition 10 is satisfied in the case considered by Earman and Norton, it suffices to consider, as they originally did, a series of “holes”  $\{H_j\}$  and, for each hole  $H_j$ , a diffeomorphism  $\phi$  that is the identity on  $M \setminus H_j$  and is not the identity on  $H_j$ . This diffeomorphism  $\phi$  can be used to generate a model  $\mathcal{M}'$  from a given model  $\mathcal{M}$ .<sup>24</sup> Since the restriction of  $\phi$  to  $M \setminus H_j$  is the full identity, condition (1) of Definition 10 is satisfied; and since there is no full identity between  $\mathcal{M}$  and  $\mathcal{M}'$ , condition (2) is also satisfied. It should be noted that even if the metric is the same at every point of  $M$ , a non-trivial hole diffeomorphism can be constructed (see Example 1 in Manchak & Barrett, 2023 for such a construction for Minkowski spacetime).

If using Definition 9 relative to  $K$  we obtain the conclusion of radical indeterminism, we might wish to change our choice of the kind of maps that serve as the standard of comparison of models (call this new choice  $K'$ ), so that the models witnessing radical indeterminism relative to  $K$  would count as physically equivalent in light of  $K'$ . This  $K'$  should be chosen so that  $K \subseteq K'$ , and the maps that have been used to generate witnesses of radical indeterminism should also belong to  $K'$  (otherwise, we would still retain the same witnesses of radical indeterminism, which is what we wanted to eliminate). However, we are not forced to reach this conclusion. In particular, if we know that models  $\mathcal{M}$  and  $\mathcal{M}'$  are empirically inequivalent (that is, they can be distinguished by some observation or experiment), then we surely should not regard them as physically equivalent. In such a case, radical indeterminism would need to be

<sup>24</sup> The word “generate” might misleadingly suggest that in this way we create a new model, which previously was not there. However, this should rather be seen as a way of investigating the class of models of a given theory, which is determined by the theory’s equations. Therefore, by applying a diffeomorphism to a given model  $\mathcal{M}$  and generating in this way a new model  $\mathcal{M}'$ , we do not add anything to the class of models of the theory but only discover that it has such diffeomorphism-related models. For this reason, I do not think that one can “block” the Hole Argument before it starts by hypothesising that the hole diffeomorphism is a mathematical operation without physical significance and as such should not be performed on models (see Curiel, 2018, pp. 452–453). We (as researchers) can stop the activity of applying diffeomorphisms to already known models of the theory, but this would not prevent the existence of diffeomorphism-related models in the class of all models of this theory.

regarded as an empirical hypothesis and not a methodological vice of our theory.<sup>25</sup> Perhaps there might also be other obstacles to regarding such models as physically equivalent, even if they are empirically equivalent.<sup>26</sup> However, if there are no such obstacles, the overall conclusion of the argument is that models related by maps of kind  $K'$  are physically equivalent. The above reasoning might, of course, be repeated for our new choice of the standard of comparison  $K'$ . If it turns out that our theory is still radically indeterministic relative to  $K'$ , we should change  $K'$  to another kind of maps  $K''$  (provided that there are no obstacles to doing this, such as the empirical inequivalence of two models related by a map of kind  $K''$ ) and so on until we do not encounter any new witnesses of radical indeterminism. Therefore, our reasoning here is iterative and is an instance of a broader philosophical idea of reflective equilibrium.

The notion of isomorphism does not occur in the above reasoning. This (together with the results of Sect. 5.1) establishes the claim that appears in the title of this paper: namely, that the Hole Argument does not (or at least does not need to) rely on this notion. However, one can worry that even though the notion of isomorphism is not used explicitly in our reasoning, perhaps it is invoked there implicitly. For example, the following objection has been raised to my approach. The mere fact that in addition to the original model  $\mathcal{M}$  the laws of the theory allow another model  $\mathcal{M}'$  does not in itself show that this theory is indeterministic—having more than one model is not enough for indeterminism. In addition, one needs to specify the relation between  $\mathcal{M}$  and  $\mathcal{M}'$  that makes this pair of models a witness of indeterminism, and this—the argument goes—cannot be done without referring to the notion of isomorphism.

