Bekenstein-Hawking entropy as topological entanglement entropy

Lauren McGough[a] **and Herman Verlinde**[a,b]

[a]*Department of Physics, Princeton University,*
Washington Road, Princeton, NJ 08544, U.S.A.

[b]*Princeton Center for Theoretical Sciences, Princeton University,*
Washington Road, Princeton, NJ 08544, U.S.A.

E-mail: mcgough@princeton.edu, verlinde@princeton.edu

ABSTRACT: Black holes in 2+1 dimensions enjoy long range topological interactions similar to those of non-abelian anyon excitations in a topologically ordered medium. Using this observation, we compute the topological entanglement entropy of BTZ black holes via the established formula $S_{\rm top} = \log(S_0^a)$, with S_b^a the modular S-matrix of the Virasoro characters $\chi_a(\tau)$. We find a precise match with the Bekenstein-Hawking entropy. This result adds a new twist to the relationship between quantum entanglement and the interior geometry of black holes. We generalize our result to higher spin black holes, and again find a detailed match. We comment on a possible alternative interpretation of our result in terms of boundary entropy.

Contents

1 Introduction

The close relation between black hole physics and thermodynamics provides crucial guidance to the search for consistent quantum theories that incorporate gravity. In particular, it indicates that pure quantum gravity — i.e. any attempt to directly quantize the Einstein lagrangian, without the addition of any matter degrees of freedom — is unlikely to give rise to a complete theory. Metric excitations alone seem insufficient to account for the microscopic entropy of black holes, quantified via the Bekenstein-Hawking formula [1, 2]

$$S_{\mathrm{BH}} = \frac{\mathrm{Area}}{4G_N} \,. \tag{1.1}$$

A more promising perspective is that general relativity represents a long range effective theory with dynamical rules that encode the quantum information flow of underlying elementary degrees of freedom. This point of view is supported by string theory realizations of black hole space-times, in which the B-H formula (1.1) has been successfully matched with the microscopic entropy of the constituent strings, D-branes and their excitations [3].

Another powerful diagnostic tool is the geometric entanglement entropy [4–6], which has received much recent attention. Let A denote a region of space, such as the interior of a black hole, and B its complement, all of space outside of A. The density matrix associated with A is $\rho_A = \mathrm{tr}_B \left(|\psi\rangle\langle\psi| \right)$, where $|\psi\rangle$ is typically taken to be the ground state of the system, and the trace is over all states of B. The von Neumann entropy

$$S_A = -\mathrm{tr}(\rho_A \log \rho_A)$$

quantifies the total entanglement between region A and its complement B.

The importance of entanglement for the microscopic structure of space-time is only beginning to emerge. There are tantalizing hints of a deep connection, most notably the Ryu-Takayanagi formula [7–9] and the firewall debate [10, 11]. In this note, we study this relationship in 2+1-D AdS space-times. Einstein gravity in 2+1 dimensions has special characteristics, akin to Chern-Simons (CS) gauge theories [12–14] that capture the infrared properties of quantum critical systems with topological order [16, 17]. Massive spinning point particles and black holes enjoy long range interactions that generalize the braiding relations of particles with non-abelian statistics [18]. In addition, the system possesses a ground state degeneracy that is sensitive to the global space-time topology. In condensed matter systems, such as those exhibiting the fractional quantum Hall effect, these remarkable properties emerge because the ground state of the underlying medium is deeply entangled [19]. Quantum gravity in 2+1 dimensions should be thought of in the same way: as the effective theory that captures the topological Berry phases of the ground state wave function. It is through these topological interactions that the quantum order of the microphysical medium manifests itself.

Topological entanglement entropy provides a quantitative measure of this long range quantum order [16, 17]. Consider a region A with disk-like topology and a smooth boundary of length L. In a gapped quantum many-body system, the geometric entanglement entropy of A has the form

$$S_A = \alpha L + S_{\text{top}} + \dots \tag{1.2}$$

where $\dots$ indicate terms that vanish in the limit $L \to \infty$. The first term arises from short wavelength modes straddling the boundary of the entangling region. The pre-coefficient α is non-universal, and depends on the UV cut-off. The constant term S_{top} is the topological entanglement entropy; it represents a universal characteristic of the many-body vacuum state [16, 17]. In the above sign convention, it is typically ≤ 0. It can be isolated from the length term by dividing the region A into three or more segments and taking a suitable linear combination of the resulting entanglement entropies in which the boundary terms cancel. Since topological entanglement entropy survives in the long distance limit, $L \to \infty$, it can be calculated by means of the low energy topological field theory that describes the braiding properties of the quasi-particle excitations. In case the region A contains a single excitation labeled by some charge a, one finds that [16, 17, 19]

$$S_{\text{top}} = \log(d_a/\mathcal{D}) = \log(S_0^a) . \tag{1.3}$$

Here $\mathcal{D}$ and d_a are the quantum dimensions of the medium and the a excitation, respectively. S_0^a denotes a matrix element of the modular S-matrix of the 1+1-dimensional CFT that describes the edge excitations of the topologically ordered medium. The quantity S_{top} has the key property that it does not depend on the size or geometry of the region A.

