# Holography and unitarity

**Steven B. Giddings**

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

*E-mail:* giddings@ucsb.edu

ABSTRACT: If holography is an equivalence between quantum theories, one might expect it to be described by a map that is a bijective isometry between bulk and boundary Hilbert spaces, preserving the hamiltonian and symmetries. Holography has been believed to be a property of gravitational (or string) theories, but not of non-gravitational theories; specifically Marolf has argued that it originates from the gauge symmetries and constraints of gravity. These observations suggest study of the assumed holographic map as a function of the gravitational coupling $G$. The zero coupling limit gives ordinary quantum field theory, and is therefore not necessarily expected to be holographic. This, and the structure of gravity at non-zero $G$, raises important questions about the full map. In particular, construction of a holographic map appears to require as input a solution of the nonperturbative analog of the bulk gravitational constraints, that is, the unitary bulk evolution. Moreover, examination of the candidate boundary algebra, including the boundary hamiltonian, reveals commutators that don't close in the usual fashion expected for a boundary theory.

## Contents

## 1 Introduction

Holography, and in particular the AdS/CFT correspondence [1], has become a dominant theme in quantum gravity, but nonetheless its precise formulation and explanation remains controversial. The purpose of this paper is to more carefully examine possible properties of a holographic correspondence, and in particular to investigate the role of gravitational effects.

Gravity has been argued to play an essential role in holography [2–6], via gauge invariance and the gravitational constraints. This question can be examined systematically, beginning with scattering of states with weak gravitational fields, in large radius AdS. Specifically, for such states, we expect to be able to study properties of the correspondence in an expansion in small Newton's constant $G$. One interesting point of comparison, to infer the possible structure of holography, is the $G = 0$ limit, which we expect to correspond to ordinary local quantum field theory (QFT) and not be holographic. Then, the structure of gravitational effects can be investigated to higher orders in non-zero $G$. Such study in fact suggests properties of the all-orders theory.

In addition to raising some puzzles, this analysis leads to two interesting conclusions. The first is that in order to construct the candidate holographic map from Marolf's argument [2–4], which relies on the equality between bulk and boundary hamiltonians, one apparently needs to solve the all-orders gravitational constraints, which is tantamount to knowing the bulk evolution. Second, when one examines the candidate boundary algebra that follows from commutators of the boundary gravitational hamiltonian with operators

near the boundary, it does not exhibit the closure properties that are expected for constructing a holographic map to a boundary theory.

This paper begins with some attention to setting up the basic framework; the next section proposes a definition of a holographic correspondence as a bijective isometry between the "bulk" and "boundary" Hilbert spaces, formalizing common beliefs, reviews other features of the correspondence, and raises some questions. Section three outlines some expected properties of quantum gravity in AdS, particularly in the limit of weak gravitational fields. Section four describes the $G = 0$ QFT limit of the correspondence, and discusses properties that reinforce the expectation that this limit, with no gravity, is not holographic. Finally, section five investigates the correspondence beginning with a perturbative expansion in $G$, but also inferring more general properties. In particular, a key equation of holography, equality of the bulk and boundary hamiltonians, requires solution of the gravitational constraints, which in turn requires gravitational dressing of states and operators. All-orders solution of the constraints appears to be needed to give the arguments for gravitational holography. One can then inquire whether the resulting dressed operators have the desired properties; in particular, one finds that the candidate boundary algebra doesn't close in the fashion expected for a holographic description. Similar issues are argued to arise in entanglement wedge reconstruction, which is briefly discussed.

## 2 What is holography? (Formalizing beliefs, and some questions)

We begin by stating some expected features of a holographic correspondence, such as that between gravity (or string theory) in $AdS_5$ and $\mathcal{N} = 4$ super-Yang Mills; this is largely a summary, but also sets up a proposed precise statement of the correspondence.

Specifically, the correspondence is commonly believed to be an equivalence between quantum theories. The "boundary" theory is an ordinary quantum field theory (QFT), with conformal invariance, and so should have a Hilbert space $\mathcal{H}_\partial$ and operator algebra $\mathcal{A}_\partial$ with standard QFT properties. The "bulk" theory is quantum gravity; if it is indeed a quantum theory, it is also expected to have a Hilbert space, $\mathcal{H}_B$. We also expect additional mathematical structure on $\mathcal{H}_B$, e.g. an algebra of observables $\mathcal{A}_B$, etc., that is appropriate for describing gravity, although the precise form of this structure is still unknown.[1]

An equivalence between these quantum theories is expected to be described by a map between bulk and boundary states

$$\mathcal{H}_B \xrightarrow{M} \mathcal{H}_\partial \tag{2.1}$$

that is one-to-one, onto, and preserves inner products, that is, $M$ should be a *bijective isometry* (sometimes called a *global isometry*). $M$ therefore also has an inverse map. $M$ and its inverse also induce a map between the operator algebras acting on the two Hilbert spaces; for example given an operator $\mathcal{O}_B \in \mathcal{A}_B$, the corresponding boundary operator is

$$\mathcal{O}^\partial = \mathcal{M}(\mathcal{O}_B) = M\mathcal{O}_B M^{-1} \ . \tag{2.2}$$

---

[1]We thus approach quantum gravity from a "quantum-first" perspective; for some recent discussion of progress on these questions, see [7].

Alternately, such an operator map $\mathcal{M}$, together with an identification of the vacuum states, $|0\rangle_B \rightarrow |0\rangle_\partial$, can be thought of as defining a map $M$.

One also expects equivalent global symmetry group $SO(d, 2)$ acting on the two Hilbert spaces, in the case of gravity in asymptotically $AdS_{d+1}$, and an equivalence between hamiltonians $H_B$ and $H^\partial$ via (2.2). In the standard correspondence involving $\mathcal{N} = 4$ super-Yang Mills one of course expects the bulk spacetime to be $AdS_5 \times S^5$, but we follow common practice of ignoring the $S^5$ as nontrivial dependence on these directions is expected to be handled by standard Kaluza-Klein methods.

