Diffeomorphism invariance and the black hole information paradox

Ted Jacobson and Phuc Nguyen

Diffeomorphism invariance and the black hole information paradox

Ted Jacobson and Phuc Nguyen

Maryland Center for Fundamental Physics, University of Maryland, College Park, MD 20742

We argue that the resolution to the black hole information paradox lies in a proper accounting of the implications of diffeomorphism invariance for the Hilbert space and observables of quantum gravity. The setting of asymptotically Anti-de Sitter spacetime is adopted for most of the paper, but in the framework of canonical quantum gravity, without invoking AdS/CFT duality. We present Marolf's argument that boundary unitarity is a consequence of diffeomorphism invariance, and show that its failure to apply in the classical limit results from a lack of analyticity that has no quantum counterpart. We argue that boundary unitarity leads to a *boundary information paradox*, which generalizes the black hole information paradox and arises in virtually any scattering process. We propose a resolution that involves operators of the boundary algebra that redundantly encode information about physics in the bulk, and explain why such redundancy need not violate the algebraic no cloning theorem. We also argue that the infaller paradox, which has motivated the firewall hypothesis for black hole horizons, is ill-posed in quantum gravity, because it ignores essential aspects of the nature of the Hilbert space and observables in quantum gravity.

I. INTRODUCTION

The black hole information paradox holds promise for teaching us something fundamental about quantum gravity. The paradox concerns how to reconcile well-established principles of local quantum field theory with expectations from quantum gravity. We currently lack direct experimental evidence of quantum gravity, so the paradox provides a welcome challenge to our understanding of the foundations of quantum gravity, in particular of the nature of its states and observables. To be sure, despite many decades of research and discussion, it is not universally agreed that any paradox exists in the first place [1–3]. In this paper we review compelling arguments that it does exist, and we argue that it can be resolved by paying close attention to the role of diffeomorphism invariance in the theory. Our conclusion is not that local QFT or quantum gravity has some unexpected breakdown, but that we just need to better understand the quantum mechanical framework of quantum gravity.

The information paradox has many faces. The one we focus on in this paper is raised by Marolf's argument for *boundary unitarity* [4]. This argument is founded on the assumption that the diffeomorphism invariance of classical general relativity extends to quantum gravity. The argument is agnostic about the form taken by the UV completion of the theory, as long as it preserves the classical property—which follows from diffeomorphism invariance—that the Hamiltonian is a boundary term. In that case, the algebra of boundary observables includes the Hamiltonian. We consider the case of asymptotically AdS spacetime, so that the boundary is time-like. Then, the fact that the Hamiltonian is a boundary observable arguably implies that the boundary algebra evolves continuously into itself. If that is so, then information available at the boundary at one time is also available at any other time. Note that, although we focus on the example of asymptotic AdS boundary conditions, we are *not* invoking AdS/CFT duality, or any particular UV completion of the the quantum gravity theory (although it may be that such a duality is inevitably what a UV completion looks like).

One of our aims in this paper is to support this argument

that the boundary algebra evolves into itself, and to explain why the argument fails to apply in the classical case. This will hinge on a property of analyticity that we will establish. Our other aim is to resolve what we will call the "boundary information paradox" (BIP). This paradox is the apparent contradiction between unitary evolution of boundary observables, and the presence of bulk degrees of freedom, some of which propagate to the boundary while being entangled with others that remain in the bulk. A key point we shall argue is that the BIP arises in generic processes, not only ones involving black holes. It is an information paradox, but not a black hole information paradox. We regard the BIP as a training ground for understanding the Hilbert space and observables of quantum gravity.

We postpone to the end of the paper a consideration of the "infaller paradox", wherein local bulk physics apparently exhibits violation of the monogamy of entanglement in the neighborhood of a black hole horizon. Having developed some understanding of the Hilbert space and observables of quantum gravity, we will argue that the infaller paradox is actually ill-posed, so that it is not clear whether there is any such paradox at all.

We do *not* make use of reduced phase space quantization, nor do we make use of AdS/CFT duality, other than as an example. Let us explain why not. The bulk/boundary dichotomy at the center of our considerations would be obscured if we were to adopt the framework of reduced phase space quantization, where all gauge freedom is eliminated before quantization. Instead, we adopt the framework of Dirac quantization, in which the physical Hilbert space is the subspace of a kinematic Hilbert space that is annihilated by the gauge constraint operators, and in which physical (i.e. gauge invariant) observables commute with the constraints. In the reduced phase space approach, the question of boundary unitarity just doesn't even come up. We elect to consider the Dirac quantization, because it poses puzzles that are hopefully instructive concerning the nature of quantum gravity. The question whether or not Dirac and reduced phase space quantization are equivalent remains open [5–8].

In the setting of AdS/CFT duality, the usual presumption is that the algebra of boundary observables is *identical* to the

algebra of CFT observables. Moreover, the completeness of the boundary algebra might be a generic feature of quantum gravity [9], at least in a superselection sector [10]. If that is so, then, in a sense, the boundary information paradox is resolved simply by the fact that all information always resides at the boundary. However, our aim is to learn something about the Hilbert space and observables of quantum gravity, by confronting the the paradox from the bulk viewpoint: the physical Hilbert space is, according to Dirac quantization, a subspace of a kinematic Hilbert space, which itself is, as in nongravitational quantum field theory, roughly a tensor product of local factors. The boundary observables act on the boundary kinematic factor. To resolve the BIP, then, one must understand how it is that certain "bulk observables" can *also* correspond to boundary observables,[1] without violating quantum mechanics. To this end, it does not help to simply invoke completeness of the CFT observables. It should also be emphasized that, even if bulk physics is only "emergent", emergent bulk diffeomorphism invariance may nevertheless be essential to resolving the information paradox. After all, the paradox is concerned with the emergent bulk physics.

The rest of the paper is structured as follows. In Sec. II the argument for BU is recalled, and we identify analyticity of time evolution as a key distinction between the quantum and classical cases. Sec. III poses the BIP, and Sec. IV posits our proposed resolution of the paradox, which involves the notion of the "nebulon," which is our name for the nebulous operators of the boundary algebra that redundantly encode information about physics in the bulk. The purpose of Sec. V is to explore more specific descriptions of the nebulon, both nonperturbative and perturbative, as well as to answer an objection from Ref. [11] to the proposal that "nebulous bulk degrees of freedom" could resolve the black hole information paradox. In Sec. VI we explain how it could be that the nebulon avatar mechanism is not precluded by the (algebraic) no-cloning theorem. In Sec. VII we recall the infaller paradox, which is special to black holes, and we argue that as stated heretofore it is ill-posed in quantum gravity, because it ignores essential aspects of the nature of the Hilbert space and observables in quantum gravity. Finally, we close in Sec. VIII with a few remarks.