Topological entanglement entropy and black hole entropy seem unrelated. The B-H formula of 2+1-D black holes [20, 21] relates the entropy to the length of the event horizon via

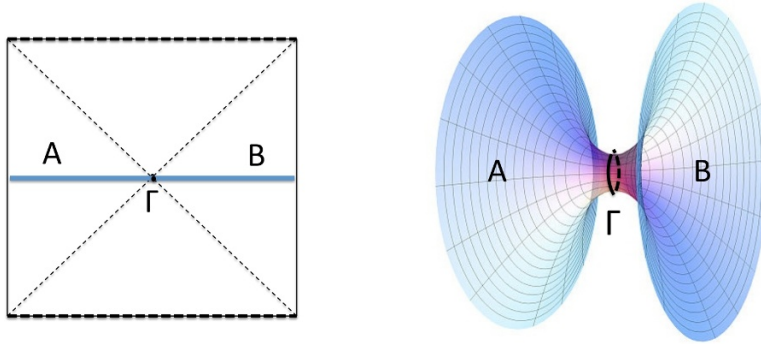$$S_{\text{BH}} = \frac{\text{Length}(\Gamma)}{4G_N} . \tag{1.4}$$

Figure 1. The black hole horizon forms a geodesic Γ. The entanglement entropy between the inside and outside regions A and B is equal to $\text{Length}(\Gamma)/4$.

This looks similar to the non-universal length term in (1.2), except that the coefficient α is now a universal constant. Because of this similarity, many authors have suggested that the B-H formula may also have an interpretation as geometric entanglement entropy [4–6]. There is growing evidence that this is indeed the case [7–9, 22]. This is an important insight. In particular, it indicates that black holes are typically in a near-maximally entangled state.

However, there is one unsatisfactory aspect to relating the length contributions in (1.4) and (1.2). Unlike the first term in (1.2), the B-H formula (1.4) is universal and robust. In this respect, black hole entropy seems more similar to the universal constant contribution in (1.2). Could it be that the 2+1-D black hole entropy (1.4) can be identified with the universal topological entanglement entropy associated with the black hole space-time?

At a first glance, this seems implausible: the B-H formula does not appear topological, for it is proportional to a length. How, then, could this be true? Figure 1 shows a Penrose diagram of an eternal BTZ black hole of mass M (and spin $J = 0$) and a spatial slice with an Einstein-Rosen bridge connecting the two sides. The horizon is a geodesic: it has minimal length for the given topology of Γ. So we can view $\text{Length}(\Gamma)$ as a common property of all loops with the same topology of Γ. In other words, $\text{Length}(\Gamma)$ should not be viewed as a geometric property of a loop, but as a quantum number of the black hole state, determined by its mass M and spin J.

Let us now view the black hole as a localized defect of a topological ordered system, and treat M and J in the same way as the charge label a in (1.3). This interpretation is natural given that 2+1-D gravity can be written as a $\mathcal{G} = \text{SL}(2,\mathbb{R}) \times \text{SL}(2,\mathbb{R})$ Chern-Simons theory [12, 13], in which the black hole state represents a heavy particle with a large $\mathcal{G}$ charge. The edge states of 2+1-D gravity are described by Liouville theory [23–27], the universal conformal field theory associated with the Virasoro algebra. Although Liouville theory is a non-rational CFT with central charge $c = \frac{3}{2}\ell \gg 1$, it shares many features with rational CFTs [26, 27]. In particular, the Virasoro conformal blocks form a unitary representation of the modular and braid group, characterized by the quantum group $\mathcal{U}_q\big(\mathfrak{sl}(2,\mathbb{R}) \times \mathfrak{sl}(2,\mathbb{R})\big)$. This representation is infinite-dimensional, and modular and fusion relations are expressed as integrals rather than finite sums. Nonetheless, one can identify analogs of quantum dimensions and of the modular S-matrix S_b^a.

We can thus apply the same formulas (1.3) to compute the topological entanglement entropy associated with the black hole excitation. Using the proper identification of a black hole of mass M and J with a superselection label a of Liouville CFT, we find a precise match

$$S_{\text{BH}} = \log \left(S_0^a \right), \qquad a = (M, J). \tag{1.5}$$

We describe the details of this calculation in the following sections. To test the robustness of our result, we also consider the higher spin black holes, and find an encouraging match with known results.

This identification and interpretation of the Bekenstein-Hawking entropy as topological entanglement entropy raises many conceptual questions. Why does the computation of the topological entanglement entropy reproduce the microscopic entropy? What does our computation say about the applicability and validity of pure quantum gravity in 2+1 dimensions? What is entangled with what? What does the calculation imply for the firewall controversy? We address these questions in the concluding section.

2 The BTZ black hole

We briefly summarize the main properties of the BTZ black hole [20, 21, 24, 25]. From now on we put $G_N = 1$, so ℓ denotes the AdS$_3$ curvature radius in Planck units.

AdS$_3$ can be identified with the universal covering space of the group SL(2,$\mathbb{R}$), and has isometry group $\mathcal{G} = \text{SL}(2,\mathbb{R}) \times \text{SL}(2,\mathbb{R})$. The BTZ black hole space-time is obtained by taking the quotient of AdS$_3$ with a hyperbolic group element $(h_+, h_-) \in \mathcal{G}$, acting via

$$g \sim h_+ g h_-, \qquad h_\pm = e^{\pi(r_+ \pm r_-)\sigma_3/\ell}. \tag{2.1}$$

The quotient describes a stationary and axially symmetric black hole with an outer event horizon at r_+ and an inner Cauchy horizon at r_-. The BTZ metric can be written as

$$ds^2 = -4\ell(\Delta_+ du^2 + \Delta_- dv^2) + d\rho^2 + (\ell^2 e^{2\rho} + \Delta_+ \Delta_-) du dv. \tag{2.2}$$

The two radii $r_\pm$ and the constants $\Delta_\pm$ are related to the black hole mass and spin via

$$M = \frac{r_+^2 + r_-^2}{8\ell^2}, \qquad J = \frac{r_+ r_-}{4\ell}, \qquad \Delta_\pm = \frac{(r_+ \pm r_-)^2}{16\ell} = \frac{1}{2}(\ell M \pm J). \tag{2.3}$$

Einstein gravity in 2+1 dimensions can be formulated as a CS-type gauge theory by introducing the dreibein e^a and spin connection ω^a. The linear combinations $A_\pm^a = \omega^a \pm \frac{1}{\ell} e^a$ form two SL(2,$\mathbb{R}$) connections, in terms of which the torsion constraint and Einstein equation take the form of flatness constraints [12, 13]. The group elements $h_\pm$ in equation (2.1) coincide with the holonomies of $A_\pm^a$ around the black hole. In general, SL(2,$\mathbb{R}$) holonomies come in three types, depending on whether the conjugacy class of the group element is hyperbolic, parabolic, or elliptic. For a black hole, both holonomies are in a hyperbolic conjugacy class.

