Microcanonical action and the entropy of Hawking radiation

Juan F. Pedraza[1], Andrew Svesko[2], Watse Sybesma[3], and Manus R. Visser[4]

[1]*Departament de Física Quàntica i Astrofísica and Institut de Ciéncies del Cosmos,*
Universitat de Barcelona, 08028 Barcelona, Spain
[2]*Department of Physics and Astronomy, University College London, London WC1E 6BT, United Kingdom*
[3]*Science Institute, University of Iceland, Dunhaga 3, 107 Reykjavík, Iceland*
[4]*Department of Theoretical Physics, University of Geneva,*
24 quai Ernest-Ansermet, 1211 Genève 4, Switzerland

The island formula—an extremization prescription for generalized entropy—is known to result in a unitary Page curve for the entropy of Hawking radiation. This semiclassical entropy formula has been derived for Jackiw-Teitelboim (JT) gravity coupled to conformal matter using the "replica trick" to evaluate the Euclidean path integral. Alternatively, for eternal Anti–de Sitter black holes, we derive the extremization of generalized entropy from minimizing the microcanonical action of an entanglement wedge. The on-shell action is minus the entropy and arises in the saddle-point approximation of the (nonreplicated) microcanonical path integral. When the black hole is coupled to a bath, islands emerge from maximizing the entropy at fixed energy, consistent with the island formula. Our method applies to JT gravity as well as other two-dimensional dilaton gravity theories.

I. INTRODUCTION

Hawking's discovery [1,2] of thermal radiation emitted by black holes has led to a puzzle quantum gravity is expected to resolve: the information paradox [3]. A solution will provide a better understanding of the quantum mechanical evolution of black holes. Either black holes evolve unitarily, such that the von Neumann (vN) entropy of radiation S_{vN}^{rad} follows a Page curve [4,5], or they do not. For dynamical black holes formed from gravitational collapse, the Page curve has the generic feature that S_{vN}^{rad} increases from zero (as the radiation begins in a pure state), until it reaches a maximum at the Page time t_P, and then decreases to zero at late times, thus conserving information. Advances motivated by the Anti–de Sitter/Conformal Field Theory (AdS/CFT) correspondence suggest S_{vN}^{rad} indeed evolves unitarily [6–15].

Surprisingly, the unitary Page curve can be derived within semiclassical gravity. Crucial to this derivation are quantum extremal surfaces (QESs)—codimension-2 surfaces X which extremize the semiclassical generalized entropy S_{gen}—and the QES formula [16–18]

$$S_{vN}(\Sigma_X) = \min_X \operatorname*{ext}_X \left[\frac{\text{Area}(X)}{4G} + S_{vN}^{sc}(\Sigma_X) \right]. \quad (1)$$

Here, $S_{vN}(\Sigma_X)$ is the fine-grained vN entropy of Σ_X in the full quantum theory, S_{vN}^{sc} is the vN entropy of bulk quantum fields including both matter fields and the metric in the semiclassical approximation, and Σ_X is a codimension-1 slice bounded by X and a cutoff surface. The bracketed term is the generalized entropy $S_{gen}(\Sigma_X)$, and obeys a generalized second law [19–21].

The QES formula (1) also holds for the vN entropy of Hawking radiation S_{vN}^{rad}, where it is known as the "island formula" [9]. Indeed, applying (1) to evaporating or eternal black holes reveals a Page curve [6–8]. In this case Σ_X may be disconnected, $\Sigma_X = \Sigma_{rad} \cup I$, where Σ_{rad} is the region outside the black hole and I is an "island" with $X = \partial I$. For evaporating black holes I lies inside the black hole [9,14,15], while for eternal black holes islands extend outside the classical horizon [8,22–25]. Motivated by [17,26–29], the island formula has been derived using the "replica trick" in the context of Jackiw-Teitelboim (JT) gravity [10–12]. The Page curve arises from a competition between two saddle point geometries dominating the Euclidean gravitational path integral (PI). At early times the (replicated) black hole solution, or "Hawking saddle", dominates the PI and is responsible for the rise in S_{vN}^{rad}. At late times, "replica wormholes" overtake the black hole leading to the "island saddle" and, for evaporating black holes, a decrease in S_{vN}^{rad}.

Working in the microcanonical ensemble [30,31], here we derive the extremization condition in (1) from a (non-replicated) AdS$_2$ gravitational PI. Indeed, for black holes with $U(1)$ Killing symmetry, the replica trick is not necessary to compute gravitational fine-grained entropies [32]. Rather, one may opt to use the standard thermal gravitational partition function [33]. Further, in this ensemble we show how islands arise for eternal AdS$_2$ black holes coupled to a bath, a setting with its own information paradox [8]. For definiteness, we will focus on semiclassical JT gravity but our methodology applies to more general two-dimensional (2D) dilaton theories of gravity.

II. SEMICLASSICAL JT GRAVITY AND THE WALD ENTROPY

Semiclassical JT gravity is characterized by the action

$$I = I_{\text{JT}} + I_{\text{Pol}}, \qquad (2)$$

where $I_{\text{JT}} = I_{\text{JT}}^{\text{bulk}} + I_{\text{JT}}^{\text{GHY}} + I_{\text{JT}}^{\text{ct}}$ is the classical JT action [34,35] with a Gibbons-Hawking-York (GHY) term, such that the variational principle is well posed for spacetimes M with boundary B, and a local counterterm to regulate divergences at B

$$I_{\text{JT}}^{\text{bulk}} = \frac{1}{16\pi G} \int_M d^2 x \sqrt{-g}\left((\phi_0 + \phi)R + \frac{2\phi}{L^2} \right),$$

$$I_{\text{JT}}^{\text{GHY}} + I_{\text{JT}}^{\text{ct}} = \frac{1}{8\pi G} \int_B dt \sqrt{-\gamma}\left((\phi_0 + \phi)K - \frac{\phi}{L} \right). \qquad (3)$$

Here ϕ is the dilaton arising from a spherical reduction of the parent theory, ϕ_0 a constant proportional to the extremal entropy of the higher-dimensional black hole system, L is the AdS$_2$ length scale, and K is the trace of the extrinsic curvature of B with induced metric $\gamma_{\mu\nu}$. The semiclassical contribution $I_{\text{Pol}} = I_{\text{Pol}}^{\text{bulk}} + I_{\text{Pol}}^{\text{GHY}} + I_{\text{Pol}}^{\text{ct}}$ is composed of the Polyakov action localized via an auxiliary massless scalar field χ of central charge c, a GHY term and counterterm [36]

$$I_{\text{Pol}}^{\text{bulk}} = -\frac{c}{24\pi} \int_M d^2 x \sqrt{-g}[(\nabla\chi)^2 + \chi R],$$

$$I_{\text{Pol}}^{\text{GHY}} + I_{\text{Pol}}^{\text{ct}} = -\frac{c}{12\pi} \int_{B'} dt \sqrt{-\gamma}\left(\chi K + \frac{1}{2L} \right). \qquad (4)$$

Here B' represents a cutoff surface near infinity in the flat space regions, which are sewn at the AdS$_2$ boundary B with transparent boundary conditions, as in [8]. Notably, the Polyakov action arises as the quantum effective action associated to the conformal anomaly of a CFT$_2$ coupled to any 2D theory of gravity [37].

