



# Solvable models of quantum black holes: a review on Jackiw–Teitelboim gravity

Thomas G. Mertens<sup>1</sup> · Gustavo J. Turiaci<sup>2,3</sup>

Received: 30 January 2023 / Accepted: 21 June 2023  
© The Author(s) 2023

## Abstract

We review recent developments in Jackiw–Teitelboim gravity. This is a simple solvable model of quantum gravity in two dimensions (that arises e.g. from the s-wave sector of higher dimensional gravity systems with spherical symmetry). Due to its solvability, it has proven to be a fruitful toy model to analyze important questions such as the relation between black holes and chaos, the role of wormholes in black hole physics and holography, and the way in which information that falls into a black hole can be recovered.

**Keywords** Quantum gravity · Solvable models · Black holes · Jackiw–Teitelboim gravity

## Contents

1	Introduction	3
2	Classical Jackiw–Teitelboim gravity	4
2.1	Dilaton gravity models	5
2.1.1	First-order formulation	7
2.1.2	Motivation: near-extremal black holes	9
2.2	Classical solutions	12
2.2.1	Metric solution	12
2.2.2	Dilaton solution	16
2.3	Boundary conditions and Schwarzian dynamics	19
2.3.1	Real-time derivation	19

---

✉ Thomas G. Mertens  
thomas.mertens@ugent.be  
Gustavo J. Turiaci  
turiaci@uw.edu

<sup>1</sup> Department of Physics and Astronomy, Ghent University, Krijgslaan, 281-S9, 9000 Ghent, Belgium

<sup>2</sup> Institute for Advanced Study, Princeton, NJ, USA

<sup>3</sup> Physics Department, University of Washington, Seattle, WA, USA

2.3.2	Aside: general dilaton gravity	24
2.3.3	The Schwarzian action	25
2.3.4	Geometric derivation from the action	28
2.4	Quantum matter in classical gravity	31
2.4.1	Application: Hawking–Unruh effect and information loss	32
3	Quantum Jackiw–Teitelboim gravity	35
3.1	Spectrum of quantum black holes	35
3.1.1	Perturbative calculation	36
3.2	Quantum Jackiw–Teitelboim gravity coupled to matter	38
3.3	Correlators	40
3.3.1	Path integral representation	40
3.3.2	A first approach: Schwarzian perturbation theory	42
3.3.3	Exact two-point function	43
3.3.4	Exact four-point function	47
3.3.5	Application: quantum chaos	49
3.4	Diagrammatic rules for exact correlators	54
3.5	Pure states and end-of-the-world (EOW) branes	55
3.6	Other operator insertions in JT gravity	57
3.7	Outline of derivations in the literature	61
3.7.1	Free particle approach	61
3.7.2	Limit of Liouville theory	62
3.7.3	Boundary particle approach	64
3.7.4	Two-dimensional gauge theory	64
4	Spacetime wormholes and random matrices	67
4.1	Motivation: information loss and late time decay	67
4.2	Multiboundary higher genus amplitudes	69
4.2.1	The disk $Z_{0,1}(\beta)$	71
4.2.2	The cylinder $Z_{0,2}(\beta_1, \beta_2)$	71
4.2.3	The general case $Z_{g,n}(\beta_1, \dots, \beta_n)$	73
4.3	JT and random matrices	77
4.3.1	SSS duality	78
4.3.2	Double-scaling limit of matrix integrals	80
4.3.3	Derivation of the duality	82
4.4	Non-perturbative effects in topological expansion	83
4.4.1	Density of states	84
4.4.2	Late time decay of spectral form factor	86
4.5	Generalization	89
4.5.1	Other ensembles	89
4.5.2	2d dilaton gravity with general potential	90
4.5.3	JT gravity coupled to gauge fields	92
4.5.4	JT gravity coupled to matter	93
5	Applications and future directions	95
5.1	The entropy of Hawking radiation	95
5.1.1	Entanglement islands in JT gravity	96
5.1.2	Replica wormholes	99
5.1.3	Replica wormholes and EOW branes	101
5.2	Factorization, discreteness, and ensemble averaging in gravity	103
5.3	Near-extremal black holes	105
5.4	Supersymmetric JT	105
5.5	Two-dimensional cosmology	106
5.6	Traversable wormholes	107
5.7	Finite cut-off and $T\bar{T}$	109
5.8	Volume of black hole interior and complexity	110
5.9	Bulk correlators and observables	111
5.10	Liouville gravity and minimal string	112
5.11	Universe field theory and quantum chaos	113
	References	115

## 1 Introduction

Understanding quantum gravity and quantum black hole physics are some of the most pressing open problems in contemporary theoretical physics. In particular, deep questions arise in the problem of black hole formation and subsequent evaporation, starting with Hawking's work in the 1970s (Hawking 1975, 1976).

The road towards quantum gravity, starting with the problem of non-renormalizability of pure Einstein–Hilbert gravity in 3+1d, has in the years led us through higher dimensions, string theory, compactification, branes.... And some significant successes have been made using this approach (Strominger and Vafa 1996). However, it is safe to say that we do not have a fully satisfying understanding, and any alternative approaches that could shed new light on the problem would be most welcome.

With this goal in mind, an alternative strategy toward quantum gravity would be to work in lower dimensions (2d or 3d), where gravitational models can make sense at the level of the Euclidean path integral. If we furthermore do this in the framework of holography, then we have a preferred anchoring point, and are guided by the major breakthroughs in that field throughout the years (Maldacena 1999).

Even within this strategy, finding interesting lower-dimensional solvable models of gravity is an art on its own. Few such models exist, but the degree of solvability that we can obtain in 2d Jackiw–Teitelboim gravity is unprecedented as we aim to explain with this review. If we work in two spacetime dimensions, the simplest candidate model would be Einstein–Hilbert gravity

$$S = \frac{1}{16\pi G_N} \int d^2x \sqrt{-g} R + \dots \quad (1.1)$$

However, this model is topological since its Euclidean action is just the Euler characteristic, and the Einstein tensor vanishes identically. This model does play an important role as the 2d gravity on the string worldsheet, weighing different topologies. Additionally coupling this model to a matter action  $S_m$ , the gravitational equations lead to  $T_{\mu\nu}^m = 0$  and no energy flows exist. Hence when using it as a classical toy model for black hole formation and evaporation, this model has little value.<sup>1</sup>

To get something more interesting, the required adjustment in two dimensions is to introduce a direct coupling of the Ricci scalar to a new scalar field  $\Phi$ , called the dilaton field for historical (string-theoretic) reasons:

$$S = \frac{1}{16\pi G_N} \int d^2x \sqrt{-g} \Phi R + \dots \quad (1.2)$$

Models of this kind have been introduced in 3+1d as scalar-tensor (Brans–Dicke type) models, and provide deformations of general relativity with spacetime-varying gravitational coupling. Here we view these dilaton gravity models as quantum mechanical solvable toy models of quantum gravity and black hole physics. Jackiw–Teitelboim

<sup>1</sup> Quantum mechanically, gauge fixing this model with conformal matter, leads to Liouville gravity which is an interesting model appearing in the context of the non-critical string (Polyakov 1981; Distler and Kawai 1989; David 1988). We do not follow this route.

(JT) gravity is one such particular model that we will specify below. Interestingly, this particular model also captures the dynamics close to the horizon of near-extremal black holes in higher dimensions, see the upcoming review by Iliesiu and Turiaci (2023).

Gravitational models of this kind have attracted widespread attention over the years [see e.g. the reviews by Grumiller et al. (2002), Nojiri and Odintsov (2001)], but it has been only recently, sparked by developments in 2015 by A. Kitaev in solvable many-body models (the Sachdev–Ye–Kitaev or SYK models; Sachdev and Ye 1993; Kitaev 2014, 2015a, b), that we have reached a far deeper understanding of their quantum mechanical aspects. This work refined the original proposal relating SYK and  $\text{AdS}_2$  gravity put forward by Sachdev (2010).

Our aim here is to provide an in-depth review of these developments in JT gravity. This review is organized into four main sections. In Sect. 2 we introduce the model, and fully solve its classical equations of motion, crucially incorporating boundary conditions at the holographic boundary that allow us to map the dynamics to its boundary Schwarzian description. Section 3 proceeds with the exact quantization of the model, by computing the Euclidean gravitational path integral in the Schwarzian language. In Sect. 4 we furthermore include non-trivial topological corrections (or wormholes) to the quantum amplitudes, that modify the heavy quantum regime even further. Finally, Sect. 5 contains several applications of the exact solvability of JT gravity. We do not treat these in technical detail, but refer to the literature for more information. In particular, in the past few years significant progress has been made on Hawking’s information loss paradox, made possible to a large extent due to the solvability of JT gravity. We discuss some of these developments, but will not do justice to this topic in this work. We refer e.g. to the excellent recent review by Almheiri et al. (2021) for details.

A few reviews have been written before on the connection between JT gravity and SYK. We refer the interested reader to the excellent review on this topic by Chowdhury et al. (2022) that focuses mostly on the SYK side. Other earlier reviews are Sárosi (2018) and Rosenhaus (2019). However, no comprehensive review exists that combines these earlier developments with the current state-of-the-art of JT gravity specifically and lower-dimensional gravitational models more generally, something we hope to address with this review.

Finally, a few comments on conventions: in Lorentzian signature our metric signature convention is  $(-, +, \dots +)$ . We denote Lorentzian actions as  $S$  and Euclidean actions as  $I$ .

## 2 Classical Jackiw–Teitelboim gravity

We begin by introducing and motivating the JT model and its coupling to matter. In this section, we study the classical solution of this model including gravitational backreaction. Our endeavors will ultimately lead us to a description in terms of a boundary Schwarzian model that will be the starting point for a quantum mechanical solution in the next Sect. 3.

## 2.1 Dilaton gravity models

To start, it is instructive to consider what is the most general theory of dilaton gravity in two dimensions with a two-derivative action (Banks and O’Loughlin 1991; Louis-Martinez et al. 1994; Ikeda 1994; Grumiller et al. 2002). Working in Euclidean signature, any such theory can be written as

$$I = -\frac{1}{16\pi G_N} \int_{\mathcal{M}} d^2x \sqrt{g} \left( U_1(\tilde{\Phi})R + U_2(\tilde{\Phi})g^{\mu\nu}\partial_\mu\tilde{\Phi}\partial_\nu\tilde{\Phi} + U_3(\tilde{\Phi}) \right), \quad (2.1)$$

where  $g_{\mu\nu}$  is the two-dimensional metric and  $\tilde{\Phi}(x)$  is the dilaton field. The 2d Newton’s constant  $G_N$  is dimensionless, and hence unlike in higher dimensions does not set the scale of physics.<sup>2</sup> This theory seems to be parametrized by three functions  $U_1(\tilde{\Phi})$ ,  $U_2(\tilde{\Phi})$  and  $U_3(\tilde{\Phi})$  but two of them are redundant. To see this, first perform a field redefinition on the dilaton  $\tilde{\Phi} \rightarrow \Phi = U_1(\tilde{\Phi})$ . We assume there is no value of  $\tilde{\Phi}$  such that  $U'_1(\tilde{\Phi}) = 0$  so the field redefinition is invertible, otherwise the resulting kinetic term in  $\Phi$  will be ill-defined.

We can further set  $U_2(\Phi) = 0$  by making a Weyl transformation on the metric as follows. Under a local rescaling, the 2d Ricci scalar transforms as

$$g'_{\mu\nu} = e^{2\omega} g_{\mu\nu}, \quad \rightarrow g'^{1/2} R' = g^{1/2} (R - 2\nabla^2 \omega). \quad (2.2)$$

It is then a simple calculation to show that the choice

$$\omega(x) = \frac{1}{2} \int^{\Phi(x)} U_2(\Phi') d\Phi', \quad (2.3)$$

will cancel the term in (2.1) proportional to  $U_2$ . In dimensions greater than two, this Weyl field redefinition  $g'_{\mu\nu} = e^{2\omega} g_{\mu\nu}$ , with a suitable choice of  $\omega$ , allows us to redefine  $\Phi R \rightarrow R'$ , a procedure that is well-known in string theory to transfer between the so-called string frame and the Einstein frame. In 2d however, the above Weyl field redefinition allows us instead to set the kinetic term of the dilaton to zero, and it is impossible to get rid of the  $\Phi R$  term: dilaton gravity is (in this sense) an invariant notion in 2d. This observation also implies that calling  $\Phi$  a dilaton is not a good nomenclature in 2d since it is not related to rescalings of the metric. Regardless of this point, to avoid confusion we will follow tradition and continue calling it the dilaton field.

This leaves us with a very general class of 2d dilaton gravity models:

$$I[g, \Phi] = -\frac{1}{16\pi G_N} \int_{\mathcal{M}} d^2x \sqrt{g} (\Phi R + U(\Phi)), \quad (2.4)$$

parametrized by a single function  $U(\Phi)$  called the dilaton potential. Let us first make some comments on this class of models.

<sup>2</sup> Later on, around Eq. (2.58) in JT gravity, we will see an effective scale emerge nonetheless.

Notice that the above field transformations from (2.1) to (2.4) are done at the classical level. In the full quantum theory, care has to be taken when performing these steps, in particular with regard to the thermodynamics of the resulting models.

We started with a two-derivative action (2.1). One can relax this assumption and include higher-derivative terms. The resulting model becomes power-counting non-renormalizable, but since there are no local degrees of freedom anyway, this may not be a problem. See Grumiller et al. (2022) for investigations of this generalization of Eq. (2.1).

In some works, partially motivated by string theory, the dilaton field is denoted instead as  $\Phi \rightarrow \Phi^2$  or  $\Phi \rightarrow \exp(-2\Phi)$ , emphasizing its positivity throughout space-time. The dilaton coupling parametrizes a spacetime-dependent effective Newton's constant

$$G_{\text{eff}}(x) = \frac{G_N}{\Phi(x)}, \quad (2.5)$$

and to interpret a classical solution physically, one wants this effective gravitational coupling to remain positive everywhere.<sup>3</sup> This will also lead to another more appropriate interpretation for the dilaton: as a measure of entropy.