The two features of my formulation of the Hole Argument that are crucial for seeing that we do not need the notion of isomorphism here is its reliance on the notion of physical equivalence and its iterative nature. According to Definition 9, the relation that makes a pair of models a witness of indeterminism is the existence of a map of kind  $K$  between proper parts of  $\mathcal{M}$  and  $\mathcal{M}'$  together with the non-existence of any map of kind  $K$  between  $\mathcal{M}$  and  $\mathcal{M}'$  (taken as wholes). Since maps of kind  $K$  are taken (tentatively) at a given stage of our iterative procedure as a standard of physical equivalence, this means that proper parts of  $\mathcal{M}$  and  $\mathcal{M}'$  are physically equivalent in light of this standard, whereas  $\mathcal{M}$  and  $\mathcal{M}'$  are not physically equivalent in light of this standard. If for arbitrarily large proper parts of  $\mathcal{M}$  one can find such  $\mathcal{M}'$ , then

<sup>25</sup> For this reason, I think that the two dilemmas that Earman and Norton present as two separate motivations for regarding models related by a hole diffeomorphism as physically equivalent (i.e., the verificationist dilemma and the indeterminism dilemma) are not in fact independent: the radical indeterminism that arises in the Hole Argument (which is the subject of the indeterminism dilemma) is methodologically problematic *because* the models that are its witnesses are empirically indistinguishable (which is the subject of the verificationist dilemma). At this point, I disagree with Pooley and Read (2021), who want to separate what they call the “underdetermination version of the argument” and the “indeterminism version of the argument” (closely corresponding to Earman and Norton’s two dilemmas). I think that even though we do not need (empirical) underdetermination to establish (radical) indeterminism, in the absence of such underdetermination, this indeterminism would be just an empirical hypothesis and not something inherently problematic. If this empirical hypothesis found no support in our observations and experiments, then this, I think, would be a sign that our theory (rather than its interpretation) is wrong and that we should look for a different one (with different dynamical equations).

<sup>26</sup> For example, Møller-Nielsen (2017) suggests that two models should be regarded as physically equivalent only if there is available a perspicuous account of their shared ontology—their empirical equivalence is not sufficient.

the theory under consideration is said to be radically indeterministic. Therefore, for radical indeterminism we require much more than just the existence of two models of the same dynamical equations—they need to be related in a particular way, which is defined in terms of maps of kind  $K$ . Importantly, the notion of physical equivalence is not the notion of isomorphism in disguise: at any stage of our iterative procedure we may consider maps that are not mathematical isomorphisms of the theory's models, and there is no guarantee that at the end of this procedure  $K$  will coincide with the class of such isomorphisms. Since the procedure is iterative, we are not presupposing that a given kind of maps  $K$  coincides with the actual physical equivalence relation: we tentatively choose some  $K$ , investigate what are the consequences of this choice (among which there might be the radical indeterminism of our theory relative to  $K$ ), and then, on this basis, we adjust our choice.

A general picture that emerges from these considerations is that physical equivalence can be established by means of an iterative procedure in which we examine various candidate classes of maps, and, depending on the outcomes, we may need to broaden or narrow these classes. The exact criteria that should be used in assessing the proposed kinds of maps are a matter of debate. In this paper, one such criterion has been spelled out: if for a given  $K$  our theory  $T$  turns out to be radically indeterministic relative to  $K$  in the sense of Definition 9 and models related by maps of kind  $K$  are empirically indistinguishable, then this is a reason to replace  $K$  with some  $K'$  such that  $K \subsetneq K'$  and such that the family of models that is a witness of radical indeterminism is no longer a witness of radical indeterminism of  $T$  relative to  $K'$ .<sup>27</sup> In this iterative procedure, it might also turn out that  $K$  is too broad a class of maps. For example, if models related by  $K$  are empirically distinguishable, then we need to choose a narrower class, while taking into account the outcomes of previous steps (in order to avoid restoring the previously eliminated cases of radical indeterminism). The Hole Argument can be viewed as a particular instance of this procedure, where  $T$  is GR (or any local spacetime theory) and  $K$  is the full identity.<sup>28</sup>

Why not just choose  $K$  at the very beginning to be the class of maps that induces the empirical equivalence relation? There are several reasons for not doing this. First, it might not be obvious from the start which models are empirically equivalent and which are not, and establishing the empirical equivalence relation might be a part of a larger work that also involves various formal and conceptual analyses of the theory. Second (and relatedly), it might happen that we first encounter a witness of the radical indeterminism of our theory (at the level of formal calculations) and only after that we ask ourselves whether models constituting it are empirically equivalent or not. This order of thought is correctly captured by the procedure described here.