Semiclassical JT gravity admits eternal AdS$_2$ black holes as a solution, for which the line element of the metric $g_{\mu\nu}$ in Schwarzschild coordinates (t, r) is

$$d\ell^2 = -f(r)dt^2 + f^{-1}(r)dr^2, \qquad f(r) = \frac{r^2}{L^2} - \mu, \qquad (5)$$

with μ a mass parameter in the classical Arnowitt-Deser-Misner (ADM) energy. The horizon is at $r_H = L\sqrt{\mu}$. The remaining semiclassical equations of motion may be solved once the state of the quantum matter is specified. Requiring regularity at the horizon fixes the vacuum state to be the Hartle-Hawking state $|\text{HH}\rangle$ [38,39], for which observers in (null) static coordinates $(u, v) = (t - r_*, t + r_*)$, with tortoise coordinate $r_* \in [-\infty, 0]$, see a thermal bath of particles at Hawking temperature $T_H = \frac{\sqrt{\mu}}{2\pi L}$. The expectation value of the normal-ordered Polyakov stress tensor is $\langle\text{HH}| : T_{uu}^\chi : |\text{HH}\rangle = \frac{c\pi}{12} T_H^2$ (and similarly for T_{vv}^χ). With respect to $|\text{HH}\rangle$, the semiclassical correction to ϕ is just a constant [36,40]

$$\phi(r) = \phi_r \frac{r}{L} + \frac{Gc}{3} = \phi_r \sqrt{\mu} \coth\left(-\frac{\sqrt{\mu}}{L} r_* \right) + \frac{Gc}{3}, \qquad (6)$$

where the dimensionless parameter $\phi_r > 0$ is the boundary value of ϕ, such that at a cutoff $r = \epsilon_c^{-1}$ near the boundary $\phi \to \frac{\phi_r}{\epsilon_c L}$. Recently, moreover, we demonstrated χ is generically time dependent, such that in the Hartle-Hawking state $|\text{HH}\rangle$ [40]

$$\chi = -\frac{1}{2} \log\left[4\mu \left(1 + \frac{\mu UV}{L^2} \right)^{-2} \right]$$
$$+ k - \log\left[\left(K_U - \frac{\sqrt{\mu}U}{L} \right)\left(K_V + \frac{\sqrt{\mu}V}{L} \right) \right], \qquad (7)$$

where $(U, V) = (-\frac{L}{\sqrt{\mu}} e^{-\frac{\sqrt{\mu}u}{L}}, \frac{L}{\sqrt{\mu}} e^{\frac{\sqrt{\mu}v}{L}})$ are dimensionful Kruskal coordinates, and k, K_U, K_V are constants.

Directly from the action (2), using the Noether charge formalism [41], we can derive the Wald entropy S_W

$$S_W \equiv -2\pi \epsilon_{\mu\nu} \epsilon_{\rho\sigma} \frac{\partial \mathcal{L}}{\partial R^{\mu\nu\rho\sigma}} = \frac{1}{4G}(\phi_0 + \phi) - \frac{c}{6}\chi, \qquad (8)$$

which is generally evaluated on a bifurcate Killing horizon with binormal $\epsilon_{\mu\nu}$. Remarkably, imposing a Dirichlet boundary condition on χ, consistent with the transparent boundary condition at B, the integration constants k, K_U, K_V may be fixed such that the semiclassical correction to S_W is exactly the vN entropy of a single interval with endpoints (U_1, V_1) and (U_2, V_2) of a CFT$_2$ in $|\text{HH}\rangle$ with UV regulators δ_1, δ_2 [40]

$$S_{\text{vN}}^{\text{HH}} = -\frac{c}{6}\chi$$
$$= \frac{c}{12} \log\left[\frac{16\mu^2}{(1 + \frac{\mu U_1 V_1}{L^2})^2 (1 + \frac{\mu U_2 V_2}{L^2})^2} \right]$$
$$+ \frac{c}{12} \log\left[\frac{1}{\delta_1^2 \delta_2^2} (U_2 - U_1)^2 (V_2 - V_1)^2 \right]. \qquad (9)$$

We emphasize that this result simply follows from solving χ with Dirichlet boundary. Consequently, while the Wald entropy is evaluated on some bifurcation point, the solution χ describes the entropy associated with an interval. Thus, in semiclassical JT gravity the Wald entropy (8) equals the generalized entropy S_{gen} [40]

$$S_{\text{W}} = S_{\text{BH}} + S_{\text{vN}}^{\text{HH}} = S_{\text{gen}}, \qquad (10)$$

which is the sum of the Bekenstein-Hawking entropy $S_{\text{BH}} = \frac{1}{4G}(\phi_0 + \phi)$—corresponding to the area term in (1) in higher dimensions—and the CFT entropy $S_{\text{vN}}^{\text{HH}}$ (9). This resolves the discrepancy found in [42] between the Wald entropy and generalized entropy in two dimensions.

A key insight of [40] is that, in the microcanonical ensemble, the generalized entropy is equal to the microcanonical entropy. Further, S_{gen} is stationary at fixed energy, $\delta S_{\text{gen}}|_E = 0$, which has the basic elements of the extremization prescription in the QES formula. This observation suggests, at least for backgrounds with $U(1)$ Killing symmetry, that we may compute the generalized entropy and its extremization directly when working with Euclidean path integrals in the microcanonical ensemble.

III. MICROCANONICAL ACTION AND GENERALIZED ENTROPY

We are interested in deriving the extremization prescription in the QES formula (1) from first principles. This was accomplished in [29] using the replica trick. Now we will show for eternal black holes how to derive the extremization of S_{gen} from a gravitational Euclidean PI in the microcanonical ensemble.

We consider the saddle-point approximation to the microcanonical gravitational partition function [30,43]

$$W(E_0) = \int \mathcal{D}\psi e^{-I_{\text{E}}^{\text{mc}}(\psi)} \approx e^{-I_{\text{E}}^{\text{mc}}(\psi_0)}, \qquad (11)$$

where the path integral is taken over all dynamical fields $\psi = (g_{\mu\nu}, \phi, \chi)$ with fixed energy E_0 (specified below). In the saddle-point approximation the Euclidean microcanonical action I_{E}^{mc} is evaluated on the solutions ψ_0 of the semiclassical field equations. We will compute the action I_{E}^{mc} on the entanglement wedge [44–46] of an interval Σ in AdS$_2$, which is the domain of dependence of any achronal surface with boundary $\partial\Sigma$. The microcanonical boundary condition specifies an entanglement wedge at fixed energy E_0 in the eternal AdS$_2$ black hole background, where the external CFT$_2$ remains in the Hartle-Hawking vacuum.

The entanglement wedge consists of a rectangular causal diamond (CD) with null boundaries at $(u - u_0 = \pm a, v - v_0 = \pm b)$, where a, $b > 0$ are in principle different length scales. Since after extremizing I_{E}^{mc} the lengths a and b coincide, we set $a = b$ for purposes of clarity. Such CDs
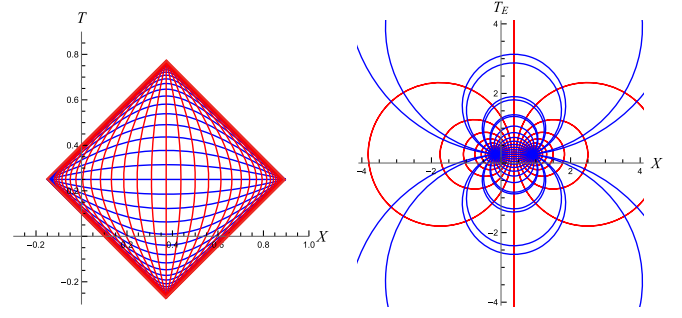

FIG. 1. Lorentzian AdS$_2$ diamond (left) and its Euclidean continuation (right) in Kruskal coordinates $(T, X) = (\frac{1}{2}(V + U), \frac{1}{2}(V - U))$. Lines of constant x (red) and lines of constant s, s_{E} (blue) are at equal intervals of 0.125. On the right, high contour density corresponds to horizon punctures at $x = \pm\infty$. We have set $a = b = 1/2$, $\mu = L = \kappa_a = \kappa_b = 1$ and $u_0 \neq v_0 \neq 0$.

have a conformal isometry generated by a conformal Killing vector ζ [47–49], which we fix uniquely by demanding ζ becomes the boost Killing vector of AdS$_2$-Rindler when the future and past vertices of the CD both lie on the AdS boundary (see Appendix A). We cover the CD in "diamond universe" coordinates (s, x), adapted to the flow of the vector field ζ [48] (left Fig. 1). The coordinate s is the conformal Killing time, satisfying $\zeta \cdot ds = 1$. The line element in these coordinates is (Appendix B)

$$d\ell^2 = C^2(s, x)(-ds^2 + dx^2), \qquad (12)$$

with $s, x \in [-\infty, \infty]$ and the conformal factor C^2 given in (B4). The null boundaries are conformal Killing horizons with constant surface gravity κ, defined by $\nabla_\mu \zeta^2 = -2\kappa\zeta_\mu$ [50], which is positive on the future horizon and negative on the past horizon (below we take $\kappa > 0$). When restricting the HH state to the CD, it becomes a thermal density matrix with temperature $T = \kappa/2\pi$. Near the null boundaries of the causal diamond at $x = \pm\infty$, the metric (12) is approximately

$$d\ell^2 \approx 4L^2\kappa^2 e^{\mp 2\kappa x}(-ds^2 + dx^2). \qquad (13)$$

This is simply the Rindler metric $d\ell^2 = -\kappa^2\varrho^2 ds^2 + d\varrho^2$, with radial coordinate $\varrho \equiv 2Le^{\mp\kappa x}$ and surface gravity $\kappa = \mp C^{-1}\partial_x C|_{x\to\pm\infty}$. Hence, $\zeta = \partial_s$ becomes an approximate boost Killing vector near $x = \pm\infty$.