The Bekenstein-Hawking entropy of the BTZ black hole is equal to
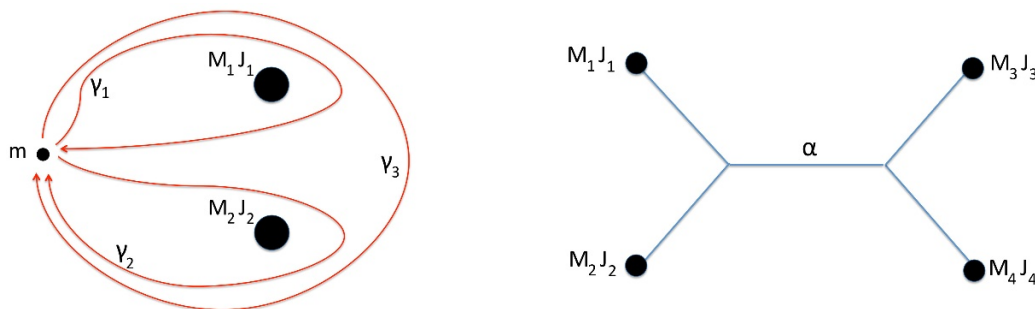
$$S_{\text{BH}} = \frac{2\pi r_+}{4}. \tag{2.4}$$

Figure 2. The classical space-time geometry is specified by the holonomies around the paths γ_i. In the quantum theory, states are identified with conformal blocks of 2D Liouville CFT.

This formula has been reproduced in numerous dual CFT realizations of string theory on AdS_3 by counting the number of states at energy M and with angular momentum J. Below we will give a new derivation and interpretation.

3 Quantum geometry

Quantum geometry arises from quantizing the phase space of space-time geometries. As an example, figure 2 indicates the geometry of two BTZ black holes, specified by the $\mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{R})$ holonomies around the paths γ_i. These holonomies are determined, up to overall conjugation, by the mass, spin, center of mass energy and total angular momentum of the two black holes. This description generalizes to any number of point particles and black holes [13]. The space of $\mathrm{SL}(2,\mathbb{R})$ holonomies is isomorphic to Teichmuller space, the space of constant negative curvature metrics on a 2-D surface. The phase space of 2+1-D Einstein gravity consists of two copies of Teichmuller space [23].

The problem of quantizing Teichmuller space has been solved [26–29]. It gives rise to a Hilbert space of states that can be identified with the linear space spanned by the chiral conformal blocks of 2-D Liouville theory [26–29]

$$S_L(\varphi) = \frac{1}{4\pi} \int d^2\xi \left[\frac{1}{2}(\partial\varphi)^2 + QR\varphi + \mu e^{b\varphi}\right], \qquad Q = b + b^{-1}. \tag{3.1}$$

This correspondence generalizes the well-known relationship [14] between Chern-Simons theories and WZW conformal field theory [15]. The dictionary is analogous. The 2-D CFT describes the massless edge excitations at the boundary of the AdS space, and supports a unitary representation of the asymptotic symmetry group of the bulk theory. For pure AdS_3 gravity, this symmetry group takes the Virasoro algebra with central charge [30]

$$c = 1 + 6Q^2 = 3\ell/2. \tag{3.2}$$

States of 2+1-D gravity with particle and black hole excitations in the bulk are identified with the product of left and right conformal blocks of Liouville CFT with corresponding vertex operator insertions. They enjoy a q-deformed version of the monodromy properties of the classical geometry, governed by the non-compact quantum group $\mathcal{U}_q\big(\mathfrak{sl}(2,\mathbb{R}) \times \mathfrak{sl}(2,\mathbb{R})\big)$

Elliptic   Hyperbolic

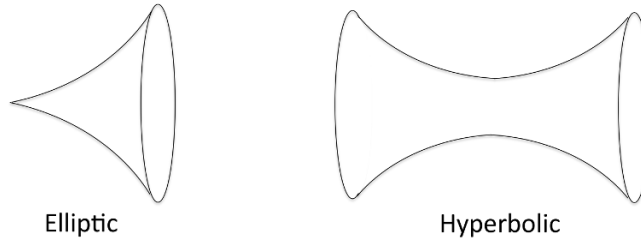**Figure 3**. Liouville vertex operators fall into two classes. Those with $\Delta < \frac{1}{4}Q^2$ create *elliptic* solutions (punctures), those with $\Delta > \frac{1}{4}Q^2$ create *hyperbolic* solutions (macroscopic holes) [31].

with $q = \exp(i\pi b^2)$. Liouville vertex operators take the general form $V_\alpha = e^{\alpha_+ \varphi_+} e^{\alpha_- \varphi_-}$, and are in one-to-one correspondence with unitary highest weight representations of the left and right Virasoro algebra with conformal weights $\Delta_\pm = \alpha_\pm(Q - \alpha_\pm)$. The physical range of positive conformal weights splits into two separate regimes of Liouville momenta

$$\alpha_\pm \in \left[0, \frac{1}{2}Q\right] \cup \left(\frac{1}{2}Q + i\mathbb{R}^+\right). \tag{3.3}$$

