Foundations of Physics (2024) 54:27
https://doi.org/10.1007/s10701-024-00758-6

RESEARCH

Beyond the Quantum Membrane Paradigm: A Philosophical Analysis of the Structure of Black Holes in Full QG

Enrico Cinti[1,2,3] · Marco Sanchioni[4]

Received: 2 November 2023 / Accepted: 16 March 2024

Abstract
This paper presents a philosophical analysis of the structure of black holes, focusing on the event horizon and its fundamental status. While black holes have been at the centre of countless paradoxes arising from the attempt to merge quantum mechanics and general relativity, recent experimental discoveries have emphasised their importance as objects for the development of Quantum Gravity. In particular, the statistical mechanical underpinning of black hole thermodynamics has been a central research topic. The Quantum Membrane Paradigm, proposed by Wallace (Stud Hist Philos Sci Part B 66:103-117, 2019), posits a real membrane made of black hole microstates at the black hole horizon to provide a statistical mechanical understanding of black hole thermodynamics from an exterior observer's point of view. However, we argue that the Quantum Membrane Paradigm is limited to low-energy Quantum Gravity and needs to be modified to avoid reference to geometric notions, such as the event horizon, which presumably do not make sense in the non-spatiotemporal context of full Quantum Gravity. Our proposal relies on the central dogma of black hole physics. It considers recent developments, such as replica wormholes and entanglement wedge reconstruction, to provide a new framework for understanding the nature of black hole horizons in full Quantum Gravity.

Keywords Quantum black holes · Gravitational path integral · Quantum membrane paradigm · Philosophy of physics

Enrico Cinti and Marco Sanchioni have contributed equally to this work.

✉ Enrico Cinti
e.cinti@uva.nl

Marco Sanchioni
marco.sanchioni@sophiauniversity.org

1 Institute of Physics, University of Amsterdam, 1090 GL Amsterdam, The Netherlands

2 Institute for Logic, Language, and Computation, University of Amsterdam, 1098 XG Amsterdam, The Netherlands

3 Dutch Institute for Emergent Phenomena, University of Amsterdam, 1090 GL Amsterdam, The Netherlands

4 Istituto Universitario Sophia, Via San Vito n.28, 50064 Loppiano, Italy

Published online: 11 May 2024

1 Introduction

Black holes are centrally situated in recent research in theoretical and experimental physics. Indeed, many important experimental discoveries concern the physics of black holes, for example, the detection of gravitational waves due to the merging of two black holes or the photo of the black hole at the centre of the Milky Way [1]. Even more importantly, black holes have been one of the main focuses of high-energy physicists searching for a consistent theory of Quantum Gravity (QG). Indeed, from [2]'s pivotal paper, black holes have been the centre of countless paradoxes,[1] arising from the attempt at putting together Quantum Mechanics (QM) and General Relativity (GR): black holes are one of the few known objects where the merging of the two theories is necessary because their dynamics display both gravitational and quantum mechanical effects. The solutions to these paradoxes are a guiding part of the research in QG, giving important clues on the structure of spacetime beyond GR. Given the importance of black holes for recent research developments, both from a theoretical and an experimental point of view, a more in-depth study of their conceptual foundation is of great importance. In particular, we should develop a consistent philosophical framework for understanding black hole physics.

This paper contributes to this goal by focusing on the structure of the black hole event horizon and its fundamental status. The first steps in this direction have been made in [5, 4]: the first paper shows that black holes can be considered thermodynamical objects in the fullest sense, while the second paper analyses the statistical mechanical underpinning of black hole thermodynamics. Both papers touch upon the structure of the event horizon of the black hole. [4] says that a phenomenological membrane called the stretched horizon, should be posited just outside the actual black hole horizon to interpret black holes as thermodynamical objects in GR. [5] argues that to have a statistical mechanical underpinning of black hole thermodynamics from the point of view of an exterior observer, a real membrane, made of black hole microstates, should be posited at the black hole horizon.[2] Wallace calls this proposal the Quantum Membrane Paradigm **(QMP)**.

This paper aims to continue the investigation of the black hole horizon by arguing that Wallace's proposal, as it stands, only makes sense in the context of low-energy QG [6]. To have an understanding of black hole thermodynamics in full QG, as we will argue, we need to modify **(QMP)** to avoid reference to geometric notions, such as the event horizon, which presumably do not make sense in the non-spatiotemporal context of full QG.[3] Note that, by full QG, we mean any UV complete high energy theory of QG to be contrasted with its low energy effective field theory.

[1] See [3, 4] for philosophical discussion of some of the black hole paradoxes.

[2] By exterior observer, we mean an observer far from the black hole. For the rest of this paper, we will interchangeably use exterior observer and asymptotic observer.

[3] For discussions on the non-spatiotemporal nature of QG and its philosophical consequence, see [7, 8].

Our proposal crucially relies on taking the so-called *central dogma* of black hole physics [9] seriously. Building on [10, 11],[4] we discuss how the semiclassical description of the black hole interior appears in this proposal, and how the standard path integral description from which [5] justifies **(QMP)** and role underpinning black hole thermodynamics, is subtly modified to include certain quantum gravitational effects and lead to the picture that we propose. Overall, our goal in this article is to clarify the limitation of Wallace's **(QMP)** proposal in the context of full QG and suggest a modification of his proposal that avoids these limitations. The paper is structured as follows: §2 clarifies the phenomenological status of the stretched horizon, while §3 is a reconstruction of [1]'s argument for positing a quantum membrane at the event horizon of a black hole. §4 discusses the limitations of **(QMP)** in the context of full QG. §5 introduces our extension of **(QMP)** to full QG based on the central dogma, while §6 discusses how our proposal avoids the full QG arguments against **(QMP)**. §7 concludes.

2 The Stretched Horizon

Black holes have been considered thermal objects among physicists, starting with [2, 13]. Indeed, when an object is thrown into a stationary black hole, i.e. a black hole which is in a certain equilibrium state, the black hole evolves dynamically to a new equilibrium state, following[5]

$$\frac{\kappa}{8\pi G_N} dA_{\text{hor}} = dM - \Omega dJ \,, \tag{1}$$

where κ is the surface gravity, A_{hor} is the area of the event horizon, M is the mass of the black hole, and Ω and J are the angular velocity and the angular momentum of the black hole respectively. On the surface, this formula appears to be very similar to the second law of thermodynamics. Indeed, if we postulate that the black hole has an entropy S_{BH} and a temperature T, given by

$$S_{\text{BH}} = \frac{A_{\text{hor}}}{4G_N} \quad , \quad T = \frac{\kappa}{2\pi} \,, \tag{2}$$

as [2, 13] did, we find exactly the first law of black holes thermodynamics, i.e.

[4] Considered, within the high-energy physics community, crucial for resolving what [4] calls the Page time paradox, initially introduced in [12], which concerns the unitarity of the black hole evaporation process.

[5] In this review, we do not consider certain complications which go beyond the scope of the paper, of which we give a lightning-quick review in this footnote. If one considers also charged black holes, one can add a term like ΦdQ in the Eq. (1), where Φ and Q are respectively, the chemical potential and the charge of the black hole. Suppose one considers a black hole in a non-flat, slowly dynamically evolving background. In that case, one can add a term like $\mathcal{T}dL$ (or PdV) in the Eq. (1), where $\mathcal{T}$ is the tension of the black hole with respect to the environment and L is the curvature of the background (while P and V are the pressure and the volume of the black hole). For a complete discussion of these additional terms in Eq. (1) and related issues, see [14].

$$dM = TdS_{\mathrm{BH}} + \Omega dJ\ . \tag{3}$$

Even though (1) is just a formal rewriting of the Einstein equations, and the definitions given in (2) are only admitted at the quantum level, i.e. when considering quantum fields in the surroundings of the black hole, as proposed by [2], interpreting (1) as the second law of thermodynamics also at the classical level is very tempting.