*Jackiw–Teitelboim (or JT) gravity* corresponds to a specific linear choice of dilaton potential in (2.4), where  $U(\Phi) = -\Lambda \Phi$  (Jackiw 1985; Teitelboim 1983):

$$I_{\text{JT}}[g, \Phi] = -\frac{1}{16\pi G_N} \int_{\mathcal{M}} \sqrt{g} \Phi (R - \Lambda). \quad (2.6)$$

The quantity  $\Lambda$  (of dimension energy squared) is the cosmological constant of the model. To focus on the Anti-de Sitter space (AdS) case, we choose negative  $\Lambda$  and set  $\Lambda = -2/L^2$ . Working in units where the AdS length  $L = 1$ , we can write down our JT model:

$$I_{\text{JT}}[g, \Phi] = -\frac{1}{16\pi G_N} \int_{\mathcal{M}} \sqrt{g} \Phi (R + 2) - \frac{1}{8\pi G_N} \oint_{\partial\mathcal{M}} \sqrt{h} \Phi (K - 1). \quad (2.7)$$

For manifolds that have a boundary  $\partial\mathcal{M}$ , we wrote explicitly the Gibbons–Hawking–York (GHY) boundary term (Gibbons and Hawking 1977; York 1972), and a holographic counterterm (the “−1”) which is needed a posteriori to make sense of the model at an asymptotically AdS boundary as will become clear later on. It will be important later to add the topological Einstein–Hilbert term as well, proportional to a parameter of the model  $S_0$ . The final action we will study throughout this review is hence

$$I[g, \Phi] = -S_0 \chi + I_{\text{JT}}[g, \Phi], \quad (2.8)$$

<sup>3</sup> This will be revisited when considering the integration contour for the dilaton in the quantum theory in Sect. 3.

where we introduced the Euler characteristic of the manifold

$$\chi = \frac{1}{4\pi} \int_{\mathcal{M}} \sqrt{g} R + \frac{1}{2\pi} \oint_{\partial\mathcal{M}} \sqrt{h} K. \quad (2.9)$$

The topological term will be important in Sect. 4, but in this section it will only produce an overall shift of the action. This term can be thought of as being produced by a shift in the dilaton field by a constant  $\Phi(x) \rightarrow \Phi_0 + \Phi(x)$  producing the term (2.8) with  $S_0 = \frac{\Phi_0}{4G_N}$ .

It is possible to study JT gravity in asymptotically de Sitter (dS) by choosing the same dilaton potential with positive  $\Lambda = +2/L^2$  and we will consider this theory briefly in Sect. 5.5. Finally, the asymptotically flat case where  $U(\Phi) = \text{constant}$ , also has interesting applications which we will briefly mention there as well. For the body of this work, we will restrict to AdS.

### 2.1.1 First-order formulation

Just as any gravitational model, 2d dilaton gravity can be written in the first-order formulation. Unlike gravity in four or higher spacetime dimensions, the resulting model is a topological gauge theory: the Poisson-sigma model (Schaller and Strobl 1994b). For the specific case of JT gravity, this gauge theory further simplifies into the topological BF model (Horowitz 1989). Here we review this connection for closed manifolds  $\mathcal{M}$ . We start with the Euclidean dilaton gravity model:

$$I = -\frac{1}{16\pi G_N} \int_{\mathcal{M}} d^2x \sqrt{g} (\Phi R + U(\Phi)). \quad (2.10)$$

Introducing the frame field  $e_\mu^a$  as:

$$g_{\mu\nu} = e_\mu^a e_\nu^b \delta_{ab}, \quad (2.11)$$

where  $a, b \in \{0, 1\}$ , and the torsion-free spin connection  $\omega^{ab} = \omega_\mu^{[ab]} dx^\mu$ , determined by  $de^a + \omega^a_b \wedge e^b = 0$ , we have the following relations in 2d:

$$\omega^{ab} = \epsilon^{ab} \omega, \quad d^2x \sqrt{g} = e^0 \wedge e^1, \quad d^2x \sqrt{g} R = 2 d\omega. \quad (2.12)$$

We can then write the gravitational action (2.10) in the first-order form:

$$I = -\frac{1}{8\pi G_N} \int_{\mathcal{M}} \left[ \Phi d\omega + \frac{1}{4} U(\Phi) \epsilon^{ab} e_a \wedge e_b + X^a (de_a + \epsilon_a^b \omega \wedge e_b) \right]. \quad (2.13)$$

Indeed, integrating over the Lagrange multipliers  $X^{0,1}$ , we produce the torsion-free conditions and the remaining action reduces back to the second-order form (2.10). This model can be identified with a topological Poisson sigma model, with 3d target space (Ikeda and Izawa 1993; Ikeda 1994). To see this, introduce a connection  $A_i = (e_0, e_1, \omega)$  and a 3d space parametrized by scalars  $X^i = (X^0, X^1, X^2 \equiv \Phi)$ . Poisson

manifolds are characterized by a (possibly degenerate) bracket  $\{X^i, X^j\}_{\text{PB}} \equiv P^{ij}(X)$ . The antisymmetric tensor  $P^{ij}$  satisfies the Jacobi identity  $\partial_l P^{[ij]} P^{l|k]} = 0$ . In terms of this data the Poisson sigma model action is given by

$$I_{\text{PSM}} = -\frac{1}{8\pi G_N} \int_{\mathcal{M}} \left( A_i \wedge dX^i + \frac{1}{2} P^{ij}(X) A_i \wedge A_j \right), \quad i = 0, 1, 2, \quad (2.14)$$

where the brackets needed to reproduce (2.13) can be read off as

$$P^{01} = \{X^0, X^1\}_{\text{PB}} = \frac{U(X^2)}{2}, \quad P^{a2} = \{X^a, X^2\}_{\text{PB}} = \epsilon^a_b X^b. \quad (2.15)$$

Upon redefining the generators as  $E^\pm \equiv -X^0 \pm X^1$  and  $H \equiv X^2$ , this non-linear Poisson algebra becomes:

$$\{H, E^\pm\}_{\text{PB}} = \pm E^\pm, \quad \{E^+, E^-\}_{\text{PB}} = U(H). \quad (2.16)$$

For the special case of JT gravity where  $U(\Phi) = 2\Phi$ , this algebra becomes the linear  $\mathfrak{sl}(2, \mathbb{R})$  Lie algebra. Starting with (2.14), one can perform an integration by parts in the first term and, since the Poisson tensor  $P^{ij}(X)$  is linear in  $X$  in this case, rewrite the action (2.13) as a BF model (Fukuyama and Kamimura 1985; Isler and Trugenberger 1989; Chamseddine and Wyler 1989; Jackiw 1992):

$$I = -\frac{1}{8\pi G_N} \int_{\mathcal{M}} X^i F_i = -\frac{1}{4\pi G_N} \int_{\mathcal{M}} \text{Tr} XF, \quad F = dA + A \wedge A, \quad (2.17)$$

where  $X = X^i \lambda_i$  (and similarly for  $F$ ) with  $\{\lambda_1, \lambda_2, \lambda_3\}$  the generators of  $\mathfrak{sl}(2, \mathbb{R})$ , normalized such that  $\text{Tr} \lambda_i \lambda^j = \delta_i^j / 2$ .

This discussion was valid for manifolds without boundaries. We will come back to the case with boundaries later on in Sect. 3.7.4. The  $\mathfrak{sl}(2, \mathbb{R})$  BF description of JT gravity is extremely useful to obtain explicit solutions, but care has to be taken: the BF description is only *locally* equivalent with JT gravity, and the relation with the global group-theoretic description is a bit more subtle, and contains several layers of complications:

- A good gauge connection  $A_\mu$  can be singular in Euclidean gravity. An example is  $A_\mu = 0$ , leading to  $g_{\mu\nu} = 0$ . These singular metrics are to be excluded from the gravitational path integral.
- Gravity contains large diffeomorphisms (such as the  $\text{SL}(2, \mathbb{Z})$  modular group of the torus; more generally in 2d this is the mapping class group) that are not contained within the gauge transformations of the BF description.
- Gravity contains a sum over topologies that is not included in the gauge theory description.

All of these items are present for any model of dilaton gravity, but they can be dealt with very explicitly even in quantum JT gravity, for which we refer to the recent works (Blommaert et al. 2019b; Fan and Mertens 2022b, a), and to some of the older literature



(Schaller and Strobl 1994a). These complications parallel the difficulties in writing 3d gravity as a Chern–Simons gauge theory, explained nicely in Witten (2007).

### 2.1.2 Motivation: near-extremal black holes

One of the attractive features of JT gravity is the fact that it universally describes the physics near the horizon of near-extremal black holes in higher dimensions. The simplest realization of this is a near-extremal charged black hole in 4d Einstein–Maxwell theory. The 4d theory has a  $U(1)$  gauge field beside the metric, and the Lorentzian action is:

$$S = \frac{1}{16\pi G_N^{(4)}} \int d^4x \sqrt{-g} (R - F^2) + S_{\text{bdy}}. \quad (2.18)$$

$S_{\text{bdy}}$  includes the necessary boundary terms and counterterms to make the variational problem well-defined, such as the Gibbons–Hawking–York one. The Reissner–Nordström black hole solution with charge  $Q$  is given by

$$ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega_2^2, \quad f(r) = 1 - \frac{2M}{r} + \frac{Q^2}{r^2}, \quad (2.19)$$

where we absorbed  $G_N^{(4)}$  into  $M$  for brevity of notation. The only non-zero component of the electric field is  $F_{rt} \propto Q/r^2$ , and  $d\Omega_2^2$  is the metric element of the transverse 2-sphere. The horizons are at  $r_{\pm} = M \pm \sqrt{M^2 - Q^2}$  and the Hawking temperature of the black hole is  $T = |f'(r_+)|/4\pi$ . The near-extremal limit corresponds to low temperatures such that  $M \sim |Q|$  and hence  $r_+ \sim r_- \sim |Q|$ . From now on we take  $Q > 0$ .

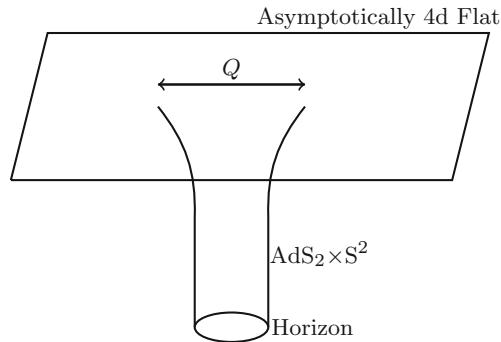
The first step to see how JT gravity arises from this 4d theory, is to identify the near-horizon near-extremal geometry. We consider the scaling limit of small  $\Delta M \equiv M - Q$ , where  $r = Q + Q^2 \tilde{r}$ , and the radial coordinate scales as  $\tilde{r} \sim \sqrt{\Delta M/Q^3}$ . In this region, the spacetime metric (2.19) becomes:

$$ds^2 \approx -Q^2 \left( \tilde{r}^2 - \frac{2\Delta M}{Q^3} \right) dt^2 + \frac{Q^2 d\tilde{r}^2}{\tilde{r}^2 - \frac{2\Delta M}{Q^3}} + Q^2 d\Omega_2^2. \quad (2.20)$$

This is the near-horizon  $\text{AdS}_2 \times \text{S}^2$  geometry, with AdS length and 2-sphere radius both  $Q$ . Rescaling lengths by  $Q^{-1}$ , we can rewrite this in the canonical form:

$$Q^{-2} ds^2 \approx -(\tilde{r}^2 - r_h^2) dt^2 + \frac{d\tilde{r}^2}{\tilde{r}^2 - r_h^2} + d\Omega_2^2, \quad r_h = \sqrt{\frac{2\Delta M}{Q^3}}, \quad (2.21)$$

with the AdS length and sphere radius equal to 1 in these units. The exactly extremal case is found by setting  $\Delta M = 0$  which implies  $r_h = 0$ . Otherwise, we have a black hole in  $\text{AdS}_2$  with a horizon at a finite distance  $r_h$ . Within this  $\text{AdS}_2$  region,  $\tilde{r} - r_h \ll r_h$



**Fig. 1** Spatial geometry of the Reissner–Nordström black hole which interpolates between flat space far away and an  $\text{AdS}_2 \times S^2$  throat near the horizon when the temperature is low. The size of the black hole is of order  $Q$  (in appropriate units) while the length of the  $\text{AdS}_2 \times S^2$  throat diverges as we approach extremality. JT gravity arises as a sector of Einstein–Maxwell in 4d and is the dominant mode controlling the low energy dynamics taking place inside the throat. This is true for a large class of black holes with  $\text{AdS}_2 \times S^2$  throats

corresponds to the Rindler region whereas  $\tilde{r} \rightarrow \infty$  corresponds to the boundary of  $\text{AdS}_2$ .

In Fig. 1, we illustrate the spatial geometry of the Reissner–Nordström black hole which interpolates between flat space far away and an  $\text{AdS}_2 \times S^2$  throat near the horizon when the temperature is low. The size of the black hole is of order  $Q$  (in appropriate units) while the length of the  $\text{AdS}_2 \times S^2$  throat diverges as we approach extremality.

Having described the classical geometry near the horizon, we now study perturbations around it. Metric fluctuations can be parametrized by Davison et al. (2017), Nayak et al. (2018), Sachdev (2019)

$$ds^2 = \frac{1}{\sqrt{\chi}} g_{\mu\nu} dx^\mu dx^\nu + \chi (dy^i + A_\mu^{ij} y^j dx^\mu) (dy^i + A_\mu^{ij} y^j dx^\mu) + \dots, \quad (2.22)$$

where  $y = (y^1, y^2, y^3)$  with  $y^i y^i = 1$  parametrize the transverse  $S^2$ . The coordinate  $x^\mu = (t, r)$  parametrizes the time and radial directions. Both  $\chi$ ,  $g_{\mu\nu}$  and  $A_\mu^{ij}$  are assumed to depend only on  $x^\mu$ . The dots in the equation above denotes other Kaluza–Klein modes that are not turned on in the Reissner–Nordström solution and whose fluctuations around the background are massive, with masses  $\sim Q^{-1}$  the AdS scale and hence not parametrically large for a near-extremal black hole. Therefore the correct low-energy description involves two-dimensional gravity with a large number of light ( $\sim Q^{-1}$ ) fields.<sup>4</sup> The one-forms  $A_\mu^{ij}$  reduce on the 2d space to an  $\text{SU}(2)$  gauge field whose charge corresponds to angular momentum  $J$ .

For simplicity, we work in an ensemble with fixed black hole charge and fixed vanishing angular momentum. Then the Einstein–Maxwell action (2.18) dimensionally

<sup>4</sup> This feature is not different from other realizations of AdS. For example in a five-dimensional description of the  $\text{AdS}_5 \times S^5$  background in type IIB string theory, there are also a large number of light fields from compactification on the sphere.

reduces to the restricted 2d dilaton-gravity dynamics (with dilaton  $\chi$ )<sup>5</sup>:

$$S[g, \chi] = \frac{1}{4G_N^{(4)}} \int d^2x \sqrt{-g} (\chi R + U(\chi)) + \dots, \quad U(\chi) = -\frac{2Q^2}{\sqrt{\chi}^3} + \frac{2}{\sqrt{\chi}}, \quad (2.23)$$

where we can define the 2d Newton's constant by  $G_N^{(4)} = 4\pi G_N^{(2)}$ . Turning on angular momentum would correct the potential by a term proportional to  $J^2$ . The classical solution of this theory, with appropriate boundary conditions, reproduces (2.19). The dots in the equation above denote contribution from either matter or other Kaluza–Klein modes of the Einstein–Maxwell sector, resulting in 2d dilaton-gravity coupled to matter.

