Wormhole calculus, replicas, and entropies

Steven B. Giddings and Gustavo J. Turiaci

Department of Physics, University of California,
Santa Barbara, CA 93106, U.S.A.

E-mail: giddings@ucsb.edu, turiaci@ucsb.edu

ABSTRACT: We investigate contributions of spacetime wormholes, describing baby universe emission and absorption, to calculations of entropies and correlation functions, for example those based on the replica method. We find that the rules of the "wormhole calculus", developed in the 1980s, together with standard quantum mechanical prescriptions for computing entropies and correlators, imply definite rules for *limited* patterns of connection between replica factors in simple calculations. These results stand in contrast with assumptions that all topologies connecting replicas should be summed over, and call into question the explanation for the latter. In a "free" approximation baby universes introduce probability distributions for coupling constants, and we review and extend arguments that successive experiments in a "parent" universe increasingly precisely fix such couplings, resulting in ultimately pure evolution. Once this has happened, the nontrivial question remains of how topology-changing effects can modify the standard description of black hole information loss.

Contents

1 Introduction

Nontrivial spacetime topologies, and in particular change in the topology of space, have long been considered to be a possible feature of dynamical gravity. Topology-changing processes were particularly intensively studied in the late 1980s, in the context of the question of their contribution to possible loss of quantum coherence [1–5]. Specifically, one can consider processes where space branches into two disconnected components; one of these may typically be comparatively small, and was called a "baby universe" (BU). In the "free BU" approximation where multiple BUs can be emitted, or rejoin, a bigger "parent universe", but where the BUs don't interact or create other large universes, it was found that the leading effect of such processes is *not* to induce an ongoing loss of quantum coherence [4, 5].[1] Instead, these processes lead to an effective probability distribution for coupling constants that multiply operators describing the effect of the BUs on the fields in the parent universe.

There has been a recent resurgence of interest in topology change, arising from suggestions that nontrivial topologies may help explain how black hole evolution can be reconciled with unitary quantum mechanical evolution [8–10].[2] Specifically, [9, 10] have argued that nontrivial topological contributions can produce expressions for BH entropies that behave as expected for unitary evolution [15, 16]. This work builds on earlier discussion [17–19] about the role of quantum extremal surfaces, and that of [8] on topologies and ensembles of couplings in Jackiw-Teitelboim gravity (see also the related work [20–24]). The topologies studied in [9, 10] involve spacetime wormhole connections, but of a somewhat

---

[1]Effects beyond this approximation were discussed in [6], and recently in [7].
[2]For earlier work in this direction, see [11, 12]. For a different but possibly related approach see [13, 14].

different kind than those studied in the 1980s. Specifically, entropies are calculated by the replica method [25], in which multiple copies of the spacetime geometry are considered. One then makes the Ansatz that wormholes, or more general nontrivial topologies, connect these replicas. While the replica wormhole contributions have not yet been shown to correspond to quantum amplitudes describing unitary evolution, they have been argued to produce entropy formulas that reflect unitary behavior, giving an appropriate form of a "Page curve"[15, 16].

The obvious possible connection between replica wormholes and the spacetime wormholes considered previously was noted in [9, 10], and further developed in [7]. However, an important question in the discussion is to better understand the precise connection, and to test and understand the correct rules for replica calculations in the presence of euclidean wormholes/BU emission. Specifically, [4, 5] previously developed a set of rules for incorporating topology change, respecting certain general quantum properties such as the composition of amplitudes; this is sometimes called the "wormhole calculus". Given the wormhole calculus, one can then perform standard quantum-mechanical calculations — such as of entropies, e.g. using replicas — and ask what the combined set of rules tells us about the contribution of nontrivial topologies connecting replicas, and regarding the question of summing over all such replica geometries.

That is one of the goals of this paper. Specifically, we find that the previously-developed rules of the wormhole calculus, which have been well studied in a framework consistent with quantum mechanics, together with basic quantum-mechanical rules for computing entropies, imply *specific limited patterns of wormhole connections in replica geometries.* These do not include sums over all connections between replicas. This runs contrary to the prevalent Ansatz that one should generally sum over all such replica topologies [9, 10], and calls into question the meaning of calculations based on such a sum. Specifically this suggests that if there is a role for replica wormholes in certain calculations, it needs to be more carefully understood; alternatively it may also be that including such contributions represents a modification of usual quantum-mechanical rules for calculating entropies, or somehow gives an effective description summarizing the contribution of other effects.

In outline, the next section gives a brief review of the wormhole calculus. Section three then turns to the question of calculating some simple entropies, as well as correlators, in the presence of nontrivial spacetime topologies and ensembles of BUs, showing that the wormhole calculus together with the usual rules dictate only certain patterns of wormhole connections between replicas. Section four discusses a related question, namely that of understanding the effects of BUs as providing a probability distribution for coupling constants, and the way in which subsequent experiments determine these couplings; this provides a generalization of the analysis of [4, 5] of these questions. Section five closes with some further discussion.

2   Review of the wormhole calculus

We begin by reviewing the basics of the wormhole calculus, developed in [4, 5]. This was based on assuming the existence of topology-changing interactions in which a universe can
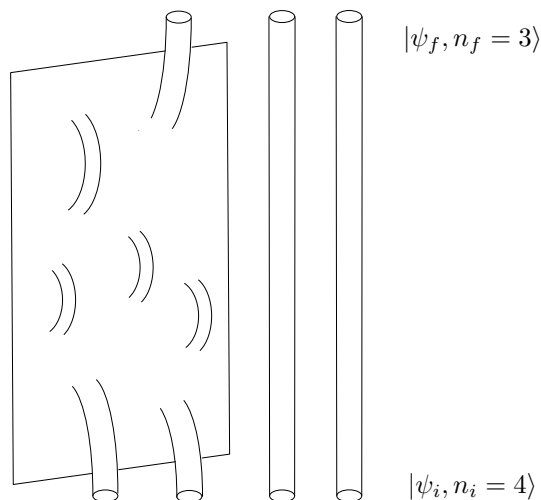
Figure 1. To compute the amplitude to go from an initial state of the parent universe, plus some number of BUs, to a similar final state, we integrate over all intermediate geometries. In this figure we sketch one particular geometry contributing to the transition amplitude between four and three BUs.

split, emitting a disconnected baby universe (BU). A simple instanton describing such processes, in the presence of a massless axionic field, was found in [3]; similar processes were also considered by [1, 2, 11, 26–31].