The parameters of the bulk theory are the $D = d + 1$-dimensional Newton's constant $G_D$ and the AdS radius $R$ determined by the cosmological constant. The parameters of a boundary $SU(N)$ CFT are the coupling $g_{YM}$, and $N$. For the correspondence with $AdS_5 \times S^5$ the parameters are related by

$$\frac{R^4}{l_p^4} = N \tag{2.3}$$

where $l_p \sim G_{10}^{1/8}$ is the Planck length. If the bulk is a string theory, one also has the string length parameter $l_s = l_p/\sqrt{g_{YM}}$; we will primarily focus on gravitational effects, for which this is not relevant.

This discussion leaves us with a number of questions. A first question is, what is the map $M$? This is of course closely connected with a second question of describing $\mathcal{H}_B$ and $\mathcal{A}_B$. Specifically, as we will outline in the next section, we expect to have an approximate description of the bulk Hilbert space and algebra in a "correspondence" limit of weak gravity. But, the problem of quantum gravity can be viewed as that of finding the complete mathematical structure underlying bulk physics, which matches onto such a weak gravity description in this limit [7, 8]. Obviously, if we *did* have a complete specification of $M$, that could furnish such structure. (Conversely, if we understood $\mathcal{H}_B$ and $\mathcal{A}_B$, we might think of that as defining a boundary theory with Hilbert space $\mathcal{H}_\partial$.)

A more recent thread in the study of holography considers whether *ensembles* of theories, with different values for the couplings, are relevant (see e.g. [9]). The simplest possibility is that an ensemble of boundary theories corresponds to an ensemble of bulk theories, through an ensemble of maps such as (2.1). However, it is also possible that a bulk theory with an enlarged Hilbert space, due to topology-changing processes, induces an effective ensemble of theories [10–13], and there is significant current discussion about the exact role of such ensembles. This paper will focus on the traditional view of holography, as a map between single theories, but is also expected to contain lessons for the case of ensembles.

## 3 Expectations for bulk quantum gravity

While we don't presently understand the complete mathematical structure on the bulk $\mathcal{H}_B$, we do expect that we have a good approximate description of its structure in a weak gravity limit. Specifically, assume that $R \gg l_p$, so $N \gg 1$. Then, for example, we expect that there are states with bulk energies $E \ll 1/l_p$ which behave just like QFT states on a

background AdS spacetime, and in particular induce very small gravitational perturbations to the geometry;[2] moreover, there are such states with $E \gg 1/R$ which are well-localized compared to $R$. Since some $\mathrm{SO}(d,2)$ transformations act like boosts, this symmetry also implies that there are equivalent states that have superplanckian energies. The space of such low-energy states, and their $\mathrm{SO}(d,2)$ images, will be called $\mathcal{H}_{\mathrm{LE}} \subset \mathcal{H}_B$.

There are additional states for which we expect a weak gravity description to be valid. For example, the gravitational scattering of Oumuamua with our solar system certainly had center-of-mass (CM) energy $E_{CM} \gg 1/l_p$, yet is expected to be well described within weak gravity, and equally so if we lived in an AdS universe with $R \gg 10^{10}ly$. Specifically, we expect a valid description of physics via weak gravity and local QFT to hold in scattering where CM energies of any subprocess (including multi-particle subprocesses) don't exceed a "locality bound [14–16]," given in terms of the CM separation between particles $\Delta x$

$$E \lesssim \frac{\Delta x^{D-3}}{G_D} \tag{3.1}$$

(for multiparticle subprocesses, we take $\Delta x$ to be the bounding distance for the collection of particles). For energies violating the locality bound (3.1), we expect strong gravity effects to be important, and a description via local QFT likely fails. Notice that this criterion also allows ultraplanckian high-energy collisions, at sufficiently large impact parameter. The space of states respecting such a locality bound, and their $\mathrm{SO}(d,2)$ images, will be called $\mathcal{H}_{\mathrm{LB}} \subset \mathcal{H}_B$. Elsewhere in the literature, the terminology [17] "code subspace" is sometimes used to describe a similarly-defined space of states.

In either $\mathcal{H}_{\mathrm{LE}}$ or $\mathcal{H}_{\mathrm{LB}}$, one expects that gravitational effects in amplitudes are accounted for via perturbative gravity, incorporating graviton exchange and radiation; in the case of $\mathcal{H}_{\mathrm{LB}}$, exchange of multiple gravitons (hence, loops) is relevant, but can be, e.g., treated as eikonalized single graviton exchange, so that strong gravity effects are not relevant. This can be explained in terms of the small momentum transfer carried by any single graviton line (due to "momentum fractionation"); for further discussion, see, e.g., [18]. This weak approximation fails when the locality bound (3.1) is violated.

We moreover expect that there are subspaces of $\mathcal{H}_{\mathrm{LE}}$ or $\mathcal{H}_{\mathrm{LB}}$ in which gravity is completely negligible; one can consider states where $l_p E$ is negligible, or in the latter case, those in which collision energies yield negligible $G_D E / \Delta x^{D-3}$. One can also formally say that these states arise from taking the limit $G_D \to 0$, though the characterization in terms of energy is more accurate since $G_D$ is a dimensionful parameter. Such states and their evolution can be thought of as being accurately described entirely within QFT.

## 4   What holography isn't

Holography has been believed to be a property of theories that are gravitational, or string theories, and so it is conversely *not* an expected property of non-gravitational local QFTs,

---

[2]To avoid string scale effects, one also may require $R \gg l_s$, or $g_{\mathrm{YM}}^2 N \gg 1$, and a corresponding condition on energies.
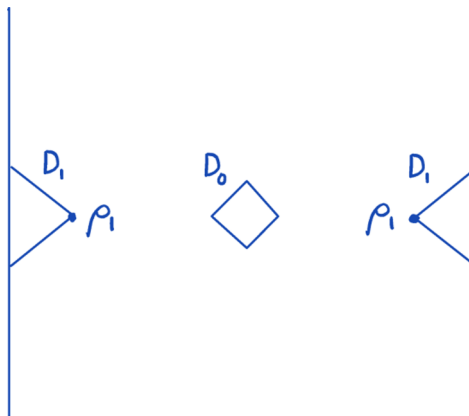
**Figure 1**. Shown is a small causal diamond $D_0$ near the center of AdS, and a causal domain $D_1$ near its boundary.

including those on an AdS background. This section will briefly explore properties of such theories that support the latter expectation.