However, there are some obstacles to this. First, a black hole in GR is just a *place* in spacetime. Indeed, a standard definition of black holes that one receives from a physics textbook is something like "a region of spacetime where gravity is so strong that nothing can escape from it." However, it is not evident how a place in spacetime can have properties, change in time, and be in equilibrium. Secondly, as seen from an asymptotic observer, the black hole seems disconnected from its exterior. Indeed, the black hole horizon is a null hypersurface in spacetime. If one tries to throw an object into a black hole, from the point of view of an exterior observer, it will never reach the horizon, and therefore it will not perturb the black hole. From the perspective of an asymptotic observer, how can an object be thrown into a black hole, perturb it, and originate the dynamical evolution given by (1)?

To provide a reasonable explanation for these worries, [15] postulated the existence of a new surface, called the stretched horizon, defined as follows (Fig. 1):

(SH) The **Stretched Horizon** is a time-like surface placed around the event horizon of the black hole at a Plank-length proper distance from it.[6]

With this definition, one can assign properties of the black hole to **(SH)**. From this point of view, **(SH)** is a two-dimensional (potentially charged) viscous fluid around the black hole, and in this way, the interpretation of (1) as the second law of thermodynamics becomes sensible. Indeed, since **(SH)** is a standard fluid, it can have properties and evolve in time, and since it is time-like, an object thrown into the black hole in equilibrium can interact with it from the point of view of an exterior observer. Thus, the second law of black hole thermodynamics follows accordingly, and **(SH)** describes the physics of a black hole as seen by an asymptotic observer.

From the point of view presented above, it seems that **(SH)** is just a fictitious surface introduced to solve the puzzle created by our intuition that black holes should be classical thermodynamical objects. It seems, indeed, that we just adopted a mathematical trick to solve the problematic interpretation of Eq. (1). It is then natural to interpret **(SH)** as a *phenomenological* surface, useful for characterising the behaviour of a black hole in GR from the point of view of an asymptotic observer, but without any fundamental physical import.

Furthermore, the idea that **(SH)** should be just a phenomenological surface is also supported by consideration of the equivalence principle. Indeed, consider a Schwarzschild black hole in GR described by the metric

[6] The distance between the stretched horizon and the true event horizon of the black hole is so small that no real particles can cross the stretched horizon and come back (only virtual particles can).

$$ds^2 = \left(1 - \frac{r_s}{r}\right)dt^2 + \frac{dr^2}{1 - \frac{r_s}{r}} + r^2 d\Omega_2^2 \, . \tag{4}$$

This black hole solution, as it is well known, has an event horizon at $r = r_s$, where r_s is the Schwarzschild radius. The metric (4) describes the black hole as seen by an asymptotic observer. However, suppose we perform a change of coordinates suitable for describing the experience of an infalling observer. In that case, it is possible to study the structure of the spacetime in a neighbourhood of the event horizon. The "zoomed-in" black hole metric, which represents the near-horizon metric of a Schwarzschild black hole, is

$$ds^2 \approx -\rho^2 d\tau^2 + d\rho^2 \, , \tag{5}$$

which is conformally equivalent to ordinary flat Minkowski spacetime $ds^2 = -dt^2 + dx^2$. Thus, within GR, an observer free-falling into a black hole does not encounter anything special at the event horizon $r = r_s$. This observation is just an instance of the equivalence principle of GR, which states that a free-falling observer does not feel the effect of gravity.

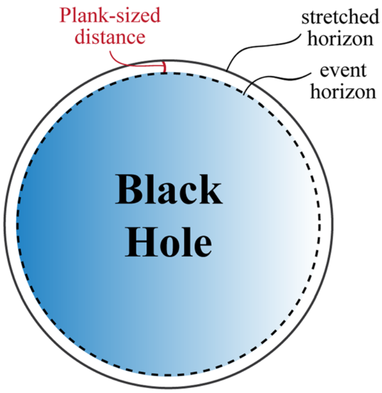
Therefore, **(SH)** should be just a phenomenological surface which describes, at a non-fundamental level, the black hole as seen from an asymptotic observer. This claim holds for two main reasons: first of all, because **(SH)** only concerns GR and does not incomporate quantum effects and, as such, cannot provide a full description of the black hole; moreover, **(SH)** can only be the non-fundamental perspective of an asymptotic observer because otherwise the equivalence principle would be violated. If that were not the case when crossing the event horizon, one would encounter **(SH)** and interact with it, thus experiencing something different from free fall, contra the equivalence principle. We will see in the following chapter how black hole complementarity attempts to solve this issue.

3 The Quantum Membrane

Considering black holes as thermodynamical objects is fully justified when considering GR coupled to Quantum Field Theory. In this framework, as shown by [2], black holes have a temperature given by (2), which fixes the proportionality constant between the entropy S_{BH} and the area A_{hor} to a value which gives the second law of thermodynamics, given by Eq. (3). Moreover, at the quantum level, black holes, as ordinary thermal objects, radiate and exchange heat with other surrounding objects. Indeed, as initially proposed in [2], black holes in semiclassical gravity emit thermal radiation, called Hawking radiation. The fact that black holes radiate makes them thermally connected with objects in their exterior via Hawking radiation. Therefore, black holes seem to be thermodynamic objects in the fullest sense only when considering quantum effects.

One could infer, from what we have said so far, that, as for ordinary thermodynamic objects, there should be a statistical mechanical underpinning of black hole

Fig. 1 An illustrative picture of a black hole with its stretched horizon



thermodynamics.[7] In other words, that there should be black hole microstates from which we can derive macroscopic thermodynamic-like properties. The proposal of [5] for such statistical mechanical underpinning is to promote **(SH)** to a quantum membrane:

(QMP) **Quantum Membrane Paradigm**: With respect to any physical process taking place outside the stretched horizon of a stationary or near-stationary black hole, **(SH)** may be treated as a quantum-mechanical system at (or near) thermal equilibrium, with density of states given approximately by $N \approx \exp S_{\text{BH}}$.

To understand the necessity of **(QMP)**, we should proceed, following [2]'s work on gravity as an effective field theory, by looking at structure of the black hole event horizon when small quantum effects are included as in pertubative quantum gravity.

Gravity, as it is well-known, is a non-renormalisable theory.[8] Indeed, if we try to apply standard quantisation techniques to the gravitation field, we get an ill-defined theory. These issues arise if we want the quantisation of Einstein's theory to be well-defined at all energy scales. Suppose we instead treat the quantisation of the gravitational field as an effective theory, i.e. a theory which is valid up to an energy cut-off scale Λ. In that case, there is no problem with renormalizability, as the loop momenta are cut-off. We can thus write a path integral for the gravitational field while keeping in mind that it gives a coherent description of physics only at scales below the cut-off scale Λ. Instead of the cut-off Λ, we can equivalently regard the cut off black hole path integral as describing the black hole in a box of surface A_{box}. The gravitational path integral is given by the following expression:

$$\log Z(\beta) = \int_{\beta} \mathcal{D}g\mathcal{D}\phi \exp\{-S[g,\phi]\}\ . \tag{6}$$

[7] Though see [16] for arguments against this line of thought.

[8] For a review of renormalisation in the context of QFT, see [17].

Even though, to study physics at energies higher than Λ, we would need a complete theory of QG, we can still draw some conclusions about the structure of black holes at the semiclassical level described by the cut off path integral.

To study quantum effects on a Schwarzschild black hole background we follow, as mentioned above, [2]. As for ordinary thermal objects, one needs to derive macroscopic thermodynamic-like properties from the thermal partition function $Z = \mathrm{Tr}\left[e^{-\beta H}\right]$. The thermal partition function of a quantum system in equilibrium is mathematically equivalent to the Euclidean time evolution of that quantum system on a circle of length β. In particular, by studying the path integral in Euclidian time on a circle of length β, one can derive the thermodynamic properties of a system in thermal equilibrium at temperature $T = 1/\beta$. We thus consider the Euclidean path integral (6) on a Schwarzschild black hole background. The metric of a Schwarzschild black hole, represented in Fig. 2, in Euclidian-time coordinates, is

$$ds^2 = \left(1 - \frac{r_s}{r}\right)dt_E^2 + \frac{dr^2}{1 - \frac{r_s}{r}} + r^2 d\Omega_2^2\,, \qquad t_E = t_E + \beta\,, \tag{7}$$

where t_E is the Euclidean time (the Wick rotation of standard Lorentzian time, $t_E = it$), r_s is the Schwarzschild radius,[9] and the last Eq. in (7) expresses the fact that the Euclidean time should be periodic (which is a consequence of the fact that we study the Euclidean time evolution on a circle of length β). The metric (7) has the feature that it shrinks to zero at the horizon radius $r = r_s$. To preserve the equivalence principle of GR, one needs the metric around $r = r_s$ to be smooth, which fixes β to be $\beta = 4\pi r_s$. This smoothness requirement fixes the black hole's temperature to be $T = \kappa/2\pi$, i.e. that of (2).