Specifically, suppose that we work in the free BU approximation where BUs can be emitted and absorbed by a single parent universe, but do not interact among themselves or create other large universes; going beyond this approximation can be described in a third-quantized framework [6]. For simplicity, consider the case where the parent universe has an asymptotic region where time can be defined, such as asymptotically flat or AdS space. Then, we can consider finite-time transitions between states of the parent universe, but at the same time there can be transitions in the number of BUs.

In general, the BUs can have different internal states, but for simplicity consider the case where there is a single internal state, or "species", of BU. Then, one can consider transitions between an initial state of the parent universe, together with some number of BUs, and a final state of the parent universe together with some typically different number of BUs. The amplitudes for such processes can be calculated by summing over geometries such as in figure 1, in analogy to other instanton sums in physics.

As shown in [4, 5], at scales large as compared to the typical BU size (which may be set by a microscopic scale), these amplitudes can be reproduced from a simple hamiltonian. This takes the form

$$H = H(\phi_i) + \int d^3x \, \mathcal{O}(x)(a + a^\dagger) \ . \tag{2.1}$$

Here ϕ_i are the fields on the parent universe (which may also include the metric), $H(\phi_i)$ is their hamiltonian, and $\mathcal{O}(x)$ is an operator that describes the effect of the BU emission on these fields. The operators $a^\dagger$ and a act on a BU Fock space, to create/annihilate BUs;

for example, an n BU state is given by

$$|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle \qquad (2.2)$$

where $|0\rangle$ is the BU vacuum. The form of the hamiltonian (2.1) is dictated by various considerations: the fact that BU emission conserves energy/momentum, since BUs are closed and carry no net energy/momentum, indistinguishability of BUs, and the requirement that the basic amplitudes, of the form

$$\langle \psi_f, n_f | e^{-iHT} | \psi_i, n_i \rangle , \qquad (2.3)$$

satisfy a composition law,

$$\sum_{n,\psi_I} \langle \psi_f, n_f | e^{-iHT_2} | \psi_I, n\rangle \langle \psi_I, n | e^{-iHT_1} | \psi_i, n_i \rangle = \langle \psi_f, n_f | e^{-iH(T_1+T_2)} | \psi_i, n_i \rangle \qquad (2.4)$$

where the sum includes that over a basis ψ_I of intermediate parent universe states. The discussion readily generalizes to multiple species of BUs, and can be summarized by introducing operators a_i, $a_i^\dagger$ for the different species, together with different operators $\mathcal{O}_i$ summarizing their couplings to the parent.

The form of the BU couplings (2.1) implies that, in the free BU approximation, there is a simple relation between BU states and couplings. Specifically, consider the states

$$|\alpha\rangle = \mathcal{N}\, e^{-\frac{1}{2}(a^\dagger - \alpha)^2}|0\rangle \; ; \qquad (2.5)$$

these diagonalize $a + a^\dagger$, with eigenvalue α, and if we normalize them as $\mathcal{N}^2 = e^{\alpha^2/2}/\sqrt{2\pi}$ then they satisfy the normalization convention

$$\langle \alpha | \alpha' \rangle = \delta(\alpha - \alpha') . \qquad (2.6)$$

Such a state is then an eigenstate of the hamiltonian (2.1), which takes the form

$$H_\alpha = H(\phi_i) + \alpha \int d^3x\, \mathcal{O}(x) . \qquad (2.7)$$

Thus, the evolution is that of a theory with a new coupling constant multiplying the operator $\mathcal{O}(x)$. A more general BU state can be written as a superposition of the α eigenstates, and so can be thought of as describing an ensemble of such couplings. This, together with weighting factors arising from disconnected parent universes, was for example proposed in [29] to solve the (then) cosmological constant problem, by arguing that the weighting factors overwhelmingly prefer $\Lambda = 0$.

3 Renyis, replicas, and wormhole connections

We next turn to a discussion of what the rules of the wormhole calculus, combined with basic rules of quantum mechanics, imply in the context of computing quantities such as entropies that characterize the distribution of information in the system, as well as correlators.
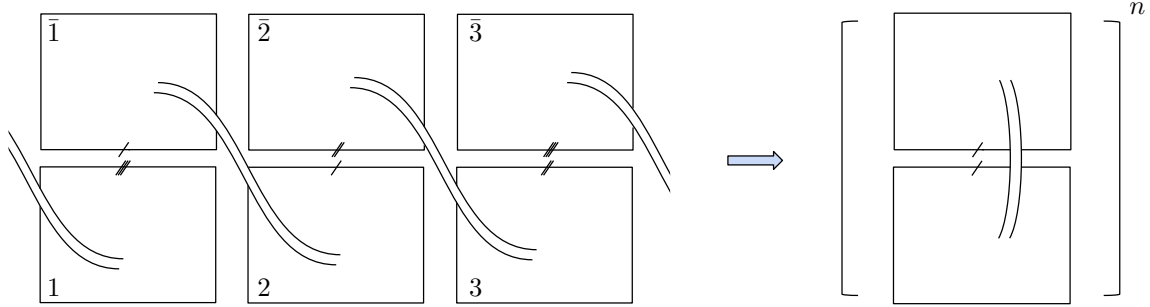
Figure 2. Shown is a sketch of the geometry used in a replica method calculation of the nth Renyi entropy of the density matrix (3.2). Time runs upwards (downwards) in the lower (upper) copies. This calculation produces only wormholes that connect different replicas in the pattern $\bar{1} - 2, \bar{2} - 3,$ $\cdots, \bar{n} - 1$. We also indicate how the parent universes are identified at time T. The wormhole joining at 1 is emitted from $\bar{n}$ while the one emitted from $\bar{3}$ joins 4, etc.. Wormholes connecting $1 - 1, \bar{1} - \bar{1}, 2 - 2,$ etc. are present, but not shown. The right panel shows a rearrangement of the diagrams making the purity of (3.2) manifest.

3.1 Entropies

Suppose that one begins with an initial state $|\Psi_i\rangle$ for the combined parent/BU system, which in time T then evolves by the hamiltonian (2.1) to

$$|\Psi, T\rangle = e^{-iHT}|\Psi_i\rangle \,, \tag{3.1}$$

with corresponding density matrix

$$\rho(T) = |\Psi, T\rangle\langle\Psi, T| \,. \tag{3.2}$$

A first simple problem is to compute Renyi entropies of this density matrix. These are given by the standard formula

$$S_n = \frac{1}{1-n} \log \mathrm{Tr}(\rho^n) \,, \tag{3.3}$$

with the trace (over both parent and BU Hilbert spaces) given by

$$\mathrm{Tr}(\rho^n) = \mathrm{Tr}\left(|\Psi, T\rangle\langle\Psi, T| \cdots |\Psi, T\rangle\langle\Psi, T|\right) = 1 \,, \tag{3.4}$$

if states are properly normalized.

