This is a repository copy of Better late than never: Information retrieval from black holes.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/75302/

Version: Published Version

Article:
Braunstein, Sam orcid.org/0000-0003-4790-136X, Pirandola, Stefano orcid.org/0000-0001-6165-5615 and Zyczkowski, Karol (2013) Better late than never: Information retrieval from black holes. Physical Review Letters. 101301. ISSN 1079-7114

https://doi.org/10.1103/PhysRevLett.110.101301

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Better Late than Never: Information Retrieval from Black Holes

Samuel L. Braunstein and Stefano Pirandola
Department of Computer Science, University of York, York YO10 5GH, United Kingdom

Karol Życzkowski
Institute of Physics, Jagiellonian University, 30-059 Krakow, Poland and Center for Theoretical Physics, Polish Academy of Science, 02-668 Warszawa, Poland

(Received 17 April 2012; published 5 March 2013)

We show that, in order to preserve the equivalence principle until late times in unitarily evaporating black holes, the thermodynamic entropy of a black hole must be primarily entropy of entanglement across the event horizon. For such black holes, we show that the information entering a black hole becomes encoded in correlations within a tripartite quantum state, the quantum analogue of a one-time pad, and is only decoded into the outgoing radiation very late in the evaporation. This behavior generically describes the unitary evaporation of highly entangled black holes and requires no specially designed evolution. Our work suggests the existence of a matter-field sum rule for any fundamental theory.

DOI: 10.1103/PhysRevLett.110.101301

PACS numbers: 04.70.Dy, 03.65.Xp, 03.67.3

Black hole evaporation as tunneling.—Although pair creation provides the conventional heuristic picture of the microscopic process by which a black hole evaporates [1], it has come under increasing suspicion due to intrinsic difficulties. In particular, pair creation necessarily requires the dimensionality of the interior Hilbert space of a black hole to be increasing while simultaneously its physical size is decreasing [2,3].

By contrast, quantum tunneling, which operates by moving quantum subsystems across the classically forbidden barrier of the event horizon, naturally avoids this difficulty [3]. Furthermore, quantum tunneling invites an elegant Hilbert space description of the evaporation process across event horizons [3]: We start with the standard decomposition of a black hole Hilbert space into a tensor product between the interior (int) and exterior (ext) by $\mathcal{H}_{\text{int}} \otimes \mathcal{H}_{\text{ext}}$ [4] and note that an event horizon’s tensor product structure in no way implies that its spatial location cannot be fuzzy [3].

Tunneling now operates [3] by selecting some subsystem from the black hole interior and moving it to the exterior $\mathcal{H}_{\text{int}} \to \mathcal{H}_{R} \otimes \mathcal{H}_{R}$ by

$$|i\rangle_{\text{int}} \to (U|i\rangle)_{BR}, \quad (1)$$

where $U$ denotes the unitary process that might be thought of as “selecting” the subsystem to eject, $|i\rangle$ is the initial state of the black hole interior, $B$ denotes the reduced size subsystem corresponding to the remaining interior after evaporation, and $R$ denotes the subsystem that escapes as radiation [3,5,6].

Equation (1) has been used before to study black hole evaporation [3,5,6]; however, with the exception of Ref. [3], it has not been used as a process associated with any underlying physical mechanism. Indeed, Ref. [3] showed that the symmetries implicit in this equation, in conjunction with global conservation laws for the no-hair quantities (energy, charge, and angular momentum), suffice to completely determine black hole tunneling probabilities for any black hole and particle type, reproducing and even extending the predictions of field theory on curved spacetime. This work therefore strongly supports Eq. (1) as a pertinent microscopic formulation of unitary black hole evaporation. Its implications for the retrieval of information about in-fallen matter will be further studied here.