We just derived from the smoothness condition that the black hole has a temperature. We can thus consider the black hole to be an ordinary thermal object. If this is the case, then we can associate an entropy to the black hole given by $dS = dE/T = dM/T$, as we do for all other ordinary objects with a temperature.[10] Integrating this equation and using the relations for r_s and β, we can derive the standard formula for black hole entropy:

$$S_{\text{BH}} = \int \frac{dM}{R} = \frac{A_{\text{hor}}}{4G_N}\,. \tag{8}$$

In conclusion, the requirement of the smooth horizon in the path integral formulation of the partition function for a Schwarzschild black hole fixes the temperature and the entropy to be the one of (2).

We may ask now if the entropy we just introduced can be derived directly from the path integral, i.e. from the partition function. If this is the case, we would have the statistical mechanical underpinning we desired. We thus turn to consider an

[9] Which can be expressed in terms of the mass of the black hole as $r_s = 2GM/c^2$. This quantity is the only feature of the black hole present in the metric (7).

[10] This is a standard procedure in experimental physics when dealing with objects with a temperature.

expansion of the path integral (6) around the classical gravitational background given by (7), i.e. the Schwarzschild black hole:

$$\log Z(\beta) = -S_{\mathrm{EH}}^{\mathrm{classical}} + \log Z_{\mathrm{quantum}} , \tag{9}$$

where $S_{\mathrm{EH}}^{\mathrm{classical}}$ is the Einstein-Hilbert action evaluated on the classical background (7), which is the gravitational part of the path integral, and Z_{quantum} is the quantum fields contribution, obtained by computing the partition function of the quantum fields on the fixed classical background (7). One can then compute the entropy of the black hole by:

$$S_{\mathrm{BH}} = \left(1 - \beta\partial_{\beta}\right) \log Z(\beta) = \frac{A_{\mathrm{hor}}}{4G_N} + S_{\mathrm{matter}} , \tag{10}$$

using the saddle-point approximation.[11] S_{matter} represents the von-Neumann entropy of the quantum fields in the surroundings of the black hole.[12] The entropy of the black hole computed by (10) is fine-grained since it is calculated from the partition function. By fine-grained, we mean that the computation of the black hole entropy through (10) has information on the black hole microstates.

In conclusion, since the entropy (10) is independent of the cut-off area A_{box}, it is proportional to the horizon area A_{hor}, and should have information on the black hole microstates, [1] concludes that the microscopic degrees of freedom of the black hole should be localised at the horizon. The upshot of [1]'s proposal is that we can promote **(SH)** to a fundamental membrane by moving to **(QMP)**. In this way, we can see the relation between the phenomenological membrane **(SH)** and the fundamental quantum membrane **(QMP)** at the horizon as analogues to the relation between thermodynamics **(SH)** and statistical mechanics **(QMP)**. This fact follows because the degrees of freedom of the black hole, responsible for its thermodynamics, are localised via **(QMP)** at the horizon.

4 Limitations of (QMP)

Having introduced **(QMP)**, we can now discuss some reasons why it does not extend smoothly to full QG. In particular, our arguments aim to show that ultimately **(QMP)** relies in a non-trivial manner on a semiclassical picture of spacetime which is not available in the non-spatiotemporal realm of QG. A rough and intuitive way

[11] As we will see in §4, the implementation of the saddle-point approximation is the main problem with this derivation.

[12] The S_{matter} term in the derivation of the black hole entropy from the path integral was already present in [13]'s description. However, [18] understood that also Hawking radiation should contribute to the matter entropy. Indeed, even if the entropy of the vacuum in Quantum Field Theory is zero, if one divides the space into parts, as the horizon of the black hole does, then the vacuum, when restricted in a region of space, say the exterior of the black hole, carries some entropy which is the entanglement entropy between the interior and the exterior of the black hole. Also, this entropy should contribute to S_{matter}, and the entropy of the Hawking radiation precisely gives this contribution.

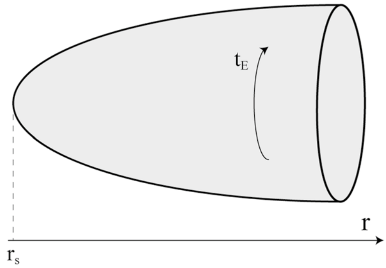


Fig. 2 A Schwarzschild black hole in Euclidean time

to see why this is the case is by noticing that **(QMP)** claims that the black hole's degrees of freedom are located at the horizon. However, the black hole's horizon is a geometric notion that we have no a priori reason to expect to hold in QG. Indeed, the idea of a smooth membrane exactly located one Planck length from the horizon is exactly the sort of geometric notion one expect to not make sense in QG in which spacetime structure is not well-defined but rather quantum and fuzzy. Let us briefly see three arguments for why **(QMP)** does not extend to full QG.

(i) **Conventionality.** Wallace's justification for **(QMP)** mostly comes, as we saw in §3, from a calculation done via the Euclidean path-integral for gravity. However, this calculation relies explicitly on using a saddle-point approximation to evaluate the path-integral. As such, it is a purely semiclassical calculation since saddle-point approximations are equivalent to semiclassical reasoning in the form of a choice of GR spacetime background given by the saddle-point metric. Indeed, this fact is evident from reasoning analogous to the one above, that **(QMP)** relies on geometric notions which can only make sense in the semiclassical approximation of QG, encoded here in the saddle-point approximation to the Euclidean path-integral.

What would be required to claim that **(QMP)** applies to full QG would be an argument of the following sort:[13] in the full path-integral (6) with boundary conditions appropriate to a Schwarzschild black hole, a calculation of the entropy returns a result proportional to the area of the black hole. This fact, the argument continues, is enough to establish that the black hole's degrees of freedom are located at the horizon of the Schwarzschild metric.

It is straightforward to see why this sort of argument cannot work. In the full path-integral (6), we are summing over all possible geometries compatible with a given boundary condition of the problem. This includes geometries that do not have a Schwarzschild metric and even geometries with different, possibly disconnected, topologies. It would be a cosmic coincidence if all these different geometries were all to agree on the location of the horizon! As such, it seems pretty challenging to say that, in a full non-semiclassical path-

[13] We here proceed formally by considering a path-integral of the form (6). Strictly speaking, this expression is not well-defined. However, most physicists expect that the full QG theory, such as String Theory or Loop Quantum Gravity, will be a well-defined finite avatar for (6). We consider (6) because our argument relies on very general features of a sum over geometries and, as such, does not require consideration of the details of any specific approach.

integral computation, the black hole's degrees of freedom are located in any specific place, let alone the horizon. In other words, the choice of a particular semiclassical background, which is necessary for **(QMP)**'s claim that the black hole's degrees of freedom are exactly located at the horizon, is ultimately a *conventional* choice which does not reflect any fundamental feature of the full QG theory, but at most the structure of its semiclassical limit.

(ii) **Inconsistency.** Given (i)'s argument that **(QMP)** relies on a semiclassical approximation for its justification, there is an immediate problem with **(QMP)**'s claim that black hole physics is unitary. As is well-known, semiclassical calculations of the kind rehearsed in §3 lead to violations of unitarity. A vivid example is the so-called *Page-time Paradox* [19], which shows that semiclassical gravity is incompatible with unitary black hole evaporation. As such, the path-integral calculation which Wallace relies upon to justify **(QMP)** is *inconsistent* with **(QMP)**'s claim that black hole physics is unitary[14] since the same calculation shows that black hole evaporation and, hence, the black hole's dynamics is not unitary.

