

# Hawking radiation and analogue experiments: A Bayesian analysis

Radin Dardashti <sup>a</sup>, Stephan Hartmann <sup>b</sup>, Karim Thébault <sup>c,\*</sup>, Eric Winsberg <sup>d</sup>

<sup>a</sup> Bergische Universität Wuppertal, Germany

<sup>b</sup> Ludwig-Maximilians-Universität München, Germany

<sup>c</sup> University of Bristol, United Kingdom

<sup>d</sup> University of South Florida, Florida, USA



## ARTICLE INFO

### Article history:

Received 25 June 2018

Received in revised form

21 January 2019

Accepted 9 April 2019

Available online 11 May 2019

## ABSTRACT

We present a Bayesian analysis of the epistemology of analogue experiments with particular reference to Hawking radiation. First, we prove that such experiments can be confirmatory in Bayesian terms based upon appeal to 'universality arguments'. Second, we provide a formal model for the scaling behaviour of the confirmation measure for multiple distinct realisations of the analogue system and isolate a generic saturation feature. Finally, we demonstrate that different potential analogue realisations could provide different levels of confirmation. Our results provide a basis both to formalise the epistemic value of analogue experiments that have been conducted and to advise scientists as to the respective epistemic value of future analogue experiments.

© 2019 Elsevier Ltd. All rights reserved.

## 1. Introduction

Recent years have seen a proliferation of experiments designed to probe the phenomenon of Hawking radiation via analogue black hole systems. Reports on these experiments include claims of observation of classical, thermal aspects of Hawking radiation in an analogue white hole created using surface water waves (Weinfurter et al. 2011, 2013) and observation of quantum entanglement across an analogue horizon, as is expected from Hawking radiation, using a Bose-Einstein Condensate (BEC) (Steinhauer, 2016a). Despite these successes, an uncontested and repeated experimental demonstration of the full analogue Hawking phenomenon is still to be provided.<sup>1</sup> The not fully settled experimental status of analogue Hawking radiation notwithstanding, there is a clear need for methodological appraisal of what, in principle, can be established about black hole Hawking radiation based upon analogue experiments. The most pressing questions is whether or not there are circumstances in which analogue

experiments can be taken to provide inductive support for conclusions about astrophysical black holes. Can analogue experiments provide 'experimental confirmation of Hawking's prediction' (Jeff Steinhauer quoted in Haaretz (Efrati 2016)), or are they simply 'amusing feat[s] of engineering' that 'won't teach us anything about black holes' (Daniel Harlow quoted in *Quanta* (Shmahalo 2016)).

In this paper we will substantially extend previous philosophical work characterising analogue black hole experiments as a form of 'analogue simulation' (Dardashti, Thébault, & Winsberg, 2017; Thébault, 2019) via application of Bayesian confirmation theory. In that previous analysis, emphasis was placed upon the *qualitative* claim that certain 'universality arguments' can be used to link evidence about the 'source' dumb hole system to the 'target' black hole system. The results of this paper are *quantitative* in nature and licence normatively valuable conclusions regarding the structure of such inferences.

We first present a Theorem that demonstrates how the confirmation claim can be qualitatively characterised in Bayesian terms. The role of the universality arguments is understood in terms of support for a background assumption that is common between the source and target models. This means that there is a binary variable that can be assumed to be positively correlated with the empirical adequacy of both the source and target models. Evidence in favour of the model of the source system can thus be used to make inferences about the target system. Although not in-and-of-itself a validation of the argument for confirmation via analogue simulation, the Bayesian analysis demonstrates the internal consistency of the informal

\* Corresponding author.

E-mail address: [karim.thebault@bristol.ac.uk](mailto:karim.thebault@bristol.ac.uk) (K. Thébault).

<sup>1</sup> For more on surface water wave experiments see Rousseaux, Mathis, Maïssa, Philbin, and Leonhardt (2008, 2010); Michel and Parentani (2014); Unruh (2014); Euvé, Michel, Parentani, Philbin, and Rousseaux (2016); Torres et al. (2017); Euvé et al. (2018). For further results and discussion of the Steinhauer's BEC experiments see Steinhauer (2014, 2015); Finke, Jain, and Weinfurter (2016); Steinhauer (2016b); de Nova et al. (2018); Leonhardt (2018).

arguments given in (Dardashti et al., 2017). Furthermore, the formal model provides a powerful evaluative and heuristic tool for the further analysis of the structure of the inferences involved in cases of analogue simulation. In particular, in the remains of the paper, we present two further results that we expect will be of interest to contemporary analogue black hole experimentalists.

Our second principal result relates to the behaviour of the confirmation measure in the context of experimental realisations of the analogue system in different types of media. This is of particular relevance to contemporary analogue black hole research where a diverse range of analogue realisations of Hawking radiation are being pursued: e.g., surface waterwaves, BECs, superfluid helium-3, moving optical media. The immediate question in this context is how the number of distinct types of analogue system one constructs relates to the confidence one should have in the astrophysical effect. The second key result of this paper is a formal model for 'multiple source' analogue simulation displaying the generic feature of 'saturation' in confirmatory power with an increase in the number of sources. Significantly, the saturation in confirmation indicates that, under plausible assignments of priors, even an extraordinarily large range of diverse analogue experiments will not lead to conclusive confirmation of astrophysical Hawking radiation. This is in tune with the scientific intuition that there is a limit to what can be learned about astrophysical Hawking radiation via analogue experiments.

Finally, and perhaps most intriguingly, there is the question of whether different potential analogue realisations could provide different levels of confirmation. Would we learn more about astrophysical black holes from an analogue experiment based upon liquid helium or BECs? The third key result of the paper is a Theorem proving that analogue experiments in which we are more confident about the fundamental physics (e.g. BECs) teach us less about the target system than those about which we are less confident (e.g. superfluid helium-3). Our results thus provide a basis to both formalise the epistemic value of analogue experiments that have been conducted and to advise scientists as to the respective epistemic value of future analogue experiments. As such, our work demonstrates the enduring value of the Bayesian framework as a tool for analysing the protean forms of scientific inference.

## 2. Confirmation, analogy and experiment

The literature on analogical reasoning in science is fairly extensive, with particularly noteworthy contributions by Keynes (1921), Hesse (1963; 1964; 1974), Bailer-Jones (2009) and Bartha (2010; 2013).<sup>2</sup> Typically analogical arguments have the form of abstract speculative inferences regarding possible features of one system ('the target') based on known features of another system ('the source'). Classic examples are Reid's argument for the existence of life on other planets based upon life on earth (Reid & Hamilton, 1850) or Hume's argument for animal consciousness based upon human consciousness (Hume, 1738/1978).

Analogical arguments evidently play an important *heuristic role* in scientific practice in that they provide 'cognitive strategies for creative discovery' (Bailer-Jones, 2009, p. 56). The *epistemic role* of analogical arguments in science is, however, more controversial. In particular, the literature contains diverging answers regarding whether analogical arguments can provide Bayesian confirmation of a hypothesis regarding the target system. From a Bayesian

perspective on confirmation, evidence for a hypothesis can count as confirmatory only if the probability of the hypothesis given the evidence together with certain background assumptions is larger than the probability of the hypothesis given only the background assumptions. In a detailed and nuanced treatment of the issue, Hesse (1974, pp. 208–19) suggests that analogical arguments *can in some cases* be confirmatory in a Bayesian sense, so long as the analogical relationship that holds is in terms of what she calls a 'material analogy'. That is, where there is a similarity relation of sharing at least one predicate between the target and source systems.<sup>3</sup>

Contrastingly, Bartha, 2010; 2013 offers arguments that analogical arguments *cannot in principle* be confirmatory in a Bayesian sense. In particular, he suggests that because the information encapsulated in an analogical argument is taken to already be part of the 'background knowledge', the probability of a hypothesis regarding the target system must be identical before and after including the analogical argument. Bartha takes this instance of the familiar problem of old evidence (Glymour, 1980) to be significant enough to bar analogical arguments from being confirmatory in Bayesian terms. Rather, following Salmon, 1967; 1990, Bartha argues that arguments by analogy can establish only the *plausibility* of a conclusion in the technical sense of justifying the assignment of a non-negligible prior probability assignment (Bartha, 2010, §8.5). On this analysis, it is not *in principle* possible for analogical arguments to confer inductive support for a hypothesis. That is, although analogical arguments can certainly be stronger or weaker, even the strongest possible analogical argument cannot confer confirmation in a Bayesian sense: they are abstract inferences that can only ever support plausibility claims rather than providing inductive evidence.