A classical solution with constant sphere area has  $U(\chi = \Phi_0) = 0$  and therefore  $\Phi_0 = Q^2$ , which corresponds to rigid  $\text{AdS}_2 \times S^2$  spacetime. This is the geometry very close to the horizon of the near-extremal black hole as presented above. JT gravity arises as we take the  $S^2$  area large, and allow it to fluctuate slightly:

$$\chi(x) = \Phi_0 + \Phi(x), \quad \Phi \ll \Phi_0. \quad (2.24)$$

Under this assumption the 4d Einstein–Maxwell theory reduces near the horizon to JT gravity, after rescaling the 2d metric  $g \rightarrow Q^3 g$ :

$$S[g, \Phi] = \frac{\Phi_0}{4G_N^{(4)}} \int \sqrt{-g} R + \frac{1}{4G_N^{(4)}} \int \sqrt{-g} \Phi (R + 2) + \dots \quad (2.25)$$

Corrections to this action are suppressed in the regime (2.24) and we end up with the JT action (2.7) coupled to matter arising both from 4d matter as well as metric KK modes we ignored in (2.22). (To leading order in the large  $\Phi_0$  limit, the matter fields do not couple to the dilaton  $\Phi$ , a feature that will be important later when studying quantum effects.) Hence JT gravity arises as a sector in 4d Einstein–Maxwell theory, and it is the dominant mode controlling the low-energy dynamics taking place inside the throat.

This derivation also gives a nice interpretation of each term that appears in the lower-dimensional gravity action. We can identify the parameter  $S_0$  in (2.8) with

$$S_0 = \frac{\pi \Phi_0}{G_N^{(4)}} = \frac{\pi Q^2}{G_N^{(4)}}, \quad (2.26)$$

which is the Bekenstein–Hawking area term corresponding to the 4d extremal black hole. Note that from the 2d perspective, this is  $A/4G_N^{(2)}$  where  $\Phi_0$  plays the role of the area  $A$ . We will see more evidence later that the dilaton should indeed be identified with the entropy of the black hole. We might be tempted to refer to the extremal Bekenstein–Hawking area  $S_0$  as extremal degeneracy of black hole microstates, although we will

<sup>5</sup> The coefficient of the Einstein term in two dimensions vanishes when  $\chi = 0$ . This corresponds to the singularity of the 4d black hole. Since we will focus on the physics near the horizon we will not worry about that, although it is important to keep in mind.

see this is an incorrect interpretation both in JT gravity (Stanford and Witten 2017) and for the higher-dimensional black hole. This point was made for the near-extremal BTZ black hole in Ghosh et al. (2020) and then generalized to the Reissner–Nordström black hole in Iliesiu and Turiaci (2021), Heydeman et al. (2022). The reason is that quantum effects from the JT mode get enhanced at low enough temperatures and, in a fixed charge ensemble, only the JT mode affects the temperature-dependence of the free energy, even at the quantum level.

The 4d Reissner–Nordström black hole is not the only higher-dimensional black hole whose near-horizon near-extremal dynamics is governed by JT gravity, but there is in fact a large class. We will discuss this further in Sect. 5.3. The universality of JT gravity in this kinematic regime of black hole physics is hence a strong motivation for studying the JT model in more detail.

## 2.2 Classical solutions

In this section, we will solve and study the classical equations of motion of Lorentzian JT gravity, including its coupling to a matter action:

$$S = \frac{\Phi_0}{16\pi G_N} \left[ \int \sqrt{-g} R + 2 \oint \sqrt{-h} K \right] + S_{\text{JT}}[g, \Phi] + S_m[\phi, g]. \quad (2.27)$$

The first term corresponds to a constant dilaton background  $\Phi_0$  and is the Lorentzian version of Eq. (2.8) where  $S_0 = \Phi_0/(4G_N)$ . The second term is the JT action itself with a varying dilaton

$$S_{\text{JT}}[g, \Phi] = \frac{1}{16\pi G_N} \left[ \int \sqrt{-g} \Phi (R + 2) + 2 \oint \sqrt{-h} \Phi (K - 1) \right]. \quad (2.28)$$

Finally,  $S_m[\phi, g]$  is the matter action, which we assume only couples to the metric, and not directly to the dilaton field  $\Phi$ . This assumption will allow for an explicit and analytic solution.<sup>6</sup> We will not have to be explicit about the matter action in order to write down the most general solution of Einstein's equations, which is one of the benefits of working in lower dimensions.

### 2.2.1 Metric solution

Now we describe general classical solutions of the matter-coupled model (2.27). The equation of motion one obtains by varying (2.27) with respect to the dilaton  $\Phi$  is given by

$$R(x) = -2. \quad (2.29)$$

<sup>6</sup> This is true when applying JT gravity to the dynamics of gravity and matter near the horizon of a near-extremal black hole in Einstein–Maxwell gravity. Another example is given by extremal dilatonic black holes in four or higher dimensions in superstring theories, where the relevant matter action coming from Ramond–Ramond fields, describing the radial dynamics, is just a free scalar field in 2d, and would indeed not have a direct coupling to the dilaton field (Callan et al. 1992).

In two dimensions, knowing the value of the local Ricci scalar is sufficient to know the metric itself up to coordinate transformations, in this case some patch of the  $\text{AdS}_2$  manifold. Let us make this statement more explicit.

It is convenient to describe the 2d metric in conformal gauge, which can always be reached by a coordinate transformation, where:

$$ds^2 = -e^{2\omega(u,v)} du dv, \quad (2.30)$$

where the only metric degree of freedom is  $\omega(u, v)$ . The coordinates  $u$  and  $v$  are lightcone coordinates defined as  $u = t + z$  and  $v = t - z$ . For a conformal gauge metric (2.30), we have  $R = 8e^{-2\omega}\partial_u\partial_v\omega$ , and (2.29) boils down to Liouville's equation  $4\partial_u\partial_v\omega + e^{2\omega} = 0$ . The general solution to Liouville's equation is well-known:

$$e^{2\omega(u,v)} = \frac{4\partial_u U(u)\partial_v V(v)}{(U(u) - V(v))^2}, \quad (2.31)$$

in terms of two chiral functions  $U(u)$  and  $V(v)$ . This solution can be restricted when imposing boundary conditions as we will implement further on in Sect. 2.3. With this solution, the metric (2.30) is

$$ds^2 = -\frac{4\partial_u U(u)\partial_v V(v) du dv}{(U(u) - V(v))^2} = -\frac{4dU dV}{(U - V)^2}, \quad (2.32)$$

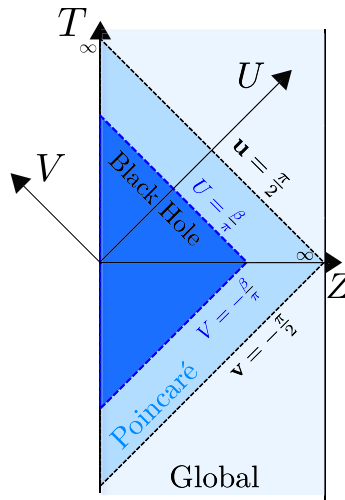
and represents the Poincaré patch of  $\text{AdS}_2$  with lightcone coordinate  $(U, V)$ . The two chiral functions  $U(u)$  and  $V(v)$  are interpreted as the diffeomorphisms that preserve conformal gauge, and relate the lightcone coordinates  $U$  and  $V$  of the Poincaré patch to new lightcone coordinates  $u$  and  $v$ . This proves that all metric solutions are just different coordinate frames on the  $\text{AdS}_2$  manifold.

It is useful to summarize the geometric properties of several important classical frames of  $\text{AdS}_2$ . The Poincaré patch itself is described by the metric

$$ds^2 = \frac{-dT^2 + dZ^2}{Z^2} = -R^2 dT^2 + \frac{1}{R^2} dR^2 = -\frac{4dU dV}{(U - V)^2}, \quad U > V, \quad (2.33)$$

where we defined  $R = 1/Z$  in the first equality. The isometry group of  $\text{AdS}_2$  is  $\text{PSL}(2, \mathbb{R}) \simeq \text{SO}(2, 1)$ , which acts as Möbius transformations on the Poincaré lightcone coordinates  $U \rightarrow \frac{aU+b}{cU+d}$ ,  $V \rightarrow \frac{aV+b}{cV+d}$ , where  $a, b, c, d \in \mathbb{R}$  and  $ad - bc = 1$ .  $\text{PSL}(2, \mathbb{R})$  is the projective subgroup of  $\text{SL}(2, \mathbb{R})$ , the group of invertible  $2 \times 2$  matrices with determinant 1, where we identify the matrix with minus itself.

We denote the Poincaré time coordinate as  $T = (U + V)/2$  and the spatial coordinate as  $Z = (U - V)/2$ . It has a single boundary at  $U = V$  (or  $Z = 0$ ). It has horizons at  $U - V \rightarrow \infty$ , or  $U \rightarrow +\infty$  (future horizon) and at  $V \rightarrow -\infty$  (past horizon). This geometry is found in the near-horizon limit of a higher-dimensional extremal black hole, the  $r_h = 0$  limit of Eq. (2.21). The horizons are at an infinite proper distance ( $\int^{+\infty} \frac{dZ}{Z} \rightarrow \infty$ ) at the end of an infinite throat. Starting with the Poincaré patch, we



**Fig. 2** Penrose diagram with the different classical patches of global  $\text{AdS}_2$

can reach other important frames that we discuss next. We illustrate these in Fig. 2, and the reader can consult this figure throughout.

The global frame is found by performing the following coordinate transformation

$$U(\mathbf{u}) = \tan(\mathbf{u}), \quad V(\mathbf{v}) = \tan(\mathbf{v}), \quad (2.34)$$

defining new lightcone coordinates  $(\mathbf{u}, \mathbf{v})$ . The metric (2.33) is transformed into:

$$ds^2 = -\frac{4}{\sin^2(\mathbf{u} - \mathbf{v})} d\mathbf{u}d\mathbf{v}, \quad (2.35)$$

and can be continued beyond the original Poincaré patch. We define the global spatial coordinate  $\mathbf{z} = (\mathbf{u} - \mathbf{v})/2$  and time coordinate  $\mathbf{t} = (\mathbf{u} + \mathbf{v})/2$ . There are two boundaries where the metric blows up: this is at  $\mathbf{u} = \mathbf{v}$  (or  $\mathbf{z} = 0$ ), and at  $\mathbf{u} = \mathbf{v} + \pi$  (or  $\mathbf{z} = \pi/2$ ). These coordinates cover a strip-region with two boundaries containing the Poincaré patch. Usually in  $\text{AdS}_D/\text{CFT}^{D-1}$ , the boundary in global coordinates is the  $\mathbb{R} \times S^{D-2}$  manifold, which in this case gives a 0-sphere for the spatial manifold, or two disjoint points.

Finally, the black hole patch can be found by the following coordinate transformation:

$$U(u) = \frac{\beta}{\pi} \tanh\left(\frac{\pi}{\beta} u\right), \quad V(v) = \frac{\beta}{\pi} \tanh\left(\frac{\pi}{\beta} v\right), \quad (2.36)$$

with parameter  $\beta$ . This leads to the metric

$$ds^2 = -\frac{\pi^2}{\beta^2} \frac{4}{\sinh^2\left(\frac{\pi}{\beta}(u - v)\right)} du dv. \quad (2.37)$$

We define also the coordinates  $z = (u - v)/2$  and  $t = (u + v)/2$ . The black hole patch (2.37) is contained within the Poincaré patch (2.33) since the Poincaré coordinates only range over the restricted range  $-\frac{\beta}{\pi} < U, V < \frac{\beta}{\pi}$  (with  $U > V$ ). This geometry has horizons where  $u - v \rightarrow \infty$  at finite distance ( $\int^{+\infty} \frac{dz}{\sinh \frac{2\pi}{\beta} z} < \infty$ ). Using the radial coordinate  $r = r_h \coth \frac{2\pi z}{\beta}$  where  $r_h = 2\pi/\beta$ , the metric can be rewritten as

$$ds^2 = -(r^2 - r_h^2)dt^2 + \frac{dr^2}{r^2 - r_h^2}. \quad (2.38)$$

This patch is directly found in the near-horizon regime of a higher-dimensional near-extremal black hole, as in (2.21). From the near-horizon region where  $r \approx r_h$  of the Euclidean section, we can deduce the Hawking temperature  $T = \beta^{-1} = r_h/2\pi$ , justifying the parameter label  $\beta$  as the inverse temperature. This is simultaneously the temperature of the higher-dimensional near-extremal black hole. One further way of rewriting this patch is to use the proper distance coordinate  $\rho$  to the black hole horizon, which allows us to write the same patch as

$$\sinh \rho = \frac{1}{\sinh \frac{\pi}{\beta}(u - v)} \rightarrow ds^2 = -\frac{4\pi^2}{\beta^2} \sinh^2 \rho dt^2 + d\rho^2. \quad (2.39)$$

Near the horizon where  $\rho \ll 1$ , this becomes the Rindler geometry.<sup>7</sup> Some early references on the black hole solution are (Lemos and Sa 1994a, b; Lemos 1996).

Since the metric solution is always a patch of the  $\text{AdS}_2$  manifold, our model can be interpreted in a holographic context with holographic boundary at  $Z = 0$ . However,  $\text{AdS}_2/\text{CFT}_1$  has been notoriously tricky to make sense of. Resolving the tensions around it, has been one of the chief successes of JT gravity as we illustrate below. For now, let us review some arguments why this version of holography is subtle.

Firstly, for a CFT in any number of dimensions, the stress tensor is traceless:  $T_\mu{}^\mu = 0$ .<sup>8</sup> But in 0+1d there is only one index, and hence  $T_{tt} = 0$ . Hence the energy is always zero, and we are only able to describe a theory of ground states, or extremal states. Whereas this is consistent, it is not our main interest when considering dynamical models of black hole formation and evaporation. Even though the metric is locally  $\text{AdS}_2$ , we have seen the physics in each of the patches described above is very different. Whether there is a black hole horizon or not, and the global structure of spacetime, can vary from patch to patch. So far we have not given a physical mechanism within the theory that fixes this ambiguity, something one cannot do without a (non-zero) boundary Hamiltonian  $T_{tt}$ . In the example in Sect. 2.1.2, the black hole patch is fixed by the way the  $\text{AdS}_2$  region is glued to the asymptotically flat space. We will see in the next subsection how to phrase this intrinsically in terms of JT gravity, and the price to pay is the loss of conformal invariance.