To properly claim that **(QMP)** encodes the unitary dynamics of black holes, one would need to rely, as in (i), on a full path-integral calculation, rather than a semiclassical approximation. However, to make sense of **(QMP)**, as we argue above, one does indeed need the semiclassical approximation to define the geometric notions on which **(QMP)** relies. As such, there is no easy way out of the inconsistency between **(QMP)** reliance on semiclassical approximations and unitary black hole dynamics.

(iii) **Incompleteness.** Finally, it is unclear how to account for the black hole interior within **(QMP)** since **(QMP)** only deals with what happens at the horizon, not inside of it. [19] solves this issue by appealing to black hole complementarity. Black hole complementarity describes the interior by noticing that, by an appropriate change of coordinates, we can connect the exterior horizon description given by **(QMP)** and expressed in terms of the Schwarzschild coordinates, which are singular at the horizon, with a description in terms of a coordinate system where the interior is manifest (such as a coordinate system adapted to an infalling observer, e.g. Eddington-Finkelstein coordinates).

There are various problems with this move. First, black hole complementarity relies explicitly on a semiclassical approximation, as is clear in our talk about the change of coordinate systems for relativistic spacetimes. Hence, again, **(QMP)** accounting for the black hole interior relies on a semiclassical approximation. Moreover, insofar as **(QMP)** is supposed to encode the fundamental quantum description of a black hole, appeals to black hole complementarity provide, at best, an *incomplete* account of the black hole interior within **(QMP)**. This is because black hole complementarity relies explicitly on a semiclassical approximation and does not say anything about how this semiclassical description, and hence the black hole interior, is encoded within full QG and, therefore, **(QMP)**.

[14] As assumed in the definition of **(QMP)** where the degrees of freedom at the event horizon are treated as an ordinary quantum mechanical system, and hence as unitary.

Therefore, **(QMP)** crucially relies on a semiclassical approximation, and it is unclear how to extend its formulation given in terms of a semiclassical approximation to full QG. Indeed, trying to do so leads us to both difficulties in defining **(QMP)** itself (i) and to various paradoxes in the connection between semiclassical and fundamental physics in QG (ii-iii), which point to the inconsistency and incompleteness of a straightforward extension of **(QMP)** to full QG.

5 Central Dogma, a.k.a. The Exterior Description

While a straightforward extension of **(QMP)** to full QG is not viable, **(QMP)**'s effectiveness in accounting for black hole statistical mechanics still suggests that finding an appropriate extension of **(QMP)** to full QG would be a worthwhile project. Our goal in this section is to make exactly such a proposal.

We want to argue that, such an extension, is provided by the *central dogma*:[15]. To understand the central dogma, we start by noticing that the thermodynamics of black holes has a statistical mechanical underpinning is a strong motivation for interpreting them as ordinary quantum systems, with N degrees of freedom given by $N \approx \log S_{BH}$, obeying ordinary quantum mechanical rules. In particular, N would represent the dimension of the Hilbert space which describes the black hole. The proposal of interpreting black holes as ordinary quantum systems is the essence of the central dogma:

(CD) **Central Dogma:** As seen from the outside, a black hole can be described in terms of a quantum system with $\frac{A}{4G_N}$ (+ higher corrections) degrees of freedom, which evolves unitarily under time evolution (Fig. 3).

A formal derivation of the *central dogma* is given in the context of string theory in [20] and in the context of AdS/CFT correspondence in [21]. More generally, any theory for which the black hole entropy is given by the Bekenstein-Hawking formula can be reasonably expected to satisfy **(CD)** [9].

(CD) act as a full QG extension of **(QMP)** since it preserves the basic insights of the statistical mechanical underpinning of black hole thermodynamics given by **(QMP)**, i.e. that the black hole ultimately is a unitary evolving quantum systems with $\frac{A}{4G_N}$ degrees of freedom, while avoiding its problematic features, i.e. the requirement that those degrees of freedom are located at the black hole horizon.

(CD) states that the gravitational system composed by the black hole and the matter located in its exterior, up to a cut-off surface of area A, is *equivalent* to a quantum mechanical system with N degrees of freedom. In AdS/CFT, for example, this picture is realised by saying that the gravitational system is *dual* to a quantum

[15] The name Central Dogma comes from biology, where it refers to the information transfer from DNA and RNA to proteins. Here the statement of the central dogma is also about (quantum) information, the information transfer from the quantum system to the gravitational system [9].

mechanical system living at the boundary of the AdS spacetime. To understand the picture behind **(CD)** beyond AdS/CFT, we can say that these two systems are equivalent: the gravitational system of the black hole and its surroundings and the quantum mechanical system with $\frac{A}{4G_N}$ degrees of freedom. In other words, **(CD)** does not demand any specific (geometric) location for the degrees of freedom in the gravity description. In particular, we do not need to locate those degrees of freedom on the horizon.

Note that the degrees of freedom of the central dogma are not manifest in the gravity description. In particular, the fact that the entropy of these degrees of freedom is proportional to an area should not be taken as evidence that they are anywhere located on the surface having that area. In particular, this avoids the problematic reliance on a semiclassical approximation of **(QMP)**. While **(QMP)** required that the entropy of the black hole be proportional to the area and the degrees of freedom be located on the surface of that area, all **(CD)** requires is that the black hole entropy is proportional to the area which is a straightforward consequence of a path-integral computation and does not require any sort of semiclassical approximation (up to the inclusion of appropriate higher order corrections). In particular, note that nowhere in the definition of **(CD)**, we did not have to appeal to any sort of saddle-point metric, as was instead the case for **(QMP)**, as we highlight in (i).

We can now see how **(CD)** acts as a full QG extension of **(QMP)** and avoids the argument of (i). **(QMP)** was first introduced to have a statistical mechanical underpinning of black hole thermodynamics. However, as discussed above, one can have the same statistical mechanical underpinning of black holes as a simple consequence of **(CD)**, without relying on a semiclassical approximation as **(QMP)** did. Indeed, **(CD)** states that the black hole system, together with its surroundings up to the cut-off surface, is an ordinary quantum mechanical system with $N = \frac{A}{4G_N}$ degrees of freedom, which is a result that can be straightforwardly derived from the path-integral computation reviewed in §3 with or without any semiclassical approximation. If **(CD)** is true, the black hole system carries an entropy given by $S = \log N$ as any ordinary quantum mechanical system. Therefore accepting **(CD)** entails, without any further assumptions, a statistical mechanical underpinning of black hole thermodynamics. At the same time, **(CD)** does not make the problematic assumption that black hole's degrees of freedom are located at the horizon, which, contrary to the Bekenstein-Hawking entropy formula, is not presumably a well-defined notion in the full path-integral calculation and hence in full QG. Rather, this sort of localisation of degrees of freedom requires, at the very least, as we argued in §4, a commitment to a semiclassical approximation, which is the problem raised by (i). Hence, **(CD)** provides an extension of **(QMP)** to full QG avoiding the problems raised by (i).

Up to this point, nothing has been said about the interior structure of the black hole. In particular, we do not know if the degrees of freedom of **(CD)** also describe the interior of a black hole. In the next section, we will see how **(CD)** deals with the black hole interior and with the gravitational path integral in a way that avoids the problem raised in (ii) and (iii), thus leading to a satisfactory extension of **(QMP)** to the full QG regime.