<sup>27</sup> This formulation is cautious in the sense that this conjunction of radical indeterminism and empirical indistinguishability is said to be only a reason for changing  $K$ , which is consistent with this reason being outweighed by other reasons for not changing  $K$ . It is a further question whether this reason is always sufficient.

<sup>28</sup> Another instance can be found, arguably, in classical electromagnetism. If we start with a theory formulated in terms of the electromagnetic potential  $A_\mu$  and choose  $K$  to be the full identity, then we can construct a family of models that are empirically indistinguishable and are a witness of radical indeterminism in the sense of Definition 9. This suggests another choice of  $K$ : namely, models related by a gauge transformation  $A_\mu \mapsto A'_\mu = A_\mu + \partial\phi$  should also be regarded as physically equivalent (even though they are not isomorphic because they differ in the value of  $A_\mu$ ).

Third, there might be reasons for regarding empirically equivalent models as physically inequivalent (cf. footnote 26), or at least we should not exclude this in advance.

An important question is where the above iterative procedure ends up in the case of GR. Should we regard any two diffeomorphism-related models (in the sense of Definition 4) as physically equivalent? I do not attempt to settle this issue here. There are suggestions in the literature that our standard of physical equivalence should be a class of maps narrower than all diffeomorphisms. For example, Johns (2019) claims that only diffeomorphisms smoothly connected to the identity can be used to run the Hole Argument; and Belot (2013) argues that in certain contexts models related by a diffeomorphism are physically inequivalent if this diffeomorphism does not leave invariant a certain fixed structure (but see Luc, 2022 for the criticism of this view in the case of models considered as representing entire possible worlds).

If indeed the physical equivalence of GR models is given by the class of all diffeomorphisms (or all diffeomorphisms smoothly connected to the identity), then the final result of the iterative procedure formulated in this section is not at all revolutionary. However, it was not my aim to suggest a new answer to the question about which GR models are physically equivalent but rather to explicitly formulate a systematic way of settling questions of this kind. I believe that something close to what I called here an “iterative procedure” is in fact used implicitly by many researchers interested in physical equivalence, but I find it valuable to make this procedure explicit and more precise.

## 6 Some further comments on papers by Halvorson and Manchak (2022) and Weatherall (2018)

In this section, I collect a few remarks that further clarify the relationship between the view on the Hole Argument developed in Sect. 5 and papers by Halvorson and Manchak (2022) and Weatherall (2018). First, I will analyse the relationship between Theorem 1 by Halvorson and Manchak and the existence of the hole diffeomorphism that leads to radical indeterminism relative to the full identity (in the sense of Definition 10). Second, I will compare my partial definition of radical indeterminism (Definition 9) with Halvorson and Manchak’s definition of indeterminism (Definition 7). Third, I will show how my approach can be used to reply to an earlier criticism of the Hole Argument by Weatherall. Finally, I will make some comments on the usage of category theory in the discussion of the Hole Argument.

It has been suggested to me that the main technical result of Halvorson and Manchak, their Theorem 1, establishes that there is no hole diffeomorphism that acts in the way needed to satisfy Definition 10. However, the content of that theorem is tangential to what I have argued for in Sect. 5. To recall, Theorem 1 (Halvorson & Manchak, 2022, p. 18) states that if  $\phi$  and  $\psi$  are isometries (in the sense of Definition 3) between models  $\mathcal{M} = \langle M, g \rangle$  and  $\mathcal{M}' = \langle M', g' \rangle$  such that  $\phi|_O = \psi|_O$  for some nonempty subset  $O$  of  $M$ , then  $\phi = \psi$ . The models considered in this theorem in general might have different base sets, in which case we cannot use Definitions 9 and 10 (and the full identity map cannot be defined). Therefore, in order to relate Theorem 1 to my considerations, we need to assume that  $M = M'$ . The next difference is