II. BOUNDARY UNITARITY

The Hamiltonian of a classical theory with diffeomorphism symmetry, such as general relativity, is a linear combination of constraints, plus a boundary term if the spatial manifold Σ has a boundary $\partial\Sigma$ [12],

$$H = \int_\Sigma N_\mu C^\mu + H_\partial \,. \tag{1}$$

Here N_μ are the lapse and shift functions, C^μ are the constraints, and H_∂ is the boundary term. If N_μ approaches an asymptotic time translation at the boundary, H generates that time translation via Hamilton's equations. The constraints arise because of bulk diffeomorphism symmetry.[2] The Poisson brackets of the constraints close on the constraints (the constraints are "first class"), hence the constraints generate "gauge" transformations corresponding to diffeomorphisms. Any gauge-invariant observable $\mathcal{O}$ Poisson-commutes with the constraints $\{\mathcal{O}, C^\mu\} = 0$ (up to terms that vanish with the constraints), so the evolution of observables with respect to asymptotic time-translations is governed entirely by the boundary term H_∂:

$$\frac{d\mathcal{O}}{dt} = \{\mathcal{O}, H\} = \{\mathcal{O}, H_\partial\}. \tag{2}$$

So far this is all classical physics.

In the quantized version of the foregoing structure, the canonical variables become operators, the Poisson bracket relations become commutator relations, and *presumably*, if the diffeomorphism symmetry survives quantization, the algebra of the constraints still closes on the constraints. Of course it is well known that the quantization of general relativity cannot by itself make sense, without some "UV completion" of the theory. But if the diffeomorphism symmetry survives this UV completion it is plausible that the Hamiltonian is still a boundary term, and that gauge invariant observables evolve by commutation with this Hamiltonian, according to the equation

$$\frac{d\mathcal{O}}{dt} = \frac{1}{i\hbar}[\mathcal{O}, H_\partial], \tag{3}$$

using the Heisenberg picture. In Sec. IV we discuss the possibility that in the UV completion of the gravity theory the Hamiltonian is no longer localized at the boundary, and we argue that even if so, that in itself likely does not eliminate the BIP.

A. Boundary observables and unitarity

Boundary unitarity refers to the time evolution of boundary observables, by which we mean (gauge invariant, self-adjoint) observables that can be built from the fields in the intersection of any neighborhood of a bulk Cauchy slice with any neighborhood of the boundary. We refer to the algebra generated by these as the *boundary algebra* $\mathcal{A}_\partial$ associated with the given (asymptotic) Cauchy slice. The Hamiltonian—at least at the level of EFT—is in the boundary algebra, as is the commutator of the Hamiltonian with anything in the algebra. The time derivative (3) is therefore in $\mathcal{A}_\partial$, which leads to the expectation that $\mathcal{A}_\partial$ evolves unitarily into itself. This property was termed "boundary unitarity" (BU) in [4]. BU implies that the information contained in boundary observables is time independent.

---

[1] And how, perhaps, boundary observables could be complete.

[2] Throughout this paper "diffeomorphism" refers to bulk diffeomorphism. Diffeomorphisms that act nontrivially at the boundary must preserve the boundary conditions, and that nontrivial action gives rise to symmetry rather than gauge transformations.

Note that we are *not* assuming that the boundary algebra includes all observables, although that might be the case.[3] Moreover, if there is more than one boundary of the space-time, then there is more than one boundary algebra, so the observables at one boundary are clearly not complete, but this does not change anything essential in our considerations. There would be one term in the Hamiltonian for each boundary, and the algebra of observables at any one boundary would evolve into itself.

But before getting carried away with the implications of BU and the paradox it raises, we must examine more closely the justification for the expectation that $\mathcal{A}_\partial$ evolves unitarily. Something more than the argument already given is clearly required, because the corresponding *classical* statement is *not* true [4, 13]. There are solutions of the Einstein equation that agree *exactly*, for example, with the Schwarzschild-AdS solution outside some radius on some time slice, and which nevertheless have disturbances which propagate out to the boundary at a later time [14] (see also [15–17]). The arrival of such disturbances would clearly constitute new information in the boundary observables, so would violate the classical version of BU. In fact, the quantum case has a key property that the classical case lacks: analyticity of time evolution.

B. Classical case

The classical evolution equation (2) is closely analogous to the quantum one (3), and the reasoning we applied above for the latter applies to the former: the rate of change of an observable is another element of the boundary algebra. However, as just noted, that must not in general be sufficient to infer that this property is shared by the time-dependent solution to this equation. That solution can be written formally, as

$$\mathcal{O}(t) = e^{t\{\cdot,H\}}\mathcal{O}(0), \qquad (4)$$

where $\{\cdot, H\}$ denotes the operation of taking the Poisson bracket with the function to its right, placed in the first entry of the bracket. Expanding the exponential, (4) becomes an infinite series of nested Poisson brackets. The question, then, is whether this series converges, for all t, to the actual solution to (2), on each integral curve of the Hamiltonian vector field on phase space. The answer is evidently no, not in general: not all solutions are analytic in t with infinite radius of convergence, so in general the series does not converge to the solution for all t. Thus, although it *formally* appears from (2) that the values of the boundary observables at one time determines the values at any other time, that fails to be the case.

Let us illustrate this concretely using a massless scalar field $\phi(x,t)$ in $1+1$ spacetime dimensions. The Hamiltonian is $H = \int dx[\frac{1}{2}\pi^2 + \frac{1}{2}(\partial_x\phi)^2]$, the conjugate momentum is $\pi = \partial_t\phi$, and for the observable we take $\mathcal{O} = \phi(x_o)$, for some fixed x_o. Consider now a right-moving solution: $\phi(x,t) = f(x-t)$, for some function f. This solution is determined by the initial data $\phi(x,0) = f(x)$ and $\pi(x,0) = -\partial_x f(x)$. For n nested commutators, $\{\ldots\{\{\phi(x),H\},H\}\ldots,H\}$ is equal to $\partial_x^{n-1}\pi(x)$ for odd n, and $\partial_x^n\phi(x)$ for even n. Using the initial condition we thus find, on this particular trajectory,

$$e^{t\{\cdot,H\}}\phi(x_o,0) = \sum_{n=0}^{\infty}\frac{1}{n!}(-t)^n\partial_x^n f(x_o), \qquad (5)$$

which is the Taylor series for $f(x_o-t)$. If $f(x)$ is not analytic with infinite radius of convergence, then for some t the series will fail to agree with $\phi(x_o,t)$. For example, suppose $f(x)$ is a bump function, $\exp[-1/(x^2-1)]$ for $|x| < 1$ and 0 for $|x| \geq 1$, and choose $x_o = 2$. Then the series converges to zero for all t, whereas $\phi(2,t)$ is nonzero for $1 < t < 3$. Alternatively, if we choose $x_o = 0$, then the series converges to the correct solution only for $|t| < 1$.