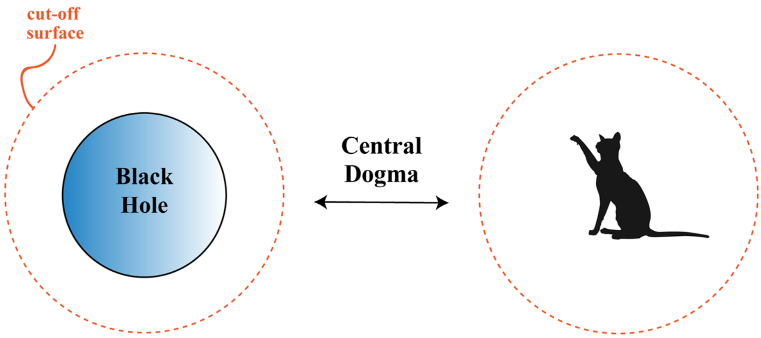


Fig. 3 The Central Dogma. There is a one-to-one correspondence between a system of a black hole with its surroundings (up to the cut-off surface, represented by the orange dashed line) and a quantum mechanical system (pictorially represented by a Schrödinger cat) (Color figure online)

6 Beyond (QMP)

In this section, our goal is to show how **(CD)** avoids the problems raised by (ii), i.e. the compatibility between path integral computation underpinning **(QMP)** and **(CD)** and unitarity of black hole dynamics, and (iii), i.e. the description of the interior starting from the fundamental degrees of freedom of **(QMP)** and **(CD)**. Since both problems ultimately rely on the structure of a full theory of QG rather than simply its semiclassical approximation, one needs to rely on a specific theory of QG to deal with them precisely. As a proof-of-concept of how **(CD)** might resolve the problems raised by (ii) and (iii), we discuss its implementation within the AdS/CFT correspondence where these questions have been worked out explicitly and exactly.

Before starting, let us first address a possible issue concerning the generality of our arguments. This section will mostly rely on *entanglement wedge reconstruction*. Entanglement wedge reconstruction is a theorem within the AdS/CFT correspondence [22, 23]. Therefore, the results of this section are well-grounded within the AdS/CFT correspondence. At the same time, following recent works of [10, 11], the basic consequences of entanglement wedge reconstruction for black holes can be derived directly from the gravitational path integral, which is crucial for avoiding (iii), without reference to holography. Hence, those results should generally be extendible to non-holographic theories of QG.[16] In this section, we first explain in §6.1 how a slight modification of the gravitational path integral required to accommodate **(CD)** leads to the appearance of new non-perturbative saddles (*replica wormholes*), which avoid the conflict between unitary evolution and path-integral methods highlighted by (ii). Then, in §6.2, we discuss how via entanglement wedge

[16] The main drawback of [10, 11] is that, even if the overall structure of their arguments is fully general and works for generally evaporating black holes, they present their calculation in a simple two-dimensional gravitational model, called JT gravity [24]. However, at least among string theorists, the explicit calculation done in JT gravity of [10, 11] is expected to generalise to other situations.

reconstruction **(CD)** can avoid the problems raised by (iii). In particular, we see how to recover the interior from the fundamental description of **(CD)**.

6.1 The Gravitational Path Integral

The problem raised by (ii) amounts to the fact that there is an incompatibility between the path-integral computation of the black hole entropy and unitary black hole evolution. This incompatibility is manifested by the Page-time paradox, which somewhat roughly consists of the divergence between the two entropy curves of Fig. 4. In particular, in Fig. 4, the light blue line represents the entropy of the black hole, which is given by the Bekenstein-Hawking area formula; the green line represents the entropy of the Hawking radiation, as computed in Hawking's original calculation: if the composite black hole plus radiation system evolves unitarily, then the composite system state should be pure at all times, and the entropy of Hawking radiation should never exceed the black hole entropy, hence following the red curve in Fig. 5 which is called the Page curve (the point where the light blue and the green lines cross is called the *Page time* t_P). However, suppose the entropy of the radiation follows from Hawking's original path-integral calculation reviewed in Sect. §3. In that case, the radiation's entropy will exceed the black hole thermodynamic entropy, leading to the following dilemma: either the black hole entropy is not given by the Bekenstein-Hawking area formula, or the black hole does not evolve unitarily.

Insofar as **(CD)** assumes the validity of both unitarity and the Bekenstein-Hawking area formula, then it seems to be inconsistent by an application of standard path-integral arguments, on which **(CD)** itself originally relied to justify its statistical mechanical underpinning of black hole thermodynamics.

To escape this problem, a natural move is to correct the path-integral prescription employed by Hawking in a way that leads to unitary evaporation and the Bekenstein-Hawking area formula, thus restoring the validity of **(CD)**. A way to accomplish this by adding non-perturbative corrections to the standard approach to the gravitational path-integral has been developed by [10, 11]. The starting point of [10, 11] is the same as the one developed in Sect. §3: the path integral (9), which describes the partition function of a Schwarzschild black hole. The problem concerns the computation of the partition function and the entropy of the black hole from the partition function, using the saddle-point approximation.[17] A useful method for computing the von Neumann entropy $S_{\text{BH}} = -\text{Tr}\left[\rho \log \rho\right]$ of the black hole from the partition function is called the *replica trick*. It consists in first computing the Rényi entropy $S_n = \text{Tr}[\rho^n]$ for integer n, then performing an analytic continuation for real n, and finally computing the von Neumann entropy as

$$S_{\text{VN}} = \lim_{n \to 1} \frac{1}{n-1} S_n \, . \tag{11}$$

[17] Note that for the saddle-point approximation to be reliable, one needs specific conditions on the matter content. This is discussed in [11].

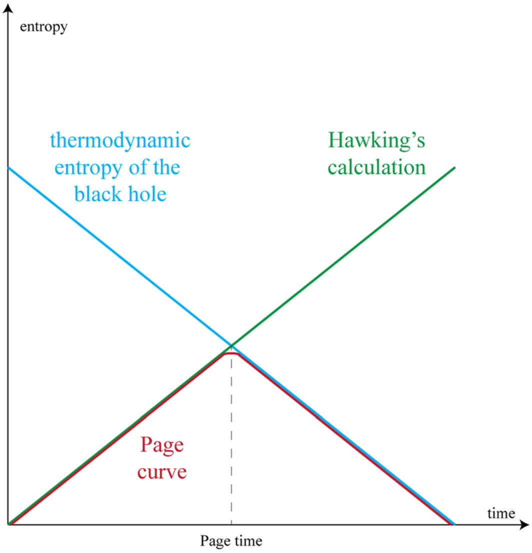


Fig. 4 The Page curve of an evaporating black hole. The light blue line represents the thermodynamic entropy of the black hole; the green line represents the entanglement entropy of the black hole in the original Hawking calculation; the red line represents the the Page curve. The point where the light blue and the green lines cross is called the *Page time* t_P (Color figure online)

For example, with outstanding results, this method has been used by [25] for computing the entanglement entropy in two-dimensional conformal field theory. In this setup, the computation of $\mathrm{Tr}[\rho^n]$ can be seen as the computation of a single observable $\mathrm{Tr}\rho$ in n copies (or replicas) of the original system, choosing appropriate boundary conditions that connect the various replicas.

Without delving into unnecessary details, the upshot of the work of [10, 11] is the following: there are two different saddle-points that contribute to the computation of $\mathrm{Tr}[\rho^n]$ and therefore of S_{VN}. One of them is the saddle-point considered by [2], in which the different replicas are sewn together along with their branch points, as is usually done in Quantum Field Theory calculations (see Fig. 5a). There is, however, another saddle point where gravity dynamically glues together the different replicas through a wormhole geometry (see Fig. 5b). These new geometrical structures responsible for a new saddle-point of the path integral (9) are called *replica wormholes*.