<sup>7</sup> Note that these two new radial coordinates  $r$  and  $\rho$  are found by doing purely spatial coordinate transformations, which hence preserve the causality structure, the thermodynamics, and the black hole interpretation of the spacetime.

<sup>8</sup> In an even number of spacetime dimensions, there is the famous trace anomaly that violates this equality at the quantum level, but here we are in 0+1 dimensions.

Secondly, due to a lack of dimensionful parameters, the density of states of a 1d CFT is by dimensional analysis of the form (Jensen et al. 2011)

$$\rho(E) = A\delta(E) + \frac{B}{E}, \quad (2.40)$$

where  $A$  and  $B$  are dimensionless.<sup>9</sup> Now, either  $B = 0$ , in which case we again have a theory of ground states as before, or  $B \neq 0$  in which case the number of low-energy states diverges, and the model does not work without an additional IR cutoff.

We summarize that genuine  $\text{AdS}_2/\text{CFT}_1$  only makes sense as a model of zero-energy states (as topological QM), but to describe dynamics, we would have to modify this framework. This is precisely what JT gravity achieves as will become apparent further on. Moreover, in examples where a genuine  $\text{AdS}_2/\text{CFT}_1$  limit exists the bulk theory is strongly coupled and does not have a simple description (Heydeman et al. 2022; Boruch et al. 2022; Lin et al. 2023a, b).

### 2.2.2 Dilaton solution

Let us continue with the classical solution of the JT model. Varying (2.27) with respect to the metric, we find the equations of motion for the dilaton field:

$$\nabla_\mu \nabla_\nu \Phi - g_{\mu\nu} \nabla^2 \Phi + g_{\mu\nu} \Phi = -8\pi G_N T_{\mu\nu}, \quad (2.41)$$

or when written out in conformal gauge:

$$\begin{aligned} -e^{2\omega} \partial_u \left( e^{-2\omega} \partial_u \Phi \right) &= 8\pi G_N T_{uu}, \\ -e^{2\omega} \partial_v \left( e^{-2\omega} \partial_v \Phi \right) &= 8\pi G_N T_{vv}, \\ 2\partial_u \partial_v \Phi + e^{2\omega} \Phi &= 16\pi G_N T_{uv}, \end{aligned} \quad (2.42)$$

where  $T_{\mu\nu}$  are the stress tensor components of the matter sector in (2.27). Notice that in this model all backreaction of matter fields is contained in the dilaton field  $\Phi$  and *not* in the actual geometry itself.

These coupled equations of motion (2.42) can be solved systematically by integrating the first and second equation twice, and then inserting it in the last equation to find a consistency requirement.

Let us first describe the vacuum solutions where  $T_{uu} = T_{vv} = T_{uv} = 0$ . It is easy to show by directly using (2.41) that the quantity  $\zeta^\mu = \epsilon^{\mu\nu} \partial_\nu \Phi$  is a Killing vector of the entire dilaton gravity system:  $\nabla_\mu \zeta_\nu + \nabla_\nu \zeta_\mu = 0$ . This is an important observation for later on: in Euclidean signature, hyperbolic surfaces ( $R = -2$ ) at higher genus  $g \geq 2$  do not have Killing vectors. So there will be no classical solutions to JT gravity on higher genus surfaces.

<sup>9</sup> In higher dimensions, one can have the volume of space as a further dimensionful factor, which allows for other possibilities.



The most general solution of Eq. (2.42) is in this case described by three integration constants  $a$ ,  $b$  and  $\mu$ :

$$\Phi(u, v) = \frac{a + b(U + V) - \mu UV}{U - V}. \quad (2.43)$$

We note the following: the dilaton solution blows up on the holographic boundary ( $U = V$ ). The integration constant  $a$  has the dimension of length, and will be important to understand this divergent nature of the solution. The parameter  $\mu$  has the dimension of energy, and will be identified with the mass of a black hole later on. The parameter  $b$  will turn out to be not physically meaningful (one can set  $b = 0$  by using an  $\text{PSL}(2, \mathbb{R})$  isometry of the metric as discussed below (2.33), and we assume this has been done from here on).

Next, we consider non-zero matter sources. For simplicity of the expressions, we restrict to a 2d conformal field theory (CFT) as our matter sector. We will make a brief comment on the more generic case later on, but we point out that for the purposes of finding good models of black hole formation and evaporation, any (conserved) matter will do. The solution with conformal classical matter, for which  $T_{uv} = 0$  and by energy conservation  $T_{uu}(u)$  and  $T_{vv}(v)$  are chiral functions as denoted, is given by (Almheiri and Polchinski 2015)

$$\Phi(u, v) = \frac{a}{U - V} \left( 1 - \frac{\mu}{a} UV - \frac{8\pi G_N}{a} (I_+ + I_-) \right), \quad (2.44)$$

where  $I_{\pm}$  are given in terms of the matter stress-tensor by

$$\begin{aligned} I_+(u, v) &= \int_U^{+\infty} ds (s - U)(s - V) T_{UU}(s), \\ I_-(u, v) &= \int_{-\infty}^V ds (s - U)(s - V) T_{VV}(s), \end{aligned} \quad (2.45)$$

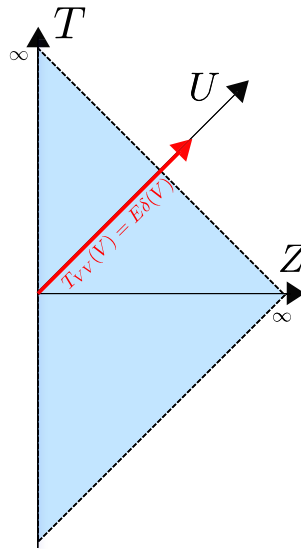
and  $T_{UU}dU^2 = T_{uu}du^2$  and  $T_{VV}dV^2 = T_{vv}dv^2$ . In the rest of this subsection we point out some important features of this solution. It is useful to first work out a simple example.

**Example:**

Consider an infalling energy pulse of energy  $E > 0$  in the Poincaré patch, modeled as  $T_{VV}(V) = E\delta(V)$  as in Fig. 3. This immediately leads to  $I_+ = 0$  and  $I_- = EUV\theta(V)$ , and hence the dilaton profile

$$\Phi = \frac{a - 8\pi G_N EUV\theta(V)}{U - V}. \quad (2.46)$$

Comparing with (2.43), we see that after the pulse has passed, we have created the vacuum solution where one identifies the parameter  $\mu = 8\pi G_N E$ .



**Fig. 3** Penrose diagram of  $\text{AdS}_2$  in the Poincaré patch. An energy pulse is injected at time  $T = 0$ , indicated in red

We will retake this specific example at multiple instances further on as we develop the classical model. We have interpreted the parameter  $\mu$  of Eq. (2.43) in terms of the total spacetime energy  $E$ .

It will be useful to describe this finite  $E$  solution using black hole coordinates introduced in Sect. 2.2.1, and it is convenient to rewrite it in several ways as:

$$\Phi = \frac{a - \mu U(u)V(v)}{U(u) - V(v)} = \sqrt{a\mu} \coth \left( \sqrt{\frac{\mu}{a}} 2z \right) = \frac{a}{2} r, \quad (2.47)$$

using the coordinates of Eqs. (2.37) and (2.38).

—o—

Imagine we set  $a = 0$  a priori in (2.44), such that initially  $\Phi(u, v) = 0$ . Then sending in some matter through  $T_{uu}$  or  $T_{vv}$  suddenly causes  $\Phi$  at  $Z = 0$  to diverge. This runaway backreaction at the holographic boundary was first noticed in Maldacena et al. (1999), where it was argued to ruin Maldacena's decoupling limit (Maldacena 1999) for  $\text{AdS}_2/\text{CFT}_1$ . This is one of the bulk avatars of the difficulties with  $\text{AdS}_2/\text{CFT}_1$ . We see here that the resolution is to consider backgrounds with  $a \neq 0$ , where the dilaton always diverges at the boundary. The rate of divergence of the solution then contains the information on backreaction. We elaborate on this point in the next subsection.

Moreover, in order to assess the regime of validity of the classical approximation of this section, it is useful to determine the location of possible singularities. Since the metric is always  $\text{AdS}_2$ , there are no curvature singularities. However, tracking the dilaton profile, there can be places where the effective Newton's constant blows up. This corresponds to points where  $\Phi_0 + (2.44) = 0$ . One expects the classical

approximation (studied in this section) to break down far before reaching this region.<sup>10</sup> Demanding this locus to be far away from the holographic  $\text{AdS}_2$  boundary for the solution (2.44), restricts us to  $a > 0$ , providing an independent reason to regularize the backreaction by introducing the non-zero quantity  $a$ .

An interesting point that emphasizes the connection between the dilaton and entanglement entropy is the following: the quantities  $I_+/(U - V)$  and  $I_-/(U - V)$  introduced in (2.44) match with expressions of modular Hamiltonians of a 2d CFT in an interval between the point  $(U, V)$  and a boundary at  $Z = 0$ . This observation was explored in Callebaut and Verlinde (2019), Callebaut (2019) to show that JT gravity can be viewed as an emergent model describing the entanglement dynamics of a 2d CFT on a halfspace.

The generic dilaton solution with arbitrary non-conformal matter is also known (Joshi et al. 2020), albeit is less elegant.

## 2.3 Boundary conditions and Schwarzian dynamics

In order to proceed further, we need to implement suitable boundary conditions at the holographic boundary  $u = v$ . Several equivalent perspectives have been developed to both motivate and implement these efficiently. An approach rooted in hydrodynamic effective actions can be found in Jensen (2016). In parallel, an insightful Euclidean geometric argument was constructed in Maldacena et al. (2016b), while in Engelsöy et al. (2016), generalizing the set-up of Almheiri and Polchinski (2015), a real-time perspective at the level of the equations of motion was developed. These arguments have been reviewed in many works by now. Here we focus first on the approach of Engelsöy et al. (2016), and then summarize the geometric route taken by Maldacena et al. (2016b).

### 2.3.1 Real-time derivation

We impose two boundary conditions that will restrict the freedom in the coordinate transformations at the boundary as follows.

*1. The geometry is asymptotically AdS.* We demand any bulk solution to have the leading asymptotics in Fefferman–Graham gauge:

$$ds^2 = \frac{-dt^2 + dz^2}{z^2} + (\text{subleading as } z \rightarrow 0). \quad (2.48)$$

If we now perform an arbitrary (non-chiral) coordinate transformation in lightcone coordinates  $U(u, v)$  and  $V(u, v)$ , then it is easy to show that in order to preserve the asymptotics (2.48), we require: (i)  $\partial_u V = 0 = \partial_v U$  at leading order as  $z \rightarrow 0$ , (ii)  $U(u, v) = V(u, v)$  at leading order as  $z \rightarrow 0$ . The first identity leads to chiral functions, and the second shows that both functions need to be the same. Translating

<sup>10</sup> In the case considered in Sect. 2.1.2, where JT gravity arises from a near-extremal black hole in 4d, the locus  $\Phi_0 + \Phi = 0$  corresponds to  $r = 0$  in coordinates (2.19), i.e. the black hole singularity. At this point, the expansion studied in Sect. 2.1.2 breaks down as well.

this to new time and radial coordinates through  $U \equiv F + Z$  and  $V \equiv F - Z$ , we obtain the near-boundary relations:

$$F(t) \equiv \frac{1}{2} (U(t + \epsilon) + V(t - \epsilon)) = U(t) + \mathcal{O}(\epsilon), \quad (2.49)$$

$$Z(t) \equiv \frac{1}{2} (U(t + \epsilon) - V(t - \epsilon)) = \epsilon F'(t) + \mathcal{O}(\epsilon^2). \quad (2.50)$$

where  $\epsilon$  is viewed as the near-boundary cutoff  $z = \frac{u-v}{2} \approx \epsilon$ , regularizing the holographic boundary by moving it slightly inwards from  $z = 0$  to  $z = \epsilon$ .

In Poincaré coordinates, this leads to a boundary trajectory determined in terms of a single function  $F(t)$ , as  $(T = F(t), Z = \epsilon F'(t))$ . This function  $F(t)$  parametrizes the subset of large diffeomorphisms that preserves the asymptotics (2.48), and by the usual dictionary in AdS/CFT, these generate the group of conformal transformations (i.e. arbitrary time reparametrizations in 1d) in the dual QM. However, this is not the end of the story since for all classical solutions, just like the metric, the dilaton field (2.44) blows up as  $z \rightarrow 0$  as well.

2. *The dilaton field asymptotics is fixed.* We have argued that a solution with finite  $\Phi$  at the boundary is not consistent with backreaction. To avoid this we choose the divergent dilaton asymptotics:

$$\Phi = \frac{a}{2z} + (\text{subleading as } z \rightarrow 0). \quad (2.51)$$

The parameter  $a$  of dimension length defines the particular model we are studying.

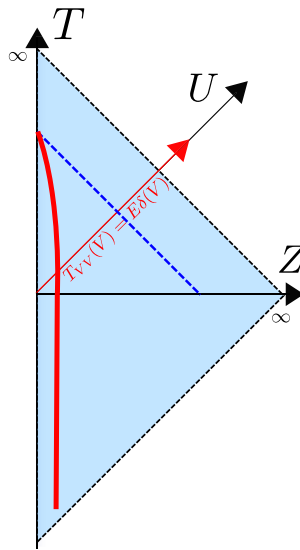
A physical way to motivate this boundary condition, is that to define the timelike trajectory of the boundary curve, one has to impose a scalar condition. Making the boundary follow a fixed value of the dilaton field  $\Phi$  is a natural coordinate-invariant way of defining this line in a dynamical situation. From a higher-dimensional perspective, the dilaton field descends from the transverse area of the fixed radial hypersurface, as discussed in Sect. 2.1.2. Keeping fixed the size of this surface is an invariant way of defining a radial location  $r$  in a dynamical bulk geometry.