Consider a QFT on a background spacetime; for concreteness, one could think of a scalar field theory with lagrangian $\mathcal{L}(\phi)$, or generalize to include other fields. Such a theory is expected to define a Hilbert space $\mathcal{H}_B$ with algebra of observables $\mathcal{A}_B$. In a local QFT, this algebra has additional structure, corresponding to the background spacetime. Namely, in the algebraic approach[3] one studies the *net* of subalgebras of $\mathcal{A}_B$ that are associated to open spatial regions, or to corresponding causal diamonds. Via inclusion, intersection, etc. relationships, these "mirror" the topological structure of the spacetime manifold. They also capture the local/causal structure; locality is the statement that subalgebras associated to spacelike separated regions commute.

Consider specifically a background AdS spacetime. It has been expected that in holographic theories, the operators of the boundary theory correspond to limits of bulk operators as they approach the boundary, as was described for example in [20, 21] (the "extrapolate dictionary").[4] This is, however, at odds with the locality property of QFT. Consider a small causal diamond $D_0$ near the center of AdS. For example, suppose we work in the global coordinates where AdS takes the form

$$ds^2 = \frac{R^2}{\cos^2 \rho} \left( -d\tau^2 + d\rho^2 + \sin^2 \rho \, d\Omega_{D-2}^2 \right) , \tag{4.1}$$

and let $D_0$ be centered on $\tau = \rho = 0$. Consider also the domain of dependence[5] $D_1$ of the annular region $\rho > \rho_1 < \pi/2$, as in figure 1; the union of diamonds contained in this domain corresponds to a subalgebra, which one therefore expects to contain a subalgebra

---

[3]See, for example, [19] for a review.

[4]The expectation that boundary operators arise from limits of bulk operators extends to nonlocal boundary operators; for example, in string theory, Wilson loops have been argued to correspond to operators that create corresponding long bulk strings.

[5]Given a closed spatial region this is the union of the future and past Cauchy development; for such an AdS region extending to the boundary, one also needs data on the boundary of AdS.

equivalent to the boundary algebra $\mathcal{A}_\partial$ via (2.2). However, there are states in $\mathcal{H}_B$ which are indistinguishable from $|0\rangle_B$ via such a subalgebra $\mathcal{A}_\partial$. Concretely, if one considers a bulk scalar field $\phi(x)$, and a c-number source $J(x)$ with support only in $D_0$, then the state

$$|J\rangle = e^{i \int d^D x \sqrt{-g} J(x) \phi(x)} |0\rangle_B \tag{4.2}$$

has the property that

$$\langle J | A' | J \rangle = {}_B\langle 0 | A' | 0 \rangle_B \tag{4.3}$$

for any operator $A' \in \mathcal{A}_{D_1}$. No operators (including composite nonlocal operators) defined from a boundary limit of bulk operators in the region $D_1$ register the nontrivial state $|J\rangle$. Put differently, the operator $\exp\{i \int J\phi\}$ in (4.2) doesn't obey the time slice axiom with respect to the hypothesized boundary algebra — it commutes with boundary operators, but is nontrivial.

In other words, locality implies that there is bulk information that is not present in any such "boundary theory." One expects to also be able to explain this by construction of a split vacuum,[6] which is a bulk state $|U_{D_0}\rangle$ such that

$$\langle U_{D_0} | A A' | U_{D_0} \rangle = {}_B\langle 0 | A | 0 \rangle_B \; {}_B\langle 0 | A' | 0 \rangle_B \tag{4.4}$$

for any $A \in \mathcal{A}_{D_0}$ and $A' \in \mathcal{A}_{D_1}$; thus any two states $A_1 |U_{D_0}\rangle$, $A_2 |U_{D_0}\rangle$, with $A_1, A_2 \in \mathcal{A}_{D_0}$, are also indistinguishable in the boundary algebra.

While these statements are expected to capture the essential contrast between holography and locality, it is also worth examining other kinds of statements made about holographic theories, in this non-holographic context.

One approach to holography is to consider evolution to the boundary. For example, once the causal future of $D_0$ intersects the AdS boundary, one expects that it is possible for near-boundary operators to distinguish the state $|J\rangle$ of (4.2) from the vacuum. Then, if there are corresponding boundary operators in a boundary theory, one might also naïvely expect that the boundary evolution could be used to evolve the pertinent boundary operators back in time to construct boundary operators at time $\tau = 0$ that register the state $|J\rangle$, in contrast to the preceding discussion.

However, this fails in QFT because the bulk hamiltonian $H_B$ does not map to a boundary hamiltonian $H^\partial \in \mathcal{A}_\partial$ that preserves the boundary algebra. A simplified version of this, capturing the essential point, can be illustrated by considering a free scalar hamiltonian in flat 3d space,

$$H_B = \frac{1}{2} \int d^3 x [\pi^2 + (\nabla \phi)^2], \tag{4.5}$$

and, say, a boundary given by the plane $z = 0$. Here, we do not expect a holographic description of the theory at $z < 0$. Indeed, suppose that $\phi(t, x, y, 0)$ is in a boundary algebra. Then $[H_B, \phi(t, x, y, 0)] = -i\pi(t, x, y, 0)$, and so $[H_B, [H_B, \phi(t, x, y, 0)]] = -\sum_i \partial_i^2 \phi(t, x, y, 0)$. While $x$ and $y$ derivatives of $\phi$ of course lie in the boundary algebra, iteration shows that commutators with $H$ do not close on any finite number of $z$

---

[6]See [19], and references therein; we assume such constructions extend to nontrivial spacetimes such as AdS.
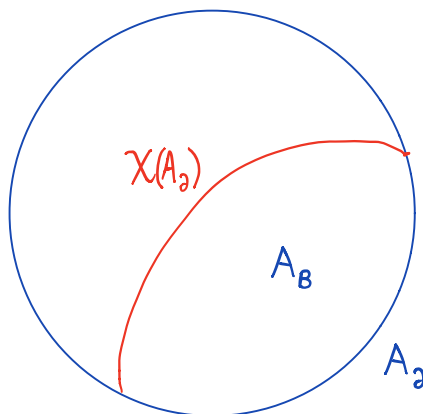
**Figure 2**. A boundary region $A_\partial$, and its corresponding Ryu-Takayanagi surface $\chi(A_\partial)$ and bulk domain $A_B$.

derivatives at the boundary. This reflects the fact that information can reach $z = 0$ from $z < 0$. Similar statements hold for QFT in an AdS background.