This seemingly trivial calculation already carries an important lesson regarding replicas and wormholes. In the replica method [25], each factor of $|\Psi, T\rangle\langle\Psi, T|$ may be represented in terms of a functional integral involving a replica copy of the geometry; the replica parent universes are pictured in figure 2. Then, when we calculate $\mathrm{Tr}(\rho^n)$, that implies a cyclic identification of final time slices of each factor, in the pattern $\bar{1} - 2, \bar{2} - 3, \cdots, \bar{n} - 1$. This *also* applies for the identification of the BUs of figure 1. That is: *The rules of the wormhole calculus, combined with the standard quantum-mechanical rules for calculating the entropy S_n, imply wormhole connections only between neighboring replicas, in the preceding pattern — they do* not *imply that one should sum over geometries with wormhole connections between all replicas, in the way that is commonly conjectured* [9, 10].
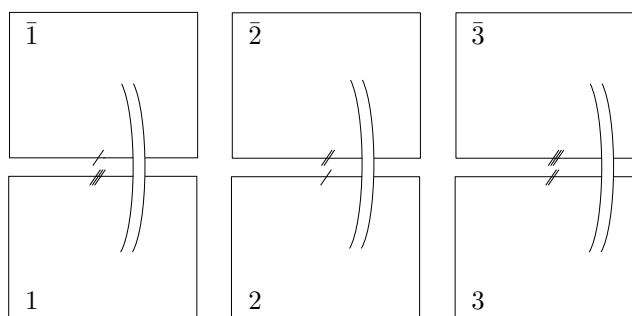
Figure 3. Shown is a sketch of the replica method calculation of the nth Renyi entropy of the reduced density matrix $\rho_p = \mathrm{Tr_{BU}} \, \rho$. Time runs upwards (downwards) in the lower (upper) copies. Here there are no wormholes connecting different replicas, and the connections have the pattern $\bar{1} - 1$, $\bar{2} - 2$, $\cdots$, $\bar{n} - n$. We also indicate how the parent universes are identified. Wormholes connecting $1 - 1$, $\bar{1} - \bar{1}$, $2 - 2$, etc. are present, but not shown.

The wormhole calculus, together with standard quantum mechanical rules, dictate where replicas should be connected by wormholes. Specifically, the wormhole connections follow from the contraction of indices between bra and ket factors, arising from either taking traces, or multiplying density matrices. This principle is expected to generalize to restrict replica topologies in cases where one has more complicated geometries contributing to amplitudes than simple BU emission/absorption. This conclusion does not change if we trace over a subregion of the parent universe (we comment on this below).

As one simple check, we show in figure 2 the pattern above allows us to rearrange the diagrams in a way that makes manifest that $\mathrm{Tr}(\rho^n) = (\mathrm{Tr}\rho)^n$. This implies that ρ is pure, which is consistent with the fact that we started from a pure state in the total Hilbert space (parent plus BUs) and the evolution is unitary.

The preceding principle can be illustrated by a different calculation. Suppose that we instead consider the density matrix of the parent universe,

$$\rho_p = \mathrm{Tr_{BU}}|\Psi, T\rangle\langle\Psi, T| \, , \tag{3.5}$$

and consider its Renyi entropies, given in terms of $\mathrm{Tr}(\rho_p^n)$. The BU trace in (3.5) connects BUs in the bra and ket. Then, when one calculates $\mathrm{Tr}(\rho_p^n)$, the final time slices on the parent universes are identified in the preceding pattern. The BU connections instead form the pattern $\bar{1} - 1$, $\bar{2} - 2$, $\cdots$, $\bar{n} - n$, as illustrated in figure 3, but once again one does not sum over topologies with BUs connecting all replicas. One also sees that, from the perspective of usual quantum mechanical rules, the latter kind of sum would appear rather unusual — that would correspond to contracting various BU indices in a product such as (3.4) between *all* different factors. Figure 3 also shows that due to the way the parent universe degrees of freedom are identified, now $\mathrm{Tr}(\rho_p^n) \neq (\mathrm{Tr}\rho_p)^n$. This implies that ρ_p is not pure, consistent with the QM interpretation.

It is also informative to examine the corresponding expressions written in terms of the α vacua. Consider, for example, an initial uncorrelated (product) state of the BUs and

parent; after evolution by T, (3.1) then gives

$$|\Psi, T\rangle = \int d\alpha \, \psi(\alpha) U_\alpha(T) |\psi_i\rangle |\alpha\rangle \, , \qquad (3.6)$$

where $U_\alpha(T) = \exp\{-iH_\alpha T\}$ is the evolution operator for a given α. Then ρ_p becomes

$$\rho_p = \int d\alpha |\psi(\alpha)|^2 \, U_\alpha(T) |\psi_i\rangle \langle \psi_i| U_\alpha^\dagger(T) \qquad (3.7)$$

and the nth Renyi entropy is given by

$$\mathrm{Tr}(\rho_p^n) = \int \prod_{k=1}^n d\alpha_k |\psi(\alpha_k)|^2 \, \langle \psi_i| U_{\alpha_k}^\dagger(T) U_{\alpha_{k+1}}(T) |\psi_i\rangle \, , \qquad (3.8)$$

where we identify $\alpha_{n+1} = \alpha_1$.

In contrast, a sum over all possible wormhole connections between replicas (as suggested by [9, 10]) would correspond to the expression

$$\mathrm{Tr}(\rho_p^n) = \int d\alpha |\psi(\alpha)|^2 \, \left(\langle \psi_i| U_\alpha^\dagger(T) U_\alpha(T) |\psi_i\rangle \right)^n \, , \qquad (3.9)$$

or, for evolution of an initial parent density matrix $\rho_{p,i}$,

$$\mathrm{Tr}(\rho_p^n) = \int d\alpha |\psi(\alpha)|^2 \mathrm{Tr} \left(U_\alpha(T) \rho_{p,i} U_\alpha^\dagger(T) \right)^n \, . \qquad (3.10)$$

While this behaves like an average over an ensemble of couplings with probability distribution $|\psi(\alpha)|^2$, it does not follow in a straightforward way from combining the rules for summing over topologies in amplitudes with a standard quantum-mechanical calculation.

