



# Spacetime: function and approximation

Sam Baron<sup>1</sup>

Received: 29 August 2021 / Accepted: 7 November 2021 / Published online: 15 April 2022  
© The Author(s) 2022

## Abstract

Several approaches to quantum gravity (QG) signal the loss of spacetime at some level. According to spacetime functionalism, spacetime is functionally realised by a more fundamental structure. According to one version of spacetime functionalism, the spacetime role is specified by Ramsifying general relativity (GR). In some approaches to QG, however, there does not appear to be anything that exactly realises the functional role defined by a Ramsey sentence for GR. The spacetime role is approximately realised. It is open to the spacetime functionalist to adopt a ‘near enough is good enough’ attitude to functional realisation, and maintain that spacetime is functionally realised nonetheless. In this paper I present a challenge for such an ‘approximate’ spacetime functionalism. The challenge, in brief, is to provide an account of how ‘close’ is close enough for approximate realisation to occur. I canvass a range of options for spelling out a similarity relation of the relevant kind, and argue that none are successful. In light of the challenge, I recommend giving up on the functional realisation of spacetime. I argue, however, that even if spacetime as a whole is not functionally realised, some of the functions of spacetime may still be performed.

**Keywords** Spacetime · Functionalism · Emergence · Observation · Ontology · Time · Space

## 1 Introduction

Several approaches to quantum gravity (QG) signal the loss of spacetime at some level of description.<sup>1</sup> This has been flagged for loop quantum gravity, causal set theory,

---

<sup>1</sup> Quantum Gravity is the name of a broad research program that aims to reconcile our two best physical theories at the moment: general relativity and quantum mechanics.

---

✉ Sam Baron  
samuel.baron@acu.edu.au

<sup>1</sup> Dianioia Institute of Philosophy, Australian Catholic University, East Melbourne, Australia

canonical quantum gravity and, to a certain extent, string theory.<sup>2</sup> Without spacetime, such approaches face a potential threat of empirical incoherence (cf. Huggett & Wüthrich, 2013). The threat arises because empirical confirmation can be linked to a certain notion of observation, one that relies on the detection of entities with a spatiotemporal location. The concern, then, is that without spacetime it won't be possible to conduct observations that confirm a given approach to QG. Indeed, the very fact that we can conduct observations at all seems to suggest that spacetime must emerge at some level. If spacetime is not recoverable from a theory of QG then such a theory would seem to be falsified by the mere fact that observation occurs.<sup>3</sup>

Spacetime functionalism has been proposed as one solution to the threat of empirical incoherence (cf. Lam & Wüthrich, 2018, 2020). According to spacetime functionalism, spacetime is functionally realised by a more fundamental structure.<sup>4</sup> Just as, say, pain emerges because a certain neural structure plays the pain role; so too can we say that spacetime emerges because some structure posited by a given approach to QG plays the spacetime role. The empirical coherence of a given approach to QG is thus secured via the functional realisation of spacetime by non-spatiotemporal entities.

In the case of pain, the pain role is typically specified via reference to a particular theory of pain. So too, in the case of spacetime, the spacetime role can be specified via a specific theory. The most obvious candidate for specifying the spacetime role is the general theory of relativity (GR). Spacetime as described by GR, however, is widely expected to be at best an approximate description of the physical structure posited by QG. So the full functional role for GR may not be exactly realised. This need not spell the end of spacetime functionalism, however. For it is open to the spacetime functionalist to adopt a 'near enough is good enough' attitude to functional realisation, and maintain that spacetime is functionally realised nonetheless.

My goal in this paper is to first pose a challenge to this 'approximate' spacetime functionalism before suggesting a way forward. I begin, in Sect. 2, by motivating and then more precisely formulating a version of approximate spacetime functionalism. In Sect. 3, I present the challenge. The challenge, in brief, is to provide an account of how 'close' is close enough for realisation to occur. I canvass a range of options for spelling out a similarity relation of the relevant kind, and argue that none are successful. In light of the challenge, I recommend giving up the functional realisation of spacetime in Sect. 4. I argue, however, that even if spacetime as a whole is not functionally realised, some of the functions of spacetime may still be performed, and that is enough to secure the empirical coherence of a given approach to QG.

---

<sup>2</sup> Huggett and Wüthrich (2013) make the case for loop quantum gravity, causal set theory and string theory. For canonical quantum gravity, see Butterfield and Isham (1999), Barbour (1994). The emergent gravity program has also been flagged as one in which spacetime is not fundamental, see Carlip (2014), Hu (2009). See Crowther (2016) for general discussion.

<sup>3</sup> This version of the problem focuses on location. Healey (2002) offers a related version of the problem involving experience, and Braddon-Mitchell and Miller (2019) outline related difficulties involving mental content. See Ney (2015) for a related discussion in quantum mechanics.

<sup>4</sup> Versions of spacetime functionalism have been defended by Knox (2019), Lam and Wüthrich (2018, 2020) and Chalmers (forthcoming). An alternative to functionalism is defended by Le Bihan (2018a, b). For critical discussion of spacetime functionalism see Le Bihan (2021), Yates (2021) and for an overview of some recent work see Crowther et al. (2020).

## 2 Approximate spacetime functionalism

According to Lam and Wüthrich, spacetime functionalism proceeds via the following two steps:

(FR-1) The higher-level properties or entities, which are the target of the reduction, are ‘functionalised’, that is, they are given a functional definition in terms of their causal or functional role.

(FR-2) An explanation is provided of how the lower-level properties or entities can fill this functional role. (Lam & Wüthrich, 2018, p. 43)

Spacetime functionalism is thus modelled closely on functionalism in the philosophy of mind. In the philosophy of mind, the standard method for conducting the first stage of the functionalist analysis involves Ramsification: the process of producing a Ramsey sentence for a given theory (see Lewis, 1972; 1970, based on Ramsey, 1931). The same method will be applied here. To grasp the broad procedure in question, it is useful to consider the mental state case in a bit of detail.

We start with a theory, such as a simple theory of pain. According to this theory, let us suppose, pain is caused by pinches and punches and causes yells and yelps. We then draw a distinction between the t-terms and the o-terms that appear within the theory. The t-terms are typically the terms that a theory is about. They are often new terms, in the sense that they are defined within the theory for the first time (though this is not always the case). The o-terms, by contrast, are familiar terms that are well-defined outside the scope of the specific theory in question and are used to define the t-terms. In the case of our simple pain theory, the t-term is just ‘pain’, whereas the o-terms include ‘pinch’, ‘punch’, ‘yell’ and ‘yelp’ but also ‘cause’.

The t-terms in a given theory are defined both in terms of the o-terms and in terms of one another. A theory, in general, corresponds to a single sentence that specifies each t-term introduced by the theory, along with the o-terms that are used to define them. This is the theoretical postulate,  $T(t)$ . By replacing each of the t-terms in  $T(t)$  with free variables, we end up with the realisation formula  $T(x)$ . We can produce a Ramsey sentence for the theory  $T$  by binding the variables in the realisation formula within the scope of one or more quantifiers. The Ramsey sentence thus has the general form  $\exists xT(x)$ .

So, to take our simple pain theory again, we produce a Ramsey sentence for that theory by replacing the predicate ‘pain’ with a variable, bound in the scope of a quantifier. This gives us the following:  $\exists x$  ( $x$  is caused by pinches and punches and  $x$  causes yells and yelps). The Ramsey sentence provides an abstract characterisation of pain via a system of relations and, importantly, eschews any mental state terms. In this way, the Ramsey sentence defines the pain role and thereby lays down the conditions under which pain is functionally realised. The Ramsey sentence is thus a functional characterisation of pain in our simple pain theory.

Having isolated the functional role in this way, we can then look to another theory, and see if it posits some n-tuple of entities  $\langle e_1 \dots e_n \rangle$  that plays the pain role as defined by the simple pain theory’s Ramsey sentence. An n-tuple plays the pain role when its

elements possess each of the properties and stand in each of the relations defined by the o-terms in the relevant Ramsey sentence. So, for instance, consider a neurophysiological theory, which posits a range of neural states. Suppose, moreover, that there is some neural state posited by the theory and that the presence of that state is caused by pinches and punches and causes yells and yelps. Then it follows that there is an n-tuple of entities in the neurophysiological theory at issue that satisfies the pain role defined by the Ramsey sentence for our simple pain theory. We can thus conclude that the relevant neural state functionally realises pain.

We can extend the same basic picture to the case of spacetime with one important caveat. In the pain case, the functional role defined by the Ramsey sentence is a *causal* role. Pain is whatever stands in a certain set of causal relations defined by the simple pain theory. It would be very controversial, however, to treat functionalism in the case of spacetime in causal terms. Whether spacetime does causal work in GR remains unclear (see Vassallo, 2020 for useful discussion). Spacetime functionalism must therefore be understood in terms of non-causal functional roles. As Lam and Wüthrich (2018, p. 42) note, this is not a significant departure from functionalism since “functional properties need not be exclusively characterized in causal terms”. As Polger (2007) argues, there is good reason to adopt a view of functionalism of this kind even for the case of mental properties.