The Liouville equation prescribes that the metric have constant negative curvature everywhere except at the location of the vertex operators. Vertex operators with real Liouville momentum in the interval $[0, \frac{1}{2}Q]$ create elliptic solutions, which are local cusps specified by a patching function in the elliptic conjugacy class of the isometry group $\mathcal{G}$. Vertex operators with complex momenta of the form $\frac{1}{2}Q + i\mathbb{R}^+$ create hyperbolic solutions, which are macroscopic holes in 2-D space identified with the spatial section of BTZ black hole geometries (as shown in figures 1 and 3). We may parametrize the Liouville momenta in this range as

$$\alpha_\pm = \frac{1}{2}Q + ip_\pm\,, \qquad \Delta_\pm = p_\pm^2 + \frac{1}{4}Q^2. \tag{3.4}$$

These relations, combined with equations (2.2)–(2.3), specify a precise dictionary between the classical data of the BTZ black hole and the quantum data of Liouville theory. For later reference, we make note that in the semiclassical regime $p_\pm \gg b \gg 1$, the relations between the Liouville momenta $p_\pm$ and the conjugacy class of the holonomies $h_\pm$ in (2.1) simplify to

$$r_\pm = 4b(p_+ \pm p_-)\,, \qquad b^2 = \ell/4\,. \tag{3.5}$$

Most of the above dictionary was known before the discovery of gauge/gravity duality. An important insight from AdS/CFT is that the bulk theory can not be pure gravity. Gravity in 2+1 dimensions describes how massive localized excitations interact at long distances, but it does not specify the hyperfine structure of the excitation spectrum of the bulk string theory.

The situation in the 1+1-D boundary theory is analogous. Liouville theory has a continuous spectrum of conformal dimensions, and is therefore capable of describing any set of Virasoro representations. However, it does not prescribe the spectrum of some given unitary CFT. Liouville theory is similar to a non-compact space with a continuous spectrum of wave solutions; choosing a specific CFT realization of AdS$_3$ is like putting the wave solutions in a finite box, so that the spectrum becomes discrete and countable.

4 Quantum dimension

An important ingredient of our story is the *quantum dimension* associated with a local excitation in a topological quantum field theory. We first recall the definition and properties of the quantum dimension of a topological QFT associated to a rational CFT. We then generalize to the case of interest, the non-rational $c > 25$ Virasoro CFTs.

The most physical definition of the quantum dimension is as follows. Let $\mathcal{H}_a(N)$ denote the Hilbert space of the 2+1-dimensional topological QFT spanned by all states that contain N local excitations of charge a. It can be shown that the dimension of this Hilbert space grows exponentially at large N according to

$$\dim \mathcal{H}_a(N) \propto (d_a)^N. \tag{4.1}$$

The number d_a defines the quantum dimension of the excitation a.

Quantum dimensions are linked with the fusion algebra [32–34]. A local excitation with charge a corresponds to a primary vertex operator V_a in the CFT. The operator product of V_a and V_b can be expanded as a sum of operators V_c. For rational CFTs, the fusion coefficients $N_{ab}{}^c$ are integers that specify the multiplicity of V_c in this expansion. The fusion algebra is commutative and associative, and admits a one-dimensional representation $d_a d_b = \sum_c N_{ab}{}^c d_c$. This relation can be used to prove the result (4.1).

Quantum dimensions can be thought of as the character of the superselection sector $\mathcal{H}_a$ associated with the primary vertex operator V_a. States in $\mathcal{H}_a$ are obtained by acting with symmetry generators on the primary state $|a\rangle = V_a|0\rangle$. The partition function

$$\chi_a(\tau) = \operatorname{tr}_{\mathcal{H}_a}\left(e^{i\pi\tau L_0}\right) \tag{4.2}$$

is called the character of the sector $\mathcal{H}_a$. The quantum dimension d_a is obtained by taking the $\tau \to 0$ limit of the ratio of $\chi_a(\tau)$ with the identity character [32–34]

$$d_a = \lim_{\tau \to 0} \frac{\chi_a(\tau)}{\chi_0(\tau)}\,. \tag{4.3}$$

This definition naturally explains why the quantum dimensions generate the fusion algebra. It also allows us to re-express d_a in terms of the modular S-matrix, which describes the transformation properties of the characters under the modular transformation $\tau \to -1/\tau$

$$\chi_a(-1/\tau) = \sum_b S_a^b \chi_b(\tau)\,. \tag{4.4}$$

Applying the modular transformation (4.4) to (4.3), and using that for $\tau \to 0$ the dominant term in the sum comes from the identity character, one finds that

$$d_a = \frac{S_0^a}{S_0^0}\,. \tag{4.5}$$

This formula for the quantum dimension holds for rational CFTs and plays a key role in the computation of topological entanglement entropy. We will use this connection momentarily.

First, we need to generalize the above formulas to the case of non-rational CFTs relevant to 2+1-D gravity. The modular geometry of Liouville theory is by now quite well-developed [26–29], and many of the RCFT formulas have found direct non-rational analogs. There are two main differences. Since the spectrum of allowed conformal dimensions is continuous, modular transformations and fusion coefficients are no longer described by discrete sums and finite matrices but by integrals and continuous distributions. Another important difference is that the identity representation plays a rather distinct role. In spite of these dissimilarities, there still exists a natural analog of the notion of quantum dimension.