C. Quantum case

The classical analysis suggests that, if BU is to hold, then there must be something about quantum mechanics that enforces analyticity, or something like it, for the time dependence of the boundary observables. This is indeed the case. The solution to (3) takes the form

$$\mathcal{O}(t) = e^{itH/\hbar}\mathcal{O}(0)e^{-itH/\hbar}. \qquad (6)$$

We assume that H is self-adjoint, so that the exponential $e^{-itH/\hbar}$ is well-defined in terms of the spectral projections of H [18].[4] Moreover, the exponential is a bounded operator. To ensure that the operator on the right hand side of (6) is defined everywhere, we assume that also $\mathcal{O}$ is bounded. Then (6) is indeed the solution to (3) wherever the latter is defined.[5] And, since it is a product of three operators each of which lies in $\mathcal{A}_\partial$, it too lies in $\mathcal{A}_\partial$.

While the preceding argument for BU is logically sufficient, we can give a more constructive account by considering expectation values $\langle\psi|\mathcal{O}(t)|\psi\rangle$, and restricting to energy-bounded observables. This will also exhibit clearly the difference between the classical and quantum cases.

By inserting complete sets of H eigenstates on either side of $\mathcal{O}(0)$, we obtain using (6) an expression for the expectation value as a double sum,

$$\langle\psi|\mathcal{O}(t)|\psi\rangle = \sum_{m,n} e^{i(\omega_m-\omega_n)t}\mathcal{O}_{mn}\langle\psi|m\rangle\langle n|\psi\rangle, \qquad (7)$$

---

[3] It was recently argued in Ref. [9] that in canonical quantum gravity, if a Reeh-Schlieder property holds for boundary operators, and if the projection on the vacuum lies in $\mathcal{A}_\partial$, then indeed $\mathcal{A}_\partial$ would coincide with the full algebra of observables. Also, in the setting of AdS/CFT duality, it is tempting to identify the CFT observables, which are complete for the theory, with the boundary observables.

[4] Of course we do not know whether or not H is in fact self-adjoint in (UV completed) quantum gravity. Here we assume that it is, because our main interest is in exploring whether BU can be reconciled with local quantum field theory, rather than in establishing BU itself. We note as well that, if AdS/CFT duality is valid, H is identified with the Hamiltonian of the CFT, which is presumably self-adjoint.

[5] And (6) replaces (3) wherever the rhs of (3) is *not* well defined, as is usual in quantum mechanics.

where $H|n\rangle = \hbar\omega_n|n\rangle$ and $\mathcal{O}_{mn} = \langle m|\mathcal{O}(0)|n\rangle$. If the spectrum of H includes a continuum, then the summation over the eigenstates includes integration. Since the exponentials are analytic functions of t, the expectation value $\langle\psi|\mathcal{O}(t)|\psi\rangle$ is a sum of analytic functions. Were there a finite number of terms in the sum, it too would be analytic, but there are in general an infinite number of terms, so we must be more careful.

If, as we shall assume, the Hamiltonian is bounded below, the range of the frequencies summed over in (7) is bounded below. To ensure analyticity, we can cut off the energy at some upper bound Λ, and consider the energy-projected observable $\mathcal{O}_\Lambda := P_\Lambda \mathcal{O} P_\Lambda$. Since the Hamiltonian is in $\mathcal{A}_\partial$, so is P_Λ, because P_Λ is a function of the Hamiltonian,

$$P_\Lambda = \Theta(\Lambda - H), \tag{8}$$

where Θ is the Heaviside step function and Λ is the energy cutoff. Hence the projected observable $\mathcal{O}_\Lambda$ is also in $\mathcal{A}_\partial$, provided that $\mathcal{O}$ is. Moreover, the expectation value $\langle\psi|\mathcal{O}_\Lambda(t)|\psi\rangle$ is analytic in t. To see this, just note that it is the diagonal value of a double Fourier transform, $F(t, t') = \int d\omega \int \delta\omega' e^{-i\omega t} e^{-i\omega' t'} f(\omega, \omega')$, where $f(\omega, \omega')$ has compact support in each argument. Such a function is analytic in each argument, and hence the function $G(t) := F(t, t)$ is analytic in t.

For each energy cutoff Λ we thus have an algebra of boundary observables $\mathcal{A}_\partial^\Lambda$ whose expectation values in any state evolve analytically. It follows that the collection of all time derivatives of the expectation values at one time determines the values at any other time. This is a concrete formulation of boundary unitarity. The statement holds for any cutoff Λ, hence for each Λ there is a BU property. Moreover, we can take Λ arbitrarily large, and so come arbitrarily close to encompassing all boundary observable values for finite energy states. *It is the unitarity of time evolution in quantum mechanics that leads to the analyticity property noted here.* In contrast, the classical argument for boundary unitarity fails precisely on account of the lack of analyticity in time.

It is worth noting that the fact that the energy-projected observable $P_\Lambda \mathcal{O} P_\Lambda$ is in the boundary algebra is possible only in the presence of gravity. In local, relativistic field theory without gravity, by contrast, an energy-projected version $P_\Lambda A P_\Lambda$ of a local observable A is *never* localized to any open proper subregion U of spacetime. To see this, suppose $P_\Lambda A P_\Lambda$ were localized to some region U, and choose any observable B localized in the causal complement U^c of U. Microcausality then implies

$$(P_\Lambda A P_\Lambda)B|\psi\rangle = \pm B(P_\Lambda A P_\Lambda)|\psi\rangle, \tag{9}$$

where $|\psi\rangle$ is any state, and the minus sign applies when both of the operators are fermionic. Equation (9) implies in particular that, for all B localized in U^c, the state $B(P_\Lambda A P_\Lambda)|\psi\rangle$ has energy content bounded by Λ. This contradicts the generalized Reeh-Schlieder theorem which asserts that, by acting on a state of finite energy content with operators localized in U^c, we can approximate *any* other state arbitrarily well [19, 20], including ones with energy content greater than Λ.

III. BOUNDARY INFORMATION PARADOX

The fact that bulk particles can propagate to the boundary raises a puzzle for BU [21]: how can the boundary algebra evolve unitarily into itself, when "new information" can arrive at the boundary? When such a particle arrives at the boundary and influences the boundary observables, it appears that the boundary algebra is not evolving autonomously. We shall call this puzzle the *boundary information paradox* (BIP). An example of it arises when a black hole emits Hawking radiation which is correlated with the field behind the horizon, yet which reaches the boundary. However, the BIP arises even when no black hole is present. If a bulk particle propagates to the boundary after having been dynamically correlated with bulk degrees of freedom, this puzzle is like the Hawking radiation one. The entropy of the boundary algebra seems to increase. Moreover, there is even a puzzle without such correlation, because the arrival of a bulk particle at the boundary seems to introduce new information there. The BIP focuses attention on an aspect of the black hole information paradox that deserves more attention than it has received, namely, the puzzle of *continuous unitary evolution of boundary observables*.