At any rate, functionalism on the Ramsey method does not impute a requirement for functional roles to be causal. All that matters is that t-terms can be characterised in terms of their relations to other features within a theory, even if those relations are not causal in nature. To apply the Ramsey method to spacetime, then, we start with a theory in which ‘spacetime’ is introduced as a t-term. Following a number of others, I will focus on GR, though I will briefly consider some alternative functional specifications later on. I focus on GR because it is our best available theory of spacetime (Lam & Wüthrich, 2018, 2020; Le Bihan, 2021, Yates, 2021 take a similar line). Moreover, as Chalmers points out, the concept of spacetime has its roots in relativity, so GR is a natural place to start. As he puts the point:

Spacetime as understood here is an essentially theoretical concept, one that emerges especially from the general theory of relativity. This concept can be applied to other theories, including Newtonian theories and theories of quantum gravity, but the concept itself did not exist (at least in the relevant sense) prior to relativity. (Chalmers forthcoming)

We thus take a definition of what spacetime is within the context of GR and Ram-sify it by replacing any mention of spacetime with variables bound in the scope of quantifiers. Exactly what the Ramsey sentence might be depends on what we take spacetime in GR to be. This is a controversial issue but, for now at least, I will assume that spacetime in GR can be represented by a manifold equipped with a certain metric.

The metric in GR corresponds to a metric tensor field: a mapping from a smooth manifold to a fibre bundle (a mathematical object constituted by the sum total of metric tensors associated with each point in the manifold). Each metric tensor in the field defines the metric at a point, and thus describes both distance and curvature. The metric tensor field in GR assigns metric tensors to points of a manifold based, in part, on the

mass-energy distribution. In this way, the metric tensor field and the stress-energy field together describe a complex geometrical object that exhibits continuous, point-wise variation in metric structure. The relationship between the metric tensor field and the mass-energy distribution is captured by the coupling relationship between the metric tensor and the stress-energy tensor in the Einstein field equations:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu} \quad (1)$$

where  $G_{\mu\nu}$  is the Einstein tensor,  $\Lambda$  is the cosmological constant,  $g_{\mu\nu}$  is the metric tensor,  $\kappa$  is the Einstein gravitational constant and  $T_{\mu\nu}$  is the stress-energy tensor.

We can thus define spacetime in terms of an isometry between a set  $S$  of physical elements and a manifold equipped with the metric field  $g_{\mu\nu}$ . The set  $S$  can be a set of physical distance relations, or a set of point-like physical elements, thus leaving the relational or substantival status of spacetime open. Isometry ensures that the metric functions on  $S$  are the same as those on the manifold, and thus that the behaviour of the metric tensor field specified by the field equations is preserved for the physical structure at issue. This, in turn, ensures that the relevant physical structure exhibits the geometric structure of spacetime. To Ramsify this account we need a way of quantifying over the physical elements in  $S$ . For this, we can use a plural quantifier ‘ $\exists xx$ ’. Putting this together, then, we have the following candidate Ramsey sentence for spacetime in GR:

$\exists xx$  (the set containing  $xx$  is isometric to a manifold  $M$  equipped with the metric tensor field  $g_{\mu\nu}$ ).

This may not be the full Ramsey sentence for GR. One might, for instance, wish to strengthen the sentence with a general mapping requirement: namely the requirement to preserve all functions, not just metric ones (this might be desired if, for instance, one wishes to capture the matter field as well as the metric field in one’s functional specification). But even if the above Ramsey sentence is not complete, it is likely to be a necessary part of the complete sentence. The metric structure of spacetime and the manifold representation are important aspects of what spacetime is.

I also recognise that the Ramsey–Lewis procedure for specifying the functional role for spacetime is not the only way to develop spacetime functionalism. As Chalmers (forthcoming) discusses, however, there are reasons to prefer it. Chief among these is that there is no need to try and sort through the various roles that spacetime might play, and select one or more as its necessary and sufficient features. Instead, we can simply focus on the entire functional specification delivered by the Ramsey sentence, and use that to locate spacetime in QG. As it turns out, however, we may be forced to adopt a more piecemeal approach than the one Chalmers has in mind. More on this later.

In order to determine whether anything plays the spacetime role in a given theory of QG we must see whether that theory posits anything that satisfies the conditions laid down by the Ramsey sentence. We thus need to identify some set of entities posited by a given approach to QG and show that it is isometric to the metric space associated with GR. For a number of approaches to QG, however, it is unlikely that anything fits

the bill exactly. For it is widely expected that the continuous metric tensor field in GR will, at best, be approximated by a more fundamental structure in QG.

This is expected for one prominent approach in particular, namely: loop quantum gravity (LQG).<sup>5</sup> The fundamental objects of study in LQG are spin-networks (see Rovelli, 2011b; 2004). In the kinematical form of LQG, a spin-network can be modelled as a graph in which each edge represents a three-dimensional volume of space, and each node represents a two-dimensional spatial surface. The standard dynamical form of the theory involves extruding the spin-network through a higher dimension. The result is the so-called spinfoam structure (see Rovelli & Vidotto, 2014). The fundamental structure of LQG in both its kinematical and dynamical forms is thought to be *discrete*. Because the metric structure of GR spacetime is *continuous*, it can therefore only be approximated by an underlying spinfoam.

The way that spinfoam structures approximate GR spacetime is something like the discrete approximation of the spacetime metric yielded by the Regge (1961) calculus (cf. Rovelli, 2011a). The Regge calculus can be used to produce a simplicial approximation which, while not isometric to the metric structure of spacetime, does seem to be empirically equivalent to some spacetime models. The Regge construction involves laying down the various elements in a simplicial complex to approximate both distance and curvature. One way to get a feel for the idea is to consider the construction of a geodesic dome from triangles. The face of each triangle is flat, and so there is no curvature to be found if we attend to individual tiles. However, if we look down the edges where triangles meet, we can see something that approximates curvature, though is not curvature strictly speaking. In the case of spacetime, the Regge construction use simplexes to mimic the four-dimensional curvature captured by the metric tensor field in much the same way. Curvature is approximated along the edges where simplexes meet—called the ‘bones’. In very broad terms, the LQG approximation of GR metric structure involves wrapping the elements of a spinfoam around a structure of Regge bones (see §4 of Rovelli, 2011a).<sup>6</sup>

There is, however, a very important difference between the simplicial structures found in the ordinary Regge calculus and the spinfoam structure. In the case of the ordinary Regge calculus, we can operate on the simplexes asymptotically by making them infinitely small and infinite in number. When we do this, there is a hope that we can recover the full continuum structure of the GR metric. In the case of the spinfoam structure, by contrast, we can’t perform a similar operation. While we can treat spinfoams simplicially as networks of tetrahedral structures “there is no limit in which the tetrahedra become infinitely small” (Rovelli & Vidotto, 2014, p. 139). This, Rovelli and Vidotto (2014, p. 139) note, is the “central point of LQG”: spinfoam structures

<sup>5</sup> For an overview of LQG see Rovelli (2004), Rovelli and Vidotto (2014). For overviews aimed at a more philosophical audience, see the discussion of LQG in Huggett and Wüthrich (2013), Wüthrich (2019), Wüthrich (2017).

<sup>6</sup> Rovelli (2011a) sums up this association with the Regge construction as follows:

The graphs and the two-complexes of loop gravity are essentially a description of a triangulated Regge manifold. In 3d (space), the graph is the 1-skeleton of the dual to the 3d spatial manifold with defects: loops in the graph wrap around the Regge bones. In 4d (spacetime) the two-complex is the two-skeleton of the dual of a 4d cellular decomposition.

display a ‘quantum’ discreteness. This quantum discreteness signals a discreteness in the fundamental physical structure of reality. It is because of this discreteness that the spinfoam structure can only ever approximate GR spacetime.

The connection between spinfoams and GR spacetime is further complicated by a second feature of the spinfoam structure. The fundamental structure is thought to be in a superposition of spinfoam states (Rovelli & Vidotto, 2014, pp. 109–110). Contrast this with the purely classical nature of GR spacetime, which does not exhibit the same quantum properties. The fact that spinfoams are in a superposition makes the process of recovering GR spacetime difficult. For even if it is correct that certain spinfoams provide discrete approximations of GR spacetime, an account is still needed of how a particular spinfoam structure gets selected out of the quantum superposition at the fundamental level. If no such account is forthcoming, then it may be difficult to say that spacetime is functionally realised by a spinfoam structure (or it may mean that it is realised many times over, depending on how you look at it).