To reiterate, when the connections involved in the replica trick are of the form described as in (3.4) or (3.8), the expressions can still be interpreted as quantum-mechanical calculations where the parent Hilbert space is enlarged with a baby universe Hilbert space. However, the calculations corresponding to the rule that one should sum over all connections, through (3.10), are not of standard quantum-mechanical form and instead behave as an ensemble average

$$\langle \mathrm{Tr}(\rho^n) \rangle = \int d\alpha \, P(\alpha) \mathrm{Tr}(\rho_\alpha^n), \qquad (3.11)$$

where $P(\alpha) = |\psi(\alpha)|^2$ is the probably distribution in the ensemble and ρ_α indicates, within a member of the ensemble, the state evolved with a definite Hamiltonian H_α.

We can use a similar analysis to treat the entropy calculation for density matrices after tracing over excitations in a region $\mathcal{R}$ inside the parent universe. Specifically, we can compute the Renyi entropy associated to two different types of density matrix,

$$\rho^{\mathcal{R}} = \mathrm{Tr}_{\mathcal{R}} \, \rho \, , \qquad \text{and} \qquad \rho_p^{\mathcal{R}} = \mathrm{Tr}_{\mathcal{R}} \, \rho_p \, . \qquad (3.12)$$

The first option does not include a trace over the BU Hilbert space while the second option does. When describing the wormhole connections coming from the sum over intermediate states, the nth Renyi entropy for the first and second options correspond to the connections
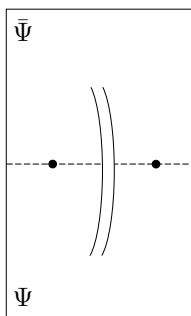
Figure 4. In this diagram we show the in-in calculation of a two point function (represented by the black dots). The bottom (top) represent the time evolution upwards (downwards) creating the bra (ket) at time T along the dashed line, in a diagram like those described in [32]. The state is glued at the dashed line including the operator insertions and the rules of QM would require us to include wormholes between them. We depict one of these wormholes.

shown in figure 2 and 3 respectively. The only difference in the calculation is that now we will also identify the parent universe degrees of freedom corresponding to region $\mathcal{R}$ between the ith and $\bar{i}$th copies, with $i = 1, \ldots, n$. The third option in which we sum over all possible wormholes (which is different from the quantum-mechanical rules we have been describing) is again given by a sum over a single α parameter weighted by $|\psi(\alpha)|^2$, analogous to (3.10).

One further comment is that the results in this section can be also reproduced using the methods of [26]. This is based on the fact that one can replace microscopic wormholes by a bilocal coupling between local operators inserted at their mouths. This interaction can be made local by introducing α parameters, which act as random coupling of local operators. In the cases studied above, when a wormhole is present between replicas, one can check (keeping track of combinatorics and phases) that the effect is to identify their α parameters. In the extreme case described last, where one includes all possible wormholes between any copies, the end result is to identify all α parameters, reproducing equation (3.10).

3.2 Correlators

As another application of these ideas we can analyze the computation of real time in-in correlators when summing over wormholes. We want to compute the expectation value of some operator $\mathcal{O}(\phi_i)$ acting on the parent universe degrees of freedom (but not acting on BU Hilbert space) at some time T. Then the rules of QM applied to this problem would give

$$\begin{aligned}
\langle \mathcal{O}(\phi_i) \rangle &= \langle \Psi, T | \mathcal{O}(\phi_i) | \Psi, T \rangle, \\
&= \int d\alpha |\psi(\alpha)|^2 \langle \psi_i | U_\alpha^\dagger(T) \mathcal{O}(\phi_i) U_\alpha(T) | \psi_i \rangle
\end{aligned} \tag{3.13}$$

where we wrote the initial state as a linear combination of α-states as in the previous section. The sum over intermediate states at time T, including BUs, imposes that the alpha parameters are the same for the bra and the ket, giving the final formula above. Geometrically this comes from including wormholes between the bra and the ket as shown in figure 4 (it is important in deriving this result that the operator does not act on the BU

Hilbert space). Wormholes of a similar type that arise in calculating the density matrix were considered in [33] (and more recently also in [20]).

In this example we see that thanks to the wormhole connecting the bra and ket the norm of the state is preserved under time evolution (if we set $\mathcal{O}(\phi_i) = 1$ then the evolution operators cancel). This would fail had we not included these wormholes, giving instead

$$\int d\alpha d\alpha' \, \psi^*(\alpha)\psi(\alpha') \, \langle\psi_i|U_\alpha^\dagger(T)\mathcal{O}(\phi_i)U_{\alpha'}(T)|\psi_i\rangle \tag{3.14}$$

which does not preserve the norm under time evolution.

Some models of natural inflation are based on a non-perturbative axion potential generated by Euclidean wormholes [28, 34, 35]. The considerations above would be relevant to compute, for example, the power spectrum or non-gaussianities in these models.

4 Determination of wormhole-induced couplings

In [4, 5] it was argued that the growth of entropy that we see from the perspective of a parent universe if we begin in a generic BU state is not a good model for the kind of information loss originally proposed by Hawking to arise from black holes [36]. Specifically, models in [4, 5] showed that this information loss is not *repeatable*: if repeated experiments are performed, the entropy increase per experiment declines as their number increases. This, together with the superselection rule for the α vacua, tell us that the entropy growth is associated with lack of knowledge of the specific value of the eigenvalue α, or effective coupling constant, within the effective ensemble with probability distribution $|\psi(\alpha)|^2$. We can use the preceding discussion to extend this argument, generalizing the argument of [4, 5], and also making contact with the question of replica wormholes.

Specifically, consider the evolution (3.6), in the case where the parent universe wavefunction describes a number s of independent systems, so

$$|\psi_i\rangle = |\tilde{\psi}_i\rangle^{\otimes s} \, . \tag{4.1}$$

Suppose that these evolve as independent noninteracting systems (aside from wormhole connections), in which case (3.6) takes the form

$$|\Psi, T\rangle = \int d\alpha \, \psi(\alpha) \left(\tilde{U}_\alpha(T)|\tilde{\psi}_i\rangle\right)^{\otimes s} |\alpha\rangle \tag{4.2}$$

with independent evolution operators $\tilde{U}_\alpha(T)$, and the parent density matrix becomes

$$\rho_{p,s} = \int d\alpha |\psi(\alpha)|^2 \left(\tilde{U}_\alpha(T)|\tilde{\psi}_i\rangle\right)^{\otimes s} \left(\langle\tilde{\psi}_i|\tilde{U}_\alpha^\dagger(T)\right)^{\otimes s} \, . \tag{4.3}$$

The Renyi entropies are now given by

$$\mathrm{Tr}\rho_{p,s}^n = \int \prod_{k=1}^n d\alpha_k |\psi(\alpha_k)|^2 \left(\langle\tilde{\psi}_i|\tilde{U}_{\alpha_k}^\dagger(T)U_{\alpha_{k+1}}(T)|\tilde{\psi}_i\rangle\right)^s \, . \tag{4.4}$$

– 9 –

At this stage, we again find that the underlying wormhole connections have a pattern like in (3.5) and in figure 3, once again connecting $1 - \bar{1}, \cdots, n - \bar{n}$.