Let us follow the naive route and simply apply the formula (4.5). The $c > 25$ Virasoro characters for the continuous representations of conformal weight $\Delta > \frac{1}{4}Q^2$ are given by

$$\chi_p(\tau) = \frac{e^{i\pi\tau p^2}}{\eta(\tau)}, \qquad \Delta_p = p^2 + \frac{1}{4}Q^2, \tag{4.6}$$

where the Dedekind η-function $\eta(\tau) = q^{c/24} \prod_{n>0}(1 - q^n)$ with $q \equiv e^{2\pi i\tau}$. The identity character

$$\chi_0(\tau) = \frac{e^{-i\pi\tau\frac{Q^2}{4}}(1 - e^{i\pi\tau})}{\eta(\tau)}, \qquad \Delta = 0 \tag{4.7}$$

follows the following modular transformation property [26]

$$\chi_0(-1/\tau) = \int_0^\infty dp\, S_0^p \chi_p(\tau) \tag{4.8}$$

$$S_0^p = 2\sqrt{2}\sinh(2\pi bp)\sinh(2\pi b^{-1}p). \tag{4.9}$$

Note that i) S_0^p is not a matrix entry of a finite matrix, but a measure on the continuous series of Virasoro representations, ii) S_0^p grows exponentially with p, and iii) the identity representation itself does not appear on the right-hand side of (4.8).

Boldly applying the formula (4.5), we find that, up to an irrelevant overall constant, the quantum dimension of the representation with Liouville momentum $\alpha = \frac{1}{2}Q + ip$ is given by

$$d(\alpha) = \sinh(2\pi bp)\sinh(2\pi b^{-1}p). \tag{4.10}$$

This quantity $d(\alpha)$ indeed plays a special role in Liouville modular geometry [28, 29]. As mentioned above, the representation theory and modular geometry of the Virasoro conformal blocks with $c > 25$ is associated with the representation theory of the quantum group $\mathcal{U}_q\big(\mathfrak{sl}(2,\mathbb{R})\big)$. This quantity (4.10) naturally appears in this context as the weight of a representation in the Peter-Weyl or Plancherel decomposition of the space of functions on the quantum group. This so-called Plancherel measure is the most natural counterpart of the quantum dimension in the nonrational case.

5 Topological entanglement entropy

Gravity is topological in the sense that every observable must be coordinate invariant. In 2+1 dimensions, this topological nature is enhanced by the fact that there are no graviton excitations, and that the metric, outside of matter distributions, locally always looks

the same. Our proposal is that from a microscopic perspective, these properties emerge because gravity is the long distance description of the highly entangled ground state of a topologically ordered system close to a quantum critical point. For analogous condensed matter systems, the natural diagnostic for the presence of topological order is the topological entanglement entropy introduced in [16, 17]. Let us briefly recall its definition.

To compute the topological entanglement entropy S_{top} of a disk-shaped region R with the outside $D = R^c$, one first divides the interior of R into three sectors A, B and C. Let $S_A = -\operatorname{tr} \rho_A \log \rho_A$ denote the von Neumann entropy of the density matrix ρ_A associated with subregion A, and analogously for S_B, S_C. Similarly, let S_{AB} denote the entropy associated with $A \cup B$, etc. The topological entanglement entropy of R is then defined as [16] $S_{\text{top}} = S_A + S_B + S_C - S_{AB} - S_{AC} - S_{BC} + S_{ABC}$. This linear combination has the special property that all non-universal perimeter-law contributions cancel out. Moreover, any local deformation of the entangling boundary does not alter the final result.

Applying this definition to a topological field theory associated with a 2-D rational CFT, one finds [16, 17] that an empty region of space has $S_{\text{top}}(0) = \log(1/D) = \log(S_0^0)$. Here D is the total quantum dimension of the medium, and related to the quantum dimension d_a of individual excitations via $D^2 = \sum_a d_a^2$. In case the region contains a quasi-particle excitation of charge a, the topological entanglement entropy is given by the formula [19]

$$S_{\text{top}}(a) = \log(d_a/D) = \log(S_0^a). \tag{5.1}$$

Note that for a rational CFT the topological entanglement is always negative since $d_a < D$. By itself, this would not make sense, as the definition of the entanglement entropy is a manifestly positive quantity. Once we include the non-universal contribution proportional to the length L of the entangling boundary, however, the total result is positive.

We make the assumption that the relationship between the quantum dimension and the topological entanglement entropy remains mostly unchanged in going from rational to non-rational CFTs. One important difference is that there no longer exists an analog of the total quantum dimension D, and hence there is no obvious notion of topological entanglement entropy of an empty region of space. However, there does exist a natural formula for the topological entanglement entropy of a black hole excitation. Applying the formula (5.1) to the hyperbolic Virasoro representation associated with a BTZ black hole, we find

$$S_{\text{top}}(M, J) = \log(S_0^{p_+} S_0^{p_-}). \tag{5.2}$$

The relation between the mass M and spin J and the Liouville momenta is given in equation (3.4) with $\Delta_\pm$ defined in (2.3). Plugging in the explicit modular S-matrix element (4.9) gives

$$S_{\text{top}}(M, J) = \log\left(8 \sinh(2\pi b p_+) \sinh(2\pi b^{-1} p_+) \sinh(2\pi b p_-) \sinh(2\pi b^{-1} p_-)\right). \tag{5.3}$$

Note that unlike the rational CFT case, the right-hand side is positive. Moreover, it grows unboundedly for large $p_\pm$. In the limit where $\ell M \pm J$ and b are all large, it reduces to

$$S_{\text{top}}(M, J) = 2\pi b (p_+ + p_-) = \frac{2\pi r_+}{4}, \tag{5.4}$$

which exactly matches the Bekenstein-Hawking entropy. This is our main result.

6 Higher spin black hole entropy

As a test of our proposal, let us consider black holes in 2+1-D higher spin gravity [35, 36]. Luckily, all the necessary technology is available. Our presentation will be brief.