Although worthy and insightful, the treatments of Hesse (1974) and Bartha (2010) do not extend to the analysis of analogue experiments. This is because analogue experiments are unlike arguments by analogy in exactly the respects that are crucial for either the Hesse or the Bartha analysis to go through. The question of whether one agrees with Hesse or Bartha about the confirmatory power of arguments by analogy is simply tangential to the analysis of the confirmatory power of analogue experiments. In the case of Hesse, this is indicated by the fact that her notion of material analogy is too strict to accommodate the subtle relation that the model of the target has to the model of the source in the case of analogue black hole experiments. We are dealing with an *analogue simulation* that does not involve a material analogy in the sense of Hesse since there is not a *physical property* common between the target and source systems. Rather than a material relation between systems, we have a syntactic isomorphism between models.<sup>4,5</sup>

Bartha's negative analysis of the prospect for confirmation via analogical argument is similarly inapplicable to analogue experiments. This is because analogue experiments, unlike analogical arguments, are essentially empirical activities: they involve learning about the world by manipulating it. In analogue black hole experiments we manipulate the source such that certain explicit modelling assumptions matching those for the target obtain.

<sup>3</sup> Hesse (1974, p. 216) explicitly rules out the possibility of confirmation obtaining in cases where there is purely a 'formal analogy'. That is, where target and source are both interpretations of the same formal calculus but do not share material similarities.

<sup>4</sup> Here 'syntactic isomorphism' is a natural generalisation of Hempel's, 1965 notion of nomic isomorphism to the case of a relation between models rather than laws.

<sup>5</sup> Hesse (1963), does in fact, rather presciently, consider the relevance of simulators in her account of models and analogies in science. Tantalisingly, she says that analogue machines (i.e. simulators): 'are useful and necessary as predictive models precisely in those cases where the material substance of parts of the analogue is *not essential* to the model, but where the mutual relations of the parts are essential' (p. 102) This connection is unfortunately not taken up in the 1974 Bayesian analysis.

<sup>2</sup> See also Norton (2011) for an importantly different take on analogical arguments. Norton's analysis focuses on analogical arguments that proceed via subsumption of the target system into a larger class of entities, including the source system. There are broad parallels between the structure of such inferences and our analysis.

Analogue simulation thus resembles a form of experimentation, involving the ‘programming’ of a physical system such that it can be used to ‘simulate’ another physical system. Thus, we see that conclusions from the philosophical analysis of analogical argument should not be taken to be readily extendible to cases of analogue simulation in contemporary science. In particular, it is self-evidently the case that the old evidence problem for the Bayesian analysis of traditional arguments by analogy à la Hesse, is no longer relevant for analogue experiments. Analogue experiments unlike analogical reasoning explicitly involve the collection of new evidence and there are not good grounds for relegating their significance to mere plausibility.

This collection of new evidence motivates us to consider the ‘epistemology of analogue experimentation’ in parallel with the epistemology of conventional experiments. As has been noted by various authors (Franklin, 1989; Winsberg, 2010; Franklin & Perovic, 2016) conventional experiments are generally only of epistemological significance when supplemented by arguments that imply that the information learned about the system being manipulated (‘the source’) is relevantly probative about the class of systems that are of interest to the experimenters (‘the target’). A nice illustration of this point is provided by experiments designed to learn about the thermal conductivity of the iron in Earth’s core (Dobson, 2016; Konôpková, McWilliams, Gómez-Pérez, & Goncharov, 2016). The experiments were carried out in the lab using samples of iron that are placed in a laser-heated diamond-anvil cell. The pressure and temperature that the iron samples were subjected to were specifically matched to those relevant to the cores of Mercury-sized to Earth-sized planets. For the experiments to achieve their epistemic purpose, they must be supplemented with arguments that the measurements are be ‘relevantly probative’ of the thermal conductivity of iron in the core’s of Mercury-sized to Earth-sized planets. That is, there must be a basis to generalise from the observations regarding the lab based system (the ‘source’), to relevant systems outside the lab (the ‘target’). In parallel, what we take to be the key question in the epistemology of analogue experimentation is whether we can provide arguments that the relevant source systems ‘stand-in’ for the target systems to which the analogical relationship refers. Can we find arguments that analogue black hole experiments are relevantly probative of the thermal properties of astrophysical black holes?<sup>6</sup>

### 3. Hawking radiation and universality

Hawking radiation (Hawking, 1975) is a thermal phenomenon that is predicted to be generically associated with black holes. Despite the absence of either a clear causal process behind the radiation or experimental evidence, it is widely believed in by theoretical physicists. In fact, the ability to recover Hawking radiation could even be taken as a theoretical test of prospective theories of quantum gravity, much like the recovery of the Bekenstein-Hawking formula for black hole entropy (Wüthrich, 2019). There are two connected reasons why physicists are so confident in the prediction of Hawking radiation. First, given the Unruh effect (Unruh, 1976), which associates a temperature with acceleration, Hawking radiation seems to be directly implied by the equivalence principle.<sup>7</sup> Second, starting from Hawking’s original calculation a

remarkable number of different derivations of the effect have been given.<sup>8</sup> The consensus is that the effect is ‘remarkably robust’ to the addition of complicating factors to the original derivation. The overall implication is that very general theoretical constraints coming from quantum field theory and general relativity (two well tested theories) necessitate that something like Hawking radiation must exist. The purpose of this paper is not to address the evidential import of such theoretical considerations. Rather, our focus is on the potential for analogue experiments to provide confirmatory evidence of a form akin to conventional experiments. This notwithstanding, questions of robustness will return to the fore in the context of a particular form of universality argument that will be found to be central for questions of confirmation. Before then, it will be instructive to consider the basic elements of the original Hawking derivation of a radiative flux for astrophysical black holes in comparison with their sonic analogues.

Hawking’s analysis is performed in the context of a semi-classical approach to gravity. That is, we consider matter as described by quantum field theory and spacetime as described by a continuous classical geometry. Crucially, although the spacetime in question can have non-trivial curvature, it is not coupled to the quantum field. That is, there can be no ‘back-reaction’ between the quantum matter and classical geometry. For this modelling framework to be valid it is assumed that we are considering quanta of wavelengths much larger than the Planck length. Quanta of the order of the Planck length could be expected to ‘see’ the (presumed) non-classical and non-continuous structure of spacetime and would necessitate a quantum theory of gravity in their description. Quite general formal considerations can be used to show that in the semi-classical framework the vacuum state of a quantum scalar field defined at past null infinity need not appear as a vacuum state to observers at positive null infinity. In particular, it may contain a ‘particle flux’. What Hawking shows in his original paper is that for spacetime in which an astrophysical black hole forms there will be a particle flux which observers at positive null infinity will associate with the black hole horizon. The asymptotic form of the expression for the particles flux is shown to depend only upon the *surface gravity* of the black hole denoted by  $\kappa_G$ . Surface gravity is essentially the force per unit mass that must be applied at infinity in order to hold a stationary zero angular momentum particle just outside the horizon (Jacobson, 1996). Hawking’s calculation implies that a black hole has intrinsic properties that are connected to a non-zero *thermal* particle flux at late times. The precise relation takes the form:

$$\left\langle \widehat{N}_\omega \right\rangle = \frac{1}{e^{\frac{2\pi\omega}{\hbar\kappa_G}} - 1} \quad \text{with} \quad T_{BH} := \hbar\kappa_G/2\pi, \quad (1)$$

where  $\widehat{N}_\omega$  is the number operator for modes detected at late times with frequency  $\omega$  and  $\hbar$  is Planck’s constant divided by  $2\pi$ .

One key feature of the derivation of the temperature is worth noting here since it will be very important in what follows. In the derivation of Hawking radiation an exponential gravitational redshift means that the black hole radiation detected at late times must be taken to correspond to extremely high frequency radiation at the horizon. These ‘*trans*-Planckian’ modes are of wavelengths that are of precisely the kind that we presumed to exclude in using the semi-classical framework. There is thus a tension between the initial modelling assumptions and the details of the calculation. We will return to this issue shortly.

Not long after the derivation of Hawking’s radiation, it was proposed by Unruh that a similar thermal effect might exist in the context of sound in fluid systems (Unruh, 1981). In particular, Unruh showed that the key elements of Hawking’s calculation

<sup>6</sup> See Thébaud (2019) for an analysis of the connection between the epistemology of analogue experiments and the notions of internal and external validation as discussed in the philosophy of experimentation.

<sup>7</sup> Such a conclusion is, in fact, a little too quick since the equivalence principle holds only locally and Hawking radiation is a global effect. See Helfer (2010).

<sup>8</sup> See Leonhardt and Philbin (2008); Thompson and Ford (2008); Wallace (2018).

could be re-applied in the context of a semi-classical model of sound in fluids. The role of the spacetime geometry is now played by a ‘bulk’ fluid flow as described by continuum hydrodynamics. The role of the quantum field is then played by a quantized linear fluctuation within the fluid, a phonon. The modelling framework of continuum hydrodynamics is only valid provided fluid density fluctuations of the order of molecular lengths can be ignored. So for this semi-classical description to be adequate the wavelengths of the phonons must be much larger than the intermolecular distances. Unruh’s brilliant insight was to recognise that there is a special class of analogue fluid systems for which the equations of semi-classical continuum hydrodynamics take a form isomorphic to those of semi-classical gravity. The role of the black hole event horizon is now played by an effective acoustic horizon where the inward flowing magnitude of the radial velocity of the fluid exceeds the speed of sound. The black hole is replaced by a *dumb hole*. Just as in the gravitational Hawking effect a black hole event horizon is associated with a late time *thermal photonic flux*, in the hydrodynamic Hawking effect Unruh showed that a dumb hole sonic horizon can be associated with a late time *thermal phononic flux*.