These new connections between different replicas, absent in the semiclassical Hawking calculation, have higher-order topology and are non-perturbatively small. Before the Page time, the new saddles are heavily suppressed, and the Hawking saddle dominates the path integral. After the Page time, however, the new replica wormhole saddle becomes important and contributes significantly to the final result of the entropy. The consequence of the presence of the new saddle is that the von-Neumann entropy of the black hole can be computed as

$$S_{\mathrm{BH}} = \min_{\chi}\left\{ \mathrm{ext}_{\chi}\left[\frac{A(\chi)}{4G_N\hbar} + S_{\text{semi-cl}}\left(\Sigma_{\chi}\right)\right]\right\}, \tag{12}$$

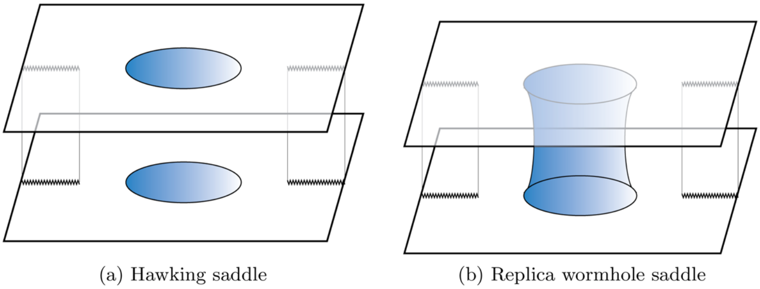


Fig. 5 **a** represents the Hawking saddle, in which the different replicas (white planes) are sewn together along the branch points (wiggly lines). **b** represents the replica wormhole saddle, in which there is a replica wormhole (blue cylinder) dynamically glueing together different replicas (white planes) (Color figure online)

where χ is a quantum extremal surface,[18] Σ_χ is the region between χ and the cut-off surface of the gravitational path integral, $A(\chi)$ is the area of the quantum extremal surface χ and $S_{\text{semi-cl}}(\Sigma_\chi)$ is the von Neumann entropy of quantum fields on Σ_χ. The quantity in square brackets is called *generalised entropy* $S_{\text{gen}}(\chi)$ of the black hole. The computation of S_{BH} consists of finding all the surfaces χ which extremise the generalised entropy and then choosing the quantum extremal surface χ, which gives the minimal generalised entropy.

Let us start by considering how to implement the formula 12 in the case of an evaporating black hole to recover the Page curve of Fig. 4. The first task for computing the black hole entropy is to find a quantum extremal surface χ which minimises the generalised entropy $S_{\text{gen}}(\chi)$. This minimal quantum extremal surface has been called the *quantum Ryu–Takayanagi* surface ([26]).[19] For an evaporating black hole, there are two quantum extremal surfaces: the empty surface $\varnothing$ and a non-vanishing surface $\tilde{\chi}$ which lies just inside the event horizon. At early times in the black hole evaporation process, i.e. before a time t_P called the Page time,[20] the generalised entropy of the surface $\varnothing$ is smaller than the generalised entropy of the surface $\tilde{\chi}$. Therefore, $\varnothing$ is the quantum Ryu–Takayanagi surface, and the black hole entropy is $S_{\text{semi-cl}}(\Sigma_\varnothing)$, which is Hawking's original result. However, at late times in the black

[18] A **quantum extremal surface** χ is defined as a surface satisfying two conditions:
(i) *Homology constraint:* given a boundary region B, a surface χ satisfies the homology constraint if, for C a space-like region, $\chi \cup B = \partial C$, i.e. the union of χ with a boundary region B is the boundary of some space-like region C. C is called *homology hypersurface*.
(ii) *Extremize the generalised entropy:* the surface χ should be a surface which extremizes the generalised entropy

$$S_{\text{gen}}(\chi) = \text{ext}\left[\frac{A(\chi)}{4G_N} + S_{\text{bulk}}(\chi)\right], \tag{13}$$

where $S_{\text{bulk}}(\chi)$ is the von Neumann entropy of the bulk fields contained in $\chi \cup B$ and $A(\chi)$ is the area of the hypersurface χ.

[19] For philosophical discussion on the quantum Ryu–Takayanagi see [27].

[20] The Page time is the time at which the semiclassical entropy of the Hawking radiation becomes larger than the black hole's thermodynamic entropy. From now on, we will use the terminology early (late) time and before (after) the Page time interchangeably.

hole evaporation process, the generalised entropy of the surface $\tilde{\chi}$ becomes larger than the generalised entropy of the surface $\varnothing$, which means that $\tilde{\chi}$ is the new quantum Ryu–Takayanagi surface. In other words, at the Page time, we have a phase transition between the two surfaces $\varnothing$ and $\tilde{\chi}$. At late times the entropy of the black hole is approximately given by $\frac{A_{\text{hor}}(\tilde{\chi})}{4G_N\hbar}$, which is equal to the thermodynamic entropy of the black hole. In conclusion, the entropy of the black hole is $S_{\text{semi-cl}}\left(\Sigma_{\varnothing}\right)$ before the Page time and $\frac{A_{\text{hor}}(\tilde{\chi})}{4G_N\hbar}$ after the Page time, i.e. it is given by

$$S_{\text{BH}} = \min\left\{ S_{\text{semi-cl}}\left(\Sigma_{\varnothing}\right), \frac{A_{\text{hor}}(\tilde{\chi})}{4G_N\hbar} \right\}, \tag{14}$$

which entails that the black hole evaporation process follows the Page curve of Fig. 4.

The formula (12) was first derived in the context of the AdS/CFT correspondence [26, 28, 29].[21] The same formula 12 has been used, under the name of *QES prescription*, by [31] in the context of an evaporating black hole showing that, according to (12), the black hole evaporation process is unitary and firewall-less.[22] The main advance of [10, 11] is that they derived (12) directly from the gravitational path integral without resorting to the AdS/CFT correspondence. Therefore, these works suggest a way out of the black hole information paradox since they derived, from the path integral approach to gravity, the conjectured formula (12), which guarantees a unitary and firewall-less black hole evaporation process.

The main conclusion of this section is that by appropriately modifying the gravitational path-integral via the inclusion of non-perturbative corrections given by *replica wormholes*, we can ensure that **(CD)** provides a consistent statistical mechanical underpinning to black hole thermodynamics which avoids the problem raised by (ii) and exemplifies most clearly by the Page time paradox.

6.2 Entanglement Wedge Reconstruction, a.k.a. The Interior Description

Having seen how to avoid the argument of (ii) including non-perturbative corrections in the gravitational path-integral, we now need to discuss how to recover the black hole interior from the degrees of freedom of **(CD)** and hence how to respond to the argument of (iii). To do so, we must introduce the concept of entanglement wedge reconstruction, the standard tool within AdS/CFT to encode regions of semiclassical spacetime within the fundamental description mathematically formulated in terms of the dual CFT. Since the degrees of freedom described by **(CD)** live in this fundamental description, then the task of responding to the argument of (iii) is equivalent to recovering the black hole interior via entanglement wedge reconstruction from the degrees of freedom of **(CD)** as encoded in the dual CFT.

[21] The AdS/CFT correspondence was initially introduced in [21]. For a philosophical discussion on the AdS/CFT correspondence, see [30].

[22] For a conceptually oriented presentation of [31] and for the consequences of the resolution of the black hole information paradox on semiclassical physics see [32].

To start, recall that the entropy (14) we have just computed for the evaporating black hole is the fine-grained entropy since it has been computed by a path-integral expression analogous to (10). Thus, it should thus have some information on the degrees of freedom of **(CD)**. The entropy (14) has been calculated by combining the geometrical information (area term of (12)) and the state of quantum fields (entanglement term of (12)) included in the region bounded by the quantum Ryu–Takayanagi surface χ (which is $\varnothing$ at early times and $\tilde{\chi}$ at late times). For this reason, it seems natural to think that the degrees of freedom of **(CD)** describe the physics within the region bounded by the minimal surface χ, i.e. the description of the black hole system and its surroundings of an exterior observer. To make this intuition more precise, it is useful to introduce the notion of the *entanglement wedge* of the quantum Ryu–Takayanagi surface χ, which is the region of spacetime bounded by the cut-off surface and the quantum Ryu–Takayanagi surface χ. The fact that the degrees of freedom of **(CD)** describe the physics up to the quantum Ryu–Takayanagi surface, an idea supported by the entropy computation of the black hole from the path integral, is the second main hypothesis of this paper. It is called the entanglement wedge reconstruction conjecture:[23]

(EWR) **Entanglement wedge reconstruction** says that all physical quantities in the entanglement wedge of a Ryu–Takayanagi surface χ are represented by operators in the quantum system which represent that spacetime region (think for example of how **(CD)** encodes information regarding the spacetime region accessible to an exterior observer).