Dynamical evaporation with entanglement.—It is now well accepted that entanglement across boundaries is generic [7]. Therefore, our key point of departure from previous work [3,5,6,8,9] will be to allow for entanglement across the event horizon. Incorporated into the evaporative dynamics of Eq. (1), but making no assumption of how much or how little trans-event horizon entanglement there may be, this entanglement gives

$$\sum_{i=1}^{N}\sqrt{p_{i}}|i\rangle_{\text{int}} \otimes |\tilde{i}\rangle_{\text{ext}} \to \sum_{i=1}^{N}\sqrt{p_{i}}(U|i\rangle)_{BR} \otimes |\tilde{i}\rangle_{\text{ext}}. \quad (2)$$

Now, the nature of the black hole as a compact object of a given mass constrains any interior evolution to only access an effectively finite dimensional Hilbert space [3]. Quantities defined within (the support of) this finite Hilbert space will similarly be finite, including, for example, any von Neumann entropies, measures of entanglement, etc. Indeed, it has been argued [10] that the dimensionality for...
the initial black hole Hilbert space should be well approximated by the thermodynamic entropy \( S_{\text{BH}} = {\mathcal{A}}/(4\ln 2) \) for a black hole of area \( \mathcal{A} \), giving a dimensionality \( N \equiv \dim({\mathcal{H}}_{\text{BH}}) = BR = 2^{S_{\text{BH}}} \), where we reuse subsystem labels for Hilbert space dimensionalities and for later convenience we evaluate entropies using base-two logarithms. We might say that the black hole interior comprises \( \log_2 N = S_{\text{BH}} \) qubits. (Throughout, the term “qubits” is used merely as a unit of information content and does not literally imply a set of two-level systems.) That the number of qubits initially within the black hole is well approximated by \( S_{\text{BH}} \) is supported by the holographic principle \([11]\) and independently by the amount of Hawking radiation that would be generated, consistent with energy conservation.

Naively, to make quantitative predictions based on this description, we would need to know the detailed dynamics \( U \) within the black hole. In fact, the behavior of information flow in a high-dimensional system under a specific unitary will be in excellent agreement with the Haar average over all unitaries acting on dimension \( N \). This follows from Levy’s lemma \([12]\), which states that the logarithm of the probability of any such difference \( \epsilon \) scales as \( -N \epsilon^2 \). For a stellar mass black hole, such dimensionalities \( N \) must be at least \( 10^{107} \), so even the smallest deviations from the average behavior should occur with vanishingly small probability. Numerical simulations in even very low dimensions show this to be well supported, and similar results are well known beyond black hole physics \([13]\). Thus, here we replace the behavior of the specific unitary in Eq. (2) by the Haar average.

**Vanishing of transevent horizon entanglement.**—Moving to the average behavior allows one to rigorously interpret the evaporative dynamics of a black hole in terms of the properties of random quantum error correcting codes \([6]\). In this interpretation, one-half of an entangled state is encoded into a larger Hilbert space via a random unitary encoding. Decoupling theorems \([14]\) tell us how much (how many qubits) of the encoded state one must have access to, in principle, in order to reconstruct the original unencoded state, including its entanglement. We derive a generalized decoupling theorem and use it to address a broader set of questions. (See the Supplemental Material in Ref. \([15]\) for proofs and a discussion of both quantum and classical decoupling theorems.)

For example, for an entangled black hole evolving via Eq. (2), this generalized decoupling theorem shows that, for any positive number \( c \), once

\[
\log_2 R = \frac{1}{2} S_{\text{BH}} + \frac{1}{2} H^{(1/2)}(\rho_{\text{ext}}) + c
\]

qubits have radiated away, the transevent horizon entanglement will have vanished, appearing instead, with virtually unit fidelity (at least \( 1 - 2^{-c} \)), as entanglement between the external neighborhood and radiation. Here, the entropy of entanglement is quantified by a Rényi entropy \( H^{(q)}(\rho) \equiv \log_2(\text{tr} \rho^q)/(1-q) \) with \( q \) of order unity, for the reduced density matrix of the (ext) state \( \rho_{\text{ext}} = \sum_{i} \rho_i |i\rangle_{\text{ext}} \langle i| \) neighboring the event horizon.

**Entanglement and the equivalence principle.**—We will now explicitly link the presence of transevent horizon entanglement with the equivalence principle. Specifically, the equivalence principle is expected to be preserved for black holes larger than the Planck scale. We will argue below that the presence of this entanglement must be similarly preserved until such scales. We then use the tunneling dynamics to calculate the initial amount of transevent horizon entanglement.