Consider, for example, a scattering type process in which particles are injected from the boundary and interact in the bulk, forming a state that emits a massless particle while leaving behind a long lived resonance that is correlated to the early emitted particle. The arrival of the decay particle at the boundary can be detected by the measurement of some boundary observable $\mathcal{O}$, which is then correlated with the remaining resonance. This appearance of correlation between a boundary observable and a non-boundary one violates the fact that the boundary algebra evolves unitarily into itself.

This scenario is rather similar to one in which the long lived resonance is a black hole. The key difference is that, in the case of the black hole, the early Hawking radiation is correlated not just to the remaining resonance, a.k.a. the black hole, but specifically to partner field modes behind the horizon of the black hole. Also, the infinite redshift at the horizon implicates a mixing of bulk UV and IR degrees of freedom in the process. However, from the viewpoint of boundary unitarity, these distinctions are immaterial.

We speak of dynamical correlation, and avoid the concept of "entanglement," because the latter refers to a state in a Hilbert space composed of two or more factors. In the setting of quantum gravity, however, a tensor factorization according to spatial localization is not available [22–25]. Moreover, the particles or field quanta referred to in the previous paragraph are not gauge invariant with respect to diffeomorphisms; and, if they are gravitationally dressed in order to become so, then they are no longer spatially localized. But we have framed the BIP so as not to presume spatial factorization of the Hilbert space or any particular gravitational dressing. The algebra of boundary observables is defined without invoking a boundary factor of the quantum gravity Hilbert space.

Despite the absence of a factorized Hilbert space, given an algebra of quantum observables one can define the *algebraic entropy* of a global quantum state restricted to the subalgebra

[26, 27], which generalizes the notion of the von Neumann entropy of a state restricted to a tensor factor. The BIP can be phrased in terms of the behavior of this algebraic entropy: If the Hamiltonian is an element of the subalgebra, then the algebraic entropy of the subalgebra cannot change in time. BU thus implies that the entropy of the boundary algebra never changes, and yet the arrival of bulk particles at the boundary appears to entail entropy change.

IV. RESOLVING THE BIP

We can see two possible routes to resolving the BIP:

1. The Hamiltonian is *not* in $\mathcal{A}_\partial$.

2. $\mathcal{A}_\partial$ contains an avatar of any bulk degree of freedom with which it can be correlated.

The argument for BU, reviewed in Section II, is that, on account of diffeomorphism invariance, the Hamiltonian *must* be in $\mathcal{A}_\partial$. However, the boundary algebra was defined using a notion of locality that makes sense within field theory but may cease to be meaningful in the UV completion of the gravitational theory. For instance, strict localization of the Hamiltonian to any neighborhood of the boundary may not be possible in string theory, because of the extended nature of strings. In this case, the Hamiltonian would couple EFT boundary observables to non-boundary observables, so observables initially in $\mathcal{A}_\partial$ would evolve to observables not contained in $\mathcal{A}_\partial$, and the BU argument as it stands would fail.[6]

That said, this failure would not necessarily provide a resolution to the BIP, because the observables involved in the BIP may be restricted to the low energy effective field theory sector, and it is not at all clear that a tiny amount of dynamical mixing with observables outside of $\mathcal{A}_\partial$ would suffice to account for the non-conservation of information in $\mathcal{A}_\partial$. Note also that the asymptotic redshift appears to help suppress such mixing. A finite Killing energy, defined with respect an asymptotic global time translation Killing vector ∂_t, corresponds to a vanishing *proper* energy in the frame defined by the unit vector $\sim r^{-1}\partial_t$, where r is the asymptotic area radius coordinate. Hence there is not enough energy to create "stringy" excitations near the boundary.

For the remainder of this paper, we will focus on the second route, namely, the possibility that $\mathcal{A}_\partial$ already contains an avatar— i.e. an image, copy, or representative—of any component of the bulk degrees of freedom needed to preserve the information available at the boundary. Related approaches to solving the black hole information paradox have been discussed before, for example in the form of black hole complementarity and "$A = R_B$" [28–31]. These have been critiqued

as non-viable, generally for reasons related to the "no-cloning theorem" of quantum mechanics, or contradictions resulting from the presence of the same quantum information in more than one location. It seems to us that, if formulated carefully, no contradictions or violations of no cloning need arise. In the next section we articulate a little more explicitly the nature of our proposed resolution, and in the following section we confront the no-cloning issue.

V. THE NEBULON

The notion that such a an avatar must be available at the boundary to purify Hawking radiation was proposed in [21], in which the continuous unitarity feature of the BIP was emphasized. This proposal was criticized in [11], which argued that "nebulous bulk degrees of freedom" could play no role in evading a version of the black hole information paradox involving an auxiliary system that extracts and converts Hawking quanta into excitations of a separate system attached to spacetime. (That criticism will be addressed below.) Inspired by that denigration, we are moved to embrace the term, and here refer to the required avatar as the *nebulon*.

A. Who is the nebulon?

Since it is diff invariance that leads to the BIP in the first place, we expect that diff invariance should lie at the root its resolution. The impact of diff invariance on the structure of the quantum gravity state space is indeed very strong. Classically, the Hamiltonian constraint builds the nonlocal aspects of gravity, like the link between sources and geometry in the bulk and multipole moments measurable at infinity, into the structure of the phase space. At the quantum level, in the setting of Dirac quantization, the Wheeler-DeWitt equation presumably does the same job, and more: it encodes also any evolution with respect to internal clocks. So a likely candidate for the nebulon mechanism is the Wheeler-DeWitt equation.

Another key aspect of the nebulon may be simply the time evolution of the boundary observables. For example, a bulk particle arriving at the boundary appears to bring new information. But, if (thanks to diff invariance) BU indeed holds, that information must have already been available at the boundary before the arrival of the particle. Long ago, in the setting of AdS/CFT, the boundary state that anticipates the arrival of a particle was referred to as a *precursor* [32–34], the precursor being a CFT state. It is easy, in a sense, to describe the precursor observable: we can "simply" time evolve the boundary particle detection observable backwards in time, to express it as a boundary observable at any earlier time. At the earlier time, it is presumably a highly "scrambled" looking observable, which would be difficult to recognize as the precursor of a simple particle detection, and yet BU guarantees its existence in $\mathcal{A}_\partial$. Although the above considerations suggest that the nebulon may exist already at the EFT level, it may be that the nebulon mechanism involves the UV completion in an essential manner.

---

[6] In the AdS/CFT context, it is plausible that there is no nontrivial subalgebra of the CFT that is invariant under time evolution. (We thank Hong Liu for a discussion on this point.) If this is the case, then either $\mathcal{A}_\partial$ must be equivalent to the entire CFT algebra, or the Hamiltonian must not lie in $\mathcal{A}_\partial$.