To evaluate the Renyi entropies for large s, note that the inner products in (4.4) can be written

$$\langle \tilde{\psi}_i | \tilde{U}_\alpha^\dagger(T) \tilde{U}_{\alpha'}(T) | \tilde{\psi}_i \rangle = e^{\gamma(\alpha,\alpha') + i\delta(\alpha,\alpha')} , \tag{4.5}$$

with real functions γ and δ; we have $\gamma(\alpha,\alpha) = \delta(\alpha,\alpha) = 0$, $\delta(\alpha',\alpha) = -\delta(\alpha,\alpha')$, and generically $\gamma(\alpha,\alpha') < 0$ for $\alpha \neq \alpha'$. This means that the integrals in (4.4) become increasingly sharply peaked at $\alpha_k = \alpha_{k+1}$ for large s. Near $\alpha = \alpha'$, we have expansions $\gamma(\alpha,\alpha') = -C(\alpha - \alpha')^2 + \cdots$ and $\delta(\alpha,\alpha') = D(\alpha - \alpha') + E(\alpha - \alpha')^3 + \cdots$. Inserting these in (4.4), the D terms cancel, the E terms contribute at subleading order in $1/s$, and we find

$$\mathrm{Tr}\rho_{p,s}^n \approx \int \prod_{k=1}^n d\alpha_k |\psi(\alpha_k)|^2 e^{-sC(\alpha_k - \alpha_{k+1})^2} . \tag{4.6}$$

For a large number s of experiments, the form of the integrals is determined by the $n-1$ gaussian factors with width $\sim 1/\sqrt{s}$ (excluding an overall "center of mass" integral), and so the entropies become

$$\mathrm{Tr}\,\rho_{p,s}^n \approx \left(\frac{1}{\sqrt{s}}\right)^{n-1} F(n), \tag{4.7}$$

for some function $F(n)$ with $F(1) = 1$. The Renyi entropies (3.3) then are

$$S_n(s) \approx \frac{1}{2} \log s + \frac{1}{1-n} \log F(n) , \tag{4.8}$$

and the change of a given Renyi entropy per experiment is

$$\frac{d}{ds} S_n(s) \approx \frac{1}{2s} . \tag{4.9}$$

The fact that the entropy grows slower than linearly with s can be traced back to the expression of the density matrix (4.3), which is neither a pure state nor a product of s density matrices. Finally, the approximation leading to (4.8) assumes the spectrum of α's is continuous. For a discrete spectrum, the entropy $S_n(s)$ would saturate for s larger than the inverse level spacing of the α's.[3]

In summary, there is a "loss of information" in subsequent experiments conducted in the parent universe, which is associated with the lack of information about the state of the BUs. However, since the BU state effectively mimics a probability distribution for coupling constants, successive experiments better and better determine the a-priori uncertain values of these couplings. In the limit of a large number of experiments, $s \to \infty$, the indeterminacy vanishes, and there is no further growth of entropy/loss of information.

---

[3]More precisely, if the spectrum of α's is discrete, then at large s the expression (4.6) would be dominated by $\alpha_k = \alpha_{k-1}$ for all k. This gives an answer independent of s up to corrections of order $e^{-sC\delta\alpha^2}$, where $\delta\alpha$ is the average level spacing. Then the entropy saturates for $s \gg 1/(C\delta\alpha^2)$. This saturation is not seen for a strictly continuous spectrum.

5 Discussion and lessons

As was first shown in [4, 5], the effect of BUs is to contribute to coupling constants multiplying operators that summarize the effect of a given kind of BU on the fields of a parent universe. A generic state of BUs leads to a probability distribution for such couplings. We have shown here, generalizing arguments in [4, 5], that successive experiments lead to an increasingly precise determination of such couplings, such that in the limit of a large number of experiments, additional experiments experience no further loss of information. One can think of this determination process as a "collapse of the wavefunction" into an α state of the BUs corresponding to a particular set of couplings. There is a well-developed set of rules, the wormhole calculus, [4, 5] that were overviewed in section two and underpin this set of observations. There are effects that go beyond the simple free BU approximation used there, and account for interactions between BUs and with other parent universes; an initial account of such effects in a more general third-quantized approach was given in [6], and some such effects were argued to lead to specific distributions effectively fixing couplings such as the cosmological constant in [29].

One may calculate quantities such as entropies and correlators, in the presence of topology change/BUs, and within the framework of the wormhole calculus, using standard quantum-mechanical rules for doing so. In particular, the wormhole calculus may be combined with replica methods [25]. When one does this, the standard quantum mechanical rules applied to the entropies or correlators we consider lead to a limited pattern of wormhole connections between replicas. These for example only produce a connection between replicas that are "nearest neighbors", and do not produce connections between different "bra" copies or "ket" copies.

The work of [9, 10] considers even more general topologies that go outside of these nearest neighbor and bra-ket constraints. An important question is how to justify such connections, based on an underlying consistent set of rules for computing amplitudes including topology change, and following standard quantum rules, e.g. based on tracing over appropriate states, for sewing amplitudes. In fact, given that the replica wormhole configurations of [9, 10] involve even more complicated topological connections between replicas than combinations of single-wormhole connections, their interpretation in terms of traces of appropriate density matrices seems even more obscure. There seem to be at least three different possibilities for explaining a role for such extended rules for replica connections. One is that they correspond to calculating other more general mathematical quantities than the simple entropies one usually considers. A second is that they represent some modification of the usual quantum-mechanical rules for composing amplitudes. A third is that they give an effective parameterization of other effects that are directly described without invoking such extended rules. It does appear, as seen in (3.9), (3.10), that some such expressions can describe certain ensemble averages for BU couplings.