An alternative medium for constructing acoustic black holes, that behaves exactly like a fluid in an appropriate limit, is given by a Bose-Einstein condensate (Garay, Anglin, Cirac, & Zoller, 2000). This is an exotic form of matter predicted in the 1920s (Bose, 1924; Einstein, 1924) but not created experimentally until 1995 (Anderson, Ensher, Matthews, Wieman, & Cornell, 1995). Crucially, in the limit of *weak coupling* and *no backreaction* density fluctuations in a BEC are described by the Gross-Pitaevskii equation. When variations in the density of the BEC occur on length scales much greater than the *healing length*, the Gross-Pitaevskii equation can be used to derive an equation of the same form as that of an irrotational fluid derived by continuum hydrodynamics. Following the same line of reasoning as Unruh’s original ideal fluid argument (Garay et al., 2000), derived a BEC Hawking Effect using appeal to this hydrodynamic approximation to a BEC.

There are now a huge number of potential analogue realisations of the Hawking effect: phonons in superfluid helium-3, ‘slow light’ in moving media, traveling refractive index interfaces in nonlinear optical media, laser pulses in nonlinear dielectric medium.<sup>9</sup> To realize the Hawking effect it seems it is sufficient to have: i) a classical (effective) background with quantum fields living on it; and ii) an (effective) geometry with an (effective) causal horizon.

A crucial feature in the derivation of all these effects is the integration over extremely high energy ‘*trans-Planckian*’ modes. As noted above, in the black hole case these modes get included in the calculation due to an exponential redshift between the horizon (where they originate) and the late time, far distance limit (where they are detected). Such a feature is generic to all models of Hawking radiation in which the modes originate near the horizon.<sup>10</sup> Since the ‘*trans-Planckian*’ regime beyond the domain of applicability of the semi-classical modelling framework we are using, this problem of exponential redshift seems rather worrying. In fact, according to some, the *trans-Planckian* problem is so serious as to cast doubt upon the Hawking calculation entirely. Unruh, for instance, even asks that ‘if the derivation relies on such absurd physical assumptions, can the result be trusted?’ (Unruh, 2014, p. 534). The problem with ‘*trans-Planckian*’ modes has a direct

analogue in both the continuum hydrodynamic and BEC derivations. In particular, the acoustic analogue of the gravitational redshift, means that in both cases we are including in our calculation phonons of wavelengths small enough to probe the regimes very far beyond the inter-molecular length and healing length respectively.

Fortunately, there are good reasons to expect that the Hawking effect in both gravitational and analogue cases will be robust to disturbance from *trans-Planckian* modes. In particular, Unruh and Schützhold (2005) have provided strong theoretical reasons to expect that, under certain conditions, any modifications to the Hawking flux by *trans-Planckian* modes will be negligible.<sup>11</sup> Unruh and Schützhold show that a wide family of *trans-Planckian* effects can be factored into the calculation of Hawking radiation via a non-trivial dispersion relation. To lowest order and given certain modelling assumptions, Hawking radiation, both astrophysical and acoustic, is independent of the details of the underlying physics.

A significant distinction that can be made in this context is between robustness and universality.<sup>12</sup> Robustness is the insensitivity of a phenomenon under a *token-level* variation with respect to different possible micro-physics in a single type of system. Universality is the insensitivity of a phenomenon under a *type-level* variation between systems with fundamentally different material constitution (e.g. BECs and a classical fluid). Given these definitions, we can plausibly take the work of Unruh and Schützhold to be an argument for *both* the robustness and the universality of the Hawking effect. We can also, on this basis, draw a suggestive parallel to the notion of ‘universality’ found in the context of condensed matter physics. That said, nothing in our analysis will hang upon the particular characterisation of universality.<sup>13</sup> Rather, inferentially speaking, what is important is that the ‘universality argument’ in question provides support for a background assumption, common between the different models, that has moderate prior probability. The common background assumption supported by the argument of Unruh and Schützhold is that in general ultra high frequency physics does not significantly affect the Hawking spectrum. We take it that the Unruh and Schützhold arguments imply that such a common background assumption should not have a very low prior, and the *trans-Planckian* problem implies that such a common background assumption should not have a very high prior. In general terms, there are good reasons to expect that universality arguments will perform something like this function, but nothing in what follows will depend upon this assumption.

## 4. Bayesian analysis

### 4.1. Single source confirmation

The key claim that we wish to investigate is whether analogue ‘dumb hole’ experiments can provide inductive support for hypotheses regarding black holes given we believe the appropriate universality arguments. In what follows we give a Bayesian network representation of the proposed inferential structure of

<sup>11</sup> For further work on these issues, using a range of different methodologies, see for example Corley (1998); Himemoto and Tanaka (2000); Barceló, Garay, and Jannes (2009); Coutant, Parentani, and Finazzi (2012). For philosophical discussion see Dardashti et al. (2017) and Gryb, Palacios, and Thébaud (2018).

<sup>12</sup> Here we are drawing upon the account of Gryb et al. (2018) who in turn are largely following the discussion of Batterman (2000).

<sup>13</sup> Gryb et al. (2018) investigates these parallels at length. Analogue experiments that might plausibly be described using our formal framework based upon Wilsonian universality arguments are Thouless (1989); Prüfer et al. (2018); Erne, Bücker, Gasenzer, Berges, and Schmiedmayer (2018); Eigen et al. (2018).

<sup>9</sup> See Jacobson and Volovik (1998); Philbin et al. (2008); Belgiorno et al. (2010); Unruh and Schützhold (2012); Liberati, Prain, and Visser (2012); Nguyen et al. (2015); Jacquet (2018).

<sup>10</sup> It is worth noting here that whilst, the non-standard derivation of Giddings (2016) does appear to allow one to avoid this feature, that of Polchinski (1995), *prima facie*, does not (Harlow, 2016, pp. 37–8).

analogue simulation defended in (Dardashti et al., 2017) and show that the evidence in the source system can provide confirmation of hypotheses regarding the target system in certain circumstances.<sup>14</sup>

Let us start with the representation of the target system  $\mathcal{T}$ . We denote by  $M$  a propositional variable that takes the two values:

$M$ : The modelling framework  $\mathcal{M}$  provides an empirically adequate description of the target system  $\mathcal{T}$  within a certain domain of conditions  $D_{\mathcal{M}}$ .

$\neg M$ : The modelling framework  $\mathcal{M}$  does not provide an empirically adequate description of the target system  $\mathcal{T}$  within a certain domain of conditions  $D_{\mathcal{M}}$ .

The adequacy of the modelling framework  $\mathcal{M}$  depends on whether the background assumptions which justify the empirical adequacy of the modelling framework obtain. We denote with  $X_M$  the random variable with the values:

$X_M$ : The background assumptions  $x_M = \{x_M^1, x_M^2, \dots, x_M^n\}$  are satisfied for target system  $\mathcal{T}$ .

$\neg X_M$ : The background assumptions  $x_M = \{x_M^1, x_M^2, \dots, x_M^n\}$  are not satisfied for target system  $\mathcal{T}$ ,

where the second statement should be read in terms of a negation of the disjuncts. The role of the background assumptions is to define and justify the domain of conditions for the model. These assumptions involve knowledge, both theoretical and empirical, which goes beyond what is encoded within the model. Such knowledge need not be in the form of a simple, unified framework. Rather the background knowledge of the people who build and use models can contain an incompletely integrated set of explicit and tacit ideas about when a particular modelling framework will be adequate for a particular purpose and to a particular desired degree of accuracy.

With this in mind, we can introduce the random variables  $A$  and  $X_A$  for the source system  $\mathcal{S}$  described by the modelling framework  $\mathcal{A}$ . The variable  $A$  is a propositional variable that takes the two values:

$A$ : The modelling framework  $\mathcal{A}$  provides an empirically adequate description of the source system  $\mathcal{S}$  within a certain domain of conditions  $D_{\mathcal{A}}$ .

$\neg A$ : The modelling framework  $\mathcal{A}$  does not provide an empirically adequate description of the source system  $\mathcal{S}$  within a certain domain of conditions  $D_{\mathcal{A}}$ .

and  $X_A$  is the random variable with the values:

$X_A$ : The background assumptions  $x_A = \{x_A^1, x_A^2, \dots, x_A^k\}$  are satisfied for source system  $\mathcal{S}$ .

$\neg X_A$ : The background assumptions  $x_A = \{x_A^1, x_A^2, \dots, x_A^k\}$  are not satisfied for source system  $\mathcal{S}$ .