We start by recalling the equivalence principle, which tells us that a freely falling observer sees no local effects due to gravity. Applied to black holes, it has been argued \([10]\) that the equivalence principle implies that an observer freely falling past the event horizon would see no Hawking radiation, only a zero temperature vacuum state—just as an unaccelerated observer in flat spacetime. Now, the well-understood quantum physics of condensed matter systems tells us that entanglement across boundaries is generic in or near the ground state \([7]\). Furthermore, in axiomatic quantum field theory, entanglement across boundaries for fields in their vacuum state is implicit in the Reeh-Schlieder theorem \([16]\). In the Supplemental Material \([15]\), we derive a lower bound for the energy of a free scalar field when the quantum state is restricted to have no entanglement across an arbitrary hypothetical boundary. This disentanglement energy diverges as a power of the UV regulator \([15]\), and hence is far above the vacuum state. Applied to black holes, this means that the loss of entanglement across the event horizon would force the quantum fields across it to be arbitrarily far from the vacuum state—an energetic curtain would have descended around the black hole \([17]\)—signaling a manifest failure of the equivalence principle.

Next, we use the epoch for the loss of transevent horizon entanglement, given by Eq. (3), to quantify how much transevent horizon entanglement was in the initial black hole. Here, we rely on the observation that a black hole’s size may be directly quantified by its area or, equivalently, its entropy. For black holes in the latter stages of evaporation via Eq. (2), their entropy is well approximated by \( S_{\text{BH}} = \log_2 R \) \([15]\). Therefore, an evaporating black hole can be said to approach the Planck scale (see Ref. \([15]\) for a detailed discussion) when, to high precision, \( \log_2 R = S_{\text{BH}} \). From Eq. (3), preserving entanglement until such late times implies that

\[
H^{(1/2)}(\rho_{\text{ext}}) \approx S_{\text{BH}}.
\]

In other words, preserving transevent horizon entanglement up until an evaporating black hole approaches the Planck scale requires that its initial entanglement entropy be almost exactly its initial thermodynamic entropy \([18]\). This result is insensitive to where we place the entry point to the Planck
hole, in the presence of in-fallen matter, gives to form a black hole to be a large fraction of a black hole’s zero vectors. This means we pad unused dimensions of the interior space by evaporative dynamics of Eq. (11). Now apply our generalized decoupling theorem to the up to overall int-local and ext-local unitaries. Here, the quantum state of the matter collapses to form a black hole. In that case, the exterior surroundings must have the form

\[ S_{\text{matter}} \ll S_{\text{BH}}. \]

This then raises the question of when and in what fashion anything but Planck scale black holes, the entropic contribution from in-fallen matter is negligible, \( S_{\text{matter}} \ll S_{\text{BH}}. \) This gives a unique description for a quantum state where information is not available within some specific subsystem. No-hiding implies that the quantum state of a newly formed black hole interior (int) and its surroundings have the form

\[ \frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle_{\text{ref}} \otimes \sum_{j} \sqrt{p_j}|i\rangle \otimes |j\rangle. \]

up to overall int-local and ext-local unitaries. Here, \( \emptyset 0 \) means we pad unused dimensions of the interior space by zero vectors [9], and \( \log_2 K \equiv S_{\text{matter}} \) is the number of qubits describing the quantum state of the matter collapsing to form the black hole.

Applying the dynamics of Eq. (1) to our entangled black hole, in the presence of in-fallen matter, gives

\[ -\frac{1}{\sqrt{K}} \sum_{i=1}^{K} |i\rangle_{\text{ref}} \otimes \sum_{j} \sqrt{p_j}[U(|i\rangle \otimes |j\rangle) \otimes 0]_{\text{BR}} \otimes |j\rangle. \]

Information retrieval from entangled black holes.—We now apply our generalized decoupling theorem to the evaporative dynamics of Eq. (5). In order to state our results, it will be convenient to roughly quantify the number of unentangled (pure) qubits within the initial black hole state in Eq. (5a); we define this “excess” as

\[ \chi^{(q)} \equiv S_{\text{BH}} - S_{\text{matter}} - H^{(q)}(\rho_{\text{ext}}) \geq 0. \]

Note that Eq. (4) implies \( \chi^{(1/2)} \ll S_{\text{BH}}. \)

We now summarize the results about information encoding and retrieval. Since, in each application of the theorem, an independent dummy variable appears \( [\epsilon \text{ in Eq. (3)}] \) that is dwarfed by other entropies, here we omit reference to them (the complete statements can be found in Ref. [15]).