Of course a simple contrast to this is a situation where there is a bona-fide (2+1)-d hamiltonian, for example

$$H = \frac{1}{2} \int d^2x [\pi^2 + (\partial_x \phi)^2 + (\partial_y \phi)^2] \,. \tag{4.6}$$

Then $[H, \phi(t, x, y)] = -i\pi(t, x, y)$ and $[H, [H, \phi(t, x, y)]] = -\partial_x^2 \phi(t, x, y) - \partial_y^2 \phi(t, x, y)$, and so the algebra closes — commutators are determined in terms of elements of the algebra. This is just a different way of saying that (2+1)-d evolution with hamiltonian (4.6) is a well-defined Cauchy problem. We would expect a general boundary theory to behave this way, and this provides a criterion for closure of the boundary algebra.

Another approach to holography is that of entanglement wedge reconstruction (EWR) [22–24], and so an obvious question is what this might say in the case of a non-holographic theory. As discussed in [22], EWR is based on the key equation relating entropies,

$$S(\rho_{A_\partial}) = S(\rho_{A_B}) + \frac{1}{4G_D} \mathrm{Tr}[\rho_{A_B} \mathrm{Area}(\chi_{A_\partial}) + \cdots] \,, \tag{4.7}$$

where $A_\partial$ is a boundary spatial region, and $A_B$ is the bulk spatial region bounded by $A_\partial$ and the Ryu-Takayanagi surface [25] $\chi(A_\partial)$ of $A_\partial$; see figure 2. The density matrices are supposed to be calculated by tracing out degrees of freedom in the complementary regions, in some fiducial state. A first question is what (4.7) states in the QFT limit of $G_D \to 0$. If $S(\rho_{A_B})$ is defined with a distance cutoff $\epsilon \gg l_p$, the second term on the right dominates, and gives an infinite entropy for $\rho_{A_\partial}$. Then the bulk QFT entropy of $A_B$ is a small, relative $\mathcal{O}(G_D) \to 0$ correction. So, we don't seem to find a particularly useful statement at $G_D = 0$. We return to the question of the nonzero $G_D$ version of EWR below.

# 5 Holography at nonzero $G_D$?

Already the preceding discussion raises a puzzle. If the holographic map $M$ of (2.1) exists for arbitrary $G_D$, we might expect it to exist in the limit $G_D \to 0$. But we have just seen that such a map is not expected in this limit, unless the behavior changes qualitatively at $G_D = 0$. More precisely, for any $G_D$, and for sufficiently large $R$, we expect to be able to define the low energy subspaces where gravity is irrelevant and all bulk physics is described by local QFT, as described in section 3. The holographic map $M$ of (2.1) must act on these subspaces, yet we have no candidate definition of a holographic boundary theory dual to the resulting QFT.

In order to investigate this puzzle more closely, as well as expand on the discussion of EWR, this section will begin by considering the case of "small but nonzero $G_D$," which we have seen is really a condition specifying a subsector of states of the $R/l_p \gg 1$ theory. In fact, Marolf [2–4] (for further discussion, see also [5, 6]) has argued that holography arises from specific features unique to gravity; we can explore the validity of this proposal, beginning with the small $G_D$ limit.

Marolf's proposed explanation of holography can be thought of as a gravitational upgrade of the argument involving propagation to the boundary described in the preceding section. Schematically, we argued that a state like $|J\rangle$ of (4.2) could be registered by boundary operators in the causal future of the support of the source $J$. However, in gravity, the hamiltonian is known to be a surface term, and thus is expected to lie in the boundary algebra. This suggests that the boundary operators that register $|J\rangle$ can now be evolved back to boundary operators at $\tau = 0$, giving operators that can detect the nontrivial structure of the state outside the causal future of the source.

## 5.1 Bulk and boundary hamiltonians

It is important to examine this argument more closely. First, the general relation between the bulk and boundary hamiltonians in AdS is given by [26]

$$H_B = H_\partial + \int_\Sigma d\Sigma \; C_\xi \tag{5.1}$$

(for an explicit derivation, see [27]). Here $\Sigma$ is a constant-$\tau$ slice, and $d\Sigma$ and $n^\mu$ are its volume element and unit normal. $C_\xi$ is a projection of the constraints

$$C_\mu = \left( T_{\mu\nu} - \frac{\Lambda}{8\pi G_D} g_{\mu\nu} - \frac{1}{8\pi G_D} G_{\mu\nu} \right) n^\nu \tag{5.2}$$

into $\xi = (1, \vec{0})$, $C_\xi = C_\mu \xi^\mu$. Here $T_{\mu\nu}$ is the matter stress tensor, $G_{\mu\nu}$ the Einstein tensor and $\Lambda$ is the cosmological constant. Note that here $H_\partial$ is a boundary limit of a bulk operator; the question of its map to $\mathcal{A}_\partial$ will be considered below. The bulk and boundary hamiltonians are equal only on states annihilated[7] by the constraint $C_\xi$; for operators, the equality of their commutators with $H_B$ or $H_\partial$ follows if the operators commute with the

---

[7]In a perturbative quantization, it appears that the correct statement is that they be weakly annihilated [28].

constraints.[8] Since the constraints generate gauge transformations (diffeomorphisms that vanish at infinity), this condition specifies gauge-invariant operators; examples of these are gravitationally dressed versions of operators of the underlying QFT that is coupled to gravity.