B. Perturbative nebulon?

The nebulon has already appeared, and not so nebulously, in the context of AdS/CFT, at the level of effective field theory, at zeroth order in Newton's constant, in the guise of expressions for bulk operators in terms of boundary operators. We have in mind for example the well known HKLL [35] reconstruction of a bulk scalar field, which uses the scalar field equation of motion to express a bulk field in terms of its boundary values in the region of the boundary that is spacelike related to the bulk point. Another example is the less well known BBPR reconstruction [36, 37], which uses the Reeh-Schlieder property [19] that arises from the strong vacuum entanglement, together with the fact that the Hamiltonian—and therefore the projection onto the ground state—is in the boundary algebra. Although both of those examples were studied in the context of AdS/CFT, and mostly with pure AdS as a background, they do not, as far as we can see, require anything beyond the general setting of asymptotically AdS quantum gravity.

As mentioned, these "bulk reconstruction" approaches work, initially, without accounting for gravitational coupling, and without enforcing diffeomorphism invariance. Perturbative corrections at lowest order in $O(1/N)$, where N is the central charge of the CFT presumed dual to the quantum gravity theory, were constructed by Kabat and Lifshitz (in a Fefferman-Graham gauge) in [38]. To the extent that any quantum gravity theory is dual to a CFT, this can presumably be viewed as a gravitational perturbation expansion in $(L_{\text{Planck}}/L_{\text{AdS}})^{D-2}$ where D is the bulk spacetime dimension.

In another approach [4], Marolf argued (without invoking AdS/CFT) that, using the bulk equations of motion, bulk operators can be expressed in terms of boundary operators lying to the past of the bulk point. Moreover he argued that the dependence on the boundary algebra can be squashed down to the infinitesimal neighborhood of a single Cauchy slice, using the fact that the Hamiltonian, and therefore the time evolution operator, lies in the boundary algebra. He inferred from this that, to all orders in perturbation theory, all bulk operators are represented in the boundary algebra at any time.[7] We view this line of reasoning (which can also be applied to the HKLL reconstruction), as well as the BBPR reconstruction [36, 37], as strongly supporting the nebulon hypothesis. However, since the boundary Hamiltonian (when expressed in terms of the canonically normalized metric perturbation) is inversely proportional to the gravitational coupling κ, the operation of squashing to a single time slice, and the projection onto the ground state, are non-perturbative in the coupling, so that the realization of the nebulon on a single time slice may be essentially nonperturbative.

It should be further noted that the perturbative analysis of

"gravitational splitting" in [39, 40] casts doubt on the existence of a perturbative nebulon.[8] It was shown there that, to first order in κ, inequivalent states can be indistinguishable outside some compact set U. These states are constructed in [39] by acting on a "split vacuum" with operators that are dressed so as to be gauge invariant to first order in κ, and in such a way that observables localized outside of U are sensitive to only the Poincaré charges of these states. This seems to imply that the algebra of operators outside U, and in particular the boundary algebra, does not contain observables that can distinguish these different states, and hence that the boundary algebra must *not* be complete. This conclusion is in tension with BU, and with what is known about bulk reconstruction in AdS/CFT, but there is is no contradiction: the splitting analysis was perturbative on one time slice, whereas the boundary time evolution that plays a role in BU and bulk reconstruction, or translation of bulk operators to the boundary [24], is nonperturbative.

A caveat should be attached to any perturbative discussion of this sort: the notion of a bulk operator satisfying bulk microcausality, i.e. commuting with operators at spacelike related points, apparently cannot be defined on the entire Hilbert space of the theory [27, 41]. Instead, it was argued in [41] that local bulk observables are defined only on a semiclassical subspace of the full quantum gravity theory Hilbert space. This subspace, was called the *code subspace*, because it appears that the encoding of bulk quantum information in the boundary algebra is redundant, in the manner of a quantum error-correcting code (QEC). In particular, the HKLL construction can be carried out restricting to a single "causal wedge" of AdS (and this can be generalized to a reconstruction in an entanglement wedge of a deformed spacetime [42]). Since a given bulk point can lie in many, overlapping causal wedges, one obtains many inequivalent boundary representatives for the same bulk operator. Be that as it may, this redundant encoding aspect of holography presents no problem for BU.

C. AUX vs. Nebulon

Ref. [11], " An Apologia for Firewalls," gave an argument involving an auxiliary system, AUX, purporting to demonstrate that purification of Hawking radiation by nebulous degrees of freedom could not solve the BIP, since it would be inconsistent with unitary evolution. The AUX system interacts with asymptotically AdS spacetime at the boundary, coupling to it only via the boundary value of a bulk scalar field, i.e. AUX couples only to a single trace, scalar primary operator in the CFT dual to the gravity theory. Via this interaction, a scalar Hawking quantum can be absorbed, its energy being transferred to an infinitely large AUX reservoir from which it will never return. If then a large black hole is formed from a pure state, it can evaporate completely, depositing all of the initial energy in the AUX reservoir. Since the entire evolution

---

[7] He allowed, however, that beyond perturbation theory it might not be the case that the boundary algebra is complete in this sense. But such incompleteness, were it to be the case, would not be incompatible with BU, which refers only to the unitary evolution of the boundary algebra, be it compete or not.

[8] We thank Steve Giddings for calling our attention to this issue.

is unitary, the entropy of the AUX state must vanish.[9]

The argument in [11] was that since AUX couples only to the scalar Hawking quanta, and not to any nebulous degrees of freedom that might have purified the Hawking quanta, the AUX quanta must inherit the entropy of the Hawking quanta, and therefore AUX can not wind up in a pure state, unless the Hawking quanta taken together, by themselves, are in a pure state. The Apologia draws the conclusion that there must be a firewall at the black hole horizon that precludes the entanglement of Hawking quanta with behind-horizon partners. However, this conclusion does not follow from the assumptions. As the black hole shrinks, and the temperature decreases, the scalar field quanta interact with the gravitational field, and hence, if only indirectly, with everything. Any entanglement between scalar and nebulon degrees of freedom will be channeled into the extracted scalar quanta as the system cools. That is, the entanglement between the scalar quanta and the nebulons will be extracted into AUX through the scalar channel. A simple analogy makes this transparent: Consider a gas mixture of helium and neon in a zero entropy state that looks like a high temperature equilibrium state, and imagine cooling the gas by coupling an auxiliary refrigerant AUX, able to absorb kinetic energy of only the neon atoms at the walls of the container. Despite the entanglement of the neon with the helium atoms, and despite the fact that AUX couples only to the neon, it is obvious that once the atoms have cooled to zero temperature (and assuming the ground state has zero entropy), AUX winds up in a pure state.