Next, we compute the on-shell microcanonical action of CDs in semiclassical JT gravity following the Hilbert action surface term method of [31,51]. Concretely, we evaluate the GHY term on the boundary of an infinitesimal disk D_ϵ of radius ϵ orthogonal to the punctures $\partial\Sigma : x = \pm\infty$ in the Euclidean diamond (right Fig. 1),

$$I_{\text{E}}^{\text{mc}} = -\lim_{\epsilon\to 0} \int_{\partial D_\epsilon \times \partial\Sigma} ds_E \sqrt{\gamma} K \left[\frac{(\phi_0 + \phi)}{8\pi G} - \frac{c\chi}{12\pi} \right]. \qquad (14)$$

Here Euclidean time $s_E = is$ is periodic, $s_E \sim s_E + \frac{2\pi}{\kappa}$, to remove the conical singularity at $x = \pm\infty$ or $\varrho = 0$. Note this regularity condition at the horizon is consistent with our choice of the Hartle-Hawking vacuum state [52]. Further, $\sqrt{\gamma} = C$ is the induced metric on constant s_E slices and the extrinsic trace of these slices is $K = \mp\, C^{-2}\partial_x C$. Crucially, in the limit $x \to \pm\infty$, the fields ϕ (6) and χ (7) are independent of s_E, and $\sqrt{\gamma}K \to \kappa$, leading to

$$I_E^{\mathrm{mc}} = -S_{\mathrm{gen}}|_{\partial\Sigma}, \tag{15}$$

where we used (8) and the nontrivial fact $S_W = S_{\mathrm{gen}}$. This is a semiclassical extension of the microcanonical action formula $I_E^{\mathrm{mc}} = -S_{\mathrm{BH}}$ obtained by [30,31], and may be interpreted as a path integral derivation of the generalized entropy of causal diamonds. Note since $\partial\Sigma$ consists of two points, the action is actually twice the entropy. We further derive (14) and (15) in Appendix C using the Noether charge method [41,53,54], and show that the microcanonical on-shell action is proportional to the Noether charge. Thus, we find the partition function (11) equals the density of states $W \approx e^{S_{\mathrm{gen}}}$ in the saddle-point approximation, with S_{gen} as the microcanonical entropy.

In standard thermodynamics the entropy is maximized at fixed energy in the microcanonical ensemble. Thence, via (15), the microcanonical action I_E^{mc} is minimized at some fixed energy E_0. For CDs we can infer E_0 from the variation of the microcanonical action on the full Euclidean diamond M_E^{CD}, which has a Euclidean time circle S^1 of period $2\pi/\kappa$. Explicitly, (see Appendix C)

$$\delta I_E^{\mathrm{mc}} = \int_{M_E^{\mathrm{CD}}} ds_E \wedge \omega(\psi, \delta\psi, \mathcal{L}_\zeta\psi) = \int_{S^1} ds_E \delta H_\zeta, \tag{16}$$

where $\omega(\psi, \delta\psi, \mathcal{L}_\zeta\psi)$ is the symplectic current 1-form evaluated on the Lie derivative $\mathcal{L}_\zeta\psi$ of ψ along ζ. In the second equality we inserted Hamilton's equations, $\delta H_\zeta = \int_{\Sigma_{s_E}} \omega(\psi, \delta\psi, \mathcal{L}_\zeta\psi)$, where H_ζ is the Hamiltonian generating evolution along the flow of ζ and Σ_{s_E} are constant s_E slices that smoothly intersect $\partial\Sigma$. Therefore, the action I_E^{mc} is stationary at fixed energy $E_0 \sim H_\zeta$.

By way of (15), the variational identity (16) is an integrated version of the semiclassical first law for AdS$_2$ CDs, $\frac{\kappa}{2\pi}\delta S_W = -\delta H_\zeta$, which can be derived using the Noether charge method [40,48]. Since the CFT is in the HH state, $T = \frac{\kappa}{2\pi}$ is identified as the temperature of the causal diamond. The first law turns into a proper thermodynamic first law $T\delta S_{\mathrm{gen}} = \delta E_0$ if we identify $E_0 = -H_\zeta$. Alternatively, one may include the minus sign into the temperature T as done in [48,55], however, this negative temperature seems inconsistent with the thermality of the HH state when reduced to the CD, see (B6). Moreover, the first law or (15)–(16) tells us the thermodynamic potential defining the microcanonical ensemble is

$S_{\mathrm{gen}}(E_0)$, obeying the equilibrium condition $\delta S_{\mathrm{gen}}|_{E_0} = 0$. This proves Jacobson's entanglement equilibrium hypothesis for JT gravity [47,56,57], which thus holds in the microcanonical ensemble.

IV. QUANTUM EXTREMAL SURFACES

The extremization prescription in the QES formula (1), for eternal AdS$_2$ black holes coupled to a heat bath, thus arises from extremizing the microcanonical action I_E^{mc}, and will lead to the existence of QESs. Specifically, for $\partial\Sigma$ consisting of one endpoint in AdS$_2$ and one in flat space near the AdS boundary, the Wald entropy (8) of the latter point vanishes, since there $\phi = 0$ (no gravity) and $\chi = 0$ (Dirichlet boundary condition). We can then extremize S_{gen} (10) using (9) with respect to the first endpoint (U_1, V_1) while holding the second point (U_2, V_2) fixed. Subtracting the extremization conditions yields $U_1/V_1 = U_2/V_2$ or $t_1 = t_2$ [40]. Substituting this into S_{gen} gives a time-independent result, whereby removing the divergence near the boundary $(r_{*,2} \to 0)$, yields

$$S_{\mathrm{gen}}(r_*) = S_{\mathrm{BH}}(r_*) + \frac{c}{6}\log\left[2\sqrt{\mu}\tanh\left(-\frac{\sqrt{\mu}r_*}{2L}\right)\right] \tag{17}$$

with $r_* = r_{*,1}$. Extremizing S_{gen} with respect to r_* leads to a QES just outside the classical horizon [8,40]

$$r_{\mathrm{QES}} = \frac{2}{3}r_H\epsilon\sqrt{1 + \frac{9}{4\epsilon^2}} \approx r_H\left(1 + \frac{2\epsilon^2}{9}\right), \quad \text{or}$$

$$r_{*,\mathrm{QES}} = -\frac{L}{\sqrt{\mu}}\mathrm{arcsinh}\left(\frac{3}{2\epsilon}\right) \approx -\frac{L}{\sqrt{\mu}}\left(\log\left(\frac{3}{\epsilon}\right) + \frac{\epsilon^2}{9}\right). \tag{18}$$

The first equality is an exact expression for the QES location, and in the second equality we expanded in terms of the small parameter $\epsilon \equiv \frac{Gc}{\sqrt{\mu}\phi_r} \ll 1$, which follows from the semiclassical regime of validity $\phi(r_H)/G \gg c \gg 1$.

V. ISLANDS AND ENTROPY OF HAWKING RADIATION

We can adapt our prescription to obtain the island and Hawking saddles corresponding to the island and Hawking phases in the Page curve for eternal AdS$_2$ black holes in equilibrium with a flat space bath at temperature T_H [8] (see Fig. 2). Radiation emitted from the black hole into the bath is modeled by a CFT$_2$ at large central charge c, entirely encoded by the semiclassical Polyakov action, and is in the HH state. To maintain equilibrium, radiation entering the baths is compensated by infalling matter, such that the total entropy S of the system is the sum of the entropies S_{BH} of the two sides of the black hole, $S = 2S_{\mathrm{BH}}$. While our derivation above is in Euclidean signature, we analytically continue to Lorentzian signature below.
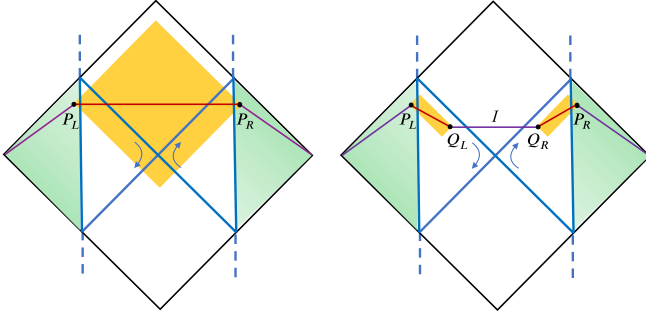
FIG. 2. Penrose diagram of AdS_2 (blue) coupled to a bath (green) displaying Hawking (left) and island phases (right). On the right, $Q_{L,R}$ are quantum extremal surfaces bounding the island I (purple). In the Hawking phase the entanglement wedge (yellow) of the black hole is the Wheeler-deWitt patch and in the island phase it consists of two identical causal diamonds.