If topological or BU effects do help in resolving the unitarity crisis associated with BH evolution, a key question is to understand the underlying transition amplitudes describing how they do so. In particular, as discussed in section four, one may perform a large number of experiments, after which evolution in our parent universe should be unitary,

with no further loss of information. Once we have "collapsed the BU wavefunction" in this fashion, we can then consider subsequent scattering experiments where BHs form and decay, and those processes should be described by unitary amplitudes. However, at this stage the net effect of the BUs is, at least neglecting higher-order effects (e.g. as in [6]), simply to contribute to various coupling constants. In the resulting effective theory, we can then ask how BH formation and decay evades the standard information-loss arguments going back to Hawking's original work [37].

It has been argued that what is needed for such a unitary description are interactions that, when viewed from the perspective of an effective field theory description of BH evolution, transfer information or entanglement from the internal state of the BH to the environment of the BH [38–43]. In particular, refs. [13, 14] give a parameterization of such interactions in an effective theory. One possibility is that the topology-changing processes somehow contribute to such interactions, which appear to be nonlocal from the effective field theory perspective. We have seen that simple connections, via a small spacetime wormhole, between two different points do not induce the right kind of nonlocal transfer of information, but possibly contributions of larger wormholes, say comparable to the BH size, could, as has been suggested in [8, 9] and [7]. If this were the case, such interactions could likely be parameterized in the general framework of [13, 14]. However, in order to justify such a picture, and even more importantly, to give a precise description of such interactions that allows one to calculate the effects on outgoing fields (and on possibly observable quantities), one needs a description of how the topology-changing processes contribute to amplitudes. This might, for example, involve finding instanton-like or other similar processes operating on scales comparable to a BH's size. A preliminary investigation reveals a number of subtleties in giving any systematic description of such effects [44], but it is worth determining whether progress can be made in this direction.

Acknowledgments

A   Puzzles for replica wormholes

In [9, 10], it has specifically been argued that replica wormholes make important contributions to the entropy, such that it follows the Page curve corresponding to unitary evolution. It is important to better understand these arguments, and in particular the question of how they might relate to underlying unitary quantum amplitudes describing black hole formation and evaporation. Since the main body of the paper has shown how amplitudes including wormholes may be combined into density matrices, and then entropies may be computed from density matrices, following basic quantum-mechanical rules for composing amplitudes and taking traces, a specific question is whether one can "reverse engineer" the
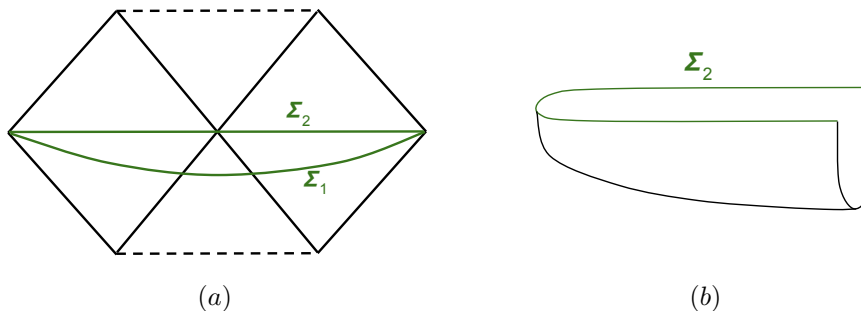
Figure 5. (a) Penrose diagram of a Schwarschild black hole. We indicate two Cauchy slices in the two-sided spacetime, Σ_1 and Σ_2. (b) Euclidean no-boundary preparation of the two-sided state on the surface Σ_2.

geometries of [9, 10] to infer the structure of underlying quantum amplitudes. In particular, this can be asked in the context of the specific rules we have found for replica geometries. One new ingredient of this discussion is that it goes beyond the free approximation. We find some interesting puzzles regarding possible connection to underlying amplitudes.

As a warmup, first consider amplitudes corresponding to states of the two-sided black hole, as shown in figure 5; for simplicity we focus on a Schwarzschild black hole. We can for example describe a state on the Cauchy surface Σ_2 by evolving from another surface Σ_1 at a previous time. This state is a functional $\Psi_{\Sigma_2}(g, \phi, \ldots)$, of the metric g on Σ_2, with ϕ denoting possible matter fields. A particularly simple state is prepared on Σ_2 through the Hartle-Hawking prescription, as pictured in figure 5(b). If the latter is viewed as a contribution to the sum over geometries, it can be thought of describing production of a two-sided black hole from "nothing". A more clearly motivated version of this geometry is when the two-sided black hole is magnetically charged; instantons that describe the Schwinger pair production of such black holes in a background magnetic field have been described in [45–47], and have near-horizon structure of precisely the same form (see, e.g., figure 1 of [47]).

Having prepared a state by one of these methods, we can calculate the entropy of a subregion $\mathcal{R}$ outside the black hole, with corresponding density matrix $\rho_{\mathcal{R}} = \mathrm{Tr}_{\mathcal{R}_c} |\Psi\rangle\langle\Psi|$, where $\mathcal{R}_c$ is the complement of the region $\mathcal{R}$ and contains the black hole. The growth of entropy with time in this region follows the original "Hawking" curve even past the Page time, when the calculation is done in the spacetimes shown in figure 5. This density matrix can be represented by the diagram figure 6(a), where we used the Lorenzian preparation of the state shown in figure 5(a). The Renyi entropy calculation is described by figure 6(b) with the identifications shown, sewing the boundaries along the identified copies of $\mathcal{R}$.