Implementing this boundary condition by inserting the dilaton solution (2.44), we obtain the following integro-differential equation for  $F(t)$ :

$$F'(t) = 1 - \frac{8\pi G_N}{a} \left( \int_{F(t)}^{+\infty} ds (s - F(t))^2 T_{UU}(s) + \int_{-\infty}^{F(t)} ds (s - F(t))^2 T_{VV}(s) \right), \quad (2.52)$$

where we have set  $\mu = 0$  since it can be created by the matter as discussed earlier in the example in Sect. 2.2.2. This procedure breaks the conformal symmetry of the boundary theory, and leads to a fixed and preferred boundary time frame  $F(t)$  determined by the above equation.<sup>11</sup> This boundary condition (2.51) resolves the tension around

<sup>11</sup> A finite-dimensional subgroup of the infinite dimensional conformal group is typically left unbroken due to the isometries of AdS<sub>2</sub>. To appreciate this, consider the metric equation  $g_{uv} = \frac{\partial_u F(u) \partial_v F(v)}{(F(u) - F(v))^2}$



**Fig. 4** Penrose diagram of  $\text{AdS}_2$  in Poincaré patch before and after injecting an energy pulse. In red we show the boundary trajectory. We see this gets kicked after the insertion in such a way that a black hole horizon appears

$\text{AdS}_2/\text{CFT}_1$  by deforming it into a  $\text{NAdS}_2/\text{NCFT}_1$  of “nearly”  $\text{AdS}_2$  (by the varying dilaton field and its asymptotics) dual to a “nearly”  $\text{CFT}_1$  (by explicit breaking of conformal symmetry in the UV).

**Example (continued):**

Let’s solve equation (2.52) for our previous example. We have

$$F < 0 \rightarrow F' = 1, \quad (2.53)$$

$$F > 0 \rightarrow F' = 1 - \frac{\mu}{a} F^2. \quad (2.54)$$

For  $t < 0$  the solution is trivial  $F(t) = t$  and  $Z(t) = \epsilon$ , while for  $t > 0$  it is

$$F(t) = \sqrt{\frac{a}{\mu}} \tanh \sqrt{\frac{\mu}{a}} t, \quad (2.55)$$

$$Z(t) = \epsilon F'(t) = \frac{\epsilon}{\cosh^2 \sqrt{\frac{\mu}{a}} t}. \quad (2.56)$$

Notice that  $Z(t) \rightarrow 0$  as  $t \rightarrow \infty$  which happens at finite Poincaré time  $F = (a/\mu)^{1/2}$ . This time reparametrization is precisely the boundary limit of the black

---

in terms of  $F$ . In transferring from  $g_{uv}$  to  $F$ , there is a redundancy since all  $F$  related by  $\text{PSL}(2, \mathbb{R})$  transformations as  $F \rightarrow \frac{aF+b}{cF+d}$  yield the same metric, and hence have to be identified. In hyperbolic spaces with other topologies, only a subgroup of  $\text{PSL}(2, \mathbb{R})$  is redundant when demanding compatibility with suitable periodicity conditions on  $F$ . This will become more clear in Sect. 3 and in particular in Sect. 3.6.

hole frame on the  $\text{AdS}_2$  manifold (2.36) we described earlier, with  $\sqrt{\frac{\mu}{a}} = \frac{\pi}{\beta}$ , relating the temperature  $\beta^{-1}$  to the parameter  $\mu$ . See Fig. 4.

—o—

In order to write Eq. (2.52) in a more manageable form, we first need to consider a different quantity in this model. The holographic stress tensor can be computed using standard techniques, supplemented here with the proper boundary coordinates (Jensen 2016; Maldacena et al. 2016b; Engelsöy et al. 2016). We will not detail the calculation itself here. In 0+1d, this quantity is the total energy contained in the spacetime at any time  $t$  and, without boundary sources for massive bulk fields, can be written suggestively as:

$$E(t) = -\frac{a}{16\pi G_N} \{F, t\}, \quad \{F, t\} \equiv \frac{F'''}{F'} - \frac{3}{2} \left( \frac{F''}{F'} \right)^2. \quad (2.57)$$

The prefactor  $\frac{16\pi G_N}{a}$  will appear a lot, and plays the role of the actual gravitational coupling constant in  $\text{NAdS}_2$ . Note that this gravitational coupling is power-counting superrenormalizable. We introduce the shorthand notation in terms of the quantity  $C$  of dimension length:

$$C \equiv \frac{a}{16\pi G_N}. \quad (2.58)$$

Let us apply the relation (2.57) to the black hole coordinates, and study black hole thermodynamics from this perspective. For the solution (2.55)  $F(t) = \sqrt{\frac{a}{\mu}} \tanh \sqrt{\frac{\mu}{a}} t$ , we immediately find the stress tensor (2.57):  $E(t) = \frac{\mu}{8\pi G_N}$ , agreeing with our previous analysis around Eq. (2.46). From this, we find the energy-temperature relation:

$$T = \frac{1}{\pi} \sqrt{\frac{E}{2C}} \quad \Rightarrow \quad E(T) = 2\pi^2 C T^2. \quad (2.59)$$

Using  $\frac{\partial S}{\partial E} = \frac{1}{T}$ , and identifying the zero-energy entropy with the parameter  $S_0$  in the action, we can find the thermal entropy contained in the system as

$$S(E) = S_0 + 2\pi \sqrt{2CE} = S_0 + \frac{\sqrt{a\mu}}{4G_N}. \quad (2.60)$$

In the canonical ensemble, the entropy has a behavior linear with the temperature given by

$$S(T) = S_0 + 4\pi^2 C T. \quad (2.61)$$

We can compare this entropy to the Bekenstein–Hawking (BH) entropy of a black hole  $S_{\text{BH}} = \frac{A}{4G_N}$ , where  $A$  is the horizon area. The analogous formula in 1+1d requires the following modifications. Firstly, the area of a  $d-1$  sphere is  $S^{d-1} = \frac{2\pi^{d/2}}{\Gamma(d/2)}$ , and hence  $A \rightarrow 2$  when  $d \rightarrow 1$ . Because the 0-sphere consists of 2 disjoint points, we

obtain  $A = 1$  for a single horizon. Secondly, in our model the effective Newton's constant in dilaton gravity is spacetime-dependent:  $1/G_{N,\text{eff}}(x) = (\Phi_0 + \Phi(x))/G_N$ , which evaluates at the horizon to  $(\Phi_0 + \Phi_h)/G_N$ , where  $\Phi_h$  is the horizon value of the dilaton field. Hence the 1+1d analog of the Bekenstein–Hawking entropy is

$$S_{\text{BH}} = \frac{\Phi_0 + \Phi_h}{4G_N}. \quad (2.62)$$

We remind the reader that for convenience we separated the constant background piece of the dilaton  $\Phi_0$  from the varying piece  $\Phi$ . The entropy is simply proportional to the total value of the dilaton evaluated at the horizon. Using now the dilaton profile for the black hole solution (2.47)

$$\Phi = \frac{a - \mu U(u)V(v)}{U(u) - V(v)} = \sqrt{a\mu} \coth\left(\sqrt{\frac{\mu}{a}} 2z\right), \quad (2.63)$$

we obtain  $\Phi_h = \sqrt{a\mu}$ , and we see that the resulting BH entropy (2.62) matches the thermal entropy (2.60). Finally, we note that the entropy can be equally computed from the Ryu–Takayanagi (RT) formula (Ryu and Takayanagi 2006) as:

$$S = \text{Min}_z \frac{\Phi_0 + \Phi(z)}{4G_N}, \quad (2.64)$$

where the minimum occurs at the black hole horizon  $z \rightarrow +\infty$ . Non-trivial checks of this 2d version of the RT formula in more complicated geometries can be found in Goel et al. (2019).

Now let's continue our main analysis, where we come to the main point. Equipped with the formula (2.57), taking suitable derivatives of Eq. (2.52), we can massage (2.52) into an ODE in terms of the energy  $E(t) = -C\{F, t\}$ <sup>12</sup>

$$\frac{dE(t)}{dt} = (T_{VV}(t) - T_{UU}(t))F'^2|_{\partial\mathcal{M}} = T_{vv}(t) - T_{uu}(t)|_{\partial\mathcal{M}}, \quad (2.65)$$

where the right hand side is the net influx of matter from the holographic boundary in the proper coordinates  $(u, v)$ . Hence the differential equation that determines the frame  $F(t)$  is just energy conservation, where the wiggly boundary curve  $(F(t), Z(t))$  reacts to a net energy influx from the holographic boundary. In particular, injecting positive energy causes the boundary trajectory to bend towards the actual boundary as illustrated before around Eq. (2.56) and vice versa. With this interpretation, the right hand side of Eq. (2.65) can be immediately generalized to arbitrary non-conformal matter as well. This equation of motion is the main result of the classical analysis of the JT equations of motion.<sup>13</sup>

<sup>12</sup> The sequence applied to (2.52) is: differentiate, divide by  $F'$ , differentiate, divide by  $F'$ , differentiate, multiply by  $F'$ . Then use that  $\{F, t\}' := F'((F''/F')'/F')'$ .

<sup>13</sup> This is reminiscent of hydrodynamics where the only conserved quantity is the energy (Jensen 2016). A related similarity is that, as we will point out in Sect. 3.6, the only local observable operator in the resulting quantum theory is constructed from the stress tensor.

Some of the above mysterious properties of 2d dilaton gravity are naturally explained when embedding the model within higher-dimensional gravity. For near extremal charged black holes in 4d considered in Sect. 2.1.2, the boundary condition on the dilaton arises from gluing the  $\text{AdS}_2$  throat to the exterior region (Almheiri and Kang 2016; Nayak et al. 2018). Another example is the observation that the dynamics of  $s$ -waves of 3d gravity are governed by 2d JT gravity (Achucarro and Ortiz 1993). Indeed, parametrizing a 3d spherically symmetric geometry by the metric ansatz:

$$ds^2 = g_{\mu\nu}^{(2)}(x^\mu)dx^\mu dx^\nu + \Phi^2(x^\mu)d\varphi^2, \quad \mu, \nu = t, r, \quad \varphi \sim \varphi + 2\pi, \quad (2.66)$$

the 3d pure gravity action reduces to JT gravity:

$$\frac{1}{16\pi G_N^{(3)}} \int d^3x \sqrt{-g}(R^{(3)} - \Lambda) = \frac{2\pi}{16\pi G_N^{(3)}} \int d^2x \sqrt{-g}\Phi(R^{(2)} - \Lambda), \quad (2.67)$$

where  $G_N^{(2)} = G_N^{(3)}/2\pi$ , and the dilaton field  $\Phi$  originated from the  $g_{\varphi\varphi}$  metric component. This immediately explains the BH entropy formula (2.62) since  $S_{\text{BH}} = 2\pi\Phi_h/4G_N^{(3)} = \Phi_h/4G_N^{(2)}$  where  $2\pi\Phi_h$  is the horizon circumference in 3d gravity. Moreover, our dilaton boundary condition (2.51) can be seen to directly arise from the usual Fefferman–Graham gauge asymptotic expansion of the full 3d metric. In this sense, JT gravity resolves the tension in  $\text{AdS}_2/\text{CFT}_1$  by being secretly a 3d gravity model in disguise.

### 2.3.2 Aside: general dilaton gravity

It is important to know that some of this story can be generalized to arbitrary dilaton gravity models with action (2.4), and  $\Phi$  in that expression denotes the total dilaton without separating into a background  $\Phi_0$  and a varying piece. Starting with (2.4), one can write down the generic black hole (sourceless) solution, generalizing the JT solution (2.38). We can partially fix coordinates in the bulk by identifying  $\Phi(r) = \frac{a}{2}r$ , for a radial coordinate  $r$  with  $r \rightarrow \infty$  the boundary of our system. In this gauge, the dilaton can be thought of as a ruler providing a definition of proper radial distance.

The generic black hole solution after further gauge fixing can be written as

$$ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)}, \quad f(r) = \frac{4}{a^2} \int_{\Phi_h}^{\Phi(r)} d\Phi U(\Phi). \quad (2.68)$$

The location  $r = r_h$  is the black hole horizon and  $\Phi_h \equiv \Phi(r_h) = \frac{a}{2}r_h$ . The solution near  $r_h$  only has the right spacetime signature if  $U(\Phi_h) \geq 0$ , and the dilaton potential for  $\Phi > \Phi_h$  should be such that  $f(r)$  is positive everywhere outside the horizon. Expanding near  $r_h$  one obtains the Hawking temperature:

$$T = \frac{U(\Phi_h)}{2\pi a}, \quad (2.69)$$

conjugate to the choice of time coordinate  $t$  in (2.68).



The location  $r = 0$  is the would-be singularity where  $\Phi \rightarrow 0$ . From the embedding of dilaton gravity models in higher dimensions, since the dilaton is proportional to the area of the transverse sphere, this is the curvature singularity of the higher-dimensional model. In 2d however, this is a strongly coupled singularity where the effective Newton's constant diverges. Since there is only a single curvature invariant, and since the  $\Phi$  equation of motion yields

$$R = -U'(\Phi), \quad (2.70)$$

a curvature singularity in 2d can only arise when  $U'(\Phi_{\text{sing.}}) \rightarrow \pm\infty$  for some  $\Phi_{\text{sing.}}$ , a situation that is usually excluded by choice of dilaton potential. However, one interesting example where this does happen is in the dilaton gravity description of Liouville gravity, for which we will make some comments in Sect. 5.10.

The general  $E(T)$  and  $S(T)$  black hole first law equations for (2.68) are:

$$E(T) = \frac{1}{8\pi G_N a} \int^{U^{-1}(2\pi a T)} U(\Phi) d\Phi, \quad S(T) = \frac{U^{-1}(2\pi a T)}{4G_N} = \frac{\Phi_h}{4G_N}, \quad (2.71)$$

where  $U^{-1}(\cdot)$  is the inverse function of  $U(\Phi)$ . This gives a physical interpretation to the dilaton potential:  $T(S) = \frac{1}{2\pi a} U(\Phi = 4G_N S)$  is the equation of state, with the entropy identified with the dilaton and the temperature with the value of the potential. The connection with the entropy is the correct interpretation of the dilaton field in 2d dilaton-gravity.

We leave it as an exercise to apply the above relations to the JT case, where in the conventions of this subsection  $U(\Phi) = 2(\Phi - \Phi_0)$ . We note that a linear dilaton potential is the only one where the resulting thermodynamical relations only depend on the ratio  $a/G_N \sim C$ .