The systems  $\mathcal{S}$  and  $\mathcal{T}$  are assumed to represent different types of systems and thus have fundamentally different material constitutions. This means that the background assumptions behind the models  $\mathcal{M}$  and  $\mathcal{A}$  can reasonably be assumed to be very different. For example, compare the assumption of a curved but fixed

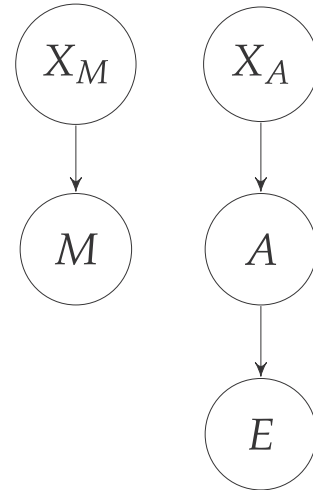


Fig. 1. The one-source system without universality.

spacetime for a black hole with the assumption of a flat spacetime for an analogue black hole; or compare the assumption of a continuous non-back-reacting spacetime for the black hole with the assumption of a continuous non-back-reacting bulk fluid flow for the analogue black hole. Given this, it is justified, *prima facie*, to assume that  $X_M$  and  $X_A$  are probabilistically independent (we will discuss the status of this assumption further shortly). Furthermore, we have assumed that the source system is empirically accessible meaning we can gain empirical evidence regarding (at least) some of its consequences. We can encode this by introducing a variable  $E$  corresponding to the two values,  $E$ , the empirical evidence obtains, and  $\neg E$ , the empirical evidence does not obtain.

We can represent all the variables introduced thus far as well as the probabilistic dependencies using a Bayesian network (Bovens & Hartmann, 2003). The random variables are represented as ‘nodes’ in the network (i.e. circles) and the probabilistic dependencies as directed edges (i.e. arrows). We draw an arrow between two nodes when the variable in the ‘parent node’ has a direct influence on the variable in the ‘child node’. Unconditional probabilistic independence is represented implicitly by the absence of an arrow between two nodes. The entire set up thus far is represented by the Bayesian network in Fig. 1.

Consider the case of an analogue black hole experiment, leading an agent to learn  $E$ , in a context where the agent does not fully believe in any form of universality argument. Clearly, due to the probabilistic independence of  $X_M$  and  $X_A$  and of  $M$  and  $A$ , we expect the independence of  $M$  and  $E$  and thus have that  $P(M|E) = P(M)$ . We therefore do not have confirmation. This is despite the fact that the equations of the source model are syntactically isomorphic to the equations that are supposed to govern the behaviour of the target model. The isomorphism is somewhat surprising, given that one arrives at the equations starting from very different areas of physics (i.e. semi-classical gravity and semi-classical continuum hydrodynamics) and by making different background assumptions. However, the agent has no reason to believe that the isomorphism represents a deeper fact about nature: on its own, the syntactic relation between the models gives us no reason to doubt the probabilistic independence of  $X_M$  and  $X_A$ .

There will, of course, always be available *some* basis to connect certain of the background assumptions. For instance, common features such as the dimensionality of spacetime or the

<sup>14</sup> For models of confirmation in terms of the Bayesian framework see Hartmann and Sprenger (2010); Bovens and Hartmann (2003) or for the hypothetic-deductive framework see Betz (2013). Throughout this paper, we follow the convention that propositional variables are printed in italic script, and that the instantiations of these variables are printed in roman script.

masslessness of relevant force carrying particles.<sup>15</sup> Thus, strictly speaking, background assumptions regarding different systems are never probabilistically independent. The reason why it is safe to neglect such common background assumptions for the purposes of confirmation is that they will generally be generic statements about which we have high confidence. This means that they cannot lead to any significant confirmation being transferred between  $E$  and  $M$  and can thus be neglected (see (Howson & Urbach, 2006, §4.c)).

Now, consider a context where the agent *does* believe in a universality argument for Hawking radiation, such as that provided by Unruh and Schützhold. The universality arguments give a physical justification for interpreting the syntactic isomorphism as indicative of a deeper fact about nature. In particular, it implies that we should assign a moderate prior to the common background assumption that ultra high frequency physics does not significantly affect the Hawking spectrum. It is crucial here that without the universality argument none of the *type-specific* forms of this background assumption already have high priors. Consider a situation where without the universality argument we already had strong independent grounds on which to assign high probability to the background assumption that for a particular type of analogue system the ultra high frequency physics does not significantly affect the Hawking spectrum. This would mean that the relevant member of  $X_A$ , corresponding to the specific statement of independence from ultra high frequency physics for that system, would have very high prior. In such a case, not much inductive support would be added to the generic form of the statement by learning  $E$ . This would block the transfer of confirmation from  $E$  to  $M$ . It is thus particularly significant that due to the generic nature of the *trans-Planckian* problem, the member of  $X_A$  corresponding to the specific statement of independence from ultra high frequency physics will for every type of system have a relatively low prior. There is thus a clear route for strong inductive support to flow through the generic statement of independence from ultra high frequency physics.

With these considerations in mind, let us then introduce a new binary variable  $X$  that has the values:

$X$ : Ultra high frequency physics does not significantly effect the Hawking spectrum.

$\neg X$ : Ultra high frequency physics does significantly effect the Hawking spectrum.

$X$  expresses a rather general claim, which can plausibly be assumed to have neither very high nor very low credence (if we were certain about  $X$ , the inference from  $\mathcal{S}$  to  $\mathcal{M}$  would be blocked. We will say more about this later).

For the sake of simplicity, let us now subsume the background assumptions, that are not addressed by  $X$  (or any other generic statement) under the relevant nodes. That is, we will subsume background assumptions exclusively concerning the target system  $\mathcal{T}$  under  $M$  and background assumptions exclusively concerning the source system  $\mathcal{S}$  under  $A$ .

Under the conditions of our assumptions, the simplified Bayesian network given in Fig. 2 will then adequately model the chain of inferences involved in analogue simulation supported by universality arguments. We would like to show that  $E$  confirms  $M$  within a Bayesian theory of confirmation. This requires that one proves that  $P(M|E) > P(M)$ . For this purpose we need to specify the prior probability of the ‘parent node’ in the Bayesian network, i.e.

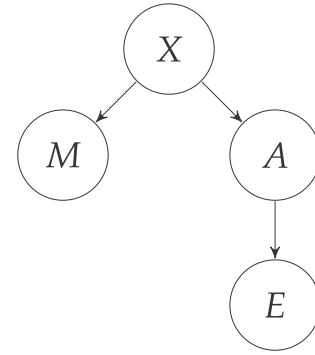


Fig. 2. The one-source system with universality.

$$P(X) =: x, \quad (2)$$

and the conditional probabilities of all ‘child nodes’, given the values of their parents, i.e.

$$\begin{aligned} P(M|X) =: m_x, P(M|\neg X) =: m_{\bar{x}} \\ P(A|X) =: a_x, P(A|\neg X) =: a_{\bar{x}} \end{aligned} \quad (3)$$

$$P(E|A) =: e_a, P(E|\neg A) =: e_{\bar{a}}.$$

The first central assumption is that the prior probability of  $X$  lies in the open interval  $(0, 1)$ . However, as we discussed in Sect. 3, we have theoretical arguments, in particular the result by Unruh and Schützhold, in favour of  $X$ . So a rational agent would assign

$$1/2 < x < 1. \quad (4)$$

The conditional probabilities are then constrained by the following conditions:

$$m_x > m_{\bar{x}} \quad (5)$$

$$a_x > a_{\bar{x}} \quad (6)$$

$$e_a > e_{\bar{a}} \quad (7)$$

Conditions (4) to (6) encode probabilistically the relevant elements of the universality argument since they allow for the possibility of a background assumption that supports both  $M$  and  $A$ . The statement (7) encodes probabilistically that the empirical evidence actually plays the role of evidence in favour of the modelling framework  $\mathcal{S}$ .

With this, the following Theorem holds (the proof is in Appendix A.1):

**Theorem 1.** Consider the Bayesian network depicted in Fig. 2 with the prior probability distribution  $P$  defined in eqs. (2) and (3). Then  $P(M|E) > P(M)$ , if conditions (4) to (7) are satisfied.

Theorem 1 implies that  $E$  confirms  $M$  within a Bayesian analysis of confirmation. Note that Theorem 1 also holds if condition (4) is replaced by the weaker condition  $0 < x < 1$ . Hence, within the framework of analogue simulation, provided we have universality arguments with prior probability that is neither unity or zero, confirmation of a hypothesis regarding the target system can obtain based upon evidence relating to the source system.

It is important to note again that having independent grounds on which to support one of the common background assumptions will ‘block’ the inductive support  $E$  can give for  $M$  as that background assumption already has a large marginal probability. This

<sup>15</sup> See (Dardashti et al., 2017, §3.2) for discussions of such common background assumptions in the context of  $1/r^2$  force laws.

does not pose a problem for this account but offers a way to distinguish between those circumstances in which the novel empirical evidence  $E$  can provide substantial inductive support for  $M$  and those circumstances it cannot be used for that purpose.