## 5.2 Dressed operators

Specifically, in a theory describing some matter coupled to gravity, we expect the constraints (5.2) to follow from the more complete theory, in the weak-gravity regime appropriate to $\mathcal{H}_{\mathrm{LE}}$ or $\mathcal{H}_{\mathrm{LB}}$. One can work perturbatively in $\kappa^2 = 32\pi G_D$, and construct dressed operators. In particular, to leading order, explicit expressions have been found in [29, 27, 8] (for earlier related discussion see [30, 31]). Considering the case of a scalar $\phi(x)$ coupled to gravity, a dressed version takes the form

$$\Phi(\bar{x}) = \phi(X_g(\bar{x})) \ . \tag{5.3}$$

Here $X_g^\mu(\bar{x})$ are functionals of the metric, which also depend on gauge-invariant parameters $\bar{x}$ that specify the location of the operator. Under a diffeomorphism $f(x)$, $\phi \to \phi(f^{-1}(x))$; for an infinitesimal diffeomorphism $f(x) = x + \kappa\xi(x)$,

$$\delta_{\kappa\xi}\phi = -\kappa\xi^\mu\partial_\mu\phi \ . \tag{5.4}$$

In order for $\Phi(\bar{x})$ to be gauge invariant, we need the transformation $X_g(\bar{x}) \to f(X_g(\bar{x}))$, or infinitesimally,

$$\delta_{\kappa\xi}X_g(\bar{x}) = \kappa\xi^\mu(\bar{x}) \ . \tag{5.5}$$

For example, such dressings, which we write as $X_g(\bar{x}) = \bar{x} + V(\bar{x})$, may be explicitly constructed for metrics that are perturbations about flat space,

$$g_{\mu\nu} = g_{\mu\nu}^0 + \kappa h_{\mu\nu} \ , \tag{5.6}$$

with $g^0 = \eta$ (the power of $\kappa$ is introduced so that $h$ and its commutators are canonically normalized). An example is the gravitational line dressing of [29, 8], which for a general curve $\Gamma$ connecting $\bar{x}$ to infinity, takes the form

$$V_\Gamma^\mu(\bar{x}) = \frac{\kappa}{2}\int_{\bar{x}}^\infty dx^\nu \left\{ h_{\mu\nu}(x) + \int_x^\infty dx'^\lambda \left[\partial_\mu h_{\nu\lambda}(x') - \partial_\nu h_{\mu\lambda}(x')\right] \right\} + \mathcal{O}(\kappa^2) \ . \tag{5.7}$$

Using the transformation for $h$,

$$\delta_{\kappa\xi}h_{\mu\nu} = -\partial_\mu\xi_\nu - \partial_\nu\xi_\mu \ , \tag{5.8}$$

$V_\Gamma^\mu$ is easily seen to satisfy (5.5), and, when $\Phi$ acts on a state to create a $\phi$ particle, creates a corresponding line-like gravitational field. Another such dressing [29] is found by taking an angular average of $V_\Gamma^\mu$ over straight lines to infinity, and creates a "Coulomb" configuration that behaves like a linearized Schwarzschild field of a particle created by $\phi(x)$. There are

---

[8]Eq. (5.1) involves only the temporal constraint $C_\tau = 0$, but SO$(d, 2)$ symmetry relates different time evolutions, and so in general one must also solve the spatial constraints $C_i = 0$.

an infinity of such perturbative dressings, which create gravitational fields differing by radiation (sourceless) gravitational fields.

Analogous expressions for dressings perturbing about AdS are explicitly constructed in [27]. There, for a simple form of the line-dressing, the curve $\Gamma$ should be a geodesic to infinity. Averaging over such curves gives Coulomb or more general dressings. The resulting dressed operators $\Phi_V(\bar{x}) = \phi(\bar{x} + V(\bar{x}))$ then commute with the constraints, implying

$$[H_B, \Phi_V(\bar{x})] = [H_\partial, \Phi_V(\bar{x})] . \tag{5.9}$$

## 5.3 Exploring gravitational holography

With this preparation, we can explicitly study the arguments for gravitational holography [2–6]. When we do so, various questions are encountered.[9]

A first question regards the structure of the extrapolate map, namely the form of the map $\mathcal{M}(\Phi_V(\bar{x}))$ of (2.2), as $\bar{x}$ approaches the boundary. For a given underlying QFT operator $\phi(x)$, there are many (infinitely many, at the perturbative level) corresponding dressed operators $\Phi_V(\bar{x})$, which create inequivalent physical states. In general these should correspond to different boundary operators, differing at order $G_D$, or in $N^{-2}$ in the case of $AdS_5$. For example, the dressing $V$ could connect to the boundary point that $\bar{x}$ limits to, or to a different one. The details of such a correspondence remain to be worked out. But, suppose we grant that in general there is a map (2.2) between such dressed operators as they approach the boundary, collectively denoted $\mathcal{O}_i(b)$, and a collection of possibly nonlocal boundary operators $\mathcal{O}_i^\partial(b) \in \mathcal{A}_\partial$. Here boundary points $b^\alpha = (\tau, \hat{e})$ are parameterized using a unit vector $\hat{e}$ on $S^{D-2}$.

If such a map is assumed, the next question regards how to invert it to determine $\Phi_V(\bar{x})$ for an $\bar{x}$ that locates the operator near the center of AdS, in terms of boundary operators. The Heisenberg picture operators evolve according to

$$\Phi_V(\tau, x^i) = e^{iH_B\tau}\Phi_V(0, x^i)e^{-iH_B\tau} ; \tag{5.10}$$

indeed, as an alternative to (5.1), $H_B$ can be equivalently written in the usual form

$$H_B = \int_\Sigma d\Sigma n^\tau (T_{\tau\tau} + t_{\tau\tau}) \tag{5.11}$$

where $t_{\tau\tau}$ is an effective stress tensor for the gravitational field. The evolution (5.10) implies Heisenberg equations

$$\dot{\Phi}_V = i[H_B, \Phi_V] \quad , \quad \ddot{\Phi}_V = -[H_B, [H_B, \Phi_V]] . \tag{5.12}$$

Suppose, now, that we can find appropriate Green functions that solve these and express $\Phi_V(\bar{x})$, for $\bar{x} \approx 0$, in terms of the boundary limits $\mathcal{O}_i(b)$ of the bulk operators. At the linear level, such maps were proposed by [32], and a proposed nonlinear generalization [33] takes the form

$$\Phi_V(\bar{x}) = \int db K_i(\bar{x}, b)\mathcal{O}_i(b) + \int db_1 db_2 K_{ij}(\bar{x}, b_1, b_2)\mathcal{O}_i(b_1)\mathcal{O}_j(b_2) + \cdots \tag{5.13}$$

---

[9]For earlier discussion of some of these, see [27].

(with summation convention used). There are various questions about the actual construction of such expressions, but suppose we assume the existence of such expressions (and defer these questions for separate work).