### 2.3.3 The Schwarzian action

We were led to a dynamical boundary curve described by (2.65), which in the absence of matter injections is governed by the simple equation

$$\frac{d}{dt} \{F, t\} = 0. \quad (2.72)$$

This equation of motion can be found as the Euler-Lagrange equation of a suitable action: the Schwarzian action. This action describing the dynamics of JT gravity is:

$$S = -C \int dt \{F, t\}, \quad \{F, t\} \equiv \frac{F'''}{F'} - \frac{3}{2} \left( \frac{F''}{F'} \right)^2. \quad (2.73)$$

and contains the Schwarzian derivative as the Lagrangian. As a higher-derivative model, this looks complicated at first sight. After integrating by parts using

$\int dt \delta \{F, t\} = - \int dt \frac{\{F, t\}'}{F'} \delta F$ , the equation of motion following from (2.73) is indeed (2.72) as promised.

It is now relatively straightforward to write down the effective action whose equations of motion give the full sourced Eq. (2.65). We just minimally couple a generic 2d matter Lagrangian  $\mathcal{L}_m(\phi, \partial_F \phi)$ , where we denoted its dependence on the fields  $\phi$  and the time-derivatives  $\partial_F \phi$ , to the wiggly boundary curve and extrapolate the time reparametrization  $t \rightarrow F(t)$  throughout the entire bulk. We can then write the total Lorentzian action in a hybrid way as<sup>14</sup>

$$S = -C \int dt \{F, t\} + \int dF dZ \mathcal{L}_m(\phi, \partial_F \phi) \quad (2.74)$$

$$= -C \int dt \{F, t\} + \int dt dZ F' \mathcal{L}_m(\phi, \frac{1}{F'} \partial_t \phi), \quad (2.75)$$

where in the second line we made the  $F$ -dependence of the matter part explicit. Varying with respect to  $F(t)$  yields

$$\delta S = \int dt \left[ C \frac{\{F, t\}'}{F'} \delta F - \int dZ \left( \partial_F \phi \frac{\partial \mathcal{L}_m}{\partial \partial_F \phi} - \mathcal{L}_m \right) \delta F' \right]. \quad (2.76)$$

Identifying the matter Hamiltonian density as  $\mathcal{H}_m \equiv \partial_F \phi \frac{\partial \mathcal{L}_m}{\partial \partial_F \phi} - \mathcal{L}_m$ , then leads to the sourced equation of motion:

$$-C \{F, t\}' = F'^2 \frac{dH_m}{dF}, \quad H_m = \int dZ \mathcal{H}_m. \quad (2.77)$$

The quantity on the RHS is the total change in the energy in the matter sector since by energy conservation in the matter sector  $\frac{dH_m}{dF} = T_{VV} - T_{UU}|_{\partial \mathcal{M}}$ . The explicit  $F'^2$  converts the matter energy fluxes to the  $t$  time coordinate. This is precisely the content of the sourced equation of motion (2.65).

In order to proceed, it is convenient to summarize some of the salient mathematical properties of the Schwarzian derivative:

1. By explicit calculation, we have the following representation of the Schwarzian derivative:

$$\frac{1}{6} \{F, t\} = \lim_{t' \rightarrow t} \left( \frac{F'(t') F'(t)}{((F(t') - F(t))^2} - \frac{1}{(t' - t)^2} \right). \quad (2.78)$$

2. Composition law:

$$\{F(G(t)), t\} = \{G(t), t\} + G'(t)^2 \{F(G), G\}(t). \quad (2.79)$$

<sup>14</sup> Some comments are in order here. Firstly, we can always extrapolate the boundary reparametrization into the bulk in any way we want. Here we choose a simple way. Secondly, the  $Z$ -integral only ranges to  $Z(t) = \epsilon F'(t)$  but this is infinitesimal and does not play a role in the current argument. This term would be important for massive matter boundary sources turned on, but this will not be studied here.

This can be proven easily using the representation (2.78).

3. The Schwarzian derivative itself is  $\text{PSL}(2, \mathbb{R})$  invariant: if  $F = \frac{aG+b}{cG+d}$ , then  $\{F, t\} = \{G, t\}$ . This can again be easily shown by using the above identity (2.78).
4. The solution to the equation  $\{F, t\} = 0$  is

$$F(t) = \frac{at + b}{ct + d}. \quad (2.80)$$

To see this, note that  $F(t) = t$  has vanishing Schwarzian derivative  $\{F, t\} = 0$ . By property 3, this means also  $\frac{at+b}{ct+d}$  has zero Schwarzian derivative. This is a three-parameter family of solutions. The differential equation  $\{F, t\} = 0$  itself is third order and hence requires three integration constants as initial conditions,  $F(0)$ ,  $F'(0)$  and  $F''(0)$  which can be one-to-one mapped into the three parameters of the  $\text{PSL}(2, \mathbb{R})$  transformation parametrized by  $a, b, c, d$  above.

5. A converse of property 3: if two functions have equal Schwarzian derivative, then they differ at most by a Möbius transformation:  $\{F, t\} = \{G, t\}$  if  $\exists \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  such that  $F = \frac{aG+b}{cG+d}$ . To prove this, we first use the composition law (2.79) to rewrite the proposition as  $\{F, G\} = 0$ . Then by property 4, we immediately get  $F = \frac{aG+b}{cG+d}$ .

Now we are equipped to discuss the solution of Eq. (2.72) in more detail. We immediately have:

$$\frac{d}{dt} \{F, t\} = 0 \Rightarrow \{F, t\} = \text{constant}. \quad (2.81)$$

For any fixed value of this constant, which is interpretable as proportional to the energy (2.57), this equation has a unique solution up to an  $\text{PSL}(2, \mathbb{R})$  Möbius ambiguity

$$F(t) \rightarrow \frac{aF(t) + b}{cF(t) + d}. \quad (2.82)$$

The Schwarzian action (2.73) has a  $\text{PSL}(2, \mathbb{R})$  symmetry, under  $F \rightarrow \frac{aF+b}{cF+d}$ . This leads to the following conserved Noether charges:

$$Q_- = C \left[ \frac{F'''}{F'^2} - \frac{F''^2}{F'^3} \right], \quad (2.83)$$

$$Q_0 = C \left[ \frac{F'''F}{F'^2} - \frac{FF''^2}{F'^3} - \frac{F''}{F'} \right], \quad (2.84)$$

$$Q_+ = C \left[ \frac{F'''F^2}{F'^2} - \frac{F^2F''^2}{F'^3} - \frac{2FF''}{F'} + 2F' \right], \quad (2.85)$$

associated to the one-parameter subgroups generated by  $b$  for  $Q_-$ ,  $a$  and  $d$  for  $Q_0$ , and  $c$  for  $Q_+$  respectively. One checks explicitly that these conserved charges form

an  $\mathfrak{sl}(2, \mathbb{R})$  algebra, and yield the Hamiltonian as the quadratic Casimir:

$$H \equiv -C\{F, t\} = \frac{1}{2C} \left[ Q_0^2 - \frac{1}{2} \{Q_+, Q_-\} \right], \quad (2.86)$$

which is also the Noether charge of Eq. (2.73) corresponding to time translations in the coordinate time  $t \rightarrow t + c$ .<sup>15</sup>

We would like to stress that even though the transformations generated by  $Q$  that act on  $F$  are symmetries, the transformation  $t \rightarrow \frac{at+b}{ct+d}$  is not a symmetry unless  $t \rightarrow t + b$ ; the boundary conformal symmetry is broken for  $C \neq 0$ .

As emphasized earlier, configurations that are related by Möbius transformations  $F \rightarrow \frac{aF+b}{cF+d}$ , are to be identified in the gravitational model. In order for  $F(t)$  to be a good reparametrization, we also require  $F' \geq 0$  anywhere on its domain. Depending on the precise situation, the set of all allowed  $F(t)$  can get further restrictions as we make precise in Sect. 3. For now, let us call this space  $G$ , and the Möbius subgroup (which is a redundancy) as  $H = \text{PSL}(2, \mathbb{R})$ . Then the set of all possible  $F(t)$  is described by the coset  $G/H$ . This is a situation familiar from e.g. the pion effective QFT in particle physics, where we view  $F$  as describing the (pseudo) Goldstone boson degrees of freedom in the symmetry breaking  $G \rightarrow H$ . The Schwarzian action itself (2.73) is however only invariant under the (gauged) subgroup  $H$ , illustrating that the original group  $G$  is broken both explicitly and spontaneously, hence the adverb *pseudo*. The prefactor  $C \sim a$  of the Schwarzian action hence represents the scale of explicit breaking of the 1d group of reparametrizations, i.e. the 1d conformal group. This is an interpretation we indeed encountered earlier in Sect. 2.3.1 when regularizing the dilaton asymptotics.

The Schwarzian term itself is the lowest order non-trivial local term one could possibly write down as a Lagrangian that is invariant under  $H$ . As such, the Schwarzian model exhibits universality, and is fully determined by this specific pattern of symmetry breaking.

It is in this language that the Schwarzian action first emerged in A. Kitaev's work in 2015 on the Sachdev–Ye–Kitaev (SYK) model (Sachdev and Ye 1993; Kitaev 2014, 2015a, b): the dynamics of SYK at low energies is approximated by a reparametrization  $F(t)$  describing an explicitly (by UV effects) and spontaneously (by making a specific choice of  $F(t) \bmod \text{PSL}(2, \mathbb{R})$ ) broken symmetry.

Hence since the  $\text{PSL}(2, \mathbb{R})$  symmetry is a gauge redundancy, this means only configurations with zero overall  $\text{PSL}(2, \mathbb{R})$  charges are physically observable. The correct interpretation for this is that the gravitational charges have to be compensated by other sectors, either in a two-sided configuration, or by adding additional matter sectors that can carry these charges.

### 2.3.4 Geometric derivation from the action

There is a faster derivation of the Schwarzian description that is more geometric in nature (Maldacena et al. 2016b) and stays at the level of the action. It will be our

<sup>15</sup> The expression for the energy (2.57) is to be compared to this Noether charge, which can be used to deduce that the prefactor of the Schwarzian action is indeed  $-C$  as in (2.73).



**Fig. 5** Cutout of the Poincaré upper half plane

starting point in the next sections. Let us go back to the Euclidean JT action:

$$I_{JT}[g, \Phi] = -\frac{1}{16\pi G_N} \int \sqrt{g} \Phi (R + 2) - \frac{1}{8\pi G_N} \oint \sqrt{h} \Phi (K - 1). \quad (2.87)$$

We first vary with respect to the dilaton field, to find  $R = -2$ , or a patch of the 2d hyperbolic plane  $ds^2 = \frac{dT^2 + dZ^2}{Z^2}$ .

The dynamics is now fully governed by the GHY boundary term on some wiggly boundary curve. One can draw this as the hyperbolic upper half plane, where we cut out and remove a shape that is close to the actual boundary  $Z = 0$ . This is the Euclidean counterpart of our discussion so far. We illustrate the procedure in Fig. 5.

We parametrize the boundary curve as  $(T = F(\tau), Z(\tau))$ . Fixing the boundary metric to  $\sqrt{h} = 1/\epsilon$  imposes  $Z(\tau) = \epsilon F'(\tau)$ . We can directly evaluate the extrinsic curvature trace:

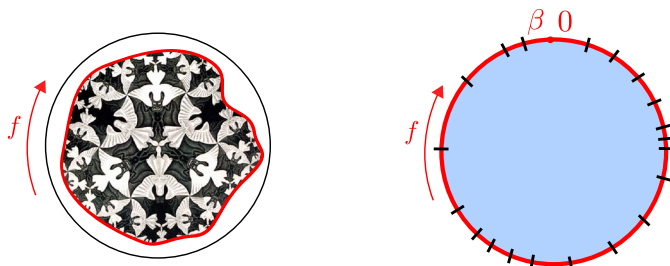
$$K = \frac{F'(F'^2 + Z'^2 + ZZ'') - ZZ'F''}{(F'^2 + Z'^2)^{3/2}} = 1 + \epsilon^2 \{F, \tau\} + \mathcal{O}(\epsilon^4), \quad (2.88)$$

using standard techniques.<sup>16</sup> We also fix the boundary dilaton to be  $\Phi_b \equiv \Phi|_{\partial} = \frac{a}{2\epsilon}$ , which appears in the GHY term in the action (2.87). Plugging this back in the Euclidean action, we immediately obtain the Schwarzian action:

$$I_{JT}[F] = -C \int d\tau \{F, \tau\}, \quad C = \frac{a}{16\pi G_N}. \quad (2.89)$$

The added benefit is that one can directly generalize this to a Euclidean gravitational path integral argument, as we will do further on in Sect. 3.

<sup>16</sup> Some technical details. The non-zero Christoffel symbols are  $\Gamma_{ZZ}^Z = -\Gamma_{TT}^Z = -\Gamma_{TZ}^T = 1/Z$ . The extrinsic curvature is computed as  $K = h^{tt} K_{tt} = h^{tt} M_t^\mu M_t^\nu \nabla_\mu n_\nu$  where  $M_t^\mu = \partial_t x^\mu$ , the pull-back metric is  $h_{tt} \equiv M_t^\mu M_t^\nu g_{\mu\nu}$ . The unit-normalized outwards-pointing normal co-vector is  $n_\mu = \left( \frac{Z'}{Z(Z'^2 + F'^2)^{1/2}}, \frac{-F'}{Z(Z'^2 + F'^2)^{1/2}} \right)$ . The  $t$ -dependence of the objects  $Z'$  and  $F'$  can be rewritten into a  $F$ -dependence using the chain rule:  $d(\cdot)/dF = (\cdot)' / F'$ , making the normal co-vector a function of  $(F, Z)$ .



**Fig. 6** The left figure depicts the wiggly thermal boundary curve as a cutout of the Poincaré disk. The right figure shows the boundary clock ticking pattern. The number of ticks on the clock  $f(\tau)$  is fixed by the periodicity constraint (2.91), but their spreading along the thermal circle is not

The Schwarzian action describes the reparametrization  $F(\tau)$  as the dynamical degree of freedom. Much like in well-known studies of 3d Chern–Simons and 2d WZW boundary CFT (Elitzur et al. 1989) (and in the BF dimensional reduction discussed briefly in Sect. 3.7.4), the physical boundary degrees of freedom are the would-be large gauge (or diff) transformations, which have become physical and observable in the presence of a boundary.<sup>17</sup>

Up to now, we have always described our reparametrization as referring to the Poincaré time coordinate  $F$ . This is not necessary, and in fact it is more natural to choose a different reference frame when considering the thermal system, that is the black hole frame. We define<sup>18</sup>:

$$F(\tau) \equiv \tan \frac{\pi}{\beta} f(\tau), \quad (2.90)$$

in terms of a new variable  $f(\tau)$ , which has the properties:

$$f(\tau + \beta) = f(\tau) + \beta, \quad f'(\tau) \geq 0. \quad (2.91)$$

These two properties give  $f(\tau)$  the interpretation of a time reparametrization of the boundary thermal circle:  $f \in \text{Diff } S^1$ . The metric boundary condition (2.48) can then be rephrased as a fixed length boundary condition, where we keep fixed the regularized boundary length to  $\frac{\beta}{\epsilon}$ . The rescaled (renormalized) boundary length is  $\beta$ , with the physical interpretation of the inverse temperature.