Higher spin gravity in 2+1 dimensions is a generalization of Einstein gravity in 2+1 dimensions that includes a collection of $n-2$ higher spin fields [35]. All the fields together can be assembled into a $\mathrm{SL}(n, \mathbb{R}) \times \mathrm{SL}(n, \mathbb{R})$ gauge connection (A_+, A_-) with a Chern-Simons action. The generalized space-time geometry of a higher spin black hole is characterized by two $\mathrm{SL}(n, \mathbb{R})$ holonomies

$$h_\pm = e^{2\pi(\lambda_+ \pm \lambda_-)/\ell}, \tag{6.1}$$

which generalize the $\mathrm{SL}(2, \mathbb{R})$ holonomies (2.1) of the BTZ black hole. Here ℓ denotes the higher spin generalization of the AdS$_3$ radius, and λ_+ and λ_- are diagonal elements of the $\mathfrak{sl}(n, \mathbb{R})$ Lie algebra. Higher spin black holes thus carry $2(n-1)$ quantum numbers, including the mass, angular momentum and $2(n-2)$ higher spin charges.

Extracting an actual space-time geometry from this general description of the higher spin black hole turns out to be a non-trivial task. In particular, there is no gauge-invariant notion of a 2+1-D space-time metric that can be used to compute a horizon area. As a result, there appears to be no immediately obvious higher spin generalization of the Bekenstein-Hawking formula. There are indeed various proposals [36–39].

A simple geometric proposal for a generalized Bekenstein-Hawking formula was put forward in [37–39]. Let e_i denote the simple roots of $\mathfrak{sl}(n)$ and $\langle , \rangle$ denote the Cartan Killing form. The Weyl vector is defined as $\rho = \frac{1}{2} \sum_{e>0} e$. The higher spin generalization of the black hole entropy formula derived in [37–39] is expressed in terms of the $\mathrm{SL}(n, \mathbb{R})$ holonomies $h_\pm$ as

$$S_{\mathrm{HSBH}} = \frac{2\pi}{4} \langle \rho, \lambda_+ \rangle. \tag{6.2}$$

This elegant proposal passes some non-trivial checks [37–39] and appears well-motivated.

Can one reproduce the generalized B-H formula (6.2) by counting states in the dual CFT? This is a non-trivial task, since one needs a generalization of the Cardy formula that keeps track of conformal dimensions and all higher spin quantum numbers. This has not been done yet. We now give a simple derivation of (6.2) via the topological entanglement entropy (1.3).

2+1-D higher spin gravity is dual to 1+1-D conformal field theory with W_n symmetry, the natural higher spin generalization of Virasoro symmetry. The universal CFT with W_n symmetry is $\mathfrak{sl}(n, \mathbb{R})$ Toda theory

$$S = \frac{1}{2\pi} \int d^2\xi \left[\langle \partial\varphi, \bar{\partial}\varphi \rangle + R\langle Q, \varphi \rangle + \mu \sum e^{b\langle e_i, \varphi \rangle} \right], \qquad Q = 2(b + b^{-1})\rho. \tag{6.3}$$

Toda theory is a non-rational CFT with central $c = n-1+3\langle Q, Q \rangle$. As before, states of the 2+1-D higher spin theory with localized excitations are identified with the tensor product of left and right conformal blocks of the CFT. Black holes states correspond to vertex operators that create macroscopic holes in the generalized space time, with holonomies (6.1) in a hyperbolic conjugacy class of $\mathrm{SL}(n, \mathbb{R})$. Their vertex operators $V = e^{\langle \alpha_+, \varphi_+ \rangle} e^{\langle \alpha_-, \varphi_- \rangle}$

have Toda momenta $\alpha_\pm = \frac{1}{2}Q + ip_\pm$ and conformal weights $\Delta_\pm = \langle \alpha_\pm, Q - \alpha_\pm \rangle = \langle p_\pm, p_\pm \rangle + \frac{1}{4}\langle Q, Q \rangle$. The semi-classical relations (3.5) naturally generalize to

$$\lambda_\pm = 4b\,(p_+ \pm p_-)\,, \qquad b^2 \langle \rho, \rho \rangle = \ell/8\,. \tag{6.4}$$

Just like their BTZ counterparts, higher spin black holes can be viewed as macroscopic quasi-particle excitations with topological interactions. We can thus compute their topological entanglement entropy in the same way as before. The relevant modular S-matrix elements of $\mathfrak{sl}(n, \mathbb{R})$ Toda field theory was computed in [40]

$$S_0^p = \Xi \prod_{e>0} 4 \sinh\left(\pi b \langle e, p \rangle\right) \sinh\left(\pi b^{-1} \langle e, p \rangle\right) \tag{6.5}$$

with Ξ some irrelevant constant. Using the formula $S_{\text{top}} = \log(S_0^{p_+} S_0^{p_-})$ and taking the semi-classical limit, we reproduce the result (6.2)

$$S_{\text{top}} = 2\pi b \left(\langle \rho, p_+ \rangle + \langle \rho, p_- \rangle \right) = \frac{2\pi}{4} \langle \rho, \lambda_+ \rangle\,. \tag{6.6}$$

7 Concluding remarks

We have put forward a new interpretation of 2+1-D quantum gravity as the effective field theory that describes the long range properties of a highly entangled ground state. In line with this interpretation, we have computed the topological entanglement entropy of a BTZ black hole. Our computation does not make use of the Bekenstein-Hawking, Ryu-Takayanagi, or Cardy formulas. It is a new and independent derivation, yet yields a leading-order result that matches all three. Our result also raises a number of questions. We briefly comment on some of them.

Does pure 2+1-D quantum gravity exist? What is its role? Via the identification with the space of left and right conformal blocks of 2-D Liouville theory, we have given a well-defined description of the Hilbert space of 2+1-D quantum gravity. Does this mean that pure 2+1-D quantum gravity exists as a UV complete theory? The answer is "No" [41]. The spectrum of Virasoro representations is continuous, and thus the level density of states of Liouville theory and pure 2+1-D gravity is strictly infinite. This is an unphysical situation. To get a well-behaved physical system, one needs to supply a specific holographic dual in the form of some unitary 2-D CFT. This CFT prescribes the allowed discrete spectrum of conformal dimensions, with a finite level density. In this note, we implicitly assumed that this CFT is maximally non-rational, i.e. that it does not have any other symmetries than conformal invariance. In this idealized case, once the spectrum of excitations is prescribed, 2+1-D gravity gives an accurate description of their long range interactions and assigns the correct quantum dimension to the black hole states.