The island formula technically computes the vN entropy of Hawking radiation using the entanglement wedge of the radiation. However, since the HH state is pure, we instead compute the action I_E^{mc} on the entanglement wedge of the black hole, as in [10,58]. The entanglement wedge of the black hole is the union of the domain of dependence of achronal surfaces with boundaries $\partial \Sigma = \mathcal{B} \cup P_R$ and $\partial \Sigma' = \mathcal{B}' \cup P_L$, where $P_{R,L}$ are points in the flat region close to the boundary and $\mathcal{B}, \mathcal{B}'$ are arbitrary points in AdS_2. After extremization, $\mathcal{B}$ becomes the QES Q_R and $\mathcal{B}'$ turns into Q_L for the island saddle, while $\mathcal{B}' = \mathcal{B} = P_L$ for the Hawking saddle. Thus, in the island phase the entanglement wedge of the black hole is given by two CDs, while in the Hawking phase the entanglement wedge is a single CD.

In the island phase we compute the on-shell action for the two identical CDs with edges $\partial \Sigma = \mathcal{B} \cup P_R$ and $\partial \Sigma' = \mathcal{B}' \cup P_L$ (right Fig. 2). We can treat each diamond separately due to large-c factorization [59,60] or because at times $t \gg t_P \gg \mu^{-1/2} L$ the vN entropy of the two CDs reduces to twice the entropy of one diamond due to an operator product expansion where $P_R(P_L)$ and $\mathcal{B}(\mathcal{B}')$ are close [8]. Focusing on the right CD, extremization of (15) fixes $t_\mathcal{B} = t_{P_R}$ (which implies $a = b$) and $\mathcal{B} = Q_R$, such that S_{gen} evaluated at the QES (18) is [40]

$$S_{gen}(r_{QES}) \approx S_{BH} + \frac{c}{6} \log(2\sqrt{\mu}) - \frac{c\epsilon}{9} + \mathcal{O}(\epsilon^2). \quad (19)$$

Here we used $S_{gen}|_{\partial \Sigma} = S_{gen}|_{Q_R}$ since at P_R ($x = \infty$) we can set $\phi = 0$ and $\chi = 0$. A similar discussion holds for the left CD. Including both CDs, since $\phi_0/G \gg \phi_r \sqrt{\mu}/G \gg c$, for $t > t_P$ we find $S_{vN}^{rad} \approx 2S_{BH}$, consistent with [8] and up to small corrections near $t \sim t_P$ [61]. Thus, we have derived the constant island phase in the Page curve, where the island is identified as the interval $[Q_L, Q_R]$.

The Hawking phase also follows from extremizing the action (15). However, since gravity is absent in the flat space bath regions, one neglects the dilaton. Extremizing

S_{vN}^{HH} with respect to (U_L, V_L) results in $U_L = V_R$ and $V_L = U_R$ (or, $t_L = -t_R \equiv -t$ and $r_{*,L} = r_{*,R} \equiv r_*$) and $r_* \ll \mu^{-1/2}L$ and $t \gg \mu^{-1/2}L$. This fixes $\mathcal{B}' = \mathcal{B} = P_L$ to be close to the left AdS boundary, such that the Polyakov term in I_E^{mc} (14) is evaluated over a Wheeler-deWitt CD (left Fig. 2). Inserting the extremal points into S_{vN}^{HH} yields a time-dependent result [8]

$$S_{vN}^{HH} = \frac{c}{3} \log \left[\frac{2L}{\delta} \frac{\cosh(\sqrt{\mu}t/L)}{\sinh(-\sqrt{\mu}r_*/L)} \right]$$

$$\approx \frac{c}{3} \frac{\sqrt{\mu}}{L} t + \cdots, \quad t \gg L/\sqrt{\mu}. \quad (20)$$

One may interpret the growth in time as arising from a sequence of Wheeler-deWitt CDs of each time slice. This linear growth leads to an information paradox [8].

Combined, we see extremizing (15) yields the Hawking and island phases of $S_{vN}^{rad}(t)$. According to the QES formula (1) the Page curve follows from a global minimization over the location of the QES, where the turnover between the Hawking and island curves occurs at the Page time $t_P = \frac{6\beta_H}{2\pi c} S_{BH} \approx \frac{3L\phi_0}{2\sqrt{\mu}Gc} \gg 1$. This global minimization is not apparent from extremizing I_E^{mc} or from the maximization condition on S_{gen} found here. This is because we are working in a microcanonical ensemble. Our choice of ensemble is akin to [62] who found the condition of dominance in the microcanonical path integrals of small eternal AdS black holes corresponds to maximizing the holographic entanglement entropy [16]. Crucially, maximizing S_{gen} with respect to the solution at fixed E_0 is consistent with the QES prescription (1) of extremizing S_{gen} with respect to the location and shape of X [62]. Thus, minimizing I_E^{mc} is consistent with the QES or island formula.

Note that maximizing S_{gen} at fixed energy suggests $S_{vN}^{rad}(E_0)$ follows a curve mirroring the Page curve, arising from the global minimization of the microcanonical action. That is, for "low" energies E_0, the entropy $S_{gen}(E_0)$ is maximal for the Hawking saddle, while at "higher" energies the entropy is maximal for the island saddle. Deriving such a curve, however, entails a more detailed knowledge of the gravitational energy of CDs.

VI. DISCUSSION

The island formula was derived for eternal and dynamical AdS_2 black holes in JT gravity using the "replica trick" [10–12,29,63,64]. It remains an open question how to generalize the derivation for other theories of gravity, including non-JT 2D dilaton theories of gravity, such as deformed JT [65], or 2D flat space analogs [66–68]. The microcanonical PI may be able to address this problem. Firstly, the above arguments for the microcanonical PI hold for generic 2D dilaton gravity models with Lagrangian

density $\mathcal{L} = L_0[Z(\phi)R + U(\phi)(\nabla\phi)^2 - V(\phi)] + \mathcal{L}_{\text{Pol}}$, for which the microcanonical on-shell action (14) generalizes to

$$I_{\text{E}}^{\text{mc}} = -\lim_{\epsilon \to 0} \int_{\partial D_\epsilon \times \partial \Sigma} ds_{\text{E}} \sqrt{\gamma} K \left[2L_0 Z(\phi) - \frac{c\chi}{12\pi} \right]. \quad (21)$$

Following the derivation above, we again find $I_{\text{E}}^{\text{mc}} = -S_{\text{W}}$ with the Wald entropy $S_{\text{W}} = 4\pi L_0 Z(\phi) - c\chi/6$. Since the Polyakov contribution may always be cast as the vN entropy, this yields $I_{\text{E}}^{\text{mc}} = -S_{\text{gen}}$ for *any* 2D dilaton gravity theory coupled to conformal matter. Then, applying our arguments here would provide a derivation of the extremization of S_{gen} for flat eternal black holes [22,23] and 2D de Sitter space [69–72], recently accomplished in [73]. This is a distinct advantage over current techniques using the replica trick.

Thus far the island formula has not been derived for theories of gravity in higher dimensions. The so-called Hilbert action boundary term method [31] or the Noether charge formalism [54] employed here holds for arbitrary theories of gravity in general spacetime dimensions [54]. In fact, our derivation of the on-shell microcanonical action in Appendix C is valid for causal diamonds in any theory of gravity. Thus, given the correct equivalent of the semi-classical Polyakov action, the microcanonical path integral may be used, in principle, to derive the island formula for higher-dimensional theories of gravity, and, correspondingly, a Page curve. An obstacle to applying our results to higher dimensions, however, is the fact that our arguments here rely on the equivalence between the Wald entropy and the generalized entropy [40]. This could be due to the fact that the Polyakov action is 1-loop exact in 2D, capturing the full effect of the conformal anomaly, which might not hold in higher dimensions.