Two useful graphical representations, emphasizing different perspectives, are the following.

The classical equation of motion of Eq. (2.89) for a thermal reparametrization (2.90) with the above periodicity requirements, leads to the unique classical solution

<sup>17</sup> We will encounter different boundary conditions later on for interior boundaries. We refer to Goel et al. (2021), Ferrari (2021) for classifications of boundary conditions.

<sup>18</sup> This is the Euclidean version of Eq. (2.36), setting both  $F \rightarrow -iF$  and  $f \rightarrow -if$ , where the prefactor  $\beta/\pi$  is removed. This prefactor is part of the  $\text{SL}(2, \mathbb{R})$  gauge isometry group and is hence not observable. It was however convenient to include it earlier because the time coordinate then has a clean  $\beta \rightarrow +\infty$  limit back to the Poincaré time coordinate  $F \rightarrow f$ .

$f(\tau) = \tau$ , up to  $\text{PSL}(2, \mathbb{R})$ . This is precisely the (Wick-rotated) black hole solution (2.36).

## 2.4 Quantum matter in classical gravity

Before starting with the quantum solution of JT gravity in the next sections, we want to present a physically interesting intermediate situation where we study quantum matter in a classical gravitational model.

The sourced Schwarzian equation of motion (2.65) is this model's Einstein equation  $G_{\mu\nu} = 8\pi G_N T_{\mu\nu}$ . In this subsection, we promote the matter sources on the right hand side to be quantum mechanical, while retaining classicality for the gravitational sector, i.e.  $G_{\mu\nu} = 8\pi G_N \langle T_{\mu\nu} \rangle$ . Such an approximation is only valid as long as the quantum effects of gravity are subdominant compared to the matter quantum effects. For a 2d matter CFT with central charge  $c$ , this can be realized when  $c \gg 1$ , and as long as any bulk black hole is not Planck-scale.

Quantum effects in a matter sector that is a 2d CFT are governed by the conformal anomaly (Davies et al. 1976; Christensen and Fulling 1977), see also the textbook by Fabbri and Navarro-Salas (2005). This leads to the following stress tensor components, to be inserted in (2.42):

$$\begin{aligned}\langle T_{uu} \rangle &= -\frac{c}{12\pi} ((\partial_u \omega)^2 - \partial_u^2 \omega) + \langle : T_{uu}(u) : \rangle, \\ \langle T_{vv} \rangle &= -\frac{c}{12\pi} ((\partial_v \omega)^2 - \partial_v^2 \omega) + \langle : T_{vv}(v) : \rangle, \\ \langle T_{uv} \rangle &= -\frac{c}{12\pi} \partial_u \partial_v \omega.\end{aligned}\tag{2.92}$$

The objects  $T_{uu}$ ,  $T_{vv}$  and  $T_{uv}$  transform as tensor components and are covariantly conserved  $\nabla_\mu T^{\mu\nu} = 0$ , as should be for consistency of Einstein's equations. This stress tensor is decomposed as shown into a Casimir contribution, that depends solely on the metric through  $\omega(u, v)$  and not on the quantum state, plus an operational part that depends on the quantum state and also on a reference state as we explain. Both transform anomalously under coordinate transformations, but their sum is a covariant tensor.

The terms on the RHS  $: T_{uu}(u) :$  and  $: T_{vv}(v) :$  are frame-dependent due to normal-ordering with respect to a specific vacuum. Let us be more explicit. Suppose we have a free boson CFT with  $c = 1$ , with bulk field equation  $\square\phi = 0$ . Then  $T_{uu} = \partial_u \phi \partial_u \phi$  is a composite operator in the quantum matter sector, and requires regularization and renormalization. We define the renormalized stress tensor by subtracting its expectation value in the vacuum state  $|0_u\rangle$  defined using positive frequency modes in the  $u$ ,  $v$ -coordinates, as:

$$: T_{uu} : \equiv \partial_u \phi \partial_u \phi - \langle 0_u | \partial_u \phi \partial_u \phi | 0_u \rangle\tag{2.93}$$

$$= \lim_{u \rightarrow u'} \left( \partial_u \phi(u) \partial_u \phi(u') + \frac{1}{4\pi} \frac{1}{(u - u')^2} \right),\tag{2.94}$$

and analogously for  $T_{vv}$ . This stress tensor is operationally defined and measured by local observers with measuring devices calibrated to their vacuum. Between different frames  $u$  and  $U$  (i.e. calibrated w.r.t. different vacua), it transforms anomalously as:

$$:T_{uu} := \left(\frac{dU}{du}\right)^2 :T_{UU} : - \frac{c}{24\pi} \{U, u\}. \quad (2.95)$$

In our specific set-up, the influence of the quantum matter sources (2.92) is the following. (1) The non-zero component  $\langle T_{uv} \rangle$  leads in the classical equation of motion to a simple shift in the dilaton  $\Phi \rightarrow \Phi + \frac{cG_N}{3}$  (Almheiri and Polchinski 2015), which has no influence on the boundary equation of motion at  $\Phi \rightarrow \infty$ . (2) For the  $\text{AdS}_2$  geometry with the specific formula (2.31) for  $\omega(u, v)$ , the  $T_{uu}$  and  $T_{vv}$  components are still holomorphic:  $T_{uu}(u)$  and  $T_{vv}(v)$ , leading to exactly the same sourced Schwarzian equation of motion (2.65).

Hence evaluating these stress tensor components (2.92) at the boundary, we have:

$$\langle T_{uu}(t) \rangle|_{\partial\mathcal{M}} = \frac{c}{24\pi} \{F, t\} + \langle :T_{uu}(t): \rangle|_{\partial\mathcal{M}} \quad (2.96)$$

$$\langle T_{vv}(t) \rangle|_{\partial\mathcal{M}} = \frac{c}{24\pi} \{F, t\} + \langle :T_{vv}(t): \rangle|_{\partial\mathcal{M}}, \quad (2.97)$$

and the equation of motion (2.65) can be written in two equivalent ways:

$$\frac{dE}{dt} = \langle T_{vv}(t) \rangle - \langle T_{uu}(t) \rangle|_{\partial\mathcal{M}} = \langle :T_{vv}(t): \rangle - \langle :T_{uu}(t): \rangle|_{\partial\mathcal{M}}, \quad (2.98)$$

in terms of either the actual net influx of energy, or the net *operationally measured* influx of energy, evaluated in the matter quantum state of interest.

### 2.4.1 Application: Hawking–Unruh effect and information loss

We can apply these equations directly to derive the 2d analog of Unruh and Hawking radiation, and probe information flows in the model.

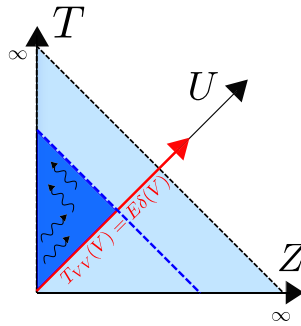
Black holes in asymptotically AdS spacetimes do not typically evaporate due to standard reflecting boundary conditions at the holographic boundary.

#### Example (continued):

We impose reflecting boundary conditions for the matter sector at the holographic boundary after the initial energy injection at  $t = 0$ , see Fig. 7. The matter sector state is the Poincaré vacuum since that is the initial vacuum state before the pulse. This leads to the boundary conditions  $\langle T_{uu}(t) \rangle|_{\partial\mathcal{M}} = 0 = \langle T_{vv}(t) \rangle|_{\partial\mathcal{M}}$ , and hence the net energy in the bulk spacetime does not change.

However, the operational stress tensor components are nonzero by the conformal anomaly (2.96). Evaluating these for the frame  $F(t) = \frac{\beta}{\pi} \tanh \frac{\pi}{\beta} t$ , with  $\beta =$





**Fig. 7** Penrose diagram of formation and subsequent Hawking radiation of a black hole, after injecting an energy pulse

$\pi\sqrt{\frac{8\pi G_N E}{a}}$  after the initial pulse, we have

$$\langle : T_{uu}(t) : \rangle|_{\partial\mathcal{M}} = \langle : T_{vv}(t) : \rangle|_{\partial\mathcal{M}} = c \frac{\pi}{12\beta^2}, \quad t > 0. \quad (2.99)$$

This is the Unruh heat bath with equal left- and right-moving energy flux, and no macroscopic flow of energy. This is the  $\text{AdS}_2$  analog of physics in the Minkowski vacuum for flat space, or the Hartle-Hawking vacuum for the Schwarzschild metric. The energy density profile is :  $T_0^0(\rho) := c \frac{\pi}{6\beta^2} \frac{1}{\sinh^2 \rho}$  where  $\rho$  is the proper distance to the horizon, as introduced around (2.39). The conformal anomaly technique provides a very efficient derivation of this observer-dependent physics (Spradlin and Strominger 1999). Notice that the thermal flux immediately sets in after the black hole is created; no transient regime is present in this model and it has instant thermalization.

—○—

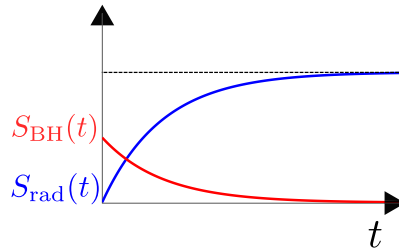
To model actual evaporation, let's instead implement absorbing boundary conditions at the holographic boundary, where the observer on the boundary line removes all Hawking radiation he detects in his local frame (detector calibrated w.r.t. the vacuum associated with his own preferred time  $t$  coordinate)  $\langle : T_{vv}(t) : \rangle|_{\partial\mathcal{M}} = 0$  (Engelsöy et al. 2016). The boundary condition for the outgoing component remains the same  $\langle T_{uu}(t) \rangle|_{\partial\mathcal{M}} = 0$ . This leads to the equation of motion:

$$\frac{dE}{dt} \equiv -C \frac{d}{dt} \{F, t\} = \langle T_{vv}(t) \rangle - \langle T_{uu}(t) \rangle|_{\partial\mathcal{M}} = \frac{c}{24\pi} \{F, t\}, \quad t > 0, \quad (2.100)$$

solvable by an exponentially decaying energy profile<sup>19</sup>:

$$E(t) = E e^{-\frac{c}{24\pi C} t} \theta(t). \quad (2.101)$$

<sup>19</sup> This profile is in agreement with a quasi-static approximation where one uses the black hole first law  $T \sim \sqrt{E}$ , combined with the 2d Stefan-Boltzman law  $\frac{dE}{dt} \sim T^2$ .



**Fig. 8** Coarse-grained entropy  $S_{\text{BH}}(t)$  and fine-grained early-late matter entanglement entropy  $S_{\text{ren}}(t)$  as a function of time  $t$  during evaporation

From this solution, one can then find the explicit time reparametrization solution

$$F(t) = \frac{2}{\alpha A} \frac{I_0(\alpha) K_0(\alpha e^{-\frac{c}{48\pi C} t}) - K_0(\alpha) I_0(\alpha e^{-\frac{c}{48\pi C} t})}{I_1(\alpha) K_0(\alpha e^{-\frac{c}{48\pi C} t}) + K_1(\alpha) I_0(\alpha e^{-\frac{c}{48\pi C} t})}, \quad (2.102)$$

in terms of modified Bessel functions  $I_0$  and  $K_0$ , and where  $\alpha = \frac{24\pi}{c} \sqrt{2CE}$ .<sup>20</sup>

From this dissipating energy profile, using (2.60), one can likewise find an instantaneous Bekenstein–Hawking entropy as (in this argument we consider entropies above extremality, and therefore  $S_0$  does not play a role)

$$S_{\text{BH}}(t) = 2\pi \sqrt{2CE} e^{-\frac{c}{48\pi C} t}, \quad (2.103)$$

which is a measure for the coarse-grained entropy of the dissipating black hole during evaporation. Using the 2d CFT entanglement entropy formula, one can also compute the fine-grained (renormalized) matter entanglement entropy between the early radiation (hitting the boundary at times  $< t$ ) and the late radiation (arriving at the boundary at times  $> t$ ) (Mertens 2019). For a macroscopic black hole where  $E \gg 1/C$ , one gets:

$$S_{\text{rad}}(t) = \frac{c}{12} \ln \frac{F(t)^2}{t^2 F'(t)} \approx 4\pi \sqrt{2CE} \left(1 - e^{-\frac{c}{48\pi C} t}\right). \quad (2.104)$$

This profile is always increasing. Both entropies are illustrated as a function of time  $t$  in Fig. 8. From the above, one sees that the matter entropy  $S_{\text{rad}}(t)$  rises twice as fast as the black hole entropy  $S_{\text{BH}}(t)$  decreases, which is no coincidence (Zurek 1982; Fiola et al. 1994). Indeed, assuming black hole evaporation is an irreversible process of radiating into empty space, we can relate the black hole first law  $dS_{\text{BH}} = dE/T$  and the 2d relativistic ideal gas law  $E = ST/2$  of the thermal atmosphere, yielding the relation  $dS_{\text{BH}} = -dS/2$ .

Since the matter entropy  $S_{\text{rad}}(t)$  never comes down to zero, this is a clean quantitative illustration of Hawking’s information loss paradox, and shows that unitarity of the full

<sup>20</sup> Generalizations of the dissipating black hole when including charges and/or supercharges can be found in De Vuyst and Mertens (2023).

quantum system is not apparent in the semi-classical gravitational approximation. This and related puzzles, motivates us to go beyond classical gravity, a problem to which we turn next.

### 3 Quantum Jackiw–Teitelboim gravity

The previous section has focused on aspects of JT gravity that can be understood from a classical gravity treatment. One of the attractive features of JT gravity is the fact that one can also study quantum gravitational effects exactly. We first study the quantum corrections to the black hole spectrum as derived from the gravitational path integral approach. Then we move on to the study of matter correlators including quantum gravity corrections.