What does the topological entanglement entropy count? This is the most important question. It is natural to interpret S_{top} as the universal contribution to the entanglement across the black hole horizon. The fact that it saturates the B-H bound is consistent with the

idea [8, 9] that entanglement is responsible for the continuity of space across the horizon. However, this interpretation immediately raises an important puzzle, closely related to the firewall paradox [10, 11].

According to the usual AdS/CFT dictionary, any typical CFT state with large enough energy describes a black hole in the bulk. The level density of the CFT indeed matches the microscopic B-H entropy. However, to write a state with entanglement entropy proportional to $S_{\rm BH}$, one needs to include *two* Hilbert space sectors each with entropy at least equal to $S_{\rm BH}$. The CFT seems to provide only one of these sectors. So where is the other sector?

Liouville vertex operators with momenta $\alpha = \frac{1}{2}Q + ip$ in fact create macroscopic holes in space, as indicated in figure 3. Based on the similarity with figure 1, it is tempting to identify both sides of the hyperboloid in figure 3 as the two sides of the eternal black hole solution. According to this interpretation, it seems that by acting with the vertex operator, one has created a completely new asymptotic region with its own holographic CFT dual. This could be where the other sector resides. But how would one create such a second asymptotic region via gravitational collapse, i.e. by acting with operators on the vacuum of one single CFT? This is one version of the firewall question.

Is there a firewall or fuzzball? Is $S_{\rm BH}$ a boundary entropy? In our view, if our proposal that the entanglement entropy of BTZ black holes saturates the B-H bound is correct, then there is no firewall. The state looks like an eternal black hole that realizes the balanced holography postulate put forward in [22]. The entanglement across the horizon is then sufficient to safeguard the continuity of space [8, 9].

There is, however, another possible interpretation[1] of our formula $S_{\rm BH} = \log S_0^a$ in terms of the Affleck-Ludwig boundary entropy [42]. Suppose that, instead of the hyperbolic solution of figure 3, we place a reflecting boundary at the location of the black hole horizon. A natural conformal boundary for Liouville CFT is the ZZ-boundary state $|ZZ\rangle$ [26]. Its overlap with the Ishibashi state $\|p\rangle\!\rangle$, the eigenstates with given Liouville momentum $\alpha = \frac{1}{2}Q + ip$, satisfies

$$\left|\Psi_{\rm ZZ}(p)\right|^2 = S_0^p, \qquad \Psi_{\rm ZZ}(p) = \langle\!\langle p \,\|\, ZZ \rangle. \tag{7.1}$$

This implies that the boundary entropy of the ZZ state in the sector with momentum p is equal to $\log(S_0^p)$. Moreover, if we identify the topological entanglement entropy with the Bekenstein-Hawking entropy of the BTZ black hole, we obtain the very suggestive relation

$$Z_{\rm BH} = \left|\Psi_{\rm ZZ}(p_+, p_-)\right|^2 \tag{7.2}$$

with $Z_{\rm BH} = e^{S_{\rm BH}}$. Could it be that, instead of topological entanglement entropy, our formula is counting the boundary entropy of a reflecting boundary at the horizon? Or are both interpretations correct? Either way, we believe that finding the answer to these questions will shed important new light on the nature of the interior geometry of black holes.

---

[1]We thank Nick Warner for emphasizing this possible alternative interpretation.

Acknowledgments

References

[1] S. Hawking, *Particle creation by black holes*, *Commun. Math. Phys.* **43** (1975) 199 [*Erratum ibid.* **46** (1976) 206] [ɪɴSPIRE].

[2] J.D. Bekenstein, *Black holes and entropy*, *Phys. Rev.* **D 7** (1973) 2333 [ɪɴSPIRE].

[3] A. Strominger and C. Vafa, *Microscopic origin of the Bekenstein-Hawking entropy*, *Phys. Lett.* **B 379** (1996) 99 [`hep-th/9601029`] [ɪɴSPIRE].

[4] L. Bombelli, R.K. Koul, J. Lee and R.D. Sorkin, *A quantum source of entropy for black holes*, *Phys. Rev.* **D 34** (1986) 373 [ɪɴSPIRE].

[5] M. Srednicki, *Entropy and area*, *Phys. Rev. Lett.* **71** (1993) 666 [`hep-th/9303048`] [ɪɴSPIRE].

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

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

[8] M. Van Raamsdonk, *Comments on quantum gravity and entanglement*, `arXiv:0907.2939` [ɪɴSPIRE].

[9] J.M. Maldacena and L. Susskind, *Cool horizons for entangled black holes*, `arXiv:1306.0533` [ɪɴSPIRE].

[10] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, *Black holes: complementarity or firewalls?*, *JHEP* **02** (2013) 062 [`arXiv:1207.3123`] [ɪɴSPIRE].

[11] S.L. Braunstein, S. Pirandola and K. Życzkowski, *Better late than never: information retrieval from black holes*, *Phys. Rev. Lett.* **110** (2013) 101301 [`arXiv:0907.1190`] [ɪɴSPIRE].

[12] A. Achucarro and P.K. Townsend, *A Chern-Simons action for three-dimensional anti-de Sitter supergravity theories*, *Phys. Lett.* **B 180** (1986) 89 [ɪɴSPIRE].

[13] E. Witten, *(2+1)-dimensional gravity as an exactly soluble system*, *Nucl. Phys.* **B 311** (1988) 46 [ɪɴSPIRE].