For present purposes, I will set aside this complication and assume for the sake of argument that we can select a spinfoam from the underlying quantum superposition. Even then, the matter of functional realisation is not straightforward. Since spinfoams display quantum discreteness, they fall short of having the full suite of properties specified by a Ramsey sentence for GR. Nonetheless, the spacetime functionalist will want to say that GR spacetime is realised well-enough because the simplicial approximation of GR spacetime is sufficiently close. The question is: how does approximate realisation of this kind work?

The approximate realisation of some functional role is not unusual for functionalism. Indeed, a similar issue arises for mental state terms, like pain. It is likely that pain is, at best, approximately realised by neural states in a relevant neurophysiological theory (see Shoemaker, 1981). In the case of pain, however, one may allow that pain is functionally realised even if the conditions laid down by the Ramsey sentence for pain are not exactly satisfied. Indeed, foreseeing this possibility, Lewis outlines some general conditions for approximating cases. Here’s Lewis:

We might want to say that the theoretical terms name the components of whichever n-tuple comes nearest to realizing the theory, if it comes near enough ... Given a theory  $T$ , we might find a slightly weaker  $T'$ , implied by but not implying  $T$ , such that an n-tuple is a realization of  $T'$  if and only if it is a near-realization of  $T$ . Then we could say that  $T'$ , not  $T$ , is the real term-introducing theory; everything we have been saying about  $T$  really ought to be taken as applying to  $T'$  instead.  $T$  itself may be recovered as the conjunction of  $T'$  with further hypotheses containing the theoretical terms already introduced by  $T'$ . (Lewis, 1970, p. 432)

The basic idea is that when the functional role specified by the Ramsey sentence for a theory  $T$  is not exactly satisfied, we can still say that the t-terms in  $T$  are functionally realised so long as (i) there is a theory  $T'$  which can be obtained by ‘slightly’ modifying  $T$ ; (ii)  $T$  implies  $T'$  but not vice versa and (iii) the conditions laid down by the Ramsey sentence for the modified theory  $T'$  are exactly satisfied by some n-tuple posited by a further theory  $T^*$ .

A ‘near enough is good enough’ approach to functional realisation is available in the case of spacetime as well. Lewis’s specific realisation conditions do not seem entirely suitable for every approach to QG, however. For it is not obvious that there is a way to ‘slightly’ modify or correct GR to produce theories that are exactly realised in those approaches to QG that appeal to a fundamental structure that is discrete. For such theories, what we would need, at a minimum, is a way to transform GR into a theory with a discrete metric structure, rather than the continuous metric structure captured by the tensor field. Now, in this vein, one might point toward certain discrete approaches to relativity, such as the Regge calculus mentioned above, which can be used to recover some of the numerical results of GR for particular spacetime models. It is far from clear, however, that such approaches can be produced by ‘slightly’ modifying GR and nor is it obvious that these approaches are implied (but do not imply) GR as Lewis’s account requires. Lewis’s account thus appears too demanding. We may well want to say that something in a theory  $T$  is approximately realised, even when we don’t have a way of modifying  $T$  to produce another theory that is both exactly realised and strictly implied by the original theory.

In what follows, then, I will work with an alternative to Lewis’s approach. My alternative approach can be stated as follows. For any theory  $T$  with Ramsey sentence  $R$ :

1. If there is an  $n$ -tuple of entities  $\langle e_1^* \dots e_n^* \rangle$  and the  $e_n^*$  exactly realise the functional role specified by  $R$ , then the  $t$ -terms in  $T$  are realised by the  $e_n^*$ .
2. If there is no  $n$ -tuple of entities that exactly realises the functional role specified by  $R$  then, if there is an  $n$ -tuple of entities  $\langle e'_1 \dots e'_n \rangle$  posited by a theory  $T'$  and the  $e'_n$  are sufficiently similar to whatever  $n$ -tuple of entities  $\langle e_1 \dots e_n \rangle$  would exactly realise the functional role specified by  $R$  were the  $e_n$  to exist, then the  $t$ -terms in  $T$  are functionally realised by the  $e'_n$ .

Under this set of realisation conditions, all that functional realisation requires is for there to be a sufficiently close relationship between whatever would exactly realise the functional role for a theory  $T$  and the entities posited by another theory  $T'$ . There is no need for a third theory  $T^*$  that constitutes a modification or correction of  $T$  to mediate the functional relationship between  $T$  and  $T'$ . The proposed set of conditions is thus less demanding than Lewis’s, at least when it comes to the relationships between theories.

Using the above realisation conditions, we can thus allow that a discrete structure in a theory of QG functionally realises spacetime because it is sufficiently similar to whatever would exactly realise the spacetime role, were it to exist. We can say this even when there is no modification to GR that yields a functional specification that is exactly realised by the theory of QG at issue. Call such a view: approximate spacetime functionalism.

### 3 The challenge

My goal in this section is to raise a challenge for approximate spacetime functionalism. The realisation conditions outlined above allow that functional realisation can occur



even when the functional role specified by the Ramsey sentence for GR is not exactly realised. All that matters is that there is something that is sufficiently similar to an exact realiser of the spacetime role. But how similar do the entities posited by a theory of QG need to be to an exact realiser before we can say that functional realisation occurs? The challenge, in brief, is to provide a plausible answer to this question. In what follows, I canvass a range of potential accounts of the similarity relation and argue that none is satisfactory. In this way I make a case for the seriousness of the challenge. I don't take the challenge to be decisive, however. There may yet be a way to specify similarity in a way that can support approximate spacetime functionalism. I do, however, think the challenge is sufficiently difficult to motivate finding an alternative.

In recent work, Lam and Wüthrich seem to endorse some version of approximate spacetime functionalism. They acknowledge that a full functional specification for spacetime may not be satisfied by a more fundamental theory of QG. For them, however, this doesn't matter. Even if the functional specification for spacetime is not exactly satisfied, so long as that specification is satisfied well-enough, we can say that spacetime is functionally realised. Specifically, so long as there is enough structure in a more fundamental theory of QG to be able to accommodate the empirical evidence that supports GR, then that is enough for functional realisation. As they put the point:

... from a functionalist point of view, **nothing remains beyond showing how the fundamental degrees of freedom can collectively behave such that they appear spatiotemporal at macroscopic scales in all relevant and empirically testable ways.** This turn out to be a hard task in quantum gravity. Functionalism can be seen as the assertion that once this task is completed, no unfinished business lingers on. (emphasis added, Lam & Wüthrich, 2018, p. 44)

We can formulate the basic idea behind Lam and Wüthrich's version of approximate spacetime functionalism as follows. First, let us say that a theory  $T'$  recaptures the empirical results of a theory  $T$  when every confirmed empirical prediction of  $T$  is also an empirical prediction of  $T'$  but not vice versa. We can then supplement the realisation conditions detailed at the end of Sect. 2 with a third clause, [ER], that specifies the operative notion of similarity in terms of empirical recovery. This yields the following full set of realisation conditions. For a theory  $T$  with Ramsey sentence  $R$ :

1. If there is an  $n$ -tuple of entities  $\langle e_1 * \dots e_n * \rangle$  and the  $e_n *$  exactly realise the functional role specified by  $R$ , then the  $t$ -terms in  $T$  are realised by the  $e_n *$ .
2. If there is no  $n$ -tuple of entities that exactly realises the functional role specified by  $R$  then, if there is an  $n$ -tuple of entities  $\langle e'_1 \dots e'_n \rangle$  posited by a theory  $T'$  and the  $e'_n$  are sufficiently similar to whatever  $n$ -tuple of entities  $\langle e_1 \dots e_n \rangle$  would exactly realise the functional role specified by  $R$  were the  $e_n$  to exist, then the  $t$ -terms in  $T$  are functionally realised by the  $e'_n$ .
3. A theory  $T'$  posits an  $n$ -tuple of entities  $\langle e'_1 \dots e'_n \rangle$  that is sufficiently similar to whatever  $n$ -tuple of entities  $\langle e_1 \dots e_n \rangle$  would exactly realise the functional role specified by  $R$  in  $T$  just when every confirmed empirical prediction of  $T$  is an empirical prediction of  $T'$  but not vice versa.

So, for example, spacetime in GR is functionally realised by something in LQG so long as LQG can recapture the empirical results of GR and thus if something in LQG

“appears spatiotemporal at macroscopic scales in all relevant and empirically testable ways” (Lam & Wüthrich, 2018, p. 44).

While initially promising, [ER] faces a problem. To see the problem, let us consider a different example. Consider, for instance, the following three theories: quantum electrodynamics (QED), special relativity (SR) and a suitably sophisticated version of Lorentz’s aether theory. SR and Lorentz’s aether theory are empirically equivalent: every confirmed empirical prediction of SR is a prediction of the aether theory as well. Lorentz’s aether theory, however, makes use of unobservable entities that play no role in SR. Most notably, Lorentz’s theory posits a physical substance that permeates all of space, through which light waves propagate and that defines an absolute rest frame. No such entity is to be found in SR. The closest thing to the aether is the electromagnetic field, which one might consider to be aether-like, in so far as it is real and permeates all of space. However, this field does not define an absolute rest frame. Since nothing else in SR does this work either, the aether is absent from the theory.