Thermalization: Initially, one might suppose that any in-fallen matter would be well within the interior of the black hole, far inside the event horizon, and so would not be selected by \( U \) to participate in tunneling across this boundary. Only after the black hole had sufficiently “scrambled” the internal states (after what might be called the global thermalization time [6] for the black hole) would the subsystem encoding the state of the in-fallen matter be accessible for selection and ejection by tunneling [20]. Note that estimates of scrambling times vary. Some recent analyses suggest that black holes are fast scramblers [6,21] (with the scrambling time being little more than the time for a single Hawking photon to evaporate), whereas other estimates are slow [22].

Encoding: During the global thermalization time and for the next \( \frac{1}{2} \chi^{(1/2)} \) qubits radiated, all the information about the state of the in-fallen matter is encoded with virtually unit fidelity within the black hole interior. For the next \( S_{\text{matter}} + \frac{1}{2} \chi^{(2)} \) qubits radiated, this information becomes encoded into the tripartite correlations of a quantum one-time pad [9] among the black hole interior, the external neighborhood, and the radiation. In other words, it is the evaporation via tunneling (across the event horizon) that encodes the information as tripartite entanglement. After encoding and until the last \( S_{\text{matter}} + \frac{1}{2} \chi^{(5)} \) qubits radiated the information remains within this quantum one-time pad; it is inaccessible from any subsystem individually, but it is accessible from any two of them. The quantum one-time pad is a random quantum error correction code. The properties of such codes dictate the size of subsystems one must have access to in order to be able to reconstruct the original state of the in-fallen matter.

Decoding: At this point in the evaporation process, entanglement within the black hole becomes so depleted that it can no longer contain the correlations of all the in-fallen matter. The final \( S_{\text{matter}} + \frac{1}{2} \chi^{(2)} \) qubits to be radiated marks the start of information release into the radiation. From here until the final \( \frac{1}{2} \chi^{(1/2)} \) qubits radiated from the black hole, the full information about the in-fallen matter is decoded and becomes available in the outgoing radiation for the first time. This decoding takes the same amount of time as the encoding. Since typically \( \chi^{(2)} - \chi^{(1/2)} \approx O(1) \) and this quantity cannot be negative, the encoding or decoding occurs at roughly the radiation emission rate; recall, Hawking quanta typically carry around one (qubit of thermal entropy. (See Ref. [15] for a heuristic picture of the flow of information.)
This completes our analysis of information retrieval for unitarily evaporating highly entangled black holes. As is evident from this summary, decoding of the information about in-fallen matter is very brief, occurring within the final and vanishingly small fraction $S_{\text{matter}}/S_{\text{BH}} \leq O(S_{\text{BH}}^{-1/4})$ of a large black hole’s lifetime (as measured in Hawking quanta radiated). This is so late that its timing is unaffected by even very long scrambling times [22]. That said, within this very brief epoch, decoding is also very slow, occurring at the radiation emission rate; thus, information about the in-fallen matter is decoded over the time scale required for $S_{\text{matter}} \leq O(S_{\text{BH}}^{3/4})$ Hawking quanta to evaporate. Because the number of qubits radiated during decoding is so vast, essentially all the information has been retrieved long before the black hole shrinks to the Planck scale. Note that the time scales above follow from ’t Hooft’s entropic bound [11]; however, none of the mechanisms or mathematical results in this Letter rely on this bound [15].

Discussion and wider implications.—The application of information theoretic approaches to the physics of black holes is relatively new [2,3,5,6,8–10]. Here, we have shown that this approach offers a description of black holes as highly entangled, with direct consequences for the time course of information retrieval therefrom. This approach necessarily requires an explicit formulation of the microscopic evaporation process, which, here, we take to be quantum tunneling [3]. The analysis and the results are grounded in black hole physics, and hence cannot be taken to apply to arbitrary horizons, but the tunneling mechanism invoked should apply more universally.

To ground our approach in the physics of black holes, we have relied on a number of key principles and results from classical general relativity and field theory, including the implications of the equivalence principle [10] for the field-theoretic state at the event horizon, the nonexistence of a “bleaching” mechanism [19], and the requirement for some thermalization or scrambling mechanism [6,21,22] that allows information from deep inside a black hole to reach the surface before radiating away [20] (although our results are largely insensitive to the time scale and hence the underlying scrambling mechanism).