We stress that the operators $\mathcal{O}_i$ are boundary limits of bulk operators, i.e., $\mathcal{O}_i \in \mathcal{A}_B$ (up to possible normalization issues). Thus, these, too, evolve according to

$$\mathcal{O}_i(\tau, \hat{e}) = e^{iH_B\tau}\mathcal{O}_i(0, \hat{e})e^{-iH_B\tau} \ . \tag{5.14}$$

At this point, the structure of the expressions, with $H_B$ given by (5.11), is the same as the (non-holographic) case of QFT, described in section 4. What is supposed to make it holographic, according to Marolf [2–4] is the equality $H_B = H_\partial$, and the identification of the boundary limits with a boundary algebra. We consider these points in turn.

## 5.4 Holography and unitary evolution

First, from (5.1), we recall that the equality $H_B = H_\partial$ requires solving the constraints. So, the constraints must be satisfied for any states for which we would like to have a holographic description — or for the operators creating such states.[10] Consider, for example, a state with two particles scattering at superplanckian CM energy. Initially , at $\tau = 0$, these particles may have separations $\Delta x$ respecting the locality bound (3.1), and thus be expected to have a good weak gravity description. But, suppose they then collide at an impact parameter violating the locality bound, and make a quantum black hole. Furthermore, suppose that this black hole then evaporates, with the formation and evaporation taking place on a time scale $\ll R$. A holographic description of this process then appears to require solution of the full nonperturbative generalization of the constraints $C_\mu = 0$.

This is tantamount to having a description of the bulk unitary evolution. This can be illustrated as in figure 3, which shows a "Wheeler-DeWitt" patch corresponding to a fixed boundary time, e.g. $\tau = 0$. We can associate to this time either the spacelike slice $\Sigma_1$ or $\Sigma_2$, and a single holographic description of the state would indicate that the states on these slices are equivalent. Indeed, evolution between the slices is generated by the constraints. For the process of black hole formation and decay, this must unitarily relate the incoming particles to the outgoing radiation, as indicated.

The conclusion of this discussion is that one needs to first have a description of unitary bulk evolution, as input *in order to construct the gravitational holographic map.* There is apparently not such a holographic map that exists independent of unitary bulk evolution, so that unitary bulk evolution can be derived purely from the boundary evolution.

To see why this is important, suppose that we want to use the holographic correspondence to provide a unitary description of black hole evolution. If there is a boundary theory we expect to have a description of unitary evolution in that theory, so if we are given the holographic map, that would determine the bulk evolution. However, the preceding argument tells us that finding the holographic map requires solving the problem we hoped holography plus boundary evolution would solve for us, namely first solving the problem of unitary bulk evolution.

---

[10]Since the required time evolution in (5.13) is over $\Delta\tau \sim 1$ (corresponding to a proper time $\sim R$), we require $H_B - H_\partial \ll 1$ for an accurate match.
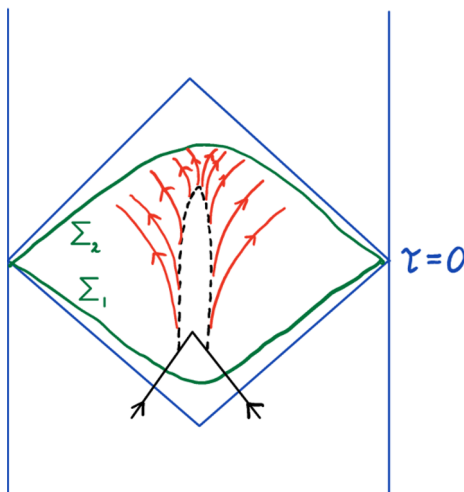
**Figure 3**. Shown is a "Wheeler-DeWitt" patch associated with a given boundary time. Since evolution between the slices $\Sigma_1$ and $\Sigma_2$ occurs at fixed boundary time, it should be generated by the constraints. This evolution could include non-perturbative processes, such as black hole formation and evaporation.

## 5.5 A boundary algebra?

The next question is that of the boundary algebra. We have assumed that limits $\mathcal{O}_i$ of bulk operators can be identified with boundary operators $\mathcal{O}_i^\partial$ via the holographic map. An important question is how this works for $H_\partial$, which we recall is in $\mathcal{A}_B$, and the role this plays in commutators. We first note that an explicit expression for $H_\partial$ in terms of the metric perturbation $h_{\mu\nu}$ was derived in [27], using covariant canonical methods:

$$H_\partial = \frac{2R^{D-4}}{\kappa} \oint d\Omega\, \xi^\tau \lim_{\rho \to \pi/2} (\tan\rho)^{D-3} \Bigg[ \cot\rho\, \hat{\nabla}^a h_{a\rho} + (D-2) h_{\rho\rho}$$
$$+ \left( \frac{2 - \cos^2\rho}{1 - \cos^2\rho} + \frac{\partial}{\partial \ln \cos\rho} \right) \hat{g}^{ab} h_{ab} \Bigg], \quad (5.15)$$

where $a, b$ denote components along the $S^{D-2}$, and $\hat{g}_{ab}$ and $\hat{\nabla}^a$ are its metric and covariant derivative. For metric perturbations with the naïve asymptotic behavior [34]

$$h_{\alpha\beta} \to \cos^{D-3}\rho\, \bar{h}_{\alpha\beta}(b^\gamma) \quad , \quad h_{\rho\rho} \to \cos^{D-3}\rho\, \bar{h}_{\rho\rho}(b^\alpha) \quad , \quad h_{\rho\alpha} \to \cos^{D-2}\rho\, \bar{h}_{\rho\alpha}(b^\beta) ,$$
$$(5.16)$$

this simplifies to

$$H_\partial = \frac{2}{\kappa} \oint d\Omega \xi^\tau\, R^{D-4} [(D-2)\bar{h}_{\rho\rho} + (D-1)\hat{g}^{ab}\bar{h}_{ab}] . \quad (5.17)$$

However, dressings such as the AdS generalization of (5.7), and its averaged Coulomb form, create metric perturbations with slower falloff than (5.16); nonetheless, these have the correct commutators with the more general expression (5.15), as shown in [27].