(EWR) has some interesting consequences on the structure of black holes. Before the Page time, the Ryu–Takayanagi surface is the empty surface $\varnothing$, and the entanglement wedge is just the region inside the cut-off surface (blue region of Fig. 6a). However, at late times in the black hole evaporation process, the Ryu–Takayanagi surface $\tilde{\chi}$ lies inside the black hole, which means, by **(EWR)**, that the degrees of freedom of **(CD)** now describe just a portion of the interior of the black hole because just a portion of the black hole interior lies within the entanglement wedge of $\tilde{\chi}$ (blue region of Fig. 6b). We could ask which system describes the other part of the black hole interior at late times.

To address this issue, let us consider the entropy of the Hawking radiation. In this setup, Hawking radiation is described as the complement system of the central dogma. The formula for computing Hawking radiation's entropy, which can be derived from the gravitational path integral, is similar to (12). The main difference is that the region Σ_χ can be disconnected. In particular, Σ_χ can be the union of two disconnected regions: Σ_{rad}, which is the region outside the cut-off surface, and Σ_{Island} which is the region between the origin of the coordinate system and the quantum extremal surface χ. The formula for computing the entropy of the Hawking radiation is given by:

[23] There are many consistency checks of this hypothesis, especially within the AdS/CFT correspondence [22, 23].

$$S_{\text{rad}} = \min_{\chi} \left\{ \text{ext}_{\chi} \left[\frac{\text{Area}(\chi)}{4G_N \hbar} + S_{\text{semi-cl}} \left(\Sigma_{\text{rad}} \cup \Sigma_{\text{Island}} \right) \right] \right\}. \tag{15}$$

Before the Page time, the quantum Ryu–Takayanagi is $\emptyset$, corresponding to the situation where there is no island, i.e. $\Sigma_{\text{Island}} = \emptyset$. In this case, the fine-grained entropy of the Hawking radiation computed by (15) gives the same result of Hawking's original calculation, i.e. $S_{\text{rad}} = S_{\text{semi-cl}}(\Sigma_{\text{rad}})$. However, after the Page time, the quantum extremal surface is $\tilde{\chi}$, there is a non-zero island contribution, and the entropy of the radiation is given by $\frac{A_{\text{hor}}(\tilde{\chi})}{4G_N \hbar}$.[24] The fine-grained entropy of the Hawking radiation, as computed by (15) is

$$S_{\text{rad}} = \min \left\{ S_{\text{semi-cl}} \left(\Sigma_{\emptyset} \right), \frac{A_{\text{hor}}(\tilde{\chi})}{4G_N \hbar} \right\}, \tag{16}$$

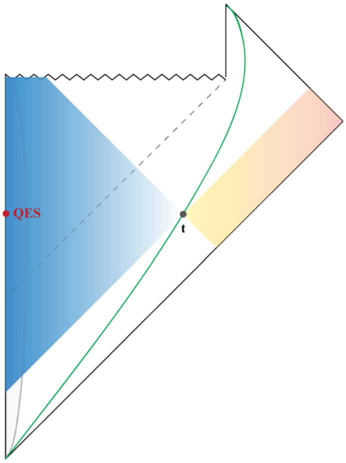
and follows the Page curve of Fig. 4. Before the Page time, the entanglement wedge of the radiation is a region in the exterior of the cut-off surface (see the orange region of Fig. 6a), while after the Page time, it is the union of a region outside the cut off-surface and the island (see the orange region of Fig. 6b).

Figure (6) gives us a clear picture of the quantum systems which describe the interior of an evaporating black hole. Before the Page time, the degrees of freedom of **(CD)** describe all the black hole interior and the surroundings of the black hole up to a cut-off surface. After the Page time, however, the degrees of freedom of **(CD)** describe only a portion of the interior. Indeed, the other portion of the interior is described by Hawking radiation since part of the black hole interior belongs to the entanglement wedge of the radiation. Hence, we now have a precise story, given in terms of entanglement wedge reconstruction, to explain how and when the degrees of freedom of **(CD)** encode the black hole interior and also they do not since the interior is encoded in the radiation. Given that the core of the argument in (iii) concerned the inability of **(QMP)** to account for the interior in terms of a fundamental set of quantum degrees of freedom, then also, in this case, we see how **(CD)** succeeds where **(QMP)** failed, by giving such an account. Hence, **(CD)** provides a natural extension of **(QMP)** to the full QG case also in the face of issues such as those raised by (iii).

Before concluding, let us underline another essential feature of the black hole spacetime structure emerging from the works of [10, 11].[25] In ordinary, General Relativistic black holes, the interior and the exterior of the black hole are two disconnected regions of spacetime. However, this is not the case when considering quantum effects, as developed in this section. Indeed, after the Page time, it is, in principle, possible to make complex operations on the Hawking radiation (accessible to an asymptotic observer) to recover information on the interior of the black hole since a portion of the interior and the Hawking radiation belong to the same quantum system. To be more precise, one could define what it means to be distinct in spacetime as:

[24] The entanglement entropy contribution in S_{gen} is tiny after the Page time since the island contains much of the interior Hawking modes, which purify the outgoing radiation.

[25] For an in-depth discussion of this topic, which we only sketch here, see [32].



(a) Before the Page time (b) After the Page time

Fig. 6 The Penrose diagrams of an evaporating black hole. The red dot represents the location of the Ryu–Takayanagi surface, the green line represents the cut-off surface, the dashed line represents the event horizon, the blue region represents the entanglement wedge of the degrees of freedom of **(CD)**, and the orange region represents the entanglement wedge of the Hawking radiation (Color figure online)

> **Spacetime Distinctness:** spacelike separated quantum systems are distinct, i.e. their observable algebras are mutually commuting.[26]

Then, in this framework, even if the interior and the exterior of a black hole are spacelike related for a fixed time slice, there are non-trivial connections between the Hawking radiation (which is in the exterior of the black hole) and the interior, where these non-trivial connections come from the fact that the radiation and the interior do not commute since they belong to the same quantum system. Therefore, after the Page time, Spacetime Distinctness is clearly violated. This further difficulty in assigning a precise geometrical location to the degrees of freedom describing a quantum black hole further reinforces **(CD)**'s basic insight that the degrees of freedom describing the black hole are not localised anywhere specifically in spacetime, let alone at the black hole horizon.

7 Conclusions and Outlooks

Black holes have been a subject of intense interest in both theoretical and experimental physics in recent years. This paper focuses on analysing the structure of the black hole event horizon and its fundamental status. Previous work by Wallace has proposed a Quantum Membrane Paradigm **(QMP)** to explain the statistical mechanical

[26] The idea that distinctness corresponds to mutual commutativity comes from Algebraic Quantum Field Theory, see [33] for a review.

underpinnings of black hole thermodynamics. According to [5]'s proposal, a real membrane made of black hole microstates should be posited at the black hole horizon to have a statistical mechanical underpinning of black hole thermodynamics. However, we have argued that [5]'s proposal, as it stands, is limited to the context of low-energy QG and needs modification to avoid reference to geometric notions, such as the event horizon, that may not make sense in the non-spatiotemporal context of full QG. In particular, we pointed out that **(QMP)** suffers from being (i) conventional, (ii) inconsistent, and (iii) incomplete.

To develop a more comprehensive and consistent philosophical framework for understanding black hole thermodynamics in the context of full QG, we propose a modification of **(QMP)** that takes the so-called central dogma of black hole physics seriously. Our proposal modifies the semiclassical description of the black hole interior to include certain quantum gravitational effects and avoids reference to geometric notions such as the event horizon.

Furthermore, building on recent work by [10, 11], we discussed how the semiclassical description of the black hole interior appears in our proposal and how the standard path integral description from which Wallace justifies **(QMP)** is subtly modified to include specific quantum gravitational effects, leading to the central-dogma-based picture that we propose. Overall, our paper provides a modified version of [5]'s proposal that overcomes the limitations of the original **(QMP)** in the context of full QG. Our proposal has important implications for understanding black hole thermodynamics and the nature of the event horizon in a non-spatiotemporal context.