[14] E. Witten, *Quantum field theory and the Jones polynomial*, *Commun. Math. Phys.* **121** (1989) 351 [ɪɴSPIRE].

[15] O. Coussaert, M. Henneaux and P. van Driel, *The asymptotic dynamics of three-dimensional Einstein gravity with a negative cosmological constant*, *Class. Quant. Grav.* **12** (1995) 2961 [`gr-qc/9506019`] [ɪɴSPIRE].

[16] A. Kitaev and J. Preskill, *Topological entanglement entropy*, *Phys. Rev. Lett.* **96** (2006) 110404 [`hep-th/0510092`] [ɪɴSPIRE].

[17] M. Levin and X.-G. Wen, *Detecting topological order in a ground state wave function*, *Phys. Rev. Lett.* **96** (2006) 110405 [ɪɴSPIRE].

[18] C. Nayak, S.H. Simon, A. Stern, M. Freedman and S. Das Sarma, *Non-Abelian anyons and topological quantum computation*, *Rev. Mod. Phys.* **80** (2008) 1083 [INSPIRE].

[19] S. Dong, E. Fradkin, R.G. Leigh and S. Nowling, *Topological entanglement entropy in Chern-Simons theories and quantum Hall fluids*, *JHEP* **05** (2008) 016 [arXiv:0802.3231] [INSPIRE].

[20] M. Bañados, C. Teitelboim and J. Zanelli, *The black hole in three-dimensional space-time*, *Phys. Rev. Lett.* **69** (1992) 1849 [hep-th/9204099] [INSPIRE].

[21] M. Bañados, M. Henneaux, C. Teitelboim and J. Zanelli, *Geometry of the (2+1) black hole*, *Phys. Rev.* **D 48** (1993) 1506 [gr-qc/9302012] [INSPIRE].

[22] E. Verlinde and H. Verlinde, *Passing through the firewall*, arXiv:1306.0515 [INSPIRE].

[23] H.L. Verlinde, *Conformal field theory, 2-d quantum gravity and quantization of Teichmüller space*, *Nucl. Phys.* **B 337** (1990) 652 [INSPIRE].

[24] S. Carlip, *Conformal field theory, (2+1)-dimensional gravity and the BTZ black hole*, *Class. Quant. Grav.* **22** (2005) R85 [gr-qc/0503022] [INSPIRE].

[25] Y.-j. Chen, *Quantum Liouville theory and BTZ black hole entropy*, *Class. Quant. Grav.* **21** (2004) 1153 [hep-th/0310234] [INSPIRE].

[26] A.B. Zamolodchikov and A.B. Zamolodchikov, *Liouville field theory on a pseudosphere*, hep-th/0101152 [INSPIRE].

[27] J. Teschner, *Liouville theory revisited*, *Class. Quant. Grav.* **18** (2001) R153 [hep-th/0104158] [INSPIRE].

[28] J. Teschner, *On the relation between quantum Liouville theory and the quantized Teichmüller spaces*, *Int. J. Mod. Phys.* **A 19S2** (2004) 459 [hep-th/0303149] [INSPIRE].

[29] J. Teschner, *Nonrational conformal field theory*, arXiv:0803.0919 [INSPIRE].

[30] J.D. Brown and M. Henneaux, *Central charges in the canonical realization of asymptotic symmetries: an example from three-dimensional gravity*, *Commun. Math. Phys.* **104** (1986) 207 [INSPIRE].

[31] N. Seiberg, *Notes on quantum Liouville theory and quantum gravity*, *Prog. Theor. Phys. Suppl.* **102** (1990) 319 [INSPIRE].

[32] E.P. Verlinde, *Fusion rules and modular transformations in 2D conformal field theory*, *Nucl. Phys.* **B 300** (1988) 360 [INSPIRE].

[33] J.L. Cardy, *Boundary conditions, fusion rules and the Verlinde formula*, *Nucl. Phys.* **B 324** (1989) 581 [INSPIRE].

[34] R. Dijkgraaf and E.P. Verlinde, *Modular invariance and the fusion algebra*, *Nucl. Phys. Proc. Suppl.* **5B** (1988) 87 [INSPIRE].

[35] M.R. Gaberdiel and R. Gopakumar, *Minimal model holography*, *J. Phys.* **A 46** (2013) 214002 [arXiv:1207.6697] [INSPIRE].

[36] M. Ammon, M. Gutperle, P. Kraus and E. Perlmutter, *Black holes in three dimensional higher spin gravity: a review*, *J. Phys.* **A 46** (2013) 214001 [arXiv:1208.5182] [INSPIRE].

[37] J. de Boer and J.I. Jottar, *Thermodynamics of higher spin black holes in AdS_3*, arXiv:1302.0816 [INSPIRE].

[38] J. de Boer and J.I. Jottar, *Entanglement entropy and higher spin holography in AdS$_3$*, arXiv:1306.4347 [INSPIRE].

[39] A. Perez, D. Tempo and R. Troncoso, *Higher spin gravity in 3D: black holes, global charges and thermodynamics*, *Phys. Lett.* **B 726** (2013) 444 [arXiv:1207.2844] [INSPIRE].

[40] N. Drukker, D. Gaiotto and J. Gomis, *The virtue of defects in 4D gauge theories and 2D CFTs*, *JHEP* **06** (2011) 025 [arXiv:1003.1112] [INSPIRE].

[41] A. Maloney and E. Witten, *Quantum gravity partition functions in three dimensions*, *JHEP* **02** (2010) 029 [arXiv:0712.0155] [INSPIRE].

[42] I. Affleck and A.W.W. Ludwig, *Universal noninteger 'ground state degeneracy' in critical quantum systems*, *Phys. Rev. Lett.* **67** (1991) 161 [INSPIRE].