An important implication of the Bayesian analysis relates to the role of the syntactic isomorphism. The structure of the Bayesian network is such that the syntactic isomorphism is not explicitly represented. Furthermore, based upon the network, even if no syntactic isomorphism obtains between the modelling frameworks  $\mathcal{M}$  and  $\mathcal{S}$ , one could sensibly talk about confirmation of  $M$  by  $E$ , provided there exists some non-empty set of shared background assumptions. The key point is that in such circumstances although confirmation of  $M$  would indeed obtain, there would be no ‘analogue simulation’. As discussed above, the role of the isomorphism is to guarantee that there will be a term within the modelling language of  $\mathcal{M}$  that is counterpart to the term within  $\mathcal{S}$  that refers to the empirical evidence specified in variable  $E$ . Without such a term within  $\mathcal{M}$  there would be no sense in which  $\mathcal{S}$  is acting as a simulator for the behaviour of  $\mathcal{T}$ . Although the syntactic isomorphism is not explicitly represented in the network, it is implicitly stipulated within the universality argument. Furthermore, as noted above, the universality argument is exactly the reason we take the syntactic isomorphism to be indicative of a deeper fact about nature. Although it cannot be used to establish confirmation, the syntactic isomorphism is a crucial, if ultimately non-fundamental, heuristic for finding the universality argument that can.

To recapitulate, in this section we have demonstrated that confirmation via analogue simulation obtains within a Bayesian analysis provided there exists an inferential connection between the conditions of applicability of the target and system models. That is, if there exists a binary variable that is assumed to be positively correlated with the empirical adequacy of both the source and target models, then evidence in favour of the model of the source system can be used to make inferences about the target system. This, in-and-of-itself, is not a particularly surprising result, and certainly the demonstration of such in principle inferential relations is not a validation of the framework for analogue simulation that is being proposed. Rather, we take the results of this section to: i) demonstrate the internal consistency of the informal arguments towards confirmation via analogue simulation given in (Dardashti et al., 2017); and ii) provide a powerful evaluative and heuristic tool for the analysis of analogue simulation as it exists within contemporary scientific practice. Two natural directions of further development are: i) the identification and evaluation of potential cases of confirmation via analogue simulation in other scientific examples; and ii) the refinement of the Bayesian model to include cases within more than one analogue system. The second of these will be pursued in the following section.

#### 4.2. Multiple source confirmation

The systems  $\mathcal{S}$  and  $\mathcal{T}$  discussed in the previous section were assumed to be different types of systems and thus have fundamentally different material constitution. In the context of analogue experiments for Hawking radiation we could think of  $\mathcal{S}$  as a BEC analogue black hole and  $\mathcal{T}$  as an astrophysical black hole. Of course, as already noted, there are in fact a variety of further types of analogue black holes in addition to the BEC. For example, rather than making use of the syntactic isomorphism between BEC and gravitational models we might draw inferences based upon analogue black holes constructed out of traveling refractive index interfaces in nonlinear optical media or ‘slow light’ in moving media (Carusotto, Fagnocchi, Recati, Balbinot, & Fabbri, 2008).

With such examples in mind, we can extend the analysis of the previous section to consider the case when we have multiple types

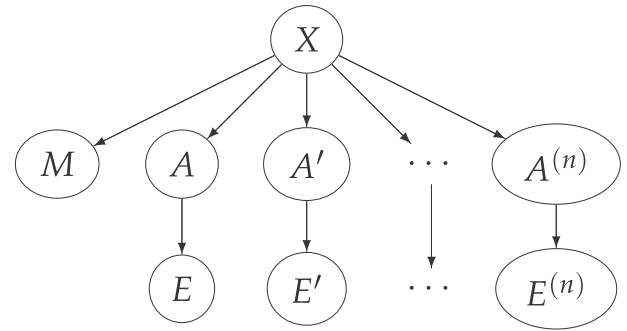


Fig. 3. The  $n$ -source system with universality.

of sources each providing independent evidence for the target system modelling framework. The expectation would be that adding more source systems should increase the degree of confirmation, but that this increase will eventually reach some ‘saturation point’ when we cannot learn any more about the target system by conducting additional analogue experiments.

Consider a Bayesian network for an  $n$ -source system (Fig. 3). The question we would like to answer is how does the confirmation measure change as one increases the number of different analogue systems providing us with evidence.

Following the same line of reasoning as the last section we assume:

$$\begin{aligned} a'_{x'} &> a'_{\bar{x}'}, e'_{a'} > e'_{\bar{a}'} \\ a''_{x''} &> a''_{\bar{x}''}, e''_{a''} > e''_{\bar{a}''} \\ &\vdots \\ a^{(n)}_{x^{(n)}} &> a^{(n)}_{\bar{x}^{(n)}}, e^{(n)}_{a^{(n)}} > e^{(n)}_{\bar{a}^{(n)}} \end{aligned} \tag{8}$$

We can now calculate the corresponding difference measure of confirmation, which is defined as

$$\Delta^{(n)} := P(M|E, E', \dots, E^{(n)}) - P(M). \tag{9}$$

It satisfies the following Theorem (the proof is in Appendix A.2):

**Theorem 2.** Consider the Bayesian Network depicted in Fig. 3 with the prior probability distribution  $P$  defined analogously to Theorem 1. Then  $\Delta^{(n)} > 0$  is a strictly increasing function of the number of source systems, if conditions (4), (5), and (8) are satisfied.

This Theorem implies that as the number of different analogue systems providing evidence increases so does the degree of confirmation.<sup>16</sup> Again, this is not a particularly surprising result. Given that confirmation via analogue simulation obtains for a single source system, one would expect that adding in more and more (independent) source systems would allow one to increase the degree of confirmation. The feature that is most interesting is not the fact that  $\Delta^{(n)}$  is strictly increasing, but rather the functional form of this increase. In particular, the natural intuition is that as the number of source systems increases the increase in the degree of confirmation would eventually saturate. One of the chief virtues of the Bayesian model for analogue simulation with multiple source systems is that it allows us to give an analytical expression for such a saturation point.

First, let us consider how  $\Delta^{(n)}$  changes in the large  $n$  limit. A little analytical work (again see Appendix A.2) allows us to show that<sup>17</sup>

<sup>16</sup> Note that also Theorem 2 holds if condition (4) is replaced by the weaker condition  $0 < x < 1$ .

<sup>17</sup> Here (and below) we use the convenient notation  $\bar{x} := 1 - x$ .

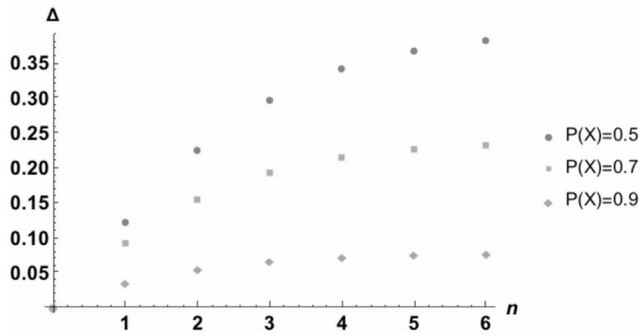


Fig. 4. The confirmation measure  $\Delta^{(n)}$  as a function of  $n$  for  $m_x - m_{\bar{x}} = 0.8$  and  $\gamma_n = 1.9^n$  and three different priors of  $X$ .

$$\lim_{n \rightarrow \infty} \Delta^{(n)} = \bar{x} (m_x - m_{\bar{x}}) =: \Delta_{sat} . \quad (10)$$

This means that the maximum amount of confirmation one can obtain by adding in more and more sources is bounded by some finite value, viz.  $\Delta_{sat}$ , determined by the prior probabilities  $x$ ,  $m_x$  and  $m_{\bar{x}}$ . Beyond this point, there is vanishingly small added value (in terms of confirmation) achieved by adding in more source systems. Two features of  $\Delta_{sat}$  are worth remarking on. First, the higher the prior probability of  $X$  the lower the saturation point will be. This makes sense because the more sure we are of  $X$  to start with, the lower the limit on the extra information we can learn from  $E, E', \dots, E^{(n)}$ . Second, the higher the relative likelihood of  $M$  given  $X$  to  $M$  given  $\neg X$  (i.e.  $m_x - m_{\bar{x}}$ ), the higher the saturation point. This makes sense because the stronger the relationship between  $X$  and  $M$  the more we can potentially learn from  $E, E', \dots, E^{(n)}$ .

A further interesting feature that we can examine is the speed with which the saturation point is approached. We can examine this ‘rate of saturation’ by plotting  $\Delta^{(n)}$  (as specified in eq. (22) in Appendix A.2) for a set of prior probabilities of  $X$ . As can be seen from Fig. 4, the higher the prior probability of  $X$ , the quicker the saturation point is approached. Strikingly, for the values of the parameters considered, we find that given a prior of greater than 0.5 for  $X$ , saturation can be approached after only three or four successful analogue experiments.