Bearing these preliminary examples in mind, we would like to understand the interpretation of the replica wormholes of [9, 10] which were argued to correct the Hawking curve and produce the Page curve. An important contribution was from the "pinwheel geometry," which for the third Renyi entropy is drawn in figure 7(a). Since we have seen how black hole geometries may be sewn to calculate entropies, the question is how this diagram may be unstitched to describe underlying amplitudes.
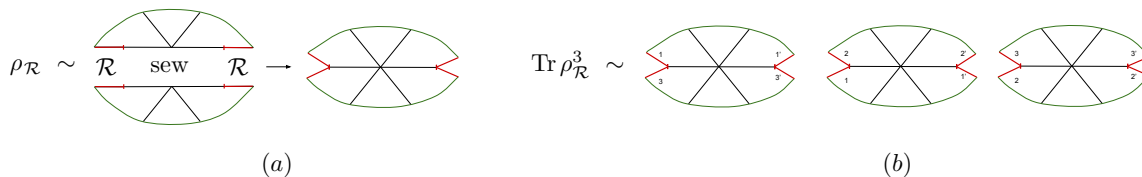
Figure 6. (a) Using the Lorenzian preparation of the state (region between Σ_1 — green curve — and Σ_2 — black and red line) on slice Σ_2 this picture denotes the operation that computes the partial trace of the density matrix $\rho_{\mathcal{R}}$ (where $\mathcal{R} \subset \Sigma_2$ are the segments in red) after tracing over the complementary region $\mathcal{R}_c$. (b) Computation of the third Renyi entropy. The states on the red segments corresponding to the copies of $\mathcal{R}$ are identified as indicated by the numbers.
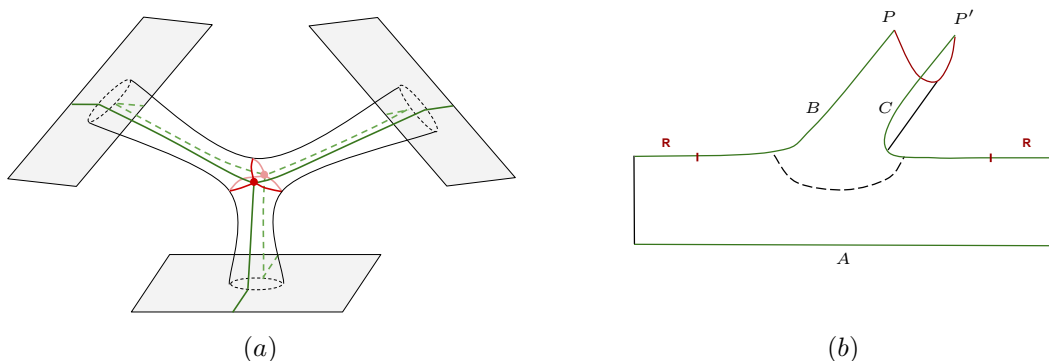

Figure 7. (a) Pinwheel amplitude for three replicas. The green lines are taken to be constant time slices. The replicas are connected through a branch cut along $\mathcal{R}$ that we do not draw to avoid clutter. (b) Building block. Constant time slice is the parent universe (green slice); the possible interpretation of the red segment is discussed in the text. We also indicate the region $\mathcal{R}$ where radiation is collected.

There are two ways to try to interpret the pinwheel geometry, and both seem to be incompatible with a quantum mechanical treatment of a parent plus baby universe Hilbert space. One approach is to view this geometry as arising from a baby universe interaction. In this interpretation each replica creates its own baby universe and the central part of the picture represents their interaction vertex. But, interactions between all replicas simultaneously are not consistent with the rules developed in section 3, even outside the free approximation.

A second possibility arises from cutting the pinwheel along the green lines in figure 7(a). This is analogous to unstitching the previous figures, figure 6(a) and (b), along the corresponding black lines, to return to the underlying amplitudes. If we do this, for each replica, then the pinwheel can be understood as gluing six copies of the building block shown in figure 7(b).

The upper green boundary in figure 7(b) is identified with the corresponding bra amplitude, i.e. $\bar{1} - 1$, $\bar{2} - 2$, etc., inside the region $\mathcal{R}_c$. The red boundary is identified instead with a different replica. Note that, if one is summing over all geometries, this identification can be to a *different* replica than the corresponding replica identifications on

region $\mathcal{R}$, analogous to those of figure 6(b). This could be made more consistent with the rules given in section 3 if the connections are given between consecutive replicas.

Even with this assumption, puzzles remain. Specifically, the interpretation of the underlying geometry of figure 7(b), is unclear; it is not obvious what lorentzian amplitude this contributes to. First, ignoring the red segment, it appears that this describes a transition from a two-sided black hole, with spatial geometry labeled A (bottom green line), to two separate spatial components B and C (top separate green segments). That leaves the question of interpreting the red segment, which is glued to a different replica. If it is not interpreted as part of the geometry of the final slice, the latter geometry remains disconnected. If, instead, it is interpreted as part of that geometry, it is not clear how to understand these gluing conditions. Indeed, the points P, P' at the junction between this segment and the components B and C are common to *all* the replicas, as seen in figure 7(a). It is not clear how such a gluing prescription is compatible with the rules of section 3.

We believe it is important to understand what unitary amplitudes, if any, underly the formal entropy calculations of [9, 10], if the underlying calculational rules can be understood without modifying quantum mechanics. The present work outlines a naïve attempt to interpret these, which leads to puzzles. We leave other efforts to understand these to future work.

Open Access.

References

[1] G.V. Lavrelashvili, V.A. Rubakov and P.G. Tinyakov, *Disruption of Quantum Coherence upon a Change in Spatial Topology in Quantum Gravity*, *JETP Lett.* **46** (1987) 167 [INSPIRE].

[2] S.W. Hawking, *Quantum Coherence Down the Wormhole*, *Phys. Lett. B* **195** (1987) 337 [INSPIRE].

[3] S.B. Giddings and A. Strominger, *Axion Induced Topology Change in Quantum Gravity and String Theory*, *Nucl. Phys. B* **306** (1988) 890 [INSPIRE].

[4] S.R. Coleman, *Black Holes as Red Herrings: Topological Fluctuations and the Loss of Quantum Coherence*, *Nucl. Phys. B* **307** (1988) 867 [INSPIRE].

[5] S.B. Giddings and A. Strominger, *Loss of Incoherence and Determination of Coupling Constants in Quantum Gravity*, *Nucl. Phys. B* **307** (1988) 854 [INSPIRE].

[6] S.B. Giddings and A. Strominger, *Baby Universes, Third Quantization and the Cosmological Constant*, *Nucl. Phys. B* **321** (1989) 481 [INSPIRE].

[7] D. Marolf and H. Maxfield, *Transcending the ensemble: baby universes, spacetime wormholes, and the order and disorder of black hole information*, *JHEP* **08** (2020) 044 [arXiv:2002.08950] [INSPIRE].

[8] P. Saad, S.H. Shenker and D. Stanford, *JT gravity as a matrix integral*, arXiv:1903.11115 [INSPIRE].

[9] G. Penington, S.H. Shenker, D. Stanford and Z. Yang, *Replica wormholes and the black hole interior*, arXiv:1911.11977 [INSPIRE].

[10] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, *Replica Wormholes and the Entropy of Hawking Radiation*, *JHEP* **05** (2020) 013 [arXiv:1911.12333] [INSPIRE].

[11] J. Polchinski and A. Strominger, *A Possible resolution of the black hole information puzzle*, *Phys. Rev. D* **50** (1994) 7403 [hep-th/9407008] [INSPIRE].

[12] A. Strominger, *Les Houches lectures on black holes*, in *NATO Advanced Study Institute: Les Houches Summer School, Session 62: Fluctuating Geometries in Statistical Mechanics and Field Theory*, (1994) [hep-th/9501071] [INSPIRE].