VI. NO-CLONING CONSTRAINTS

The nebulon avatar mechanism for resolving the BIP posits that the same quantum information can "simultaneously" take more than one form in a single global quantum state. For example, a particle can be correlated with two different subsystems, using either of which the same correlation information can be measured. The idea behind this is that the Wheeler-DeWitt equation may enforce redundant encoding of the quantum information. This mechanism was proposed in [21], and illustrated with an example involving four spins. However, that example involves only abelian observable algebras, which are inadequate for quantum state tomography (and for describing the physics at hand), and linear algebra imposes very strong constraints on the structure of redundant encoding for nonabelian algebras. Such constraints are captured by the *algebraic no-cloning theorem* (ANC) [41]. In this section we review the ANC, and then explain how the nebulon avatar mechanism might evade it.

A. The algebraic no-cloning theorem

In the framework of Dirac constrained quantization, the physical Hilbert space $\mathcal{H}_{phys}$ is the subspace of a "kinematical" Hilbert space $\mathcal{H}_{kin}$ that is annihilated by the diffeomorphism constraints,

$$\mathcal{H}_{phys} = \left\{ |\psi\rangle \in \mathcal{H}_{kin} \middle| C|\psi\rangle = 0 \right\}, \qquad (10)$$

where C represents the constraints, which generate gauge transformations. Now suppose $\mathcal{H}_{kin}$ is composed of two tensor factors:

$$\mathcal{H}_{kin} = \mathcal{H}_A \otimes \mathcal{H}_B. \qquad (11)$$

In the gravity setting, we have in mind that one factor corresponds to the kinematic degrees of freedom in the near boundary region, and the other corresponds to the degrees of freedom deeper in the bulk. But at the moment we are just concerned with the algebraic structure.

We are interested in an algebra $\mathcal{A}$ of physical operators, i.e. operators that commute with the constraints C and therefore map the space of physical states to itself. The ANC states:

If $\mathcal{A}$ is measurable on both tensor factors, then it can only be an abelian algebra.

What it means for $\mathcal{A}$ to be measurable on the tensor factor $\mathcal{H}_A$ is that, for every operator $\mathcal{O} \in \mathcal{A}$, there exists an operator $\mathcal{O}_A$ acting on $\mathcal{H}_A$ such that

$$\mathcal{O}|\psi\rangle = (\mathcal{O}_A \otimes I_B)|\psi\rangle \qquad (12)$$

for all physical states $|\psi\rangle \in \mathcal{H}_{phys}$.[10] A similar definition applies to measurability in the other tensor factor, $\mathcal{H}_B$. The proof of the ANC is trivial (as is that of the classic no-cloning theorem of quantum mechanics [43]): Consider any two operators in the algebra, and represent one of them on one tensor factor as in (12), and the other one on the other tensor factor. The two representatives commute, simply because they act on different tensor factors. By hypothesis, the action of the representatives on the physical states is the same as that of the original operators, so this implies that all of the operators in $\mathcal{A}$ must commute when acting on $\mathcal{H}_{phys}$.

B. How gravity bypasses the ANC

The ANC appears to present an insurmountable obstruction to our proposed resolution of However, the situation in quantum gravity does not conform to all of the assumptions of the the ANC theorem. To identify the wrong assumption, we consider the algebra of Poincaré charges (ADM energy,

---

[9] In the scenario of Ref. [11] the black hole was only allowed to evaporate down to another stable black hole, much smaller than the initial one. For rhetorical simplicity, here we allow the black hole to evaporate completely.

[10] This is equivalent to the requirement that the expectation values of $\mathcal{O}$ and $\mathcal{O}_A \otimes I_B$ agree for all physical states.

momentum etc.) of asymptotically flat spacetime. This algebra is nonabelian, but every operator in it can be expressed both as a bulk integral and as a surface integral. It would thus seem that the Poincaré charges are measurable in both the near-boundary tensor factor and the bulk tensor factor. Indeed, this fact was recently exploited by Donnelly [44], in the linearized theory, to implement quantum state tomography of single-particle bulk states via boundary measurements.

How is the above observation about the Poincaré charges to be reconciled with the ANC? We suppose that the reconciliation must be that the "second copy" of the observables, i.e. the bulk version, is not actually fully contained on a separate (kinematic) tensor factor, because the bulk integral must actually go all the way to infinity, since otherwise it is not even gauge invariant. Similarly, in the case of the nebulon, we suppose it is because of gravitational dressing required to turn the resonance into a gauge-invariant observable. It has been proved by Donnelly and Giddings [23], using a perturbation expansion in Newton's constant,[11] that for any operator with nonzero Poincaré charge such gravitational dressing always extends to infinity.[12] Put slightly differently, this dressing theorem—which has also been extended to asymptotically AdS gravity [24]—implies that the division of kinematic Hilbert space into a near-boundary tensor factor and bulk tensor factor does not pass to the level of the algebra of gauge invariant observables. The theorem asserts that there do not exist any gauge-invariant observables (with nonzero asymptotic charges) acting on the bulk tensor factor alone, since all such observables extend to the boundary.

VII. INFALLER PARADOX

We turn in this section to a brief consideration and critique of the infaller paradox [11, 45]. The infaller paradox is based on the supposition that black hole evaporation is a unitary process, and the argument goes as follows. After a black hole has radiated more than half of the Hawking radiation (i.e. past the *Page time*), each subsequent Hawking quantum q must be essentially maximally entangled with the earlier radiation. Moreover, Alice, who is hanging around outside the black hole, can distill from the earlier radiation a special qubit s that has extremely high probability of being maximally entangled with a particular subsequent Hawking quantum q. With this special qubit s in hand, Alice can choose to measure either the joint state of sq or the state of $\tilde{q}q$, where $\tilde{q}$ is the partner behind the horizon with which q is strongly entangled. This measurement takes place in a semiclassical region of spacetime, and in a setting where gravitational effects are

---

[11] The expansion is really in the ratio of a relevant energy to the Planck energy. As such, the result applies to the observables acting on a class of low energy states.

[12] Poincaré charges are given by surface integrals at infinity, so that if an operator does not extend to infinity, it will commute with all the Poincaré charges and hence, since the Poincaré group has no abelian invariant subgroup, the operator cannot carry any Poincaré charge.

negligible, and quantum gravity subtleties can be ignored, so ordinary local quantum field theory should correctly describe the results of Alice's measurements. Subtleties of diffeomorphism invariance, dressed operators, and the quantum gravity Hilbert space, seem to be irrelevant for this measurement part of the argument. But the double entanglement of q with both s and $\tilde{q}$ is impossible in quantum mechanics. So, assuming Alice can indeed possess the special qubit s, and barring a breakdown in local, low energy effective field theory (such as that explored in [46]), one must conclude that q cannot after all be strongly entangled with $\tilde{q}$. But such entanglement is inevitable in the regular vacuum, so it follows that there must be a "firewall" rather than a smooth vacuum at the horizon.