This result is in tune with scientific intuitions regarding analogue simulation in the context of universality arguments. Consider, in particular, the dumb hole Hawking radiation case. Let us assume that an uncontested and repeated experiment demonstrating BEC Hawking radiation has already been carried out. Given initial confidence in the universality arguments, another different implementation of a source system displaying the Hawking effect (say in nonlinear optics) would strongly increase the belief in the astrophysical Hawking effect. However, once a few such examples were constructed, one would quickly stop gaining new insight. Conversely, given initial skepticism regarding the universality arguments, a second implementation of the dumb hole source system would not radically increase the belief in the astrophysical Hawking effect. Furthermore, in such circumstances it would only be after a diverse and extensive range of implementations of source systems that one would stop believing that new examples gave new information. Significantly, the saturation in confirmation indicates that, under plausible assignments of priors, even an extraordinarily large range of diverse analogue experiments will not lead to conclusive confirmation of astrophysical Hawking radiation. This is in tune with the scientific intuitions that there is a limit to what can be learned about astrophysical Hawking radiation via analogue experiments.

### 4.3. Confirmation dependence on source system

Although most of the proposed analogue models of black hole Hawking radiation have not been tested yet, there can be significant differences in our prior belief regarding the adequacy of these models. One reason for this is that the modelling may rely on a strong theoretical basis in one system but a rather conjectural basis in the other. For example, the fundamental physics of BECs is better understood, via the Bogoliubov theory, than that of superfluid helium-3. Another reason is that one may have good control of the experimental setup, such that one has more reason to rely on the adequate realization of the various idealizing assumptions involved in the derivation of the model.

Such differences imply that we may plausibly assign significantly different marginal probabilities to the analogue models. In such circumstances, it is then interesting to consider the degree to which variations in the marginal probabilities one assigns to the adequacy of the analogue model may affect the confirmation behaviour of the analogue setup. The following Theorem addresses this question for the one-source model (see Appendix A.3 for details and the proof):

**Theorem 3.** Consider the Bayesian Network depicted in Fig. 2 with the prior probability distribution  $P$  defined in eqs. (2) and (3). Let  $\Delta := P(M|E) - P(M)$  and  $a := P(A)$ . Then  $d\Delta/da < 0$ , if conditions (4) to (7) hold.

This plausible result implies that an assignment of a higher probability to the adequacy of the analogue model will have the effect of a decrease in the confirmation of the adequacy of the target model by the observation of the analogue Hawking effect. Or to put it differently: the more certain we are about the adequacy of the analogue model that describes the source system we are experimenting on, the less effective is the evidence obtained there in confirming the adequacy of the model of the target system. Significantly, this result has direct implications for the respective epistemic value of future analogue experiments. In particular, all else being equal, it implies that scientists will learn more about the target system by conducting future analogue experiments using media about which we are *less certain* regarding their fundamental physics (i.e. marginal of  $A$  is lower), than those using media about which we are more confident regarding their fundamental physics (i.e. marginal of  $A$  is higher).

## 5. Conclusion and prospectus

History is replete with examples of ‘transformative’ technology having a profound and lasting impact on the methodological foundations of science. Much recent literature in the philosophy of science has focused on the sense in which computer simulation should be taken to have had such an impact.<sup>18</sup> Analogue simulation is a new inferential tool found at the cutting edge of modern science that we see good reasons to take as potentially transformative. Building upon the initial analysis of (Dardashti et al., 2017), in this paper we have applied a Bayesian analysis to explicate the structure of inferences that analogue simulation can and cannot allow us to make.

Our three principal results are: i) that ‘single source’ confirmation via analogue simulation can obtain under certain conditions; ii) that ‘multiple source’ confirmation via analogue simulation displays the generic feature of saturation in confirmatory power; and iii) that analogue experiments in which we are more confident

<sup>18</sup> See for example Hartmann (1996); Humphreys (1995, 2004); Parker (2009); Winsberg (1999, 2010).



about the fundamental physics provide less confirmation via analogue simulation than those about which we are more confident. Our results provide a basis to both formalise the epistemic value of analogue experiments that have been conducted and to advise scientists as to the respective epistemic value of future analogue experiments.

### Acknowledgement

KT would like to acknowledge support of the Arts and Humanities Research Council, UK (Grant Ref. AH/P004415/1).

### A Proofs

#### A.1 Theorem 1

We have to show that

$$P(M|E) = \frac{P(M, E)}{P(E)} > P(M). \quad (11)$$

To do so, we consider the Bayesian network depicted in Fig. 1 and follow the general methodology described in (Bovens & Hartmann, 2003, Sect. 3.5). First, we calculate

$$\begin{aligned} P(M, E) &= \sum_{A, X} P(X, M, A, E) \\ &= \sum_{A, X} P(X) P(M|X) P(A|X) P(E|A) \\ &= x m_x (a_x e_a + \bar{a}_x e_{\bar{a}}) + \bar{x} m_{\bar{x}} (a_{\bar{x}} e_a + \bar{a}_{\bar{x}} e_{\bar{a}}) \\ &= x m_x \alpha + \bar{x} m_{\bar{x}} \beta, \end{aligned} \quad (12)$$

with

$$\alpha := a_x e_a + \bar{a}_x e_{\bar{a}} \quad , \quad \beta := a_{\bar{x}} e_a + \bar{a}_{\bar{x}} e_{\bar{a}}. \quad (13)$$

Note that  $\alpha = P(E|X)$  and  $\beta = P(E|\bar{X})$ . Similarly we obtain

$$\begin{aligned} P(E) &= \sum_{A, M, X} P(X, M, A, E) \\ &= x \alpha + \bar{x} \beta \end{aligned} \quad (14)$$

and

$$\begin{aligned} P(M) &= \sum_{A, E, X} P(X, M, A, E) \\ &= x m_x + \bar{x} m_{\bar{x}}. \end{aligned} \quad (15)$$

Next, we define the difference measure  $\Delta := P(M|E) - P(M)$  and obtain after some algebra:

$$\begin{aligned} \Delta &= \frac{x m_x \alpha + \bar{x} m_{\bar{x}} \beta - (x m_x + \bar{x} m_{\bar{x}}) (x \alpha + \bar{x} \beta)}{x \alpha + \bar{x} \beta} \\ &= \frac{x \bar{x}}{x \alpha + \bar{x} \beta} \cdot (m_x - m_{\bar{x}}) (\alpha - \beta) \end{aligned} \quad (16)$$

With

$$\alpha - \beta = (a_x - a_{\bar{x}}) (e_a - e_{\bar{a}}) \quad (17)$$

it follows that

$$\Delta = \frac{x \bar{x}}{x \alpha + \bar{x} \beta} \cdot (a_x - a_{\bar{x}}) (e_a - e_{\bar{a}}) (m_x - m_{\bar{x}}). \quad (18)$$

Hence,  $\Delta > 0$ , if conditions (4) to (7) are satisfied. This completes the proof of Theorem 1.

#### A.2 Theorem 2

To see how Theorem 1 can be generalized to the  $n$ -source systems represented in Fig. 3 let us consider first the simplified 2-source system represented in Fig. 5. Here we need to show that  $P(M|E, E') = P(M, E, E')/P(E, E') > P(M)$ . To do so, we first calculate.

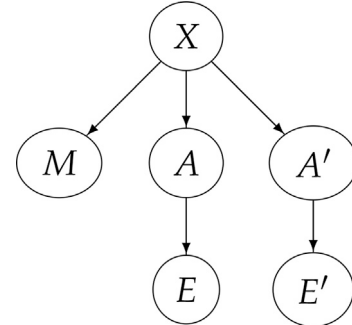


Fig. 5. The 2-source system with universality.

$$\begin{aligned} P(M, E, E') &= \sum_{A, A', X} P(X, M, A, E, A', E') \\ &= \sum_{A, A', X} P(X) P(M|X) P(A|X) P(E|A) P(A'|X) P(E'|A') \\ &= x m_x \alpha \alpha' + \bar{x} m_{\bar{x}} \beta \beta', \end{aligned} \quad (19)$$

where  $\alpha'$  and  $\beta'$  are defined analogously to eq. (13) with  $e$  and  $a$  replaced by  $e'$  and  $a'$ .

Similarly we obtain

$$\begin{aligned} P(E, E') &= \sum_{A, A', X} P(X) P(E|A) P(A|X) P(E'|A') P(A'|X) \\ &= x \alpha \alpha' + \bar{x} \beta \beta'. \end{aligned} \quad (20)$$

Defining  $\Delta^{(2)} := P(M|E, E') - P(M)$  it follows that

$$\Delta^{(2)} = \frac{x \bar{x}}{x \alpha \alpha' + \bar{x} \beta \beta'} \cdot (m_x - m_{\bar{x}}) (\alpha \alpha' - \beta \beta'). \quad (21)$$

Now it is easy to see that  $\alpha \alpha' - \beta \beta' > 0$  iff  $(\alpha/\beta) \cdot (\alpha'/\beta') > 1$ . This holds if  $\alpha > \beta$  and  $\alpha' > \beta'$ . Both conditions hold because of assumptions (8). Hence,  $\Delta^{(2)} > 0$ , if conditions (4), (5), and (8) are satisfied.