Now, suppose we produce a Ramsey sentence for Lorentz’s aether theory. Exactly what the Ramsey sentence might be is not all that important for present purposes. What matters is that the sentence will include a *t*-term for the aether. The conditions laid down by this Ramsey sentence thus won’t be exactly satisfied by anything in QED. That’s because there is no aether in QED. Suppose, however, that we adopt the realisation conditions outlined above, including [ER]. Then what happens? Well, because QED recovers all of the empirical predictions of SR, and since SR and Lorentz’s aether theory are empirically equivalent, QED manages to recover all of the empirical predictions of Lorentz’s aether theory as well. QED thus stands in a relationship of empirical recovery to Lorentz’s aether theory. Given [ER] it follows that there is an *n*-tuple in QED that is sufficiently similar to whatever would exactly realise Lorentz’s aether theory were it to exist. Given this, each *t*-term in the Ramsey sentence for Lorentz’s aether theory—including the *t*-term for the aether—manages to be realised in QED after all.

The problem is really a general problem with *t*-terms for unobservables. It is precisely because the aether makes no empirical difference that SR and Lorentz’s aether theory are empirically equivalent. Indeed, more generally, for any pair of theories *T* and *T*\*, if *T* and *T*\* differ only in unobservable respects, then any theory that empirically recovers *T* will empirically recover *T*\* as well. If we hang approximate functional realisation on empirical recovery alone, then any realisers for the *t*-terms in *T* will also be realisers for the *t*-terms in *T*\*.

Given the generality of the problem, it is straightforward to extend it to QG and GR. We can do this by producing a modified version of GR that adds an unobservable aether to that theory. Now, we have to be a bit careful here, since adding an aether to GR is likely to produce a theory with distinct empirical consequences. Adding an aether involves adding a rest frame, and so such a theory should predict situations in which Lorentz invariance is broken. Adding an aether to GR would thus likely involve undermining general covariance which, in turn, would undermine background independence (Belot, 2011, p. 2866).

Aether theories of this kind are being actively developed and they do indeed lead to empirical differences at some level (see, for instance, Jacobson & Mattingly, Jacobson and Mattingly 2001). What I have in mind, though, is a theory that modifies one of these

aetheric versions of GR by adding a mechanism through which the aether disguises itself, and wipes out all possible empirical evidence of Lorentz symmetry-breaking. For instance, we could imagine that physical systems are systematically distorted when they are in motion relative to the absolute rest-frame. The distortion prevents us from detecting any break-down of Lorentz invariance and thus prevents us from detecting the aether (this is, in fact, not so far from Lorentz's original explanation of the null results in the Michelson-Morley experiments).

The 'self-effacing' aether in this modified version of GR, call it GR+, adds no empirical consequences. GR and GR+ are thus empirically equivalent, but they have different Ramsey sentences. While the two Ramsey sentences both include a description of a manifold equipped with a metric tensor field, the Ramsey sentence for GR+ includes an additional  $t$ -term for the self-effacing aether.

If the Ramsey sentence for GR is not exactly satisfied, then the Ramsey sentence for GR+ is not exactly satisfied either, since they both include a continuous metric field. Moreover, if a theory of QG recaptures the empirical predictions of GR, then it recaptures the empirical predictions of GR+ as well. Given [ER], then, the conditions laid down by both the Ramsey sentence for GR and the Ramsey sentence for GR+ are satisfied well-enough to be able to say that the  $t$ -terms in both theories are realised by something in the more fundamental theory of QG. But then it seems we are forced to accept that the self-effacing aether is realised.

Now, one might try to avoid this outcome by further restricting the conditions under which realisation occurs in the absence of an exact realiser. In the case of QG, one might point out, there is nothing that approximates a self-effacing aether; but there is something that approximates spacetime. What we need to do, then, is take account of specific approximation relations between entities. The idea is to impose a further mapping requirement on the  $n$ -tuple that approximately realises a given functional role. It is not enough for an  $n$ -tuple to be globally similar to some  $n$ -tuple that would exactly realise a given functional role, were that  $n$ -tuple to exist. Rather, each member of the approximating  $n$ -tuple must itself approximate some member of an  $n$ -tuple that would exactly realise a given functional role. One way to capture this thought is to modify [ER] as follows. For theory  $T$  with Ramsey sentence  $R$ :

[ER\*] A theory  $T'$  posits an  $n$ -tuple of entities  $\langle e'_1 \dots e'_n \rangle$  that is sufficiently similar to whatever  $n$ -tuple of entities  $\langle e_1 \dots e_n \rangle$  would exactly realise the functional role specified by  $R$  in  $T$  just when (i) every confirmed empirical prediction of  $T$  is an empirical prediction of  $T'$  but not vice versa and (ii) for each  $e_n$  there is some entity posited by  $T'$  that approximates it.

To see how [ER\*] might help, consider the QED case again. While it is true that Lorentz's aether theory is empirically recaptured by QED, it is not the case that each member of the  $n$ -tuple that would exactly realise the functional specification supplied by the Ramsey sentence for Lorentz's aether theory is mapped, via an approximation relation, to some member of the  $n$ -tuple of entities posited by QED. In particular, there is no entity in QED that approximates the aether. So the  $t$ -terms in Lorentz's aether theory are not approximately realised in QED. The same basic reasoning applies to GR+ and QG.

The difficulty with [ER\*] however, is that it invokes a second notion of similarity: an approximation relation between pairs of entities. But then we can just ask the same question of this second notion of similarity that lead to [ER] in the first place. Namely, how similar do two entities need to be before we can say that one approximates the other? If we are too relaxed concerning the conditions for this second notion of similarity, then it will make the realisation of a self-effacing aether too easy. If we are too conservative, then it will make the realisation of spacetime too hard. It is unclear how to specify the relevant notion of similarity to achieve the right outcome. One might try to avoid the need to spell out the relevant notion of similarity at all by simply stating that each  $e_n$  is approximated by whichever  $e'_n$  is most similar to it in the underlying theory. However, this makes realisation far too easy. For if there is no lower-bound on the relevant degree of similarity, then anything, no matter how dissimilar, can qualify.

What we really need is an alternative to [ER] and [ER\*]. Specifically, we need a way to specify a similarity relation that doesn't force us to recognise the realisation of unwanted, unobservable entities. One option is to appeal to *explanation*. Let us say that a theory  $T'$  stands in a relationship of explanatory recovery to a theory  $T$  when everything that can be explained by  $T$  can also be explained by  $T'$  but not vice versa. An alternative to [ER] based on a notion of explanatory recovery can thus be stated as follows. For theory  $T$  with Ramsey sentence  $R$ :

**Explanatory Recovery [XR]** A theory  $T'$  posits an  $n$ -tuple of entities  $(e'_1 \dots e'_n)$  that is sufficiently similar to whatever  $n$ -tuple of entities  $(e_1 \dots e_n)$  would exactly realise the functional role specified by  $R$  just when everything that can be explained by  $T'$  can also be explained by  $T$  but not vice versa.

The shift from [ER]/[ER\*] to [XR] seems to help with the aether case. Consider Lorentz's aether theory once again. While SR and Lorentz's aether theory might be empirically equivalent, they may not be explanatorily equivalent. In particular, there may be things that Lorentz's aether theory can explain that SR cannot explain even though the two theories issue in the same empirical predictions.

Of course, we have to be a bit careful here, since there doesn't seem to be any observable fact that Lorentz's aether theory can explain that cannot be explained equally well by SR. But perhaps there are facts involving unobservables that are explained by the aether theory. Broadly speaking, there are three ways that unobservables might be involved in explanation. First, an explanation might use an unobservable to explain some observable fact (e.g., blackholes explain the lensing of light from distant stars). Second, an explanation might be given of why some unobservable fact holds (e.g., why are there blackholes?). Third, an explanation might be provided of why a particular unobservable fact is unobservable (e.g., blackholes are unobservable because they trap signals).

It seems unlikely that the aether theory provides explanations involving unobservables of the first two kinds. No new observable facts are explained by the aether theory and while the aether theory does include a new unobservable to be explained (the aether), there doesn't seem to be an explanation of this fact provided by the theory. The existence of the aether is, in the context of the theory, a brute fact. However, there may be a new explanation of the third kind given by Lorentz's aether theory, at least if we treat the aether as self-effacing in the sense described above. The very fact that the

aether cannot be observed is explained in terms of facts about the aether plus some further facts about the way that movement supposedly interacts with electrostatic forces that contract or expand our measuring devices. Granted, SR explains why the aether cannot be observed as well: it cannot be observed because it does not exist. However, the explanatory situation seems slightly different in so far as Lorentz's theory can explain why the aether is not observed despite the fact it exists; a fact that cannot be explained in SR.