Finally, while the replica trick derives the Page curve of Hawking radiation of dynamical black holes, it is a clear challenge for the microcanonical path integral described here. This is because our results reliably hold in equilibrium since there exists an obvious candidate for (conformal) Killing time, while a dynamical black hole lacks such a symmetry. This problem may be circumvented by instead considering the York time [74], which exists for more general spacetimes, or providing a microcanonical interpretation of dynamical horizon entropy [53,75,76].

ACKNOWLEDGMENTS

APPENDIX A: CONFORMAL ISOMETRY OF CAUSAL DIAMONDS IN AdS$_2$

We first derive the conformal isometry that preserves a causal diamond in a generic two-dimensional spacetime, for a line element in conformal gauge $d\ell^2 = -e^{2\rho} du dv$. Afterwards, we specialize to diamonds in AdS$_2$. Since the property of a conformal isometry is invariant under a Weyl rescaling of the metric, we can leave off the conformal factor and study the conformal Killing vectors of the line element $du dv$, which take the general form [47,48]

$$\zeta = A(u)\partial_u + B(v)\partial_v. \quad (A1)$$

Suppose the causal diamond consists of the intersection of the regions $[u - u_0 = -a, u - u_0 = a]$ and $[v - v_0 = -b, v - v_0 = b]$. The maximal spatial slice Σ is described by $u - u_0 = -\sqrt{(v - v_0)^2 + a^2 - b^2}$ and the line between the future and past vertices is given by $u - u_0 = \sqrt{(v - v_0)^2 + a^2 - b^2}$. The conformal Killing vector that preserves the diamond is only a function of the distance $u - u_0$ and $v - v_0$, i.e., $A = A(u - u_0)$ and $B = B(v - v_0)$. To map the diamond onto itself ζ must be tangent to the null generators on the null boundaries $u = u_0 \pm a$ and $v = v_0 \pm b$, which implies $A(\pm a) = 0$ and $B(\pm b) = 0$. Hence, $A(y) = g_a(y)[h(a) - h(y)]$ and $B(y) = m_b(y)[n(b) - n(y)]$, with $h(y) = h(-y)$ and $n(y) = n(-y)$. In addition, the flow of ζ must respect the reflection symmetries across the line between the future and past vertices and across Σ, when a and b are interchanged. In particular, replacing $u - u_0 \leftrightarrow v - v_0$ and $a \leftrightarrow b$ leaves the vector field unchanged $\zeta \to \zeta$, so $m_b(y) = g_b(y)$ and $n(y) = h(y)$. Further, if $u - u_0 \leftrightarrow -(v - v_0)$ and $a \leftrightarrow b$ we must have $\zeta \to -\zeta$, yielding $g(y) = g(-y)$. Therefore, the conformal isometry of a causal diamond is generated by

$$\zeta = A_a(u - u_0)\partial_u + A_b(v - v_0)\partial_v, \quad \text{with}$$
$$A_a(y) = g_a(y)[h(a) - h(y)], \quad g(y) = g(-y), \quad h(y) = h(-y),$$
$$(A2)$$

which holds for a generic two-dimensional spacetime. Further, the past and future null boundaries of the diamond

are conformal Killing horizons, since ζ becomes null on these boundaries. Due to the different length scales a and b, there are two (positive) surface gravities

$$\kappa_a = g_a(a)h'(a) \quad \text{and} \quad \kappa_b = g_b(b)h'(b), \quad \text{(A3)}$$

defined via $\nabla_\mu \zeta^2 = -2\kappa\zeta_\mu$ [50] evaluated on the future null boundaries $u - u_0 = a$ and $v - v_0 = b$, respectively.

Next we place the causal diamond in AdS_2 space, for which the line element in null coordinates $u = t - r_*$ and $v = t + r_*$ reads

$$d\ell^2 = -e^{2\rho}dudv, \qquad e^{2\rho} = \frac{\mu}{\sinh^2\left[\frac{\sqrt{\mu}}{2L}(v-u)\right]}. \quad \text{(A4)}$$

As a special case, we consider a Rindler wedge in AdS space which has the shape of a half causal diamond, and becomes a full diamond when AdS is glued to a Minkowski patch at the conformal boundary. We require that the conformal isometry of a generic diamond becomes the boost isometry of AdS-Rindler space, if the future and past vertices of the diamond are both located on the AdS boundary, i.e., if $b = a$ and $r_{*,0} \equiv \frac{1}{2}(v_0 - u_0) = 0$ or $t_0 \equiv \frac{1}{2}(v_0 + u_0) = v_0 = u_0$. The boost Killing vector of AdS-Rindler space is [40]

$$\xi = A(u - t_0)\partial_u + A(v - t_0)\partial_v, \quad \text{with}$$

$$A(y) = \frac{L\kappa/\sqrt{\mu}}{\sinh(\sqrt{\mu}a/L)}[\cosh(\sqrt{\mu}a/L) - \cosh(\sqrt{\mu}y/L)],$$

(A5)

for which the surface gravities coincide $\kappa_a = \kappa_b = \kappa$. Comparing (A2) and (A5), we see the requirement $\zeta \to \xi$, as $b \to a$ and $v_0, u_0 \to t_0$, restricts the functions in the conformal Killing vector (A2) to be

$$g_a(y) = \frac{L\kappa_a/\sqrt{\mu}}{\sinh(\sqrt{\mu}a/L)}, \qquad h(y) = \cosh(\sqrt{\mu}y/L), \quad \text{(A6)}$$

and similarly for $g_b(y)$. Thus, by the special case of the AdS-Rindler (half) diamond, the conformal isometry of a causal diamond in AdS_2 is uniquely fixed to be

$$\zeta = A_a(u - u_0)\partial_u + A_b(v - v_0)\partial_v, \quad \text{with}$$

$$A_a(y) = \frac{L\kappa_a/\sqrt{\mu}}{\sinh(\sqrt{\mu}a/L)}[\cosh(\sqrt{\mu}a/L) - \cosh(\sqrt{\mu}y/L)].$$

(A7)

For $a = b$ and $\mu = 1$ this is equivalent to the conformal Killing vector of a spherically symmetric causal diamond in higher-dimensional de Sitter spacetime [48].

APPENDIX B: DIAMOND UNIVERSE COORDINATES IN AdS₂

We can cover a causal diamond with the inextendible coordinates (s, x), introduced in [48] (see also [51]). This coordinate system is adapted to the flow of the conformal Killing vector ζ that preserves the diamond. In particular, the coordinate $s \in [-\infty, \infty]$ is the conformal Killing time, defined as the function that satisfies $\zeta \cdot ds = 1$ and, for $a = b$, $s = 0$ on the maximal slice Σ ($t = t_0$). Further, the coordinate $x \in [-\infty, \infty]$ is a spatial coordinate and satisfies $\zeta \cdot dx = 0$ and $|dx| = |ds|$ and, for $a = b$ we have $x = 0$ at $r_* = r_{*,0}$. It follows from these conditions that the two-dimensional line element in "diamond universe" coordinates is

$$d\ell^2 = C^2(s, x)(-ds^2 + dx^2) = -C^2(\bar{u}, \bar{v})d\bar{u}d\bar{v}. \quad \text{(B1)}$$

where $\bar{u} = s - x$ and $\bar{v} = s + x$ are null coordinates. The null boundaries of the diamond are at $\bar{u} = \pm\infty$ ($u - u_0 = \pm a$) and $\bar{v} = \pm\infty$ ($v - v_0 = \pm b$), where the plus signs corresponds to the future horizon and the minus signs to the past horizon. On any constant x slice the vertices are located at $s = \pm\infty$, and on a constant s slice the two edges are at $x = \pm\infty$. Furthermore, in these coordinates $\zeta = \partial_s = \partial_{\bar{u}} + \partial_{\bar{v}}$, which should be equivalent to the conformal Killing vector in (A2) where the functions are given by (A6) in AdS_2. From the equality of these two expressions for ζ we obtain the transformation from the null coordinates (u, v) to $(\bar{u}, \bar{v})$

$$e^{\kappa_a \bar{u}} = \frac{\sinh\left[\frac{\sqrt{\mu}}{2L}(a + u - u_0)\right]}{\sinh\left[\frac{\sqrt{\mu}}{2L}(a - u + u_0)\right]},$$

$$e^{\kappa_b \bar{v}} = \frac{\sinh\left[\frac{\sqrt{\mu}}{2L}(b + v - v_0)\right]}{\sinh\left[\frac{\sqrt{\mu}}{2L}(b - v + v_0)\right]}, \quad \text{(B2)}$$

and the inverse transformation

$$e^{\sqrt{\mu}(u-u_0)/L} = \frac{\cosh\left[(\sqrt{\mu}a/L + \kappa_a\bar{u})/2\right]}{\cosh\left[(\sqrt{\mu}a/L - \kappa_a\bar{u})/2\right]},$$

$$e^{\sqrt{\mu}(v-v_0)/L} = \frac{\cosh\left[(\sqrt{\mu}b/L + \kappa_b\bar{v})/2\right]}{\cosh\left[(\sqrt{\mu}b/L - \kappa_b\bar{v})/2\right]}. \quad \text{(B3)}$$

Kruskal coordinates $(T, X) = (\frac{1}{2}(V + U), \frac{1}{2}(V - U))$, with $(U, V) = (-\frac{L}{\sqrt{\mu}}e^{-\frac{\sqrt{\mu}u}{L}}, \frac{L}{\sqrt{\mu}}e^{\frac{\sqrt{\mu}v}{L}})$, are convenient to visualize the Euclideanized diamond [51].