It is straightforward to generalise this calculation to the  $n$ -source system represented in Fig. 3. We obtain

$$\begin{aligned} \Delta^{(n)} &= \frac{x \bar{x} (m_x - m_{\bar{x}}) \left( \prod_{k=1}^n \alpha^{(k)} - \prod_{k=1}^n \beta^{(k)} \right)}{x \prod_{k=1}^n \alpha^{(k)} + \bar{x} \prod_{k=1}^n \beta^{(k)}} \\ &= \frac{x \bar{x}}{x \gamma_n + \bar{x}} \cdot (m_x - m_{\bar{x}}) \cdot (\gamma_n - 1), \end{aligned} \quad (22)$$

with the likelihood ratio  $\gamma_n := \prod_{k=1}^n \alpha^{(k)} / \beta^{(k)}$ . Note that  $\alpha^{(k)} > \beta^{(k)}$  for all  $k$  implies that  $\gamma_n > 1$ . Hence, as before,  $\Delta^{(n)} > 0$ , if conditions (4), (5), and (8) are satisfied.

Note that  $\gamma_n$  strictly monotonically increases with  $n$ . Furthermore,

$$\frac{\partial \Delta^{(n)}}{\partial \gamma_n} = \frac{x \bar{x} (m_x - m_{\bar{x}})}{(x \gamma_n + \bar{x})^2} > 0, \quad (23)$$

if conditions (4), (5), and (8) are satisfied. Hence,  $\Delta^{(n)}$  is a positive function of  $n$ . This completes the proof of Theorem 2.

In closing this section, we investigate the limiting behaviour of  $\Delta^{(n)}$  as a function of  $n$ . Setting  $\kappa := x \bar{x} (m_x - m_{\bar{x}})$ , we obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} \Delta^{(n)} &= \lim_{x \rightarrow \infty} \kappa \frac{\gamma_n - 1}{x \gamma_n + \bar{x}} \\ &= \lim_{x \rightarrow \infty} \kappa \frac{1 - 1/\gamma_n}{x + \bar{x}/\gamma_n} \\ &= \kappa/x \\ &= \bar{x} (m_x - m_{\bar{x}}), \end{aligned} \quad (24)$$

where we have used the fact that  $\lim_{n \rightarrow \infty} \gamma_n = \infty$ .

### A.3 Theorem 3

The proof follows from a straight-forward computation starting from eq. (18):

$$\begin{aligned} \frac{d\Delta}{da} &= \frac{\partial \Delta}{\partial a_x} \cdot \frac{da_x}{da} + \frac{\partial \Delta}{\partial a_{\bar{x}}} \cdot \frac{da_{\bar{x}}}{da} + \frac{\partial \Delta}{\partial x} \cdot \frac{dx}{da} \\ &= \frac{1}{x} \cdot \frac{\partial \Delta}{\partial a_x} + \frac{1}{\bar{x}} \cdot \frac{\partial \Delta}{\partial a_{\bar{x}}} + \frac{1}{(a_x - a_{\bar{x}})} \cdot \frac{\partial \Delta}{\partial x}, \end{aligned} \quad (25)$$

where we have used that  $a := x a_x + \bar{x} a_{\bar{x}}$ . Taking the respective partial derivatives yields

$$\begin{aligned} \frac{\partial \Delta}{\partial a_x} &= \frac{k x \bar{x}}{(x \alpha + \bar{x} \beta)^2} \cdot (x \alpha + \bar{x} \beta - (a_x - a_{\bar{x}})(e_a - e_{\bar{a}}) x) \\ \frac{\partial \Delta}{\partial a_{\bar{x}}} &= -\frac{k x \bar{x}}{(x \alpha + \bar{x} \beta)^2} \cdot (x \alpha + \bar{x} \beta - (a_x - a_{\bar{x}})(e_a - e_{\bar{a}}) \bar{x}) \\ \frac{\partial \Delta}{\partial x} &= -\frac{k (a_x - a_{\bar{x}})}{(x \alpha + \bar{x} \beta)^2} \cdot (x^2 \alpha - \bar{x}^2 \beta). \end{aligned}$$

Here we have used the shorthand  $k := (e_a - e_{\bar{a}}) (m_x - m_{\bar{x}})$ . Plugging these expressions into eq. (25) one obtains after some algebra that

$$\frac{d\Delta}{da} = -\frac{k}{(x \alpha + \bar{x} \beta)^2} \cdot (x (3x - 1) \alpha - \bar{x} (3\bar{x} - 1) \beta). \quad (26)$$

Assumptions (5) and (7) imply that  $k > 0$  and assumptions (4), (6) and (7) imply that the large bracket is positive. (To see this, note that  $x (3x - 1) > \bar{x} (3\bar{x} - 1)$  for  $1/2 < x < 1$ .) Hence,  $d\Delta/da < 0$ . This completes the proof of Theorem 3.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.shpsb.2019.04.004>.

## References

Anderson, M., Ensher, J., Matthews, M., Wieman, C., & Cornell, E. (1995). Observation of Bose-Einstein condensation in a dilute atomic vapor in a dilute atomic vapor. *Science*, 269, 14.

Bailer-Jones, D. M. (2009). *Scientific models in philosophy of science*. Pittsburgh: University of Pittsburgh Press.

Barceló, C., Garay, L. J., & Jannes, G. (2009). Sensitivity of Hawking radiation to superluminal dispersion relations. *Physical Review D*, 79(2), 024016.

Bartha, P. F. (2010). *By parallel reasoning: The construction and evaluation of analogical arguments*. Oxford: Oxford University Press.

Bartha, P. (2013). Analogy and analogical reasoning. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Fall 2013 ed.).

Batterman, R. W. (2000). Multiple realizability and universality. *The British Journal for the Philosophy of Science*, 51(1), 115–145.

Belgiorio, F., Cacciatori, S., Clerici, M., Gorini, V., Ortenzi, G., Rizzi, L., et al. (2010). Hawking radiation from ultrashort laser pulse filaments. *Physical Review Letters*, 105(20), 203901.

Betz, G. (2013). Revamping hypothetico-deductivism: A dialectic account of confirmation. *Erkenntnis*, 78(5), 991–1009.

Bose, S. N. (1924). Plancks gesetz und lichtquantenhypothese. *Zeitschrift für Physik*, 26(3), 178.

Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology*. Oxford: Oxford University Press.

Carusotto, I., Fagnocchi, S., Recati, A., Balbinot, R., & Fabbri, A. (2008). Numerical observation of Hawking radiation from acoustic black holes in atomic Bose-Einstein condensates. *New Journal of Physics*, 10(10), 103001.

Corley, S. (1998). Computing the spectrum of black hole radiation in the presence of high frequency dispersion: An analytical approach. *Physical Review D*, 57(10), 6280.

Coutant, A., Parentani, R., & Finazzi, S. (2012). Black hole radiation with short distance dispersion, an analytical S-matrix approach. *Physical Review D*, 85(2), 024021.

Dardashti, R., Thébault, K. P., & Winsberg, E. (2017). Confirmation via analogue simulation: What dumb holes could tell us about gravity. *The British Journal for the Philosophy of Science*, 68(1), 55–89.

Dobson, D. (2016). Geophysics: Earth's core problem. *Nature*, 534(7605), 45–45.

Efrati, I. (2016, August). *The Israeli research that's bringing Stephen Hawking closer to the Nobel prize*. Haaretz.

Eigen, C., Glidden, J. A., Lopes, R., Cornell, E. A., Smith, R. P., & Hadzibabic, Z. (2018). *Universal prethermal dynamics of Bose gases quenched to unitarity*. arXiv preprint arXiv:1805.09802.

Einstein, A. (1924). Quantentheorie des einatomigen idealen gases. In *Sitzungsberichte der Preussische akademie der Wissenschaften* (pp. 261–267).

Erne, S., Bücker, R., Gasenzer, T., Berges, J., & Schmiedmayer, J. (2018). Universal dynamics in an isolated one-dimensional Bose gas far from equilibrium. *Nature*, 563(7730), 225.

Euvé, L.-P., Michel, F., Parentani, R., Philbin, T. G., & Rousseaux, G. (2016). Observation of noise correlated by the Hawking effect in a water tank. *Physical Review Letters*, 117(12), 121301.

Euvé, L.-P., Robertson, S., James, N., Fabbri, A., & Rousseaux, G. (2018, 06). *Scattering of surface waves on an analogue black hole*.

Finke, A., Jain, P., & Weinfurtner, S. (2016, Nov). On the observation of nonclassical excitations in Bose-Einstein condensates. *New Journal of Physics*, 18(11), 113017.

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

Franklin, A., & Perovic, S. (2016). Experiment in physics. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (winter 2016 ed.). Metaphysics Research Lab, Stanford University.

Garay, L., Anglin, J., Cirac, J., & Zoller, P. (2000). Sonic analog of gravitational black holes in Bose-Einstein condensates. *Physical Review Letters*, 85(22), 4643.

Giddings, S. B. (2016). Hawking radiation, the Stefan-Boltzmann law, and unitarization. *Physics Letters B*, 754, 39–42.

Glymour, C. (1980). *Theory and evidence*. Princeton: Princeton University Press.

Gryb, S., Palacios, P., & Thébault, K. (2018). *On the universality of Hawking radiation*. arXiv preprint arXiv:1812.07078.