The weak link in this infaller paradox reasoning is the argument that Alice can possess the special qubit s. This link goes far beyond the assumption that local quantum field theory should hold good in a local region of spacetime where Alice's measurements of sq or $\tilde{q}q$ take place. Indeed, it relies on treating the Hilbert space of black hole plus radiation like an ordinary quantum system, without gravity. The Hilbert space is taken to have the form $\mathcal{H}_{\rm bh} \otimes \mathcal{H}_{\rm rad}$, and the radiation degrees of freedom are further factorized into the early radiation an the late radiation, $\mathcal{H}_{\rm rad} = \mathcal{H}_{\rm rad,e} \otimes \mathcal{H}_{\rm rad,l}$. Invoking Page's theorem [47, 48], the argument asserts that, past the Page time (and assuming a randomness in the structure of the state), each of the later Hawking quanta in $\mathcal{H}_{\rm rad,l}$ must be essentially maximally entangled with the earlier ones in $\mathcal{H}_{\rm rad,e}$. Further, the argument supposes that Alice knows precisely the initial state that formed the black hole, and knows precisely the Hamiltonian of the system, so that in fact she knows precisely *how* q is entangled with $\mathcal{H}_{\rm rad,e}$, and therefore she can perform a unitary operation on the early radiation to distill a special qubit s that she knows must be maximally entangled with q.

Two factors related to diffeomorphism invariance render this argument dubious in the gravitating case. First, as discussed above, the Hilbert space is not factorized in the way that has been assumed. This means not only that the quantum states cannot be characterized as they are in the formulation of the paradox, but also the very assumption that the Hawking radiation must be self purifying, with no involvement of nebulons, is not justified. Second, the protocol for distillation on the early radiation must be diff invariant; in particular, the identification of the spacetime location of quanta acted on, and how they are acted on, must be diff invariant. But diff invariant observables (with nonzero asymptotic charges) require dressing that reaches to the boundary [23, 24]. This means that the distillation protocol implicates both local and global gravitational degrees of freedom other than the ideal QFT Hawking quanta to which the original argument applies, and it calls into question whether Alice could possibly carry out the distillation, even in principle. Moreover, the Hawking quantum with which a particular special distilled qubit is supposed to be perfectly entangled must also be identified in a gauge invariant fashion via some gauge invariant observable. Unlike for Alice's measurements themselves on the Hawking quanta, it is incorrect to argue that these quantum gravitational effects are irrelevant.

In short, the effects of gravity and diffeomorphism invariance turn the distillation protocol of Alice from a quasilocal to a global procedure, and although the effects that make it global may each be perturbatively small in the gravitational coupling, the protocol requires her to keep track of a huge amount of information, to very high precision. It seems perfectly possible that the global aspect of her task cannot be neglected, and that the upshot is that she simply cannot possess a special qubit that would violate local QFT near the horizon. That is, the delicate task of distilling the special qubit s and identifying the particular Hawking qubit q with which it is supposed to be maximally entangled may be not only exceedingly challenging technologically, but may simply be impossible in principle, on account of the gravitational physics.

VIII. CLOSING REMARKS

The main points we would like to emphasize are the following: 1) The boundary unitarity argument of Marolf appears sound, and the fact that it does not apply in the classical limit is understood to result from a lack of analyticity that has no quantum counterpart. 2) The black hole information paradox is just a special case of a more general boundary information paradox. 3) The resolution of the paradox(es) does not require a breakdown of known physics such as a firewall at the horizon [11] or a breakdown of local quantum field theory as in non-violent nonlocality [46], but rather just a better accounting for the impact of diffeomorphism invariance on the global structure of the Hilbert space and observables in quantum gravity.

The role of diffeomorphism invariance has been almost completely ignored in the voluminous literature on the information paradox. This is rather ironic given that, as made clear by the boundary unitarity argument, it is precisely diffeomorphism invariance that leads to the paradox in the first place. We have argued that, as in the case of other famous paradoxes of physics (the twin paradox, the Gibbs paradox, the EPR paradox...), the resolution is found *not* by forbidding the paradoxical scenario, nor by modifying the theory that led to the paradox—rather it is found by more deeply learning the lessons of the theory as to what is *observable*, and what is a *meaningful statement about the world*.

ACKNOWLEDGEMENTS

[1] William G. Unruh and Robert M. Wald, "On evolution laws taking pure states to mixed states in quantum field theory," Phys. Rev. **D52**, 2176–2182 (1995), arXiv:hep-th/9503024 [hep-th].

[2] William G. Unruh and Robert M. Wald, "Information Loss," Rept. Prog. Phys. **80**, 092002 (2017), arXiv:1703.02140 [hep-th].

[3] David Wallace, "Why Black Hole Information Loss is Paradoxical," (2017), arXiv:1710.03783 [gr-qc].

[4] Donald Marolf, "Unitarity and Holography in Gravitational Physics," Phys. Rev. **D79**, 044010 (2009), arXiv:0808.2842 [gr-qc].

[5] A. Ashtekar and G. t. Horowitz, "On the Canonical Approach to Quantum Gravity," Phys. Rev. **D26**, 3342–3353 (1982).

[6] Steven Carlip, "Six ways to quantize (2+1)-dimensional gravity," in *5th Canadian Conference on General Relativity and Relativistic Astrophysics (5CCGRRA) Waterloo, Canada, May 13-15, 1993* (1993) pp. 0215–234, arXiv:gr-qc/9305020 [gr-qc].

[7] Augustin Vanrietvelde, Philipp A. Hoehn, Flaminia Giacomini, and Esteban Castro-Ruiz, "A change of perspective: switching quantum reference frames via a perspective-neutral framework," (2018), arXiv:1809.00556 [quant-ph].

[8] Philipp A. Hoehn, "Switching internal times and a new perspective on the 'wave function of the universe'," Universe **5**, 116 (2019), arXiv:1811.00611 [gr-qc].

[9] Suvrat Raju, "Is Holography Implicit in Canonical Gravity?" (2019), arXiv:1903.11073 [gr-qc].

[10] Donald Marolf and Aron C. Wall, "Eternal Black Holes and Superselection in AdS/CFT," Class. Quant. Grav. **30**, 025001 (2013), arXiv:1210.3590 [hep-th].

[11] Ahmed Almheiri, Donald Marolf, Joseph Polchinski, Douglas Stanford, and James Sully, "An Apologia for Firewalls," JHEP **09**, 018 (2013), arXiv:1304.6483 [hep-th].

[12] Tullio Regge and Claudio Teitelboim, "Role of Surface Integrals in the Hamiltonian Formulation of General Relativity," Annals Phys. **88**, 286 (1974).

[13] R.M. Wald, personal communication.

[14] James Isenberg, David Maxwell, and Daniel Pollack, "A Gluing construction for non-vacuum solutions of the Einstein constraint equations," Adv. Theor. Math. Phys. **9**, 129–172 (2005), arXiv:gr-qc/0501083 [gr-qc].