By comparing the metrics (A4) and (B1), and using the transformation in (B3), we find

$$C^2(\bar{u}, \bar{v}) = 4\kappa_a\kappa_b L^2 (e^{\frac{\sqrt{\mu}}{L}2a} - 1)(e^{\frac{\sqrt{\mu}}{L}2b} - 1)$$
$$\times e^{\frac{\sqrt{\mu}}{L}2r_{*,0} + \kappa_a\bar{u} + \kappa_b\bar{v}} [(e^{\frac{\sqrt{\mu}}{L}b} + e^{\kappa_b\bar{v}})(e^{\frac{\sqrt{\mu}}{L}a + \kappa_a\bar{u}} + 1)$$
$$- e^{\frac{\sqrt{\mu}}{L}2r_{*,0}}(e^{\frac{\sqrt{\mu}}{L}a} + e^{\kappa_a\bar{u}})(e^{\frac{\sqrt{\mu}}{L}b + \kappa_b\bar{v}} + 1)]^{-2}, \quad \text{(B4)}$$

where $r_{*,0} \equiv \frac{1}{2}(v_0 - u_0)$. Note for $r_{*,0} = 0$ and $a = b$ we recover AdS-Rindler space since the conformal factor becomes $C^2 = \kappa^2 L^2 / \sinh^2[\frac{\kappa}{2}(\bar{v} - \bar{u})]$, where $\kappa = \kappa_a = \kappa_a$.

For a generic diamond there exist two different surface gravities, associated to the two parts of the future conformal Killing horizons, $\bar{u} = \infty$ and $\bar{v} = \infty$, given by

$$\kappa_a = -C^{-2}\partial_{\bar{u}}C^2|_{\bar{u}\to\infty}, \qquad \kappa_b = -C^{-2}\partial_{\bar{v}}C^2|_{\bar{v}\to\infty}, \quad \text{(B5)}$$

where we applied the definition $\nabla_\mu\zeta^2 = -2\kappa\zeta_\mu$ [50]. As shown in [48], surface gravities satisfying this definition are constant on a bifurcate conformal Killing horizon, so κ_a and κ_b are constant. As usual, these surface gravities can be interpreted as the temperatures corresponding to the a and b portions of the conformal Killing horizon. This is because the Hartle-Hawking state, satisfying $\langle\text{HH}|:T_{uu}^\chi:|\text{HH}\rangle = \frac{c\pi}{12}T_\text{H}^2$ with $T_\text{H} = \frac{\sqrt{\mu}}{2\pi L}$, is also thermal with respect to the $(\bar{u}, \bar{v})$ coordinates

$$\langle\text{HH}|:T_{\bar{u}\bar{u}}^\chi:|\text{HH}\rangle = \left(\frac{du}{d\bar{u}}\right)^2 \frac{c\pi}{12}T_\text{H}^2 - \frac{c}{24\pi}\{u, \bar{u}\}$$
$$= \frac{c\pi}{12}\left(\frac{\kappa_a}{2\pi}\right)^2 = \frac{c\pi}{12}T_a^2, \quad \text{(B6)}$$

where we used the anomalous transformation law for the normal-ordered stress tensor, and inserted (B3). A similar result holds for $:T_{\bar{v}\bar{v}}^\chi:$, i.e., its expectation value in $|\text{HH}\rangle$ is thermal with temperature $T_b = \kappa_b/2\pi$. This was already known for the special case of AdS-Rindler [40], but here we showed the HH state in AdS_2 is thermal with respect to any causal diamond. The temperature of the CD thus seems positive and finite, in contrast to the negative temperature interpretation in [48,55] and the infinite temperature claim in [51].

Near the bifurcation points $x = \pm\infty$ of the conformal Killing horizons of the diamond the surface gravity satisfies the relation $\nabla_\mu\zeta_\nu = \kappa n_{\mu\nu}$ [48], where $n_{\mu\nu} = 2u_{[\mu}n_{\nu]}$ is the outward and future pointing binormal, with $u = C^{-1}\partial_s$ is the future pointing timelike unit normal and $n = \pm C^{-1}\partial_x$ is the outward pointing spacelike unit normal at $x = \pm\infty$. The surface gravity can be computed to be $\kappa = \mp\frac{1}{2}C^{-2}\partial_x C^2|_{x=\pm\infty} = (\kappa_a + \kappa_b)/2$. This expression can also be obtained from the periodicity of the Euclidean time. In the Euclideanized diamond spacetime the conformal Killing horizon maps to punctures at $x \to \pm\infty$. Near $x \to \pm\infty$ the diamond universe line element (B1) becomes

$$d\ell^2 \approx 4L^2\kappa_a\kappa_b e^{\mp(\kappa_a+\kappa_b)x}(-ds^2 + dx^2), \quad \text{(B7)}$$

which is simply flat Rindler space $d\ell^2 = -\kappa^2\varrho^2 ds^2 + d\varrho^2$ for the coordinate $\varrho \equiv 4L\sqrt{\kappa_a\kappa_b}(\kappa_a + \kappa_b)^{-1}e^{\mp(\kappa_a+\kappa_b)x/2}$, and identifying the surface gravity $\kappa = (\kappa_a + \kappa_b)/2$. Note the null boundaries $x = \pm\infty$ map to the Rindler horizon $\varrho = 0$. Upon Wick rotating $s \to -is_\text{E}$, removing the conical singularity in the Euclidean spacetime located at the horizon, has us periodically identify $s_\text{E} \sim s_\text{E} + 2\pi/\kappa$.

APPENDIX C: MICROCANONICAL ACTION IN THE NOETHER CHARGE FORMALISM

In the main text we obtained the on-shell microcanonical action (15) using the Hilbert action surface term method [31,51]. However, the microcanonical action can be derived using various methods. Brown [77] showed an equivalence between the GHY surface action [31] and the microcanonical action developed by Brown and York [30]. Later, both actions were expressed in the Noether charge formalism [41,54], and shown to be equivalent to the Noether charge. These methods are typically applied to black hole spacetimes, but here we use them to define the microcanonical action for causal diamonds. Specifically, employing the Noether charge formalism, below we define the off-shell Euclidean microcanonical action of CDs, and show it equals (14) and (15) on shell.