If SR and Lorentz's aether theory are not explanatorily equivalent, then even if QED stands in a relationship of explanatory recovery to SR, there is no guarantee that it will stand in the same relationship to Lorentz's aether theory. This means that, with regard to the aether theory at least, QED will fail to satisfy a set of realisation conditions for approximate realisation that includes [XR]. At first glance, then, [XR] offers a way to avoid the difficulty with [ER]/[ER\*].

The difficulty, however, is not so easily dealt with. For even if an explanatory difference between SR and Lorentz's aether theory can be found, we can simply cook up a theory that effectively factors out explanatory differences of the relevant kind. Consider the QG case once again. Suppose we take GR+ and further stipulate that the 'self-effacing' aether does no explanatory work in that theory; at least, no work that is not already done within GR. One way to do this is to simply remove the explanation of why the aether cannot be observed. Instead of treating this as a matter of distortion to measuring devices, we simply stipulate that it is a brute fact within the theory that the aether cannot be detected. [XR] provides us with no way to prevent the aether in GR+ so construed from being functionally realised by a theory of QG.

One might object that this 'brute fact' version of GR+ is a terrible theory: one should not posit entities that are explanatorily idle. I agree! But that's no bar to our capacity to formulate a theory along these lines, and it is the fact that such a theory can be formulated that poses a problem for [XR]. Given the existence of an explanatorily idle version of GR+, [XR] incorrectly yields the result that such a theory, and its associated t-terms (including a t-term for the aether) are functionally realised in QG. [XR], then, is not a viable alternative to [ER]/[ER\*].

A second alternative to [ER] appeals to counterfactual dependence rather than to explanation. Let us say that a theory  $T'$  stands in a relation of counterfactual recovery to a theory  $T$  when every counterfactual that is true according to  $T$  is true according to  $T'$  and not vice versa. We can then say that for a theory  $T$  with Ramsey sentence  $R$ :

**Counterfactual Recovery 1 [CR1]** A theory  $T'$  posits an n-tuple of entities  $\langle e'_1 \dots e'_n \rangle$  that is sufficiently similar to whatever n-tuple of entities  $\langle e_1 \dots e_n \rangle$  would exactly realise the functional role specified by  $R$  just when every counterfactual that is true according to  $T$  is true according to  $T'$  and not vice versa.

Consider the following counterfactual: if the aether had not existed, light waves would not have propagated. This counterfactual appears to be true according to Lorentz's aether theory. The counterfactual does not appear to be true in the context of QED. Indeed, the counterfactual appears to be undefined, since the aether does not exist according to that theory. Given that there is at least one counterfactual that is true in the context of Lorentz's aether theory that is not true (because undefined) in the

context of QED, QED does not stand in a relationship of counterfactual recovery to Lorentz's aether theory, and so the t-terms in that theory are not functionally realised.

One difficulty with [CR1], however, is that the question of which counterfactuals are well-defined in the context of a theory and thus potentially true itself depends on putative facts about functional realisation. Consider, for instance, the following counterfactual: if the mass-energy distribution throughout spacetime had been different, then the curvature of spacetime would have been different. In order to demonstrate that this counterfactual is true according to some theory of QG, we must already be able to show that there is something in that theory that can play the spacetime role. Otherwise, we won't be able to even formulate the counterfactual within the context of the relevant theory of QG, let alone show that it is recoverable. But [CR1] is supposed to give content to the conditions under which something plays the spacetime role in the first place. It cannot do that work in any substantive fashion, however, if it already presupposes a way to establish functionalism. To put the point another way, if [CR1] holds then we must already have established that something in QG is sufficiently similar to an exact realiser of the spacetime role based on some other notion of similarity, at which point [CR1] becomes otiose.

Perhaps this difficulty can be avoided by Ramsifying all of the counterfactuals as well. Thus, we might take a counterfactual like "if the mass-energy distribution throughout spacetime had been different, then the curvature of spacetime would have been different" and replace any reference to t-terms like spacetime with free variables. Doing so yields the following schematised counterfactual: "if the mass-energy distribution throughout  $x$  had been different, then the curvature of  $x$  would have been different". We can then repeat this as needed for any further spatiotemporal notions, yielding a system of schematic counterfactuals that are true according to GR. We can then say that so long as there is something in QG that can be substituted in for the variables in these schematic counterfactuals to yield a true instance, then there is an n-tuple in QG that is sufficiently similar to an exact realiser of the spacetime role. Thus we can say that for a theory  $T$  with Ramsey sentence  $R$ :

**Counterfactual Recovery 2 [CR2]** A theory  $T'$  posits an n-tuple of entities  $\langle e'_1 \dots e'_n \rangle$  that is sufficiently similar to whatever n-tuple of entities  $\langle e_1 \dots e_n \rangle$  would exactly realise the functional role specified by  $R$  just when every schematised counterfactual that is true according to  $T$  is true according to  $T'$  and not vice versa.

It is not clear that [CR2] will work for QED and Lorentz's aether theory. Consider, again, the counterfactual "if the aether had not existed, light waves would not have propagated". We can Ramsify this counterfactual by removing reference to the aether, replacing it with a free variable, yielding "if  $x$  had not existed, light waves would not have propagated". The trouble, however, is that this schematised counterfactual might be true in QED. Indeed, it might be made true by whatever realises spacetime.

Perhaps, however, it is just a matter of choosing the right counterfactual to make [CR2] work. The right counterfactual will be one that, when schematised, cannot be filled in by anything posited within QED. Such a schematised counterfactual will involve the aether essentially, in this sense: only something that has the specific prop-

erties of the aether as specified by the Ramsey sentence for the Lorentzian theory can be substituted back into the schematised counterfactual to yield a true instance.

Unfortunately, we can always cook up a theory that contains a t-term for something that is ‘counterfactually idle’, in this sense: every counterfactual involving the relevant t-term is false. One way to do this is to add a t-term for something and stipulate that it is extremely modally fragile, in this sense: it cannot survive any counterfactual change to the world, and thus any counterfactual reasoning about the t-term is effectively blocked. Alternatively, we could add a t-term and stipulate that it is extremely modally robust, in this sense: it does not alter under any counterfactual variation to the world. So for instance, we could introduce a version of GR, GR+, that contains an aether and then just stipulate of the aether that it is not a difference-maker to the world. No matter how we change the world, or how we change the aether, there is no follow-on effect that results from making such a change. There is thus no counterfactual of the form ‘if the aether had not existed then ...’ that is true. Again such a theory might be a bad theory, but there doesn’t seem to be anything stopping us from inventing a theory along these lines.

As with [XR], the capacity to formulate such a theory is trouble enough for [CR2]. For such a theory will be counterfactually equivalent to GR, in this sense: every counterfactual that is true in a form of GR+ where the aether is counterfactually idle, will be true in GR as well and vice versa. That’s just because the only difference between the theories is the aether, and the aether in GR+ leads to no new counterfactuals. This means that if a theory of QG manages to satisfy [CR2] for GR, then it will satisfy it for GR+ as well, leading to the functional realisation of the aether once again.

The problem of counterfactual indolence is a problem for [CR1] as well. [CR2], however, faces a second problem that is unique to that condition. Rather than adding an aether to GR and stipulating that it is counterfactually idle, one could add the aether but imbue it with properties in such a way that it is far too easy to satisfy the schematised counterfactuals that we go on to produce. So, for instance, one might formulate a theory where the only counterfactuals that are true of the aether are like the schematised counterfactual above, namely “if x had not existed, then light waves would not have propagated”. The trouble with this counterfactual, as we saw, was that it is far too generic, running the risk that it will be satisfied within a theory of QG. If all of the counterfactuals about the aether are like this, then there won’t be a way to prevent the aether from being functionally realised via [CR2].

Now, one might object that if the only counterfactuals that are true of the aether are, when schematised, counterfactuals that can be satisfied by something in QG, then it is far from clear that we should deny that the aether is functionally realised. To be clear, however, the idea is *not* that the aether is defined solely in terms of the counterfactuals it supports. The aether is defined in terms of a Ramsey sentence, and the Ramsey sentence may attribute all kinds of bizarre properties to the aether. When we end up saying that the aether is realised, we are saying that something with these properties is functionally realised well-enough. So while we are committing to the truth of the relevant generic counterfactuals, the functionalist commitment goes beyond a commitment to the counterfactuals at issue. That’s why the realisation of the aether is a problem. We are not just admitting that something that obeys certain counterfactuals

is realised. We are saying that a specific kind of entity—namely an aether, with its attended properties—is realised well enough by some underlying theory.