A fascinating question would be finding a definition of a black hole adapted to the quantum gravitational context. Since the general relativistic definitions rely on geometric notions, and such geometric notions are not well-defined in QG, we need a new approach. At the same time, the central dogma simply tells that a black hole is a unitary evolving quantum system with a given entropy. A natural approach would be to provide a *functional* definition of a black hole, which aims to enrich the central dogma with purely quantum features of systems that behave as black holes in the appropriate limit. A promising avenue to identify a proper set of systems playing the functional role of black holes lies in the connection between black holes, chaotic systems and complexity theory. We leave the exploration of this issue to future works.

Acknowledgements For helpful comments on a previous draft, the authors wish to thank Alberto Corti and Vincenzo Fano.

Author Contributions Enrico Cinti and Marco Sanchioni contributed to the conception of the paper. Enrico Cinti and Marco Sanchioni wrote the first draft of the manuscript.Enrico Cinti and Marco Sanchioni commented on previous versions of the manuscript. Enrico Cinti and Marco Sanchioni read and approved the final manuscript.

Funding No funds, grants, or other support was received.

Declarations

Conflict of interest Enrico Cinti and Marco Sanchioni have no relevant financial or non-financial interests to disclose. The authors declare no competing interests.

References

1. Akiyama, K., et al.: First M87 event horizon telescope results. I. The shadow of the supermassive black hole. Astrophys. J. Lett. **1**, 875 (2019). https://doi.org/10.3847/2041-8213/ab0ec7
2. Hawking, S.W.: Particle creation by black holes. Commun. Math. Phys. **43**(3), 199–220 (1975)
3. Belot, G., Earman, J., Ruetsche, L.: The Hawking information loss paradox: the anatomy of controversy. Br. J. Philos. Sci. **50**(2), 189–229 (1999). https://doi.org/10.1093/bjps/50.2.189
4. Wallace, D.: 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 64, 52–67 (2018) 10.1016/j.shpsb.2018.05.002
5. Wallace, D.: The case for black hole thermodynamics part ii: statistical mechanics. Stud Hist Philos Sci Part B **66**, 103–117 (2019). https://doi.org/10.1016/j.shpsb.2018.10.006
6. Wallace, D.: Quantum gravity at low energies. Stud. Hist. Philos. Sci. **94**, 31–46 (2022)
7. Seiberg, N.: Emergent spacetime. In: The Quantum Structure Of Space And Time, pp. 163–213 (2007)
8. Huggett, N., Wüthrich, C.: Emergent spacetime and empirical (in) coherence. Stud. Hist. Philos. Sci. Part B **44**(3), 276–285 (2013)
9. Almheiri, A., Hartman, T., Maldacena, J., Shaghoulian, E., Tajdini, A.: The entropy of hawking radiation. Rev. Modern Phys. **3**, 93 (2021). https://doi.org/10.1103/revmodphys.93.035002
10. Penington, G., Shenker, S.H., Stanford, D., Yang, Z.: Replica wormholes and the black hole interior. J. High Energy Phys. **3**, 1–87 (2020)
11. Almheiri, A., Hartman, T., Maldacena, J., Shaghoulian, E., Tajdini, A.: Replica wormholes and the entropy of hawking radiation. J. High Energy Phys. **2020**(5), 13 (2020). https://doi.org/10.1007/JHEP05(2020)013
12. Page, D.N.: Information in black hole radiation. Phys. Rev. Lett. **71**(23), 3743–3746 (1993). https://doi.org/10.1103/physrevlett.71.3743
13. Bekenstein, J.D.: Black holes and entropy. Phys. Rev. D **7**, 2333–2346 (1973). https://doi.org/10.1103/PhysRevD.7.2333
14. Armas, J., Obers, N.A., Sanchioni, M.: Gravitational tension, spacetime pressure and black hole volume. J. High Energy Phys. **2016**(9), 124 (2016). https://doi.org/10.1007/JHEP09(2016)124
15. Susskind, L., Thorlacius, L., Uglum, J.: The stretched horizon and black hole complementarity. Phys. Rev. D **48**(8), 3743–3761 (1993). https://doi.org/10.1103/physrevd.48.3743
16. Wüthrich, C.: Are Black Holes about Information? In: Dardashti R., Dawid R., Thébault K. (eds.) Why Trust a Theory? Epistemology of Fundamental Physics, pp. 202–224. Cambridge University Press, Cambridge (2019)
17. Williams, P.: Renormalization group methods. In: The Routledge Companion to Philosophy of Physics, pp. 296–310. Routledge (2021)
18. Bombelli, L., Koul, R.K., Lee, J., Sorkin, R.D.: Quantum source of entropy for black holes. Phys. Rev. D **34**, 373–383 (1986). https://doi.org/10.1103/PhysRevD.34.373
19. Wallace, D.: Why black hole information loss is paradoxical. In: Matsubara k., Huggett, N., Wüthrich C. (eds.) Beyond Spacetime: The Foundations of Quantum Gravity, pp. 209–236. Cambridge University Press, Cambridge (2020)
20. Strominger, A., Vafa, C.: Microscopic origin of the Bekenstein–Hawking entropy. Phys. Lett. B **379**(1–4), 99–104 (1996). https://doi.org/10.1016/0370-2693(96)00345-0

21. Maldacena, J.: The large n limit of superconformal field theories and supergravity. Int. J. Theor. Phys. **38**(4), 1113–1133 (1999). https://doi.org/10.1023/a:1026654312961
22. Dong, X., Harlow, D., Wall, A.C.: Reconstruction of bulk operators within the entanglement wedge in gauge-gravity duality. Phys. Rev. Lett. **117**, 021601 (2016). https://doi.org/10.1103/PhysRevLett.117.021601
23. Jafferis, D.L., Lewkowycz, A., Maldacena, J., Suh, S.J.: Relative entropy equals bulk relative entropy. JHEP **6**, 004 (2016). https://doi.org/10.1007/JHEP06(2016)004
24. Saad, P., Shenker, S.H., Stanford, D.: JT gravity as a matrix integral. Preprint at https://arxiv.org/abs/1903.11115 (2019)
25. Calabrese, P., Cardy, J.: Entanglement entropy and conformal field theory. J. Phys. A **42**(50), 50400 (2009). https://doi.org/10.1088/1751-8113/42/50/504005
26. Hubeny, V.E., Rangamani, M., Takayanagi, T.: A covariant holographic entanglement entropy proposal. J. High Energy Phys. **2007**(07), 062–062 (2007). https://doi.org/10.1088/1126-6708/2007/07/062
27. Bain, J.: Spacetime as a quantum error-correcting code? Stud. Hist. Philos. Sci. Part B **71**, 26–36 (2020). https://doi.org/10.1016/j.shpsb.2020.04.002
28. Ryu, S., Takayanagi, T.: Holographic derivation of entanglement entropy from the anti–de sitter space/conformal field theory correspondence. Phys. Rev. Lett. 96, 181602 (2006) 10.1103/PhysRevLett.96.181602
29. Faulkner, T., Lewkowycz, A., Maldacena, J.: Quantum corrections to holographic entanglement entropy. J. High Energy Phys. **2013**(11), 74 (2013). https://doi.org/10.1007/JHEP11(2013)074
30. De Haro, S., Mayerson, D.R., Butterfield, J.N.: Conceptual aspects of gauge/gravity. Duality. Found. Phys. **46**(11), 1381–1425 (2016). https://doi.org/10.1007/s10701-016-0037-4
31. Penington, G.: Entanglement wedge reconstruction and the information. J High Energy Phys. **9**, 1–84 (2020)
32. Cinti, E., Sanchioni, M.: The devil in the (implicit) details. Int. J. Theor. Phys. **60**(9), 3234–3253 (2021). https://doi.org/10.1007/s10773-021-04901-1
33. Halvorson, H., Mueger, M.: Algebraic quantum field theory. In: Butterfield J., Earman J., (eds.) Philosophy of Physics. Amsterdam, North-Holland (2007)