that Theorem 1 concerns two maps between models, whereas Definitions 9 and 10 concern the existence of a map of kind  $K$  between proper parts of models and the non-existence of a map of kind  $K$  between these models taken as wholes—at no point is a comparison of two maps of kind  $K$  considered. Finally, Theorem 1 does not say anything about full identity maps; it is only about isometries. In particular, condition  $\phi|_O = \psi|_O$  does not guarantee that  $\phi$  and  $\psi$  are full identity maps on  $O$ . Incidentally, even though Halvorson and Manchak’s Theorem 1 is about isometries, it does not help to test whether Definition 9 holds for  $K$  being the class of isometries because its *presupposition* is that there exists an isometry between  $\mathcal{M}$  and  $\mathcal{M}'$ , while here we ask *whether* there exists such an isometry, given that there exists an isometry between (sufficiently large) proper parts of  $\mathcal{M}$  and  $\mathcal{M}'$ .<sup>29</sup>

This leads us to the issue of the relationship between Definition 9 (the partial definition of radical indeterminism proposed in this paper) and Definition 7 (the definition of indeterminism used by Halvorson and Manchak). I claim that these definitions are significantly different in that Definition 7 can be applied only to models between which there exists an isomorphism. This arguably results in some incorrect verdicts concerning (in)determinism. If we take isomorphisms of GR models to be isometries in the sense of Definition 3 (which is their preferred choice), then some important examples of the indeterminism of GR, such as non-isometric extensions of Taub-NUT (see Chruściel & Isenberg, 1993, Doboszewski, 2017, pp. 200–202), will be “invisible” for their definition (i.e., not classified as an instance of indeterminism) because there is no isometry between them. Non-isometric extensions of Taub-NUT will also be invisible for my Definition 9 if we choose  $K$  to be isometries (in the sense of Definition 3). However, this is a correct result because they are instances of “usual” indeterminism and not of radical indeterminism, and as such should not motivate us to change our tentative standard of physical equivalence  $K$ . I suppose that the root of the problem with Definition 7 in this context is that it relies on the intuition that in order to tell that two models differ (which is required for them to be a witness of indeterminism), we need to be able to compare them, and such a comparison can be done only by means of some map (which according to Halvorson and Manchak should be an isomorphism). However, there is another way of telling that two models are different—namely, by showing the *non-existence* of an appropriate map between them; this is (rightly, I think) captured by Butterfield (1989) Dm2 and my Definition 9 and (wrongly, I think) neglected by Halvorson and Manchak’s Definition 7. Indeterminism is then the case where the models are not the same (there is no map of kind  $K$  between them) but their sufficiently large parts are the same (there is a map of kind  $K$  between these parts).

This aspect of Halvorson and Manchak’s approach can be traced back, I think, to the paper by Weatherall (2018). Let me then briefly reconstruct Weatherall’s view here. According to him, the Hole Argument relies on a confusion arising from using two

<sup>29</sup> Corollary 1 of Theorem 1 (Halvorson & Manchak, 2022, p. 18) attempts to exclude the existence of a hole diffeomorphism, but it is not applicable to situations that are relevant for the Hole Argument, as it is understood in Sect. 5. It states that if  $\mathcal{M} = \langle M, g \rangle$  is a model of GR and  $\phi : M \rightarrow M$  is an isometry (between  $\mathcal{M}$  and itself) that is the identity on some non-empty open subset  $O$  of  $M$ , then  $\phi = \mathbb{1}_M$  (i.e.,  $\phi$  is the identity on entire  $M$ ). However, Corollary 1 is derived from Theorem 1 by taking  $M = M'$ ,  $g = g'$  and  $\psi = \mathbb{1}_{M'}$ , and in the cases relevant for the Hole Argument the last two assumptions do not hold: we have  $g \neq g'$  and  $\mathbb{1}_M$  is not an isometry between  $\mathcal{M}$  and  $\mathcal{M}'$ .



different standards of comparison of models at the same time. On the one hand, the isometry (in the sense of Definition 3) between  $\mathcal{M} = \langle M, g_{ab} \rangle$  and  $\mathcal{M}' = \langle M, g'_{ab} \rangle$  is used to establish that they are physically equivalent; and on the other hand, the identity map on  $M$  is used to establish that they are different. However, Weatherall (2018, p. 338) claims that

[...] one cannot have it both ways. Insofar as one wants to claim that these Lorentzian manifolds are physically equivalent, or agree on all observable/physical structure, one has to use [the isometry map] to establish a standard of comparison between points. And relative to this standard, the two Lorentzian manifolds agree on the metric at every point—there is no ambiguity, and no indeterminism. (This is just what it means to say that they are isometric.) Meanwhile, insofar as one wants to claim that these Lorentzian manifolds assign different values of the metric to each point, one must use a different standard of comparison. And relative to this standard—that given by [the identity map on  $M$ ]—the two Lorentzian manifolds are not equivalent. One way or the other, the hole argument seems to be blocked.