If we set aside explanation and counterfactuals, are there any other options? One possibility appeals to implication instead. Thus, one might say that for a theory  $T$  with Ramsey sentence  $R$ :

**Implication Condition [IC]** A theory  $T'$  posits an n-tuple of entities  $\langle e'_1 \dots e'_n \rangle$  that is sufficiently similar to whatever n-tuple of entities  $\langle e_1 \dots e_n \rangle$  would exactly realise the functional role specified by  $R$  in  $T$  just when  $T'$  implies  $T$  and not vice versa.

[IC] preserves one aspect of Lewis's approach to approximate realisation (introduced in the previous section). One point in favour of [IC] is that there is a widely held expectation that GR will be derivable from QG in some sense (cf. Crowther, 2018). It is important, however, to be clear about the nature of such derivability. As Butterfield and Isham (1999) make clear, the kind of derivability expected in the case of QG and GR is one that permits a certain degree of approximation. Indeed, this is already clear to some extent in the failure of exact realisation discussed in Sect. 2. As noted there, in the case of some approaches to QG, such as LQG, the metric structure of GR is only recoverable in an approximate sense. What this means is that, using various mathematical approximation procedures, we can derive numerical values that correspond very closely to certain solutions to the field equations. What we cannot clearly do yet for every approach to QG is show that GR is a logical consequence of the theory at issue.

One might concede that GR does not logically follow from QG in a strict sense but instead focus on the mathematical approximation procedures involved in the kind of derivability at issue. Thus, we might modify [IC] into a more specific version of that condition that focuses on mathematical approximation. Thus, for theory  $T$  with Ramsey sentence  $R$ :

**Mathematical Approximation Condition [MAC]** A theory  $T'$  posits an n-tuple of entities  $\langle e'_1 \dots e'_n \rangle$  that is sufficiently similar to whatever n-tuple of entities  $\langle e_1 \dots e_n \rangle$  would exactly realise the functional role specified by  $R$  just when the mathematical structure of  $T$  can be approximated by the mathematical structure of  $T'$  and not vice versa.

The difficulty with [MAC], however, is that some account must be given of the relevant notion of mathematical approximation. Mathematical approximation is another similarity relation: we can approximate the mathematical structure of one theory with the mathematical structure of another theory to a greater or lesser extent. As was the case with [ER\*] we need a further account of this second similarity relation in order for [MAC] to do its job.

One option might be to build on a suggestion made by Butterfield and Isham (1999) about approximation, and say that  $x$  mathematically approximates  $y$  when, by neglecting certain quantities, or ignoring certain degrees of freedom, we can reproduce the mathematical structure of  $y$  exactly. But this just shifts the bump in the carpet: which quantities and degrees of freedom must we hold fixed? And how many quantities and how many degrees of freedom can be neglected? We could just say 'as much



as possible, up to the point of empirical recovery' but that just brings us back to the problems with [ER] and [ER\*]. Similar considerations apply if we appeal to notions of explanatory recovery, counterfactual recovery or implication. Perhaps there is some other way to specify how close a mathematical approximation needs to be before we can say functional realisation occurs but it is unclear what the relevant conditions might be.

At this point, I have run out of ideas for how we might specify the similarity relation needed to make approximate spacetime functionalism work. We cannot obviously specify such a relation in empirical, explanatory, counterfactual, deductive or mathematical terms. T-terms for unobservables are perhaps the most significant stumbling block toward specifying a viable condition on similarity.

In §4, I will sketch out a way forward in light of this difficulty. Before doing so, however, it is worth considering some objections to the discussion thus far. First, one might worry that the challenge is simply a problem of my own making. Recall that, at the end of §2, I reformulated Lewis's conditions for approximate realisation. As we saw, according to Lewis's conditions, when the functional role specified by a Ramsey sentence  $R$  for a theory  $T$  is not exactly satisfied, the t-terms in  $T$  are functionally realised so long as there is a theory  $T'$  with Ramsey sentence  $R'$  which specifies a functional role that is exactly satisfied such that  $T'$  is a slightly modified or corrected version of  $T$  and  $T$  implies  $T'$  but not vice versa. These conditions make no mention of a similarity relation between the n-tuple that realises the role defined by  $R'$  and the n-tuple that would realise the role defined by  $R$ . Thus, if we use Lewis's original conditions, one might argue, the challenge discussed in this section simply does not arise.

But note that Lewis must explain how 'slight' a slight modification or correction to a theory is supposed to be. An answer to this question would seem to require specifying a similarity relation between theories. Since each of the conditions specified above is, primarily, a relation between theories, these are the most plausible options for spelling out what a slight modification is. Thus, Lewis might say that  $T$  is a slight modification of  $T'$  when  $T$  stands in a relation of empirical recovery to  $T'$ . Or, he might say that  $T$  is a slight modification of  $T'$  when everything that can be explained by  $T'$  can be explained by  $T$  and on it goes. In the absence of some further option for specifying a similarity relation between theories, it seems that the challenge raised here will apply equally well to Lewis's account of approximate realisation.

Second objection: it could be argued that the challenge I have raised in this section is really a general challenge for *approximate functionalism* and not really a challenge for *approximate spacetime functionalism* per se. The conditions I have formulated are, after all, very general conditions, and they seem to fail on fairly general grounds. Whatever the correct solution to the challenge turns out to be for other versions of approximate functionalism, then, one can simply invoke that solution here too, as a way to develop approximate spacetime functionalism.

There are two things to say here. First, even if the challenge is a general challenge, there is still a need for a general solution that is applicable in the case of spacetime functionalism. Without an answer to the challenge, it is unclear how to make approximate functionalism of any kind work, including approximate spacetime functionalism. Second, it is unclear that a solution to the challenge for the case of say, pain, can be

straightforwardly transported to the case of spacetime. It may turn out that the conditions under which approximate functionalism is true are domain specific. Thus, the conditions under which pain or, indeed, mental phenomena quite generally, get realised may differ from the conditions under which spacetime is realised. It may be, then, that there is no general solution to the challenge, despite the challenge being a general one. If that's right, then a specific solution for the case of spacetime is required.

The third objection focuses on the Ramsey–Lewis procedure used to specify functionalism in the first place. One might object that specifying the functional role via a Ramsey sentence involves commitment to a specific form of functionalism, namely realiser functionalism. According to realiser spacetime functionalism, spacetime is identified with whatever plays the spacetime role. Realiser spacetime functionalism is to be contrasted with role spacetime functionalism, according to which spacetime is identified with a specific higher-order property: namely, the property of a physical system being in some state or other that realises the spacetime role. The challenge that I have outlined, one might argue, is at best a challenge for a realiser version of approximate spacetime functionalism, and so can be avoided by switching to role functionalism.

True enough, the Ramsey–Lewis procedure is commonly associated with realiser functionalism. This is because, according to Lewis, the *t*-terms in a given theory *T* with Ramsey sentence *R* denote the members of whichever *n*-tuple realises the functional role specified by *R*. Thus, 'spacetime' just names a certain realiser. Lewis's account of denotation, however, is not an essential feature of the broad Ramsification procedure. One can specify a functional role in terms of a Ramsey sentence without committing to the further claim that the *t*-terms in a theory denote whatever plays the functional role specified in this way. Indeed, one can combine the Ramsey–Lewis picture with the claim that the *t*-terms denote a higher-order property of some physical system: namely, the system that is in a state that plays the relevant functional role. Strictly speaking, there is no logical requirement to adopt a particular account of the denotation of *t*-terms in combination with the Ramsification procedure. The challenge that I have raised targets the Ramsification procedure and not Lewis's account of the denotation of *t*-terms. Accordingly, the problem arises for any form of functionalism that specifies functional roles in the relevant way, including role functionalism.

Now, one might grant the point but blame the Ramsification procedure itself, rather than realiser functionalism. One might argue that the challenge I have outlined only arises if we adopt a Ramsey–Lewis approach to functionalism in the first place. In reality, however, the Ramsey–Lewis procedure is just a convenient framework for formulating the challenge. The challenge is likely to arise no matter how functionalism is specified, so long as the functional role for spacetime is linked to GR. Regardless of how we specify the functional role for spacetime in GR, it is likely that the metric structure of spacetime will be an important part of that functional specification. Since the metric is, at best, approximated within several approaches to QG, it is unlikely that the spacetime role will be exactly realised. The challenge thus arises because of the importance of continuous metric structure to spacetime in GR and the lack of continuous metric structure in some approaches to QG. It has little to do with exactly how functional roles get specified.

Finally, one might object that the challenge I have raised is a result of focusing on a specific account of the spacetime role, one that is tightly connected to GR. Shifting to an alternative specification will alleviate the challenge by obviating the need for approximate spacetime functionalism. One might, for instance, follow Knox and specify the spacetime role in inertial terms. Spacetime, according to Knox (2019), is anything that yields a structure of inertial frames. Alternatively, one might adopt an account of the spacetime role that is focused on what Chalmers (forthcoming) calls ‘phenomenal space’. Which is to say that we might try to link the spacetime role to our everyday concepts of space and time, rather than to some theoretical concept in science. Another possibility is to focus on an operational conception of spacetime, according to which the functional role for spacetime is to provide a basis for measuring devices like clocks and rods.