We consider generic semiclassical 2D dilaton gravity theories with Lagrangian 2-form $L = \epsilon L_0[RZ(\phi) + U(\phi)(\nabla\phi)^2 - V(\phi)] + L_\text{Pol}$, with ϵ the spacetime volume form on the Lorentzian CD spacetime M^CD. We foliate the CD with spacelike slices Σ_s, labeled by the conformal Killing time s, which smoothly intersect the bifurcation points $\partial\Sigma$ of the null boundaries. The Euclidean diamond spacetime M_E^CD is defined by periodically identifying the Euclidean time $s_\text{E} = is$ with $2\pi/\kappa$ to avoid a conical singularity at $\partial\Sigma$. Motivated by [30,54], we define the off-shell microcanonical Euclidean action

$$I_\text{E}^\text{mc} \equiv -i\left[\int_{M_\text{E}^\text{CD}} L - \int_{M_\text{E}^\text{CD}} ds \wedge \theta(\psi, \mathcal{L}_\zeta\psi)\right], \quad \text{(C1)}$$

where θ is the symplectic potential 1-form, and the dynamical fields are $\psi = (g_{\mu\nu}, \phi, \chi)$. Note $\theta(\psi, \mathcal{L}_\zeta\psi)$ here is non-vanishing since ζ is a conformal Killing vector instead of a Killing vector. Writing $L = ds \wedge \zeta \cdot L$, we see the two terms between brackets combine into an integral over the Noether current 1-form $j_\zeta \equiv \theta(\psi, \mathcal{L}_\zeta\psi) - \zeta \cdot L$ associated with diffeomorphisms generated by ζ. Using the on-shell identity $j_\zeta = dQ_\zeta$, with Q_ζ the Noether charge 0-form, and applying Stokes' theorem we find the on-shell Euclidean microcanonical action for CDs is equal to

$$I_{\mathrm{E}}^{\mathrm{mc}} = \int_{\partial M_{\mathrm{E}}^{\mathrm{CD}}} ds_{\mathrm{E}} \wedge Q_\zeta, \tag{C2}$$

where the total Noether charge for semiclassical 2D dilaton gravity is $Q_\zeta = Q_\zeta^\phi + Q_\zeta^{\mathrm{Pol}}$ with [40]

$$Q_\zeta^\phi = -L_0 \epsilon_{\mu\nu}[Z(\phi)\nabla^\mu\zeta^\nu + 2\zeta^\mu\nabla^\nu Z(\phi)],$$
$$Q_\zeta^{\mathrm{Pol}} = \frac{c}{24\pi}\epsilon_{\mu\nu}[\chi\nabla^\mu\zeta^\nu + 2\zeta^\mu\nabla^\nu\chi]. \tag{C3}$$

where $\epsilon_{\mu\nu} = -n_{\mu\nu}$ is the binormal volume form on $\partial\Sigma$, with $n_{\mu\nu}$ as the outward and future pointing binormal. Since $\partial M_{\mathrm{E}}^{\mathrm{CD}}$ has topology $S^1 \times \partial\Sigma$ we restrict the Noether charge to $\partial\Sigma$, where $\zeta|_{\partial\Sigma} = 0$ and $\nabla_\mu\zeta_\nu|_{\partial\Sigma} = \kappa n_{\mu\nu}$, hence it becomes $Q_\zeta|_{\partial\Sigma} = -\frac{\kappa}{2\pi}[4\pi L_0 Z(\phi) - \frac{c}{6}\chi]$. Importantly, $Q_\zeta|_{\partial\Sigma}$ is independent from s_{E}, because ϕ, χ are constant in the limit $x \to \pm\infty$. Thus, we can integrate out the Euclidean time, and obtain that the on-shell microcanonical action is equal to minus the Wald entropy

$$I_{\mathrm{E}}^{\mathrm{mc}} = \frac{2\pi}{\kappa}\oint_{\partial\Sigma} Q_\zeta = -S_{\mathrm{W}}|_{\partial\Sigma}, \tag{C4}$$

where $S_{\mathrm{W}} = 4\pi L_0 Z(\phi) - c\chi/6$. This shows the microcanonical action is proportional to the Noether charge and it proves (15), given $S_{\mathrm{W}} = S_{\mathrm{gen}}$.

In a proper microcanonical ensemble the action is minimized at fixed energy E_0. The definition of E_0 in our setup may therefore be identified from the variation of the on-shell action on $M_{\mathrm{E}}^{\mathrm{CD}}$

$$\delta I_{\mathrm{E}}^{\mathrm{mc}} = \int_{M_{\mathrm{E}}^{\mathrm{CD}}} ds_{\mathrm{E}} \wedge \omega(\psi, \delta\psi, \mathcal{L}_\zeta\psi) = \int_{S^1} ds_{\mathrm{E}}\delta H_\zeta, \tag{C5}$$

where the symplectic current 1-form is defined as $\omega(\psi, \delta_1\psi, \delta_2\psi) \equiv \delta_1\theta(\psi, \delta_2\psi) - \delta_2\theta(\psi, \delta_1\psi)$, and in the second equality we inserted $\delta H_\zeta = \int_{\Sigma_{s_{\mathrm{E}}}} \omega(\psi, \delta\psi, \mathcal{L}_\zeta\psi)$, the variation of the Hamiltonian generating evolution along the flow of ζ. The first equality follows from varying (C1) and using $\delta L = E^\psi\delta\psi + d\theta(\psi, \delta\psi)$, where $E^\psi = 0$ are the field equations for ψ, and inserting Cartan's magic formula $\zeta \cdot d\theta = \mathcal{L}_\zeta\theta - d(\zeta \cdot \theta)$. The term $\int_{\partial M_{\mathrm{E}}^{\mathrm{CD}}} ds \wedge \zeta \cdot \theta$ vanishes because ζ is zero at the edge $\partial\Sigma$. Thus, we see from (C5) the action $I_{\mathrm{E}}^{\mathrm{mc}}$ is minimized at fixed energy $E_0 \sim H_\zeta$ up to a sign and constant. We set the constant to zero and the sign

is fixed by comparing to the first law $\frac{\kappa}{2\pi}\delta S_{\mathrm{W}} = -\delta H_\zeta$ such that $E_0 = -H_\zeta$.

Lastly, let us comment on the equivalence between the on-shell microcanonical action (C4) and the GHY surface term used in the main text

$$I_{\mathrm{E}}^{\mathrm{mc}} = -\lim_{\epsilon\to 0}\int_{\partial D_\epsilon\times\mathcal{H}} ds_{\mathrm{E}}\sqrt{\gamma}K\left[2L_0 Z(\phi) - \frac{c\chi}{12\pi}\right]. \tag{C6}$$

where D_ϵ is an infinitesimal disk of radius ϵ orthogonal to the punctures $\partial\Sigma : x = \pm\infty$ in the Euclideanized diamond. While initially obscure, the equivalence of (C4) and (C6) follows from the fact that, on-shell, $H_\zeta = 0$ on the bifurcation surface and that the GHY boundary term is independent of s on $\partial\Sigma$ [54]. More explicitly, the Hamiltonian H_ζ for a theory which fixes the induced metric of the boundary ∂M of a (Lorentzian) manifold M is generically given by an integral over the codimension-2 slices $\mathcal{C}_s$ where Σ_s orthogonally intersects ∂M [40] (see also [54,78])

$$H_\zeta = \int_{\mathcal{C}_s}(Q_\zeta - \zeta \cdot b) = \oint_{\mathcal{C}_s}\epsilon_{\partial\Sigma}N\varepsilon. \tag{C7}$$

Here $b = \epsilon_B K[-2L_0 Z(\phi) + \frac{c}{12\pi}\chi]$ is the GHY boundary term 1-form, $\varepsilon = -2L_0 n^\alpha\nabla_\alpha Z(\phi)$ is the quasi-local energy density, and $N = -\zeta^\mu u_\mu$ is the lapse. Crucially, on bifurcation points $\partial\Sigma$, the lapse $N = 0$ such that $H_\zeta = 0$. Then, for a 1-parameter family of surfaces $(\partial\Sigma)_\epsilon$ in $\Sigma_{s_{\mathrm{E}}}$, where $\lim_{\epsilon\to 0}(\partial\Sigma)_\epsilon \to \partial\Sigma$, we have

$$\lim_{\epsilon\to 0}\int_{(\partial\Sigma)_\epsilon} Q_\zeta = \lim_{\epsilon\to 0}\int_{(\partial\Sigma)_\epsilon} \zeta \cdot b. \tag{C8}$$

Thus, by the definition $S_{\mathrm{W}} = -\frac{2\pi}{\kappa}\int_{\partial\Sigma} Q_\zeta$ of the Wald entropy, we find

$$S_{\mathrm{W}} = -\lim_{\epsilon\to 0}\frac{2\pi}{\kappa}\int_{(\partial\Sigma)_\epsilon} \zeta \cdot b = -\lim_{\epsilon\to 0}\int_{\partial D_\epsilon\times\partial\Sigma} ds_E \wedge \zeta \cdot b$$

$$= \lim_{\epsilon\to 0}\int_{\partial D_\epsilon\times\partial\Sigma} ds_E\sqrt{\gamma}K\left[2L_0 Z(\phi) - \frac{c\chi}{12\pi}\right], \tag{C9}$$

where $2\pi/\kappa$ was replaced with the integral $\int_{\partial D_\epsilon} ds_{\mathrm{E}}$ and we used that the GHY term b is independent of s_{E} on $(\partial\Sigma)_\epsilon$. This shows (C4) and (C6) are equal, which establishes the Hilbert action surface formula (14).