How can one respond to this criticism of the Hole Argument in light of the considerations of Sect. 5? First, I agreed there that the standard of comparison of models should be closely associated with physical equivalence. However, the former is a matter of our choice, whereas the latter is not. We cannot turn physically equivalent models into physically inequivalent ones (or the other way around) merely by changing the way of comparing them. Pooley and Read (2021) express this observation by saying that the physical equivalence of two models is a “feature that holds full stop” and not a feature that holds relative to the choice of a map used to compare these models. Whether two models are equivalent or not depends partially on our representational conventions and partially on what physical differences there are in the world. For this reason, I said (in Sect. 5.2) that the standard of comparison is a *tentative proposal* for the relation of physical equivalence and not just the relation of physical equivalence. In my approach, the (true) relation of physical equivalence is something that we are going to find by using the iterative procedure, in which the Hole Argument has an important role to play. Therefore, the choice of the full identity as  $K$  in Definition 9 does not make any two models that are not related by the full identity physically inequivalent; in our iterative procedure, we hypothetically assume that they indeed are inequivalent, but the aim of this procedure is to examine precisely this hypothesis.

Second, according to Weatherall (who is followed by Halvorson and Manchak) if maps of kind  $K$  are the standard of comparison for a given theory, then in order to compare two models of that theory,  $\mathcal{M}$  and  $\mathcal{M}'$ , we should first find a map of kind  $K$  between them and then compare  $\mathcal{M}$  and  $\mathcal{M}'$  point by point using this map. The drawback of this approach is that in many situations the models will turn out to be incomparable because there is no map of kind  $K$  between them. In fact, since Halvorson and Manchak’s Theorem 1 shows that for any pair of Lorentzian manifolds there is at most one isometry between them, if we choose isometry to be the standard of comparison (understood in the Weatherall–Halvorson–Manchak way), any two Lorentzian manifolds will turn out to be either equivalent or incomparable. However, the verdict one should expect in the case of non-isometric Lorentzian manifolds is

that they are inequivalent and not that they are incomparable. That difference between “incomparable” and “inequivalent” might seem to be purely verbal and unimportant, but when applied to the issue of (in)determinism, it leads (arguably) to some incorrect verdicts, as explained in the third paragraph of this section.

The final issue to be considered in this section is an objection that has been raised to my proposal in category-theoretic terms. According to this objection, my “suggestion that one could use different standards of equivalence than isomorphism is to deny (...) that the category used to represent the models of such theory is the correct one”, which “does nothing to show that the standard form of GR [based on the category of Lorentzian manifolds with isometries as isomorphisms] is amenable to the Hole Argument”. This is close in spirit to Weatherall (2018, p. 331), who claims that “the default sense of ‘sameness’ or ‘equivalence’ of mathematical models in physics should be the sense of equivalence given by the mathematics used in formulating those models”, which is typically “some form of isomorphism”, where the term “isomorphism” is used “in the broad sense of category theory”. In the case of GR, Weatherall (2018, p. 343) says,

once one asserts that spacetime is represent[ed] by a Lorentzian manifold, one is committed to taking isometric spacetimes to have the capacity to represent the same physical situations, since isometry is the standard of isomorphism given in the mathematical theory of Lorentzian manifolds. To deny this would be, in effect, to insist that it is some other structure—one that is not preserved by isometries [i.e., different than a Lorentzian manifold]—that represents spacetime in relativity theory.

Therefore, in the end, the two ways of blocking the Hole Argument from the previous quote by Weatherall (2018, p. 338) are not equally good according to him—only the one that relies on isometries as the standard of comparison of Lorentzian manifolds is in agreement with scientific practice.