Harlow, D. (2016). Jerusalem lectures on black holes and quantum information. *Reviews of Modern Physics*, 88(1), 015002.

Hartmann, S. (1996). The world as a process. In *Modelling and simulation in the social sciences from the philosophy of science point of view* (pp. 77–100). Springer.

Hartmann, S., & Sprenger, J. (2010). Bayesian epistemology. In S. Bernecker, & D. Pritchard (Eds.), *Routledge companion to epistemology* (pp. 609–620). London: Routledge.

Hawking, S. W. (1975). Particle creation by black holes. *Communications in Mathematical Physics*, 43(3), 199–220.

Helfer, A. D. (2010). Comment on “insensitivity of hawking radiation to an invariant planck-scale cutoff”. *Physical Review D*, 81(10), 108501.

Hempel, C. G. (1965). *Aspects of scientific explanation and other essays in the philosophy of science*. New York: Free Press.

Hesse, M. B. (1963). *Models and analogies in science* (Vol. 7). London: Sheed and Ward.

Hesse, M. (1964). Analogy and confirmation theory. *Philosophy of Science*, 31(4), 319–327.

Hesse, M. B. (1974). *The structure of scientific inference*. Univ of California Press.

Himemoto, Y., & Tanaka, T. (2000). Generalization of the model of Hawking radiation with modified high frequency dispersion relation. *Physical Review D*, 61(6), 064004.

Howson, C., & Urbach, P. (2006). *Scientific reasoning: The bayesian approach*. Open Court Publishing.

Hume, D. (1738/1978). *A treatise of human nature*. Oxford: Oxford University Press.

Humphreys, P. (1995). Computational science and scientific method. *Minds and Machines*, 5(4), 499–512.

Humphreys, P. (2004). *Extending ourselves: Computational science, empiricism, and scientific method*. Oxford University Press.

Jacobson, T. (1996). *Introductory lectures on black hole thermodynamics*. [www.physics.umd.edu/grt/taj/776b/lectures.pdf](http://www.physics.umd.edu/grt/taj/776b/lectures.pdf).

Jacobson, T., & Volovik, G. (1998). Effective spacetime and Hawking radiation from a moving domain wall in a thin film of  $^3\text{He-A}$ . *Journal of Experimental and Theoretical Physics Letters*, 68(11), 874–880.

Jacquet, M. (2018). *Negative frequency at the horizon: Theoretical study and experimental realisation of analogue gravity physics in dispersive optical media*. Springer.

Keynes, J. M. (1921). *A treatise on probability*. London: Macmillan & Co.

Konôpková, Z., McWilliams, R. S., Gómez-Pérez, N., & Goncharov, A. F. (2016). Direct measurement of thermal conductivity in solid iron at planetary core conditions. *Nature*, 534(7605), 99–101.

- Leonhardt, U. (2018, Mar). Questioning the recent observation of quantum Hawking radiation. *Annalen der Physik*, 530(5), 1700114.
- Leonhardt, U., & Philbin, T. G. (2008). The case for artificial black holes. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, 366(1877), 2851–2857.
- Liberati, S., Prain, A., & Visser, M. (2012). Quantum vacuum radiation in optical glass. *Physical Review D*, 85(8), 084014.
- Michel, F., & Parentani, R. (2014). Probing the thermal character of analogue Hawking radiation for shallow water waves? *Physical Review D*, 90(4), 044033.
- Nguyen, H., Gerace, D., Carusotto, I., Sanvitto, D., Galopin, E., Lemaître, A., et al. (2015). Acoustic black hole in a stationary hydrodynamic flow of microcavity polaritons. *Physical Review Letters*, 114(3), 036402.
- Norton, J. D. (2011). *Analogy. Manuscript*. [http://www.pitt.edu/~jdnorton/papers/material\\_theory/Analogy.pdf](http://www.pitt.edu/~jdnorton/papers/material_theory/Analogy.pdf).
- de Nova, J. R. M., Golubkov, K., Kolobov, V. I., & Steinhauer, J. (2018, 09). *Observation of thermal Hawking radiation at the Hawking temperature in an analogue black hole*. <https://arxiv.org/pdf/1809.00913>.
- Parker, W. S. (2009). Does matter really matter? Computer simulations, experiments, and materiality. *Synthese*, 169(3), 483–496.
- Philbin, T. G., Kukulwicz, C., Robertson, S., Hill, S., König, F., & Leonhardt, U. (2008). Fiber-optical analog of the event horizon. *Science*, 319(5868), 1367–1370.
- Polchinski, J. (1995). *String theory and black hole complementarity*. arXiv preprint hep-th/9507094.
- Prüfer, M., Kunkel, P., Strobel, H., Lannig, S., Linnemann, D., Schmied, C.-M., et al. (2018). Observation of universal dynamics in a spinor Bose gas far from equilibrium. *Nature*, 563(7730), 217–220.
- Reid, T., & Hamilton, W. (1850). *Essays on the intellectual powers of man*. Cambridge: J. Bartlett.
- Rousseaux, G., Maïssa, P., Mathis, C., Couillet, P., Philbin, T. G., & Leonhardt, U. (2010). Horizon effects with surface waves on moving water. *New Journal of Physics*, 12(9), 095018.
- Rousseaux, G., Mathis, C., Maïssa, P., Philbin, T. G., & Leonhardt, U. (2008). Observation of negative-frequency waves in a water tank: A classical analogue to the hawking effect? *New Journal of Physics*, 10(5), 053015.
- Salmon, W. (1967). *The foundations of scientific inference* (Vol. 28). University of Pittsburgh Press.
- Salmon, W. (1990). Rationality and objectivity in science or tom kuhn meets tom bayes. In C. Wade Savage (Ed.), *Scientific theories* (Vol. 14, pp. 175–204). Minnesota: University of Minnesota Press.
- Shmahalo, O. (2016, August). What sonic black holes say about real ones. *Quanta Magazine*. <https://www.quantamagazine.org/what-sonic-black-holes-say-about-real-ones-20161108/>.
- Steinhauer, J. (2014). Observation of self-amplifying Hawking radiation in an analogue black-hole laser. *Nature Physics*, 10, 864–869.
- Steinhauer, J. (2015). Measuring the entanglement of analogue Hawking radiation by the density-density correlation function. *Physical Review D*, 92(2), 024043.
- Steinhauer, J. (2016a). Observation of quantum Hawking radiation and its entanglement in an analogue black hole. *Nature Physics*, 12(10), 959.
- Steinhauer, J. (2016b). *Response to version 2 of the note concerning the observation of quantum Hawking radiation and its entanglement in an analogue black hole*. arXiv preprint arXiv:1609.09017.
- Thébault, K. P. Y. (2019). What can we learn from analogue experiments? In R. Dardashti, R. Dawid, & K. Thébault (Eds.), *Why trust a theory?* Cambridge University Press.
- Thompson, R., & Ford, L. (2008). Enhanced black hole horizon fluctuations. *Physical Review D*, 78(2), 024014.
- Thouless, M. D. (1989). Some mechanics for the adhesion of thin films. *Thin Solid Films*, 181(1–2), 397–406.
- Torres, T., Patrick, S., Coutant, A., Richartz, M., Tedford, E. W., & Weinfurter, S. (2017). Rotational superradiant scattering in a vortex flow. *Nature Physics*, 13(9), 833.
- Unruh, W. G. (1976). Notes on black-hole evaporation. *Physical Review D*, 14(4), 870.
- Unruh, W. (1981). Experimental black-hole evaporation? *Physical Review Letters*, 46(21), 1351–1353.
- Unruh, W. G. (2014). Has Hawking radiation been measured? *Foundations of Physics*, 44(5), 532–545.
- Unruh, W. G., & Schützhold, R. (2005). On the universality of the Hawking effect. *Physical Review D*, 71(2), 024028.
- Unruh, W., & Schützhold, R. (2012). Hawking radiation from “phase horizons” in laser filaments? *Physical Review D*, 86(6), 064006.
- Wallace, D. (2018). The case for black hole thermodynamics part i: Phenomenological thermodynamics. *Studies in history and philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* (Vol. 64., 52–67).
- Weinfurter, S., Tedford, E. W., Penrice, M. C., Unruh, W. G., & Lawrence, G. A. (2013). Classical aspects of Hawking radiation verified in analogue gravity experiment. In D. Faccio, et al. (Eds.), *Analogue gravity phenomenology, volume 870 of lecture notes in physics* (pp. 167–180). Springer.
- Weinfurter, S., Tedford, E. W., Penrice, M. C., Unruh, W. G., & Lawrence, G. A. (2011). Measurement of stimulated Hawking emission in an analogue system. *Physical Review Letters*, 106(2), 021302.
- Winsberg, E. (1999). Sanctioning models: The epistemology of simulation. *Science in Context*, 12(2), 275–292.
- Winsberg, E. (2010). *Science in the age of computer simulation*. Chicago: University of Chicago Press.
- Wüthrich, C. (2019). Are black holes about information? In R. Dardashti, R. Dawid, & K. Thébault (Eds.), *Why trust a theory?* Cambridge University Press.