I have nothing against these alternative approaches to spacetime functionalism. A difficulty emerges, however, once we are forced to try and select one as the *correct* account of the spacetime role. For how are we supposed to make this decision? One option is to rely on pre-theoretic intuitions about spacetime. But it is unclear how much weight we should put on such intuitions in general and, besides, it is unclear that we even have robust intuitions about the spacetime role in the first place. Another option might be to argue for a particular spacetime role based on GR. The thought being that such and such a specification of the spacetime role is correct because it is closest to some exact specification of the spacetime role in GR. If we take this line, however, the challenge discussed in this section is simply recreated at the level of functional roles. For now we have to specify how similar a functional specification needs to be to the Ramsey sentence for GR before we can say that it is the correct account of the spacetime role. Once again, it is unclear how to spell out the relevant notion of similarity.

Of course, we can simply give up the idea that there is a ‘correct’ spacetime role and just allow that there are multiple notions of spacetime, each with its own functional specification. As Chalmers points out, if we go this way, then the question of what the functional role for spacetime might be appears to be a largely terminological matter:

In practice this may end up as a largely verbal dispute, which we can resolve by distinguishing multiple notions of spacetime. For example we may distinguish inertial spacetime and operational spacetime, and make the case that the theory vindicates one but not the other. (Chalmers forthcoming)

To be clear, I don’t see any problem with specifying different conceptions of spacetime in this manner. Recall, however, that spacetime functionalism was introduced as one way to address a worry concerning empirical confirmation in the context of QG. If we shift our focus away from Ramsifying GR, we are more or less conceding that there is nothing that plays the spacetime role in that sense. What we need to work out, then, is whether if we give up on even approximately satisfying the functional role as specified by the Ramsey sentence for GR, we can still address lingering concerns regarding empirical confirmation. My goal in the final section is to make a preliminary case for the affirmative, thereby sidestepping the challenge facing approximate spacetime functionalism.

## 4 Functions not functionalism

As noted, I do not see the challenge presented above as decisive. But nor am I sanguine about the prospects for specifying a similarity relation that can then be used to underwrite a viable form of approximate spacetime functionalism. I thus recommend giving up on the functional realisation of spacetime, at least in so far as the functional role is tied to GR. As noted, this leaves open the possibility that spacetime in some other sense is functionally realised, though here too we face the question of what the 'correct' spacetime role might be, and no clear way to settle the matter.

What I recommend, instead, is that we focus on the functions of spacetime, rather than on a complete functional specification of spacetime. Specifically, we should focus on those functions of spacetime that play a role in grounding empirical observation and confirmation. That's because, as discussed in Sect. 1, the appeal to spacetime functionalism in QG is motivated by the need to establish the empirical coherence of specific theories. If we can show that the functions of spacetime that support observation and confirmation are performed by something in QG, then that is sufficient to address any concerns one might have regarding empirical incoherence.

Specifically, I recommend rejecting the assumption, freely granted in §1 of this paper, that observation is essentially tied to location in spacetime. If observation is essentially tied to location in spacetime then, clearly, for a given approach to QG to be empirically coherent, spacetime must be recovered. But if we give up the connection, then there is less pressure to ensure that spacetime is recovered. This is not to say that we should sever the tie between observation and *location*. The thought, rather, is that location in the context of observation need not be spatiotemporal location. What we need for observation to occur is just for entities to be located in some meaningful sense.

The point can be put in terms of a more general function, that I will call the *location function*. One of the things that spacetime does, is provide locations for entities. There are, however, other kinds of structures that are capable of providing locations for entities. Spacetime is just one structure among many that is capable of doing this kind of work. An absolute structure of space and time, for instance, can also perform the location function, which is why it is meaningful to say that entities are located in a Newtonian world. More generally, any structure that can be modelled either exactly or approximately by an  $n$ -tuple of geometrical objects defined over a manifold, including discrete geometric structures (such as those described in Brass et al., 2005; Chen, 2014), has the potential to perform the location function.

Note that this is not a complete specification of the location function. It is, at best, a sufficient condition for performing the function (and one that I offer tentatively; as I note below we likely need input from physics to settle the matter). If something along these lines is right, however, then the location function is quite broad. In particular, while it is tied to geometry, it is not tied to spacetime. To see this, consider a geometric colour space (see, for instance, Logvinenko, 2015; Griffin & Mylonas, 2019; Provenzi, 2020). Such a space allows us to specify the distance between colours, and consider smooth transformations from one colour to another. On the geometric conception of location, a colour space also provides genuine locations for colours within that space. Such colour locations are not, however, spatial, temporal or spatiotemporal in nature.

When it comes to concerns about empirical coherence, then, we only need to ensure that one of the functions that spacetime performs—that of locating entities in the above general sense—is performed by something. We can thus avoid difficult questions about what it takes to functionally realise spacetime by focusing, directly, on the location function in this way. Indeed, whether it is spacetime that performs the function of locating entities or something else matters much less than whether the function is performed at all. In the case of LQG, for instance, a discrete geometry defined directly over the spinfoam structure may perform the location function even if that geometry ultimately lacks spatiotemporal properties.

Of course, in order to accommodate the impressive empirical results that confirm GR, we must recover a notion of location that at least approximates spatiotemporal locations within the regime where GR is successful. It is, however, important to disentangle the empirical adequacy of a given approach to QG from any worries about empirical incoherence. Granted, the two issues are logically related in the following sense: an approach to QG cannot be empirically adequate without being empirically coherent. But the reverse is not true: in so far as bare empirical coherence is concerned any viable picture of location will do.<sup>7</sup>

At the end of the previous section I noted that there are ways of defining a functional role for spacetime that don't involve Ramsifying GR. One might worry, then, that my proposal is really just a version of one of these other approaches. By focusing on a particular function of spacetime—namely, that of providing locations—one might object, I am simply adopting a particular understanding of what spacetime is, and thus endorsing some brand of spacetime functionalism after all.

I concede that what it takes to support location may require the presence of structures that realise some of the other spacetime roles that have been specified in the literature. So if one wants to say that spacetime in some sense is functionally realised, then so be it. What I want to emphasise, however, is that the functional realisation of spacetime is entirely optional, at least in so far as empirical coherence is concerned. There is no need to agree on a particular conception of spacetime, and there is no need to say anything about what the functional role for spacetime might be. The advantage of my approach is that we don't need to engage in a further debate over spacetime functionalism for the purposes of understanding how a given approach to QG can be empirically coherent.

The proposal does assume, however, that there is good sense to be made of location in non-spatiotemporal terms. So how should we understand the non-spatiotemporal notion of location at issue? That, I submit, is a matter for physics. Whatever non-spatiotemporal, physical structure is needed for an approach to QG, that is the structure that we should take to yield locations for entities and thus ultimately ground empirical coherence and confirmation. In a certain sense, then, the approach I am advocating is an

---

<sup>7</sup> Even if we focus on empirical adequacy, it is not obvious that we need to recover spacetime. All we need to show is that there is something in a theory of QG that is empirically indistinguishable from spacetime under certain conditions. I suppose one could take the view that if  $x$  and  $y$  are empirically indistinguishable, then  $x$  just is  $y$ , in which case the recovery of spacetime is a requirement after all. But this would be to adopt an extremely strong positivist condition on identity, one that is difficult to defend. In a similar vein, one might argue that empirical indistinguishability is sufficient for functional realisation, and so the realisation of spacetime is required after all. This, however, is to commit to something like [ER] or [ER\*] which, as we have seen, are not very plausible conditions on functional realisation.

instance of the ‘top down’ approach to empirical coherence ultimately recommended by Huggett and Wüthrich (2013, pp. 284–285). Roughly speaking, a ‘bottom up’ approach to the issue involves coming to QG with a pre-defined understanding of what empirical confirmation requires, and then taking a theory to task for failing to satisfy that understanding. In a ‘top down’ approach, by contrast, we allow that a theory might revise the way we understand empirical coherence, and thus that the theory may be empirically coherent by its own lights. I am recommending a ‘top down’ approach to (at least) location: we should allow that our understanding of location might be updated by a new physical theory and, in this way, that the theory may secure empirical coherence on its own terms.

That being said, I suspect that we may need to use philosophy as a basis for specifying some very general constraints on the nature of location, at least as a starting point for interpreting physics. Such constraints likely include structural features like distance, contiguity and coincidence relations. This suggests a method of reflective equilibrium, whereby we start with a philosophical understanding of location and of the location function and allow that our physical theories update that understanding to a certain extent, but only up to a point. When our physical theories recommend altering the conception of location so dramatically that the notion is no-longer recognisable as location in the barest sense, we can take that as a cue to rework our understanding of the physical theories at issue. By working back and forth between philosophy and physics in this way, we can produce an understanding of location that supports empirical coherence in the absence of spacetime.