In response, let me make four points. First, as already stressed, I distinguish mathematical equivalence (encoded in the isomorphisms between the models) from physical equivalence—even if they happen to coincide, they are, from the conceptual point of view, two different equivalence relations. Therefore, establishing the mathematical equivalence relation on the class of models does not automatically lead to the conclusions concerning the physical equivalence relation on that class.

Second, the formulation of the iterative procedure does not presuppose anything about its outcome. In particular, it does not exclude that our conclusion about the physical equivalence relation appropriate for GR models obtained by means of this iterative procedure will be the same as that obtained by Weatherall, Halvorson and Manchak (and many others), as the choice of  $K$  other than the class of isomorphisms is here only tentative and subject to revision in light of the outcomes of relativising Definition 9 to this particular  $K$ . The difference is that now this conclusion will be reached using our iterative procedure, which relies on the dynamical equations of the theory and our partial definition of radical indeterminism but does not presuppose that isomorphisms are the standard of comparison and physical equivalence. Therefore, I do not “deny (...) that the category used to represent the models of such theory is the correct one”, although I also do not exclude this possibility in advance.



Third, even if we assume that the standard of comparison of models should always coincide with their isomorphisms, the iterative procedure described in Sect. 5 can still be useful but needs to be reinterpreted. Namely, instead of thinking about a fixed class of models and changing  $K$ 's, we should think about changing a category of models with the choice of  $K$  being a part of the specification of that category. Then, the question "Which kind of maps  $K$  should be the standard of comparison for the class of models of GR?" is replaced with the question "What should be the kind of maps  $K$  that partially specifies the category of models of GR?" Our iterative procedure can be used to guide our answering both of these questions, albeit with a slightly different interpretation of the result in each case.

Fourth, a way of interpreting the discussed objection is that it was a mistake to even consider the full identity as the candidate for the standard of comparison of GR models. If from the very beginning we choose the standard of comparison of GR models (i.e.,  $K$  in Definition 9) to be isometries (in the sense of Definition 3), then we would not get the verdict of radical indeterminism at any stage of our considerations. Does this mean that the Hole Argument is blocked before it starts? The answer is no because the Hole Argument relies on Definition 9 with  $K$  chosen to be the full identity; what happens for a different choice of  $K$  is irrelevant. However, one can attempt to undermine the significance of this argument by saying that we should not consider the full identity as  $K$  at all for reasons independent of the Hole Argument, which are much stronger and/or more obvious than the reason provided by the Hole Argument.<sup>30</sup> But do we have such reasons? The bare appeal to scientific practice is very unilluminating; moreover, such practice, if rational, should itself be based on some reasons, which the philosophy of science should try to uncover.

## 7 Summary

I have argued that the Hole Argument in its (what I claim to be) proper formulation does not rely on the notion of isomorphism: this notion is not needed in either of its two steps. The Gauge Theorem (step 1) establishes that if  $\mathcal{M} = \langle M, O_1, O_2, \dots, O_n \rangle$  is a model of a local spacetime theory  $T$  (satisfying the additional condition that  $O_k, \dots, O_n$  are functions of  $O_1, \dots, O_{k-1}$  together with their derivatives and are form-invariant under diffeomorphisms), then any  $\mathcal{M}'$  obtained from  $\mathcal{M}$  by applying a diffeomorphism is also a model of  $T$  because  $\mathcal{M}$  and  $\mathcal{M}'$  satisfy the same dynamical equations (and not because they are isomorphic). To conclude from the Gauge Theorem that any local spacetime theory is radically indeterministic unless we regard any two of its diffeomorphism-related models as physically equivalent (step 2), we also do not need the notion of isomorphism: we only need to tentatively choose the kind of maps that will serve as our standard of comparison of models and check whether our theory satisfies the partial definition of radical indeterminism relative to this kind of maps. If we choose a too narrow class of maps, then we will find instances of radical indeterminism, in which case we should (other things being equal) broaden this class.

<sup>30</sup> That seems to be, in the end, the point of Weatherall (2018, p. 343): "the hole argument merely supports a view one could have had independent reasons to accept".