[13] S.B. Giddings, *Nonviolent unitarization: basic postulates to soft quantum structure of black holes*, *JHEP* **12** (2017) 047 [arXiv:1701.08765] [INSPIRE].

[14] S.B. Giddings, *Black holes in the quantum universe*, *Phil. Trans. Roy. Soc. Lond. A* **377** (2019) 20190029 [arXiv:1905.08807] [INSPIRE].

[15] D.N. Page, *Average entropy of a subsystem*, *Phys. Rev. Lett.* **71** (1993) 1291 [gr-qc/9305007] [INSPIRE].

[16] D.N. Page, *Information in black hole radiation*, *Phys. Rev. Lett.* **71** (1993) 3743 [hep-th/9306083] [INSPIRE].

[17] G. Penington, *Entanglement Wedge Reconstruction and the Information Paradox*, *JHEP* **09** (2020) 002 [arXiv:1905.08255] [INSPIRE].

[18] A. Almheiri, N. Engelhardt, D. Marolf and H. Maxfield, *The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole*, *JHEP* **12** (2019) 063 [arXiv:1905.08762] [INSPIRE].

[19] A. Almheiri, R. Mahajan, J. Maldacena and Y. Zhao, *The Page curve of Hawking radiation from semiclassical geometry*, *JHEP* **03** (2020) 149 [arXiv:1908.10996] [INSPIRE].

[20] J. Maldacena, G.J. Turiaci and Z. Yang, *Two dimensional Nearly de Sitter gravity*, arXiv:1904.01911 [INSPIRE].

[21] J. Cotler, K. Jensen and A. Maloney, *Low-dimensional de Sitter quantum gravity*, *JHEP* **06** (2020) 048 [arXiv:1905.03780] [INSPIRE].

[22] D. Stanford and E. Witten, *JT Gravity and the Ensembles of Random Matrix Theory*, arXiv:1907.03363 [INSPIRE].

[23] P. Saad, *Late Time Correlation Functions, Baby Universes, and ETH in JT Gravity*, arXiv:1910.10311 [INSPIRE].

[24] A. Blommaert, T.G. Mertens and H. Verschelde, *Eigenbranes in Jackiw-Teitelboim gravity*, arXiv:1911.11603 [INSPIRE].

[25] C.G. Callan Jr. and F. Wilczek, *On geometric entropy*, *Phys. Lett. B* **333** (1994) 55 [hep-th/9401072] [INSPIRE].

[26] I.R. Klebanov, L. Susskind and T. Banks, *Wormholes and the Cosmological Constant*, *Nucl. Phys. B* **317** (1989) 665 [INSPIRE].

[27] V.A. Rubakov, *On the Third Quantization and the Cosmological Constant*, *Phys. Lett. B* **214** (1988) 503 [INSPIRE].

[28] S.-J. Rey, *The Axion Dynamics in Wormhole Background*, *Phys. Rev. D* **39** (1989) 3185 [INSPIRE].

[29] S.R. Coleman, *Why There Is Nothing Rather Than Something: A Theory of the Cosmological Constant*, *Nucl. Phys. B* **310** (1988) 643 [INSPIRE].

[30] J.M. Maldacena and L. Maoz, *Wormholes in AdS*, *JHEP* **02** (2004) 053 [hep-th/0401024] [INSPIRE].

[31] N. Arkani-Hamed, J. Orgera and J. Polchinski, *Euclidean wormholes in string theory*, *JHEP* **12** (2007) 018 [arXiv:0705.2768] [INSPIRE].

[32] S.B. Giddings and M.S. Sloth, *Cosmological diagrammatic rules*, *JCAP* **07** (2010) 015 [arXiv:1005.3287] [INSPIRE].

[33] D.N. Page, *Density Matrix of the Universe*, *Phys. Rev. D* **34** (1986) 2267 [INSPIRE].

[34] K. Freese, J.A. Frieman and A.V. Olinto, *Natural inflation with pseudo-Nambu-Goldstone bosons*, *Phys. Rev. Lett.* **65** (1990) 3233 [INSPIRE].

[35] A. Hebecker, P. Mangat, S. Theisen and L.T. Witkowski, *Can Gravitational Instantons Really Constrain Axion Inflation?*, *JHEP* **02** (2017) 097 [arXiv:1607.06814] [INSPIRE].

[36] S.W. Hawking, *Breakdown of Predictability in Gravitational Collapse*, *Phys. Rev. D* **14** (1976) 2460 [INSPIRE].

[37] S.W. Hawking, *Particle Creation by Black Holes*, *Commun. Math. Phys.* **43** (1975) 199 [*Erratum ibid.* **46** (1976) 206] [INSPIRE].

[38] S.B. Giddings, *Nonlocality versus complementarity: A Conservative approach to the information problem*, *Class. Quant. Grav.* **28** (2011) 025002 [arXiv:0911.3395] [INSPIRE].

[39] S.B. Giddings, *Models for unitary black hole disintegration*, *Phys. Rev. D* **85** (2012) 044038 [arXiv:1108.2015] [INSPIRE].

[40] S.B. Giddings, *Black holes, quantum information, and unitary evolution*, *Phys. Rev. D* **85** (2012) 124063 [arXiv:1201.1037] [INSPIRE].

[41] S.B. Giddings and Y. Shi, *Quantum information transfer and models for black hole mechanics*, *Phys. Rev. D* **87** (2013) 064031 [arXiv:1205.4732] [INSPIRE].

[42] S.B. Giddings, *Nonviolent information transfer from black holes: A field theory parametrization*, *Phys. Rev. D* **88** (2013) 024018 [arXiv:1302.2613] [INSPIRE].

[43] S.B. Giddings, *Nonviolent nonlocality*, *Phys. Rev. D* **88** (2013) 064023 [arXiv:1211.7070] [INSPIRE].

[44] S.B. Giddings and G.J. Turiaci, work in progress.

[45] G. Gibbons, *Quantized flux tubes in Einstein-Maxwell theory and noncompact internal spaces*, in *22nd Winter School of Theoretical Physics: Fields and Geometry*, pp. 597–615 (1986) [INSPIRE].

[46] D. Garfinkle and A. Strominger, *Semiclassical Wheeler wormhole production*, *Phys. Lett. B* **256** (1991) 146 [INSPIRE].

[47] D. Garfinkle, S.B. Giddings and A. Strominger, *Entropy in black hole pair production*, *Phys. Rev. D* **49** (1994) 958 [gr-qc/9306023] [INSPIRE].