[15] Justin Corvino, "Scalar Curvature Deformation and a Gluing Construction for the Einstein Constraint Equations," Communications in Mathematical Physics **214**, 137–189 (2000).

[16] Piotr T. Chrusciel and Erwann Delay, "On mapping properties of the general relativistic constraints operator in weighted function spaces, with applications," Mem. Soc. Math. France **94**, 1–103 (2003), arXiv:gr-qc/0301073 [gr-qc].

[17] Justin Corvino and Richard M. Schoen, "On the asymptotics for the vacuum Einstein constraint equations," J. Diff. Geom. **73**, 185–217 (2006), arXiv:gr-qc/0301071 [gr-qc].

[18] M. Reed and B. Simon, *I: Functional Analysis*, Methods of Modern Mathematical Physics (Elsevier Science, 1981).

[19] Stephen J. Summers, "Yet More Ado About Nothing: The Remarkable Relativistic Vacuum State," in *Symposium on Deep Beauty: Mathematical Innovations and the Search for an Underlying Intelligibility of the Quantum World Princeton, New Jersey, October 3-4, 2007* (2008) arXiv:0802.1854 [math-ph].

[20] Edward Witten, "APS Medal for Exceptional Achievement in Research: Invited article on entanglement properties of quantum field theory," Rev. Mod. Phys. **90**, 045003 (2018),

arXiv:1803.04993 [hep-th].

[21] Ted Jacobson, "Boundary unitarity and the black hole information paradox," Int. J. Mod. Phys. **D22**, 1342002 (2013), arXiv:1212.6944 [hep-th].

[22] William Donnelly and Steven B. Giddings, "Diffeomorphism-invariant observables and their nonlocal algebra," Phys. Rev. **D93**, 024030 (2016), [Erratum: Phys. Rev.D94,no.2,029903(2016)], arXiv:1507.07921 [hep-th].

[23] William Donnelly and Steven B. Giddings, "Observables, gravitational dressing, and obstructions to locality and subsystems," Phys. Rev. **D94**, 104038 (2016), arXiv:1607.01025 [hep-th].

[24] Steven B. Giddings and Alex Kinsella, "Gauge-invariant observables, gravitational dressings, and holography in AdS," JHEP **11**, 074 (2018), arXiv:1802.01602 [hep-th].

[25] Horacio Casini, Marina Huerta, and Jose Alejandro Rosabal, "Remarks on entanglement entropy for gauge fields," Phys. Rev. **D89**, 085012 (2014), arXiv:1312.1183 [hep-th].

[26] M. Ohya and D. Petz, *Quantum Entropy and Its Use*, Theoretical and Mathematical Physics (Springer Berlin Heidelberg, 2004).

[27] Daniel Harlow, "The RyuTakayanagi Formula from Quantum Error Correction," Commun. Math. Phys. **354**, 865–912 (2017), arXiv:1607.03901 [hep-th].

[28] Leonard Susskind, Larus Thorlacius, and John Uglum, "The Stretched horizon and black hole complementarity," Phys. Rev. **D48**, 3743–3761 (1993), arXiv:hep-th/9306069 [hep-th].

[29] Leonard Susskind and Larus Thorlacius, "Gedanken experiments involving black holes," Phys. Rev. **D49**, 966–974 (1994), arXiv:hep-th/9308100 [hep-th].

[30] David A. Lowe, Joseph Polchinski, Leonard Susskind, Larus Thorlacius, and John Uglum, "Black hole complementarity versus locality," Phys. Rev. **D52**, 6997–7010 (1995), arXiv:hep-th/9506138 [hep-th].

[31] Gerard 't Hooft, "On the Quantum Structure of a Black Hole," Nucl. Phys. **B256**, 727–745 (1985).

[32] Joseph Polchinski, Leonard Susskind, and Nicolaos Toumbas, "Negative energy, superluminosity and holography," Phys. Rev. **D60**, 084006 (1999), arXiv:hep-th/9903228 [hep-th].

[33] Leonard Susskind and Nicolaos Toumbas, "Wilson loops as precursors," Phys. Rev. **D61**, 044001 (2000), arXiv:hep-th/9909013 [hep-th].

[34] Ben Freivogel, Steven B. Giddings, and Matthew Lippert, "Toward a theory of precursors," Phys. Rev. **D66**, 106002 (2002), arXiv:hep-th/0207083 [hep-th].

[35] Alex Hamilton, Daniel N. Kabat, Gilad Lifschytz, and David A. Lowe, "Holographic representation of local bulk operators," Phys. Rev. **D74**, 066009 (2006), arXiv:hep-th/0606141 [hep-th].

[36] Souvik Banerjee, Jan-Willem Bryan, Kyriakos Papadodimas, and Suvrat Raju, "A toy model of black hole complementarity," JHEP **05**, 004 (2016), arXiv:1603.02812 [hep-th].

[37] Suvrat Raju, "A Toy Model of the Information Paradox in Empty Space," (2018), arXiv:1809.10154 [hep-th].

[38] Daniel Kabat and Gilad Lifschytz, "Decoding the hologram: Scalar fields interacting with gravity," Phys. Rev. **D89**, 066010 (2014), arXiv:1311.3020 [hep-th].

[39] William Donnelly and Steven B. Giddings, "Gravitational splitting at first order: Quantum information localization in gravity," Phys. Rev. **D98**, 086006 (2018), arXiv:1805.11095 [hep-th].

[40] Steven B. Giddings, "Gravitational dressing, soft charges, and perturbative gravitational splitting," (2019), arXiv:1903.06160 [hep-th].

[41] Ahmed Almheiri, Xi Dong, and Daniel Harlow, "Bulk Locality and Quantum Error Correction in AdS/CFT," JHEP **04**, 163 (2015), arXiv:1411.7041 [hep-th].

[42] Xi Dong, Daniel Harlow, and Aron C. Wall, "Reconstruction of Bulk Operators within the Entanglement Wedge in Gauge-Gravity Duality," Phys. Rev. Lett. **117**, 021601 (2016), arXiv:1601.05416 [hep-th].

[43] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," Nature **299**, 802–803 (1982).

[44] William Donnelly, "Quantum gravity tomography," (2018), arXiv:1806.05643 [hep-th].

[45] Raphael Bousso, "Complementarity Is Not Enough," Phys. Rev. **D87**, 124023 (2013), arXiv:1207.5192 [hep-th].

[46] Steven B. Giddings, "Nonviolent unitarization: basic postulates to soft quantum structure of black holes," JHEP **12**, 047 (2017), arXiv:1701.08765 [hep-th].

[47] Don N. Page, "Average entropy of a subsystem," Phys. Rev. Lett. **71**, 1291–1294 (1993), arXiv:gr-qc/9305007 [gr-qc].

[48] Daniel Harlow, "Jerusalem Lectures on Black Holes and Quantum Information," Rev. Mod. Phys. **88**, 015002 (2016), arXiv:1409.1231 [hep-th].