My conclusion that the notion of isomorphism is not needed to formulate the Hole Argument does not imply that the notion of isomorphism cannot be relevant to this argument. I think that it might be relevant: for example, if we assume that from the fact that two models are isomorphic it follows that they are physically equivalent and that diffeomorphisms (in the sense of Definition 4) are isomorphisms of GR models, then these assumptions imply that diffeomorphic models of GR are physically equivalent; therefore, the conclusion of radical indeterminism of GR is blocked. However, the role of the notion of isomorphism is here only intermediary and this notion is dispensable—one can directly consider the question of which GR models are physically equivalent without engaging in the debate about isomorphisms.

But can these two issues, the mathematical equivalence of models given by their isomorphisms and the physical equivalence of models, really be separated? Weatherall (2018, p. 337) seems to think that the only way to argue for the physical and empirical equivalence of models is to establish a mathematical isomorphism between them. If he is right, then the notion of isomorphism is not dispensable for the study of physical equivalence, as I have suggested in the previous paragraph. However, I find this view rather dubious: empirical equivalence is an empirical issue after all, so it cannot be the case that the only way to establish it is to analyse the mathematical formalism of the theory without any recourse to empirical investigations. Physical equivalence is a more subtle issue, but it also surely is not a purely formal feature of the theory, albeit it is also not the same as empirical equivalence (unless we are radical empiricists).

An example from Weatherall (2020, p. 86) can be used to support this point and to show how isomorphisms and physical equivalence can come apart.<sup>31</sup> He considers a variant of GR, the models of which are Lorentzian manifolds enriched with a smooth individuating field (its role is to label all points of spacetime and thereby make it possible to distinguish any point from the others). Diffeomorphisms are not isomorphisms of these new models because they (in general) change the individuating field, which is a part of the structure of these models. However, this does not influence the physical equivalence relation: our iterative procedure should provide the same verdict for these models as for the original ones (i.e., consisting of Lorentzian manifolds without the individuating field) because physical situations represented by these two classes of models are the same. In general, the models of our theories might possess too much or too little structure compared with what is needed to account for the physical situations they aim to describe, in which case the equivalence relation induced by their isomorphisms will be not identical with the physical equivalence relation between them.

**Acknowledgements** I would like to thank Juliusz Doboszewski, Tushar Menon, Tomasz Placek, James Read and anonymous reviewers for their valuable comments on the previous versions of this paper. This work was supported by National Science Centre grant OPUS 2022/45/B/HS1/02700 and open access has been financed by a grant from the Faculty of Philosophy under the Strategic Programme Excellence Initiative at Jagiellonian University.

## Declarations

**Conflict of interest** The authors declared that they have no conflict of interest.

<sup>31</sup> Another example of how they might come apart is given in footnote 28.