Previously, the no-hiding theorem [9,23] was used to prove that Hawking’s prediction of featureless radiation implied that the information about the in-fallen matter could not be in the radiation field but must reside in the remainder of Hilbert space—then presumed to be the black hole interior. That work presented a strong form of the black hole information paradox pitting the predictions of general relativity against those of quantum mechanics [9]. Here, we have shown that transevent horizon entanglement provides a way out, since now the “remainder of Hilbert space” comprises both the black hole interior and external neighborhood. Because the evaporating black hole actually involves three subsystems, the information may be encoded within them as pure correlations via a quantum one-time pad [9], so the information remains inaccessible from any one subsystem.

Importantly, the detailed physics of black holes (inside the event horizon) remains beyond the scope of this Letter. Thus, this Letter leaves mysterious those long-standing questions about the internal dynamics of black holes that would require knowledge of the geometry well within the black hole and extensive field-theoretic calculations or even a theory of quantum gravity to be addressed. The very assumption of unitarity is one such question. Another is our positing of a finite entanglement entropy across the event horizon, without a detailed field-theoretic description of how this should be calculated [24–28]. Similarly, the dynamics of the entangled degrees of freedom exterior to the black hole remains unclear. Finally, we assume the existence of some global thermalization process that leads to complete scrambling of the information encoded within the black hole.

The simultaneous encoding of information externally (in the combined radiation and external neighborhood) and “internally” (if one slightly stretches the horizon to envelop the bulk of the external neighborhood entanglement in addition to the black hole interior) is reminiscent of the principle of black hole complementarity [10]. This principle was introduced to account for the apparent cloning suggested by the possibility of choosing a “nice time” slice through the black hole spacetime that crosses most of the outgoing radiation as well as the collapsing body well inside the event horizon but still far from the singularity [29]. Interpreted in the context of our work here, if such slices are drawn after the encoding of the information into the tripartite quantum one-time pad, the “cloning” would be a manifestation of the multiple ways of reading out the information from the tripartite structure. If such slices are drawn before the encoding occurred, then too little of the outgoing radiation would be crossed for a potential violation of the no-cloning theorem (note that the number of qubits radiated may be used as a surrogate for a time coordinate).

Our results indicate that, except for the very final, vanishingly small fraction of a (large) black hole’s lifetime, the Hawking radiation is completely uncorrelated with the state of the in-fallen matter. Thus, the behavior Hawking found so indicative of a loss of unitarity is in fact completely generic for unitarily evolving, entangled black holes, requiring no specially designed evolution. Of course, by assuming unitarity from the outset, we cannot directly address the black hole information paradox. Rather, our result dissociates completely information-free radiation from a loss of unitarity and hence undermines the very logic used to formulate the paradox.

Finally, in light of their curious equality, it has previously been conjectured that a black hole’s thermodynamic entropy is actually entropy of entanglement [24–27]. Indeed, it unavoidably holds for some types of extremal
black holes [25,26] and even allows their entropy to be computed at the microscopic level [27]. The conventional riposte to this conjecture is made by noting that the entropy of entanglement of quantum fields piercing a black hole’s event horizon would be proportional to the number of matter fields that exist, but, since a black hole’s thermodynamic entropy is purely geometric, there should be no a priori relationship between these quantities (see, e.g., Ref. [30]; for a counterargument, see Ref. [31]). By studying dynamically evolving black holes, not merely static ones [24–27], we now counter this conventional riposte. Equating a black hole’s entropy with entropy of entanglement suggests the existence of a sum rule to constrain the number and types of matter fields in any fundamental theory.

S. L. B. acknowledges the kind hospitality of the A. Watssman Institute for Innovative Thinking, where this work was initiated. K. Z. acknowledges support by the Deutsche Forschungsgemeinschaft under Project No. SFB/TR12 and under the Project No. DEC-2011/02/A/ST1/00119, financed by the Polish National Science Center. The authors thank N. Cohen, M. Patra, and H.-J. Sommers for fruitful discussions.

Note added.—Our “energetic curtain” (first coined in Ref. [17]) appears to be the same phenomenon recently called a “firewall” [32].

[18] Our analysis is silent as to whether or not this relationship between thermodynamic and entanglement entropy also holds for acceleration or cosmological horizons.
[20] All incoming degrees of freedom are required to be scrambled for our information retrieval mechanism to apply in full. For black holes with nontrivial interiors, WKB tunneling analysis attributes a temperature to inner horizons that always exceeds that of the outer (event) horizon. Thus, the existence of an inner horizon is no absolute barrier to global scrambling.