Note that a common view is that one can work in an approximation where the backreaction of bulk operators on the metric is neglected — that is, dressing is neglected. However,

the preceding discussion shows that this leads to violation of the equation identifying the bulk and boundary hamiltonians, $H_B = H_\partial$. For example, neglecting the dressing reduces $\Phi_V$ to the operator $\phi(x)$, and the canonical commutators yield

$$[H_\partial, \phi(x)] = 0 \neq [H_B, \phi(x)] \ . \tag{5.18}$$

While the complete dressing of an operator $\Phi_V$ may require solution of the nonperturbative constraints, the asymptotic form of $V$ near the AdS boundary is expected to be dominated by the leading perturbative structure. One can investigate the commutator of $H_\partial$ with the AdS analog of the line $V$ of (5.7), or with an averaged version, and using canonical commutators for $h_{\mu\nu}$, one finds [27]

$$[H_\partial, V^\mu(\bar{x})] = -i\delta^\mu_\tau + \mathcal{O}(\kappa) \tag{5.19}$$

the $\mathcal{O}(\kappa)$ term comes from higher-order contributions to the dressing, and for consistency should describe time translation of the argument of $V$ [35]. Then, the explicit commutators of $H_\partial$ with the dressed operators $\Phi_V(\bar{x}) = \phi(\bar{x} + V(\bar{x}))$ yield

$$[H_\partial, \Phi_V(\bar{x})] = -i\partial_{\bar{\tau}}\Phi_V(\bar{x}) \quad , \quad [H_\partial, [H_\partial, \Phi_V(\bar{x})]] = -\partial^2_{\bar{\tau}}\Phi_V(\bar{x}) \quad , \quad etc. \tag{5.20}$$

In the boundary limit $\bar{x} \to \partial$, and using the map $\mathcal{M}$ of (2.2), the commutators (5.20) should map to corresponding commutators of operators in the boundary algebra $\mathcal{A}_\partial$. But, the form of these commutators can be contrasted with what is expected for commutators with the hamiltonian in a boundary theory. For example, in the simple case of a scalar boundary theory, with hamiltonian (4.6), the commutators of the hamiltonian with the boundary operators gave equations of motion that determined the $t > 0$ operators in terms of operators at $t = 0$. Specifically, these gave Cauchy evolution where the $t > 0$ operators are determined by $\phi$ and $\partial_t\phi$ at $t = 0$. Likewise, if the boundary theory is a Yang Mills theory, we expect the same behavior. Specifically, one can consider either gauge invariant operators, or consider the Yang-Mills fields and momenta in a specific gauge. The time evolution of such operators is determined by a non-trivial equation

$$\partial_\tau \mathcal{O}^\partial = i[H^\partial, \mathcal{O}^\partial], \tag{5.21}$$

which, together with its iterations, determines all time derivatives in terms of $\tau = 0$ operators with one or zero time derivatives. That is, Yang Mills theory also has a standard canonical structure and Cauchy evolution.

However, using the commutators (5.20), the Heisenberg equations (5.21) imply

$$\partial_\tau \mathcal{O}^\partial = \partial_\tau \mathcal{O}^\partial \ : \tag{5.22}$$

that is, correct but *trivial* equations which provide no information and do not determine subsequent evolution of the operators in the absence of other information. In this sense, the commutators (5.20) do not match boundary commutators that close in the expected fashion for a boundary theory.

There is an apparent logic to the failure of the boundary commutators arising from (5.20) to give an evolution equation that is independent of the bulk evolution. The time evolution of operators has already been determined by solving the constraints, as described in the preceding subsection. If there were a separate equation for boundary evolution, there is no a priori reason for the evolutions to agree. The gravitational theory can escape this potential conflict by giving the trivial equations (5.22).

## 5.6 Entanglement wedge reconstruction

As noted above, a related approach to defining holography is that of entanglement wedge reconstruction [22–24]. A simple example of this considers a boundary spatial region $A_\partial$, that defines a Rindler wedge $D(A_B)$ with spatial base $A_B$ in the bulk. (See figure 2.) A bulk operator in $D(A_B)$ is written in the form of (5.13), where the boundary integrals are over $D(A_\partial)$ (see, e.g., [24]). Then, once again the relations $H_B = H_\partial$ and between bulk and boundary algebras are argued to give an equal-time representation of the operator, in terms of boundary operators on $A_\partial$. This argument thus relies on the same assumptions, namely bulk unitary evolution and appropriate closure of a boundary algebra, that have been critically examined above.

Specifically, the general arguments of [22] for EWR rely on the relation (4.7). This can likewise be seen to rest on solving the constraints. Specifically, [36] arrive at (4.7) from an underlying relation between modular hamiltonians,

$$K_\partial = \frac{A}{4G_D} + K_B \,, \tag{5.23}$$

where this expression is supposed to generalize to more general boundary regions $A_\partial$, and the area $A$ is that of the Ryu-Takayanagi surface [25]. In the Rindler case, the modular hamiltonians can be written in the form $K = \int \xi^\mu T_{\mu\nu} n^\nu$, where now $\xi^\mu$ is the vector generating the Rindler flow. The equality of the modular hamiltonians in (5.23) thus arises, in analogy to (5.1), from the vanishing of the bulk constraints (Einstein's equations are indeed used in deriving (5.23) in section 4 of [36]). This, and the analogous issue of the algebraic structure, suggests that EWR does not evade the arguments discussed in the preceding subsection.

## Acknowledgments

# References

[1] J.M. Maldacena, *The large N limit of superconformal field theories and supergravity*, *Int. J. Theor. Phys.* **38** (1999) 1113 [*Adv. Theor. Math. Phys.* **2** (1998) 231] [hep-th/9711200] [ɪɴSPIRE].

[2] D. Marolf, *Unitarity and holography in gravitational physics*, *Phys. Rev. D* **79** (2009) 044010 [arXiv:0808.2842] [ɪɴSPIRE].

[3] D. Marolf, *Holographic thought experiments*, *Phys. Rev. D* **79** (2009) 024029 [arXiv:0808.2845] [ɪɴSPIRE].

[4] D. Marolf, *Holography without strings?*, *Class. Quant. Grav.* **31** (2014) 015008 [arXiv:1308.1977] [ɪɴSPIRE].