Given a ‘top down’ approach along these lines, the question of how a given theory manages to locate entities can really only be determined on a theory-by-theory basis. Some of the work that spacetime functionalists have already completed is relevant here. Lam and Wüthrich, for instance, argue that in the approach to QG discussed briefly in §2—LQG—there is something that plays the ‘localising function’ of spacetime. This localising function accords with what I have called the location function. Here’s what they say about the LQG case:

The ‘localizing function’ ... involves at its core the notions of coincidence and contiguity: the functional role of localizing a physical entity (what it means for a physical entity to be localized) crucially involves coincidence and contiguity relations. As a consequence, in schematic terms, the second step (FR-2) crucially involves how LQG entities (or LQG properties) can instantiate coincidence and contiguity relations in an appropriate context. (Lam & Wüthrich, 2018, p. 48)

They go on to demonstrate that in LQG there is reason to think that spin-networks can fill the location function by giving rise to coincidence and contiguity relations. They take this to be a step along the road to spacetime functionalism, and thus ultimately as part of their demonstration that spacetime metric structure is functionally realised in QG. If we set these functionalist ambitions aside, however, then we can just rest easy with the location function itself. In which case Lam and Wüthrich’s demonstration that something performs the function of locating entities in LQG is already enough to address concerns about empirical coherence. As they put the point, “no unfinished business lingers on” (Lam & Wüthrich, 2018, p. 44) but not because space-

time is functionally realised. Rather, the job is done so long as something performs the function of locating entities.

## 5 Conclusion

It is time to take stock. I have presented a challenge for a version of spacetime functionalism that anchors the spacetime role to GR. The challenge is to provide a set of realisation conditions that can accommodate the fact that the spacetime role is unlikely to be exactly realised in at least some approaches to QG. This, in turn, involves providing an account of how similar something must be to an exact realiser of the spacetime role before we are permitted to say that functional realisation occurs. I considered a number of options for specifying a viable notion of similarity, and found them all wanting. I went on to suggest a way forward: give up on the realisation of the spacetime role. Look, instead, at certain functions of spacetime, namely those implicated in observation and confirmation, with an eye to explaining how these functions can be performed by something in a given theory of QG. I focused, in particular, on a location function, suggesting that if the location function is satisfied then empirical observation is possible. If that's right, however, then the discussion of spacetime functionalism in QG is useful as a way to highlight the importance of location, but it is ultimately location that we should be seeking to understand. At present, however, we don't really understand location all that well, and nor do we have a general account of the relationship between location on the one hand, and empirical confirmation and observation on the other. What location is, and how it works, emerges as an important topic for scientifically informed philosophical investigation.

**Acknowledgements** Many thanks to Baptiste Le Bihan, David Chalmers, David Ripley and Christian Wüthrich for useful discussion of the ideas in this paper. Thanks also to audiences at the Geneva Symmetry Group and the Dianoia Work in Progress Seminar for useful feedback on earlier versions of the paper. Work on this paper was supported by a Discovery Early Career Researcher Award (DE180100414) and a Discovery Project (DP180100105) from the Australian Research Council.

**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions

**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

- Barbour, J. (1994). The timelessness of quantum gravity: I. The evidence from the classical theory. *Classical Quantum Gravity*, 11, 2853–2873.
- Belot, G. (2011). Background-independence. *General Relativity and Gravitation*, 43, 2865–2884.

- Braddon-Mitchell, D., & Miller, K. (2019). Quantum gravity, timelessness and the contents of thought. *Philosophical Studies*, 176, 1807–1829.
- Brass, P., Moser, W. O. J., & Pach, J. (2005). *Research problems in discrete geometry*. Springer.
- Butterfield, J., & Isham, C. (1999). On the emergence of time in quantum gravity. In J. Butterfield (Ed.), *The arguments of time* (pp. 111–168). Oxford University Press.
- Carlip, S. (2014). Challenges for emergent gravity. *Studies in History and Philosophy of Modern Physics*, 46, 200–208.
- Chalmers, D. J. (forthcoming). Finding space in a nonspatial world. In B. Le Bihan, & N. Huggett (Eds.), *Philosophy beyond spacetime*. Oxford University Press.
- Chen, L. (2014). *Digital and discrete geometry: Theory and algorithms*. Springer.
- Crowther, K. (2016). *Effective spacetime: Understanding emergence in effective field theory and quantum gravity*. Springer.
- Crowther, K. (2018). Inter-theory relations in quantum gravity: Correspondence, reduction, and emergence. *Studies in History and Philosophy of Modern Physics*, 63, 74–85.
- Crowther, K., Linnemann, N. S., & Wüthrich, C. (2020). Spacetime functionalism in general relativity and quantum gravity. *Synthese*. <https://doi.org/10.1007/s11229-020-02722-z>.
- Griffin, L. D., & Mylonas, D. (2019). Categorical colour geometry. *PLoS ONE*, 14, e0216296.
- Healey, R. (2002). Can physics coherently deny the reality of time? In C. Callender (Ed.), *Time, reality & experience* (pp. 293–316). Cambridge University Press.
- Hu, B.-L. (2009). Emergent/quantum gravity: Macro/micro structures of spacetime. *Journal of Physics: Conference Series*, 174, 1–16.
- Huggett, N., & Wüthrich, C. (2013). Emergent spacetime and empirical (in)coherence. *Studies in History and Philosophy of Modern Physics*, 44, 276–285.
- Jacobson, T., & Mattingly, D. (2001). Gravity with a dynamical preferred frame. *Physical Review D*, 64, 024028.
- Knox, E. (2019). Physical relativity from a functionalist perspective. *Studies in History and Philosophy of Modern Physics*. <https://doi.org/10.1016/j.shpsb.2017.09.008>.
- Lam, V., & Wüthrich, C. (2018). Spacetime is as spacetime does. *Studies in History and Philosophy of Modern Physics*, 64, 39–51.
- Lam, V., & Wüthrich, C. (2020). Spacetime functionalism from a realist perspective. *Synthese*. <https://doi.org/10.1007/s11229-020-02642-y>.
- Le Bihan, B. (2018a). Priority monism beyond spacetime. *Metaphysica*, 19, 95–111.
- Le Bihan, B. (2018b). Space emergence in contemporary physics: Why we do not need fundamentality, layers of reality and emergence. *Disputatio*, 10, 71–95.
- Le Bihan, B. (2021). Spacetime emergence in quantum gravity: Functionalism and the hard problem. *Synthese*, 199, 371–393.
- Lewis, D. (1970). How to define theoretical terms. *Journal of Philosophy*, 67, 427–446.
- Lewis, D. (1972). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*, 50, 249–258.
- Logvinenko, A. D. (2015). The geometric structure of color. *Journal of Vision*, 15, 1–9.
- Ney, A. (2015). Fundamental physical ontologies and the constraint of empirical coherence: A defense of wave function realism. *Synthese*, 192, 3105–3124.
- Polger, T. W. (2007). Realization and the metaphysics of mind. *Australasian Journal of Philosophy*, 85, 233–259.
- Provenzi, E. (2020). Geometry of color perception. Part 1: Structures and metrics of a homogeneous color space. *Journal of Mathematical Neuroscience*. <https://doi.org/10.1186/s13408-020-00084-x>.
- Ramsey, F. P. (1931). *The foundations of mathematics and other logical essays*. Routledge.
- Regge, T. E. (1961). General relativity without coordinates. *Nuovo Cimento*, 19, 558–571.
- Rovelli, C. (2004). *Quantum gravity*. Cambridge University Press.
- Rovelli, C. (2011a). Loop quantum gravity: The first 25 years. *Classical and Quantum Gravity*, 28, 153002.
- Rovelli, C. (2011b). A new look at loop quantum gravity. *Classical and Quantum Gravity*, 28, 114005.
- Rovelli, C., & Vidotto, F. (2014). *Covariant loop quantum gravity: An elementary introduction to quantum gravity and spinfoam theory*. Cambridge University Press.
- Shoemaker, S. (1981). Some varieties of functionalism. *Philosophical Topics*, 12, 93–119.
- Vassallo, A. (2020). Dependence relations in general relativity. *European Journal for Philosophy of Science*, 10, 1–28.



- Wüthrich, C. (2017). Raiders of the lost spacetime. In D. Lehmkuhl, G. Schiemann, & E. Scholz (Eds.), *Towards a theory of spacetime theories* (pp. 297–335). Birkhäuser.
- Wüthrich, C. (2019). The emergence of space and time. In S. Gibb, R. F. Hendry, & T. Lancaster (Eds.), *Routledge handbook of emergence*. Routledge.
- Yates, D. (2021). Thinking about spacetime. In C. Wüthrich, B. Le Bihan, & N. Huggett (Eds.), *Philosophy beyond spacetime*. Oxford University Press.

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