#### 3.1 Spectrum of quantum black holes

The gravitational path integral pioneered by Gibbons and Hawking (Gibbons and Hawking 1977) that computes the black hole partition function instructs us to integrate over smooth geometries with boundary conditions corresponding to fixed inverse temperature  $\beta$ , fixing the size of the thermal circle. The thermal disk or black hole partition function is given by the JT path integral:

$$Z(\beta) = e^{S_0} \int [Dg][D\Phi] e^{\frac{1}{16\pi G_N} \left[ \int_{\mathcal{M}} d^2x \sqrt{g} \Phi (R+2) + 2 \oint_{\partial\mathcal{M}} \sqrt{h} \Phi (K-1) \right]} \quad (3.1)$$

$$= e^{S_0} \int_{\text{Diff}(S^1)/\text{SL}(2,\mathbb{R})} [Df] e^{C \int_0^\beta d\tau \left\{ \tan \frac{\pi f(\tau)}{\beta}, \tau \right\}}. \quad (3.2)$$

As explained in Sect. 2.3.4, we reduce the bulk JT model to that of the Schwarzian theory on its boundary. Within the path integral, the procedure consists of first path-integrating over the dilaton along an imaginary contour, producing a functional delta constraint  $\prod_x \delta(R(x) + 2)$  reducing the integral to only hyperbolic geometries with  $R(x) = -2$ . The remaining degree of freedom, the Schwarzian mode, arises from the freedom to cut out an appropriate patch of the hyperbolic disk with a fixed length  $\beta/\epsilon$  boundary. The cost of each configuration is given by the Schwarzian action.

Within the Euclidean gravitational path integral approach to quantum gravity, an omnipresent issue is that of the unstable conformal mode, making the Euclidean gravitational action unbounded from below, and causing the path integral to diverge. The standard way of dealing with this is to define the path integral along a specific cycle of complex metrics (Gibbons et al. 1978). For JT gravity (and generic dilaton gravities), we see that the answer is to take the bulk dilaton imaginary in the Euclidean path integral. The question of unbounded Euclidean action is then transferred to the boundary Schwarzian action, which we will show explicitly does not contain negative modes.

To incorporate quantum effects, the derivation outlined in the previous paragraph is not yet complete. For Eq. (3.2) to be meaningful, it is necessary to provide a measure for the integration over  $f(\tau)$ , derived from the integral in Eq. (3.1) over the metric and

dilaton. This analysis was carried out in Sect. 3 of Saad et al. (2019). The proposal is to pick the measure of JT gravity to correspond to the symplectic measure in the BF formulation discussed in Sect. 2.1.1. This analysis leads to  $[Df] = \prod_{\tau} df(\tau)/f'(\tau)$ , which is invariant under reparametrizations of the Schwarzian mode.<sup>21</sup> We will come back to the implications of this choice later.

### 3.1.1 Perturbative calculation

We now compute the disk partition function in JT gravity. The Schwarzian action looks rather complicated, and a first approach is to use perturbation theory: we expand the reparametrization  $f(\tau)$  in terms of its saddle  $f_0(\tau) = \tau$ , corresponding to a circular boundary, plus fluctuations:

$$f(\tau) = \tau + \varepsilon(\tau), \quad \varepsilon(\tau + \beta) = \varepsilon(\tau). \quad (3.3)$$

For small  $\varepsilon$  the action becomes  $I_{\text{Sch}} \approx \frac{2\pi^2 C}{\beta} + \mathcal{O}(\varepsilon^2)$ . Setting the quantum fluctuation to zero gives the classical black hole spectrum analyzed in the previous section

$$Z(\beta) \approx e^{-I_{\text{classical}}} = e^{S_0 + \frac{2\pi^2 C}{\beta}}, \quad \Rightarrow \quad \rho(E) \approx e^{S_0 + 2\pi\sqrt{2CE}}. \quad (3.4)$$

This answer reproduces the two-dimensional analog of the Bekenstein–Hawking prescription identifying the horizon value of the dilaton with the entropy of the black hole in a classical gravity approximation.

Next we incorporate the one-loop determinant around this saddle. One of the advantages of JT gravity is the fact that this is a simple explicit computation. This is not so in higher dimensions, with the only exception being the BTZ black hole (Giombi et al. 2008; Maloney and Witten 2010). We expand fluctuations in Fourier space:

$$\varepsilon(\tau) = \frac{\beta}{2\pi} \sum_{n \neq -1, 0, 1} e^{-\frac{2\pi}{\beta} i n \tau} \varepsilon_n + \text{h.c.} \quad (3.5)$$

The action to quadratic order is given by

$$I_{\text{Sch}} = \frac{2\pi^2 C}{\beta} + \frac{4\pi^2 C}{\beta} \sum_{n>1} n^2 (n^2 - 1) \bar{\varepsilon}_n \varepsilon_n + \mathcal{O}(\varepsilon^3), \quad (3.6)$$

and computing the Gaussian integral in these variables is very simple. The measure of integration around the saddle point is given by Stanford and Witten (2017); Moitra et al. (2021):  $[Df] = \prod_{n \geq 2} 4\pi(n^3 - n) d\varepsilon_n d\bar{\varepsilon}_{-n}$ . We combine the quadratic action for the Schwarzian mode with the measure to obtain the partition function to a one-loop

<sup>21</sup> It was shown in Moitra et al. (2021) how this measure directly arises in the second order formalism.

approximation

$$Z(\beta) = e^{S_0 + \frac{2\pi^2 C}{\beta}} \prod_{n \geq 2} \frac{\beta}{Cn} = \frac{1}{4\pi^2} \left( \frac{2\pi C}{\beta} \right)^{3/2} e^{S_0 + \frac{2\pi^2 C}{\beta}}. \quad (3.7)$$

We make the following observations on this result:

- The measure is only defined up to an overall factor, which can be absorbed into a shift of  $S_0$ . This will be less trivial when considering non-perturbative contributions with spacetime wormholes in Sect. 4.
- Since there is only a single saddle, and all quadratic fluctuations are stable (positive), this means that the saddle is a global minimum and the action functional is bounded below. This is no longer true for some generalizations of the Schwarzian model to be discussed in Sect. 3.6.
- We have chosen zeta-function regularization, following standard practice. Other choices can be absorbed into shifts of  $S_0$  or the zero-point energy (which we have set to zero). The power of  $3/2$  appearing in the one-loop determinant has a nice interpretation. In zeta-function regularization we have  $\sum_{n \in \mathbb{Z}} 1 \rightarrow 0$ , and if all Fourier modes were present, the one-loop determinant would be  $\beta$ -independent. Since the Schwarzian action has a  $\mathrm{PSL}(2, \mathbb{R})$  symmetry, one has to remove the  $n = -1, 0, 1$  modes. The factor of 3 is counting the number of omitted zero-modes in the path integral.
- Potential corrections to the one-loop result appear as a Taylor expansion in  $(\beta/C)$  to (3.7). This can be seen by rescaling  $\epsilon$  by  $\epsilon \rightarrow \epsilon\beta/C$ . Therefore this ratio  $\beta/C$  can be interpreted as an effective dimensionless coupling constant.
- A somewhat miraculous statement is that the partition function  $Z(\beta)$  is in fact one-loop exact! So our expression (3.7) is the entire answer. The reason is explained by Stanford and Witten (2017) in terms of a localization argument.

The result (3.7) has interesting implications. We can compute the free energy given by the logarithm of the partition function:

$$-\beta F \equiv \log Z(\beta) = S_0 + \frac{2\pi^2 C}{\beta} + \frac{3}{2} \log \frac{2\pi C}{\beta} + \dots, \quad (3.8)$$

where the dots are temperature-independent. The first two terms on the right-hand side are classical contributions while the third is a quantum effect. The quantum effects become large as we lower the temperature  $\beta \gtrsim C$ . When applied to near-extremal black holes, it implies quantum gravity effects are large very close to extremality. We elaborate on this application in Sect. 5.3.

We can interpret (3.7) in holography, which states that a black hole can be described by a quantum system with finite entropy. The partition function of such a system with Hilbert space  $\mathcal{H}_{BH}$  and Hamiltonian  $H$  would be

$$Z(\beta) = \mathrm{Tr}_{\mathcal{H}_{BH}} \left[ e^{-\beta H} \right] = \int dE \rho(E) e^{-\beta E}, \quad \rho(E) = \sum_n \delta(E - E_n), \quad (3.9)$$

where  $E_n$  is a discrete set of states of the quantum system describing the black hole. We can use Eq. (3.7) to infer what the density of states of this quantum system should be. An inverse Laplace transform of Eq. (3.7) gives

$$\rho_{\text{JT}}(E) = \frac{C}{2\pi^2} e^{S_0} \sinh\left(2\pi\sqrt{2CE}\right). \quad (3.10)$$

Some comments on this result:

- Expression (3.10) is not a sum of delta functions and the spectrum is continuous. In quantum mechanics, this is usually associated to a non-compact space such that one considers the density of states per unit volume, but in our case there is no infinite spatial dimension in the boundary theory. A continuous spectrum implies the entropy in the microcanonical ensemble is infinite, such that information can be lost inside a black hole. To see signs of a discrete spectrum with finite entropy will require non-perturbative effects beyond those we have access to at the level of the disk.
- For large  $E \gg 1/C$ , the spectrum is  $\rho_{\text{JT}}(E) \approx e^{S_0+2\pi\sqrt{2CE}}$ , consistent with the classical Bekenstein–Hawking entropy of the JT black hole in (2.60).
- Quantum effects are large at small energies and indeed for  $E \ll 1/C$  the spectrum is strongly modified  $\rho_{\text{JT}}(E) \approx e^{S_0}\sqrt{2CE}$ . The density of states goes to zero as  $E \rightarrow 0$ ! This shows that it is wrong to interpret  $S_0$  as a ground state degeneracy.

We reduced the gravitational description in JT gravity to a solvable boundary graviton mode. There is a derivation of the path integral at the one-loop level performed by Charles and Larsen (2020) that makes the relation between the boundary Schwarzian and bulk metric fluctuations more transparent. A similar analysis can also be found in Moitra et al. (2021).

### 3.2 Quantum Jackiw–Teitelboim gravity coupled to matter

We now perform the calculation with the addition of matter (not coupled to the dilaton field  $\Phi$  as studied above in Sect. 2). For concreteness we work with a concrete example: a massive scalar field  $\phi$  propagating on the JT black hole background. Its Euclidean action is

$$I_{\text{matter}}[\phi, g] = \frac{1}{2} \int d^2x \sqrt{g} \left[ (\partial\phi)^2 + m^2\phi^2 \right]. \quad (3.11)$$

In the putative dual 1d nearly-conformal quantum mechanics, this field is dual to an operator  $\mathcal{O}$  of scaling dimension (defined at short distances or large energies where the gravitational effects that break the conformal symmetry are small) given by  $\Delta = 1/2 + \sqrt{1/4 + m^2}$ . We compute the gravitational path integral with Dirichlet boundary

conditions for the matter field

$$\begin{aligned} Z(\beta; \phi_b) &= e^{S_0} \int [Dg][D\Phi] e^{\frac{1}{16\pi G_N} [\int_{\mathcal{M}} \Phi(R+2)+2 \int_{\partial\mathcal{M}} \Phi(K-1)]} \int [D\phi] e^{-I_{\text{matter}}[\phi, g]} \\ &= e^{S_0} \int [Df] e^{-I_{\text{Sch}}[f]} Z_{\text{matter}}[\phi_b, f], \end{aligned} \quad (3.12)$$

where we defined the matter path integral in the hyperbolic disk described by a cut-off curve  $f(\tau)$  and with boundary conditions  $\phi|_{\partial} = \phi_b$  as  $Z_{\text{matter}}[\phi_b, f] = \int [D\phi] e^{-I_{\text{matter}}[\phi, g]}$ . This path integral depends in principle on both the Schwarzian mode and the boundary value of the scalar field. Notice the order we chose to do the path integration is important: first the dilaton, then the matter, and lastly the Schwarzian.

The undeformed theory corresponds to  $\phi_b \rightarrow 0$ , such that boundary sources are turned off. In the limit that the boundary curve approaches the boundary of the hyperbolic disk, i.e. its proper length goes to infinity,  $Z_{\text{matter}}[0, f]$  becomes independent of the Schwarzian mode to leading order. An explicit illustration of this fact is given in Appendix C of Yang (2019), for the case of a massless scalar field with  $m = 0$  (actually the calculation applies to any two-dimensional conformal field theory).<sup>22</sup>

The partition function (even in the presence of matter) is given by the Schwarzian density of states and therefore the quantum black hole spectrum is insensitive to the matter content.<sup>23</sup> The situation changes when we turn on a source  $\phi_b \neq 0$  for the boundary dual operators  $\mathcal{O}$ , deforming the boundary theory. The source  $\phi_r$  in the boundary's own time coordinate  $\tau$ , is defined through the relations:

$$\phi|_{\partial} = Z^{1-\Delta} \tilde{\phi}_r(T) = \epsilon^{1-\Delta} F'^{1-\Delta} \tilde{\phi}_r(T) = \epsilon^{1-\Delta} \phi_r(\tau), \quad (3.13)$$

where we have first used the Poincaré coordinates  $(T, Z)$  and its source  $\tilde{\phi}_r$ , and rewritten this in coordinates specified along the boundary curve defined by the Schwarzian mode. The result is the well-known generating functional for generalized free fields, but reparametrized to the fluctuating boundary line:

$$Z_{\text{matter}}[\phi_r, f] = e^{\frac{D}{2} \int d\tau_1 d\tau_2 \left( \frac{f'(\tau_1) f'(\tau_2)}{\frac{\beta^2}{\pi^2} \sin^2 \frac{\pi}{\beta} [f(\tau_1) - f(\tau_2)]} \right)^{\Delta} \phi_r(\tau_1) \phi_r(\tau_2)}, \quad (3.14)$$

where  $D = \frac{(\Delta - \frac{1}{2})\Gamma(\Delta)}{\sqrt{\pi}\Gamma(\Delta - \frac{1}{2})}$ . From now on we rescale the matter field to absorb the factor of  $D$ ,  $\phi_r \rightarrow \phi_r/\sqrt{D}$ . Equation (3.14) is true up to an overall coefficient independent of the Schwarzian mode and the source, and we absorb it in  $S_0$ .

<sup>22</sup> One can use conformal symmetry to map the cut-off hyperbolic disk with boundary curve labeled by  $f(\tau)$  to a unit disk with a circular boundary. The non-trivial dependence with  $f(\tau)$  arises then from the conformal anomaly involved in this transformation. An explicit calculation shows that the final answer is independent of  $f(\tau)$  up to terms suppressed in the proper length of the boundary curve.

<sup>23</sup> This requires some fine print. If we assume  $\phi_b$  is non-zero but arbitrarily small, then at low enough temperatures the contribution from the matter can be dominant in the window  $1 < \Delta < 3/2$ , see Maldacena et al. (2016b).

### 3.3 Correlators

Next, we go over the calculation of the boundary matter correlators dual to JT gravity coupled to matter fields. Some partial results were obtained in Bagrets et al. (2016, 2017) and the full solution for