[5] T. Jacobson, *Boundary unitarity and the black hole information paradox*, *Int. J. Mod. Phys. D* **22** (2013) 1342002 [arXiv:1212.6944] [ɪɴSPIRE].

[6] T. Jacobson and P. Nguyen, *Diffeomorphism invariance and the black hole information paradox*, *Phys. Rev. D* **100** (2019) 046002 [arXiv:1904.04434] [ɪɴSPIRE].

[7] S.B. Giddings, *Quantum-first gravity*, *Found. Phys.* **49** (2019) 177 [arXiv:1803.04973] [ɪɴSPIRE].

[8] S.B. Giddings, *Quantum gravity: a quantum-first approach*, *LHEP* **1** (2018) 1 [arXiv:1805.06900] [ɪɴSPIRE].

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

[10] S.R. Coleman, *Black holes as red herrings: topological fluctuations and the loss of quantum coherence*, *Nucl. Phys. B* **307** (1988) 867 [ɪɴSPIRE].

[11] S.B. Giddings and A. Strominger, *Loss of incoherence and determination of coupling constants in quantum gravity*, *Nucl. Phys. B* **307** (1988) 854 [ɪɴSPIRE].

[12] S.B. Giddings and A. Strominger, *Baby universes, third quantization and the cosmological constant*, *Nucl. Phys. B* **321** (1989) 481 [ɪɴSPIRE].

[13] 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] [ɪɴSPIRE].

[14] S.B. Giddings and M. Lippert, *Precursors, black holes, and a locality bound*, *Phys. Rev. D* **65** (2002) 024006 [hep-th/0103231] [ɪɴSPIRE].

[15] S.B. Giddings and M. Lippert, *The information paradox and the locality bound*, *Phys. Rev. D* **69** (2004) 124019 [hep-th/0402073] [ɪɴSPIRE].

[16] S.B. Giddings, *Locality in quantum gravity and string theory*, *Phys. Rev. D* **74** (2006) 106006 [hep-th/0604072] [ɪɴSPIRE].

[17] A. Almheiri, X. Dong and D. Harlow, *Bulk locality and quantum error correction in AdS/CFT*, *JHEP* **04** (2015) 163 [arXiv:1411.7041] [ɪɴSPIRE].

[18] S.B. Giddings, *The gravitational S-matrix: Erice lectures*, *Subnucl. Ser.* **48** (2013) 93 [arXiv:1105.2036] [ɪɴSPIRE].

[19] R. Haag, *Local quantum physics, fields, particles, algebras*, Springer, Berlin (1996).

[20] S.S. Gubser, I.R. Klebanov and A.M. Polyakov, *Gauge theory correlators from noncritical string theory*, *Phys. Lett. B* **428** (1998) 105 [`hep-th/9802109`] [ɪɴSPIRE].

[21] E. Witten, *Anti-de Sitter space and holography*, *Adv. Theor. Math. Phys.* **2** (1998) 253 [`hep-th/9802150`] [ɪɴSPIRE].

[22] X. Dong, D. Harlow and A.C. Wall, *Reconstruction of bulk operators within the entanglement wedge in gauge-gravity duality*, *Phys. Rev. Lett.* **117** (2016) 021601 [`arXiv:1601.05416`] [ɪɴSPIRE].

[23] J. Cotler, P. Hayden, G. Penington, G. Salton, B. Swingle and M. Walter, *Entanglement wedge reconstruction via universal recovery channels*, *Phys. Rev. X* **9** (2019) 031011 [`arXiv:1704.05839`] [ɪɴSPIRE].

[24] T. Faulkner and A. Lewkowycz, *Bulk locality from modular flow*, *JHEP* **07** (2017) 151 [`arXiv:1704.05464`] [ɪɴSPIRE].

[25] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602 [`hep-th/0603001`] [ɪɴSPIRE].

[26] T. Regge and C. Teitelboim, *Role of surface integrals in the Hamiltonian formulation of general relativity*, *Annals Phys.* **88** (1974) 286 [ɪɴSPIRE].

[27] S.B. Giddings and A. Kinsella, *Gauge-invariant observables, gravitational dressings, and holography in AdS*, *JHEP* **11** (2018) 074 [`arXiv:1802.01602`] [ɪɴSPIRE].

[28] W. Donnelly and S.B. Giddings, *Gravitational splitting at first order: quantum information localization in gravity*, *Phys. Rev. D* **98** (2018) 086006 [`arXiv:1805.11095`] [ɪɴSPIRE].

[29] W. Donnelly and S.B. Giddings, *Diffeomorphism-invariant observables and their nonlocal algebra*, *Phys. Rev. D* **93** (2016) 024030 [*Erratum ibid.* **94** (2016) 029903] [`arXiv:1507.07921`] [ɪɴSPIRE].

[30] I. Heemskerk, *Construction of bulk fields with gauge redundancy*, *JHEP* **09** (2012) 106 [`arXiv:1201.3666`] [ɪɴSPIRE].

[31] D. Kabat and G. Lifschytz, *Decoding the hologram: scalar fields interacting with gravity*, *Phys. Rev. D* **89** (2014) 066010 [`arXiv:1311.3020`] [ɪɴSPIRE].

[32] A. Hamilton, D.N. Kabat, G. Lifschytz and D.A. Lowe, *Holographic representation of local bulk operators*, *Phys. Rev. D* **74** (2006) 066009 [`hep-th/0606141`] [ɪɴSPIRE].

[33] I. Heemskerk, D. Marolf, J. Polchinski and J. Sully, *Bulk and transhorizon measurements in AdS/CFT*, *JHEP* **10** (2012) 165 [`arXiv:1201.3664`] [ɪɴSPIRE].

[34] M. Henneaux and C. Teitelboim, *Asymptotically Anti-de Sitter spaces*, *Commun. Math. Phys.* **98** (1985) 391 [ɪɴSPIRE].

[35] W. Donnelly and S.B. Giddings, *How is quantum information localized in gravity?*, *Phys. Rev. D* **96** (2017) 086013 [`arXiv:1706.03104`] [ɪɴSPIRE].

[36] D.L. Jafferis, A. Lewkowycz, J. Maldacena and S.J. Suh, *Relative entropy equals bulk relative entropy*, *JHEP* **06** (2016) 004 [`arXiv:1512.06431`] [ɪɴSPIRE].