[1] S. W. Hawking, Nature (London) **248,** 30 (1974).
[2] S. W. Hawking, Commun. Math. Phys. **43,** 199 (1975); **46,** 206(E) (1976).
[3] S. W. Hawking, Phys. Rev. D **14,** 2460 (1976).
[4] D. N. Page, Phys. Rev. Lett. **71,** 3743 (1993).
[5] D. N. Page, J. Cosmol. Astropart. Phys. 09 (2013) 028.
[6] G. Penington, J. High Energy Phys. 09 (2020) 002.
[7] A. Almheiri, N. Engelhardt, D. Marolf, and H. Maxfield, J. High Energy Phys. 12 (2019) 063.
[8] A. Almheiri, R. Mahajan, and J. Maldacena, arXiv:1910 .11077.
[9] A. Almheiri, R. Mahajan, J. Maldacena, and Y. Zhao, J. High Energy Phys. 03 (2020) 149.
[10] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, J. High Energy Phys. 05 (2020) 013.
[11] G. Penington, S. H. Shenker, D. Stanford, and Z. Yang, J. High Energy Phys. 03 (2022) 205.
[12] K. Goto, T. Hartman, and A. Tajdini, J. High Energy Phys. 04 (2021) 289.
[13] T. J. Hollowood, S. Prem Kumar, and A. Legramandi, J. Phys. A **53,** 475401 (2020).
[14] T. J. Hollowood and S. P. Kumar, J. High Energy Phys. 08 (2020) 094.
[15] X. Wang, R. Li, and J. Wang, Phys. Rev. D **103,** 126026 (2021).
[16] V. E. Hubeny, M. Rangamani, and T. Takayanagi, J. High Energy Phys. 07 (2007) 062.
[17] T. Faulkner, A. Lewkowycz, and J. Maldacena, J. High Energy Phys. 11 (2013) 074.
[18] N. Engelhardt and A. C. Wall, J. High Energy Phys. 01 (2015) 073.
[19] J. D. Bekenstein, Phys. Rev. D **9,** 3292 (1974).
[20] A. C. Wall, Phys. Rev. D **82,** 124019 (2010).
[21] A. C. Wall, Phys. Rev. D **85,** 104049 (2012); **87,** 069904(E) (2013).
[22] F. F. Gautason, L. Schneiderbauer, W. Sybesma, and L. Thorlacius, J. High Energy Phys. 05 (2020) 091.
[23] T. Hartman, E. Shaghoulian, and A. Strominger, J. High Energy Phys. 07 (2020) 022.
[24] H. Z. Chen, R. C. Myers, D. Neuenfeld, I. A. Reyes, and J. Sandor, J. High Energy Phys. 12 (2020) 025.
[25] S. He, Y. Sun, L. Zhao, and Y.-X. Zhang, arXiv:2110.07598.
[26] A. Lewkowycz and J. Maldacena, J. High Energy Phys. 08 (2013) 090.
[27] X. Dong, A. Lewkowycz, and M. Rangamani, J. High Energy Phys. 11 (2016) 028.
[28] X. Dong, Nat. Commun. **7,** 12472 (2016).
[29] X. Dong and A. Lewkowycz, J. High Energy Phys. 01 (2018) 081.
[30] J. D. Brown and J. W. York, Jr., Phys. Rev. D **47,** 1420 (1993).
[31] M. Banados, C. Teitelboim, and J. Zanelli, Phys. Rev. Lett. **72,** 957 (1994).
[32] H. Casini, M. Huerta, and R. C. Myers, J. High Energy Phys. 05 (2011) 036.
[33] G. W. Gibbons and S. W. Hawking, Phys. Rev. D **15,** 2752 (1977).
[34] R. Jackiw, Nucl. Phys. **B252,** 343 (1985).
[35] C. Teitelboim, Phys. Lett. **126B,** 41 (1983).

[36] A. Almheiri and J. Polchinski, J. High Energy Phys. 11 (2015) 014.
[37] S. M. Christensen and S. A. Fulling, Phys. Rev. D **15,** 2088 (1977).
[38] J. B. Hartle and S. W. Hawking, Phys. Rev. D **13,** 2188 (1976).
[39] M. Spradlin and A. Strominger, J. High Energy Phys. 11 (1999) 021.
[40] J. F. Pedraza, A. Svesko, W. Sybesma, and M. R. Visser, J. High Energy Phys. 12 (2021) 134.
[41] R. M. Wald, Phys. Rev. D **48,** R3427 (1993).
[42] R. C. Myers, Phys. Rev. D **50,** 6412 (1994).
[43] J. D. Brown and J. W. York, Jr., in The Black Hole 25 Years After (1994), pp. 1–24, arXiv:gr-qc/9405024.
[44] B. Czech, J. L. Karczmarek, F. Nogueira, and M. Van Raamsdonk, Classical Quantum Gravity **29,** 155009 (2012).
[45] A. C. Wall, Classical Quantum Gravity **31,** 225007 (2014).
[46] M. Headrick, V. E. Hubeny, A. Lawrence, and M. Rangamani, J. High Energy Phys. 12 (2014) 162.
[47] T. Jacobson, Phys. Rev. Lett. **116,** 201101 (2016).
[48] T. Jacobson and M. Visser, SciPost Phys. **7,** 079 (2019).
[49] M. R. Visser, Emergent gravity in a holographic universe, Ph.D. thesis, Amsterdam U., (2019).
[50] T. Jacobson and G. Kang, Classical Quantum Gravity **10,** L201 (1993).
[51] T. Banks, P. Draper, and S. Farkas, Phys. Rev. D **103,** 106022 (2021).
[52] T. Jacobson, Phys. Rev. D **50,** R6031 (1994).
[53] V. Iyer and R. M. Wald, Phys. Rev. D **50,** 846 (1994).
[54] V. Iyer and R. M. Wald, Phys. Rev. D **52,** 4430 (1995).
[55] T. Jacobson and M. Visser, Int. J. Mod. Phys. D **28,** 1944016 (2019).
[56] P. Bueno, V. S. Min, A. J. Speranza, and M. R. Visser, Phys. Rev. D **95,** 046003 (2017).
[57] N. Callebaut, J. High Energy Phys. 02 (2019) 153.
[58] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, Rev. Mod. Phys. **93,** 035002 (2021).
[59] T. Hartman, arXiv:1303.6955.
[60] A. Rolph, SciPost Phys. **12,** 079 (2022).
[61] C. Akers and G. Penington, J. High Energy Phys. 04 (2021) 062.
[62] D. Marolf, J. High Energy Phys. 09 (2018) 114.
[63] S. Colin-Ellerin, X. Dong, D. Marolf, M. Rangamani, and Z. Wang, J. High Energy Phys. 05 (2021) 117.
[64] S. Colin-Ellerin, X. Dong, D. Marolf, M. Rangamani, and Z. Wang, J. High Energy Phys. 08 (2021) 171.
[65] E. Witten, arXiv:2006.03494.
[66] C. G. Callan, Jr., S. B. Giddings, J. A. Harvey, and A. Strominger, Phys. Rev. D **45,** R1005 (1992).
[67] J. G. Russo, L. Susskind, and L. Thorlacius, Phys. Rev. D **46,** 3444 (1992).
[68] S. Bose, L. Parker, and Y. Peleg, Phys. Rev. Lett. **76,** 861 (1996).
[69] Y. Chen, V. Gorbenko, and J. Maldacena, J. High Energy Phys. 02 (2021) 009.
[70] T. Hartman, Y. Jiang, and E. Shaghoulian, J. High Energy Phys. 11 (2020) 111.

[71] V. Balasubramanian, A. Kar, and T. Ugajin, J. High Energy Phys. 02 (2021) 072.

[72] W. Sybesma, Classical Quantum Gravity **38**, 145012 (2021).

[73] A. Svesko, E. Verheijden, E. P. Verlinde, and M. R. Visser, arXiv:2203.00700.

[74] J. W. York, Jr., Phys. Rev. Lett. **28**, 1082 (1972).

[75] S. A. Hayward, S. Mukohyama, and M. C. Ashworth, Phys. Lett. A **256**, 347 (1999).

[76] A. C. Wall, Int. J. Mod. Phys. D **24**, 1544014 (2015).

[77] J. D. Brown, Phys. Rev. D **52**, 7011 (1995).

[78] D. Harlow and J.-Q. Wu, J. High Energy Phys. 10 (2020) 146.