**Ethical approval** N/A (my research does not involve human participants).

**Informed consent** N/A (my research does not involve human participants).

**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

- Belot, G. (2013). Symmetry and equivalence. In R. Batterman (Ed.), *The Oxford Handbook of Philosophy of Physics* (pp. 318–339). Oxford: Oxford University Press.
- Brading, K., & Castellani, E. (2007). Symmetries and invariances in classical physics. In J. Butterfield & J. Earman (Eds.), *Philosophy of physics* (pp. 1331–1367). Elsevier.
- Brighouse, C. (1994). Spacetime and holes. In *PSA: The Proceedings of the Biennial Meeting of the Philosophy of Science Association* (Vol. 1, pp. 117–125).
- Butterfield, J. (1989). The hole truth. *The British Journal for the Philosophy of Science*, 40(1), 1–28.
- Choquet-Bruhat, Y., & Geroch, R. (1969). Global aspects of the Cauchy problem in general relativity. *Communications in Mathematical Physics*, 14, 329–335.
- Chruściel, P. T., & Isenberg, J. (1993). Nonisometric vacuum extensions of vacuum maximal globally hyperbolic spacetimes. *Physical Review D*, 48(4), 16616–16628.
- Curiel, E. (2018). On the existence of spacetime structure. *The British Journal for the Philosophy of Science*, 69(2), 447–483.
- Dewar, N. (2019). Sophistication about symmetries. *The British Journal for the Philosophy of Science*, 70, 485–521.
- Doboszewski, J. (2017). Non-uniquely extendible maximal globally hyperbolic spacetimes in classical general relativity: A philosophical survey. In G. Hofer-Szabó & L. Wroński (Eds.), *Making it formally explicit: Probability, causality and indeterminism* (pp. 193–212). Springer.
- Doboszewski, J. (2019). Relativistic spacetimes and definitions of determinism. *European Journal for Philosophy of Science*, 9, 24.
- Earman, J., & Norton, J. (1987). What price spacetime substantivalism? The hole story. *The British Journal for the Philosophy of Science*, 38, 515–525.
- Giulini, D. (2007). Remarks on the notions of general covariance and background independence. In I.-O. Stamatescu & E. Seiler (Eds.), *Lecture notes in physics* (Vol. 721, pp. 105–120). Springer.
- Gomes, H., & Butterfield, J. (2023). The Hole Argument and beyond, Part I: The Story so Far. arXiv Preprint. [arXiv:2303.14052](https://arxiv.org/abs/2303.14052)
- Halvorson, H., & Manchak, J. B. (2022). Closing the Hole Argument. *The British Society for the Philosophy of Science*. <https://doi.org/10.1086/719193>
- Hawking, S. W., & Ellis, G. F. R. (1973). *The large scale structure of space-time*. Cambridge: Cambridge University Press.
- Hofer, C. (1996). The metaphysics of spacetime substantivalism. *Journal of Philosophy*, 93, 5–27.
- Johns, O. D. (2019). Validity of the Einstein hole argument. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 68, 62–70.
- Landsman, K. (2023). Reopening the Hole Argument. *Philosophy of Physics*, 1(1), 6, 1–24.
- Lewis, D. (1968). Counterpart theory and quantified modal logic. *Journal of Philosophy*, 65, 113–126.
- Luc, J. (2022). Arguments from scientific practice in the debate about the physical equivalence of symmetry-related models. *Synthese*, 200, 72.
- Malament, D. (2012). *Topics in the foundations of general relativity and Newtonian Gravitation Theory*. Chicago: University of Chicago Press.
- Manchak, J. B., & Barrett, T. (2023). A hierarchy of spacetime symmetries: Holes to Heraclitus. *The British Journal for the Philosophy of Science*. <https://doi.org/10.1086/727002>

- Menon, T., & Read, J. (2024). Some remarks on recent formalist responses to the Hole Argument. *Foundations of Physics*, 54, 6.
- Møller-Nielsen, T. (2017). Invariance, interpretation, and motivation. *Philosophy of Science*, 84(5), 1253–1264.
- Pooley, O. (2006). Points, particles, and structural realism. In D. Rickles, S. French, & J. Saatsi (Eds.), *The structural foundations of quantum gravity* (pp. 83–120). Oxford University Press.
- Pooley, O. (2013). Substantivalist and relationalist approaches to spacetime. In R. Batterman (Ed.), *The Oxford handbook of philosophy of physics*. Oxford University Press.
- Pooley, O. (2017). Background independence, diffeomorphism invariance and the meaning of coordinates. In D. Lehmkuhl, G. Schieman, & E. Scholz (Eds.), *Towards a theory of spacetime theories* (pp. 105–143). Birkhäuser.
- Pooley, O., & Read, J. A. M. (2021). On the mathematics and metaphysics of the hole argument. *The British Journal for the Philosophy of Science*. <https://doi.org/10.1086/718274>
- Ringström, H. (2009). *The Cauchy problem in general relativity*. European Mathematical Society.
- Roberts, B. (2020). Regarding 'Leibniz equivalence'. *Foundations of Physics*, 50, 250–269.
- Saunders, S. (2003). Physics and Leibniz's principles. In K. Brading & E. Castellani (Eds.), *Symmetries in physics: Philosophical reflections*. Cambridge: Cambridge University Press.
- Wald, R. M. (1984). *General relativity*. Chicago and London: The University of Chicago Press.
- Weatherall, J. O. (2017). Inertial motion, explanation, and the foundations of classical spacetime theories. In D. Lehmkuhl, G. Schieman, & E. Scholz (Eds.), *Towards a theory of spacetime theories* (pp. 13–42). Birkhäuser.
- Weatherall, J. O. (2018). Regarding the 'hole argument'. *The British Journal for the Philosophy of Science*, 69(2), 329–350.
- Weatherall, J. O. (2020). Some philosophical prehistory of the (Earman-Norton) hole argument. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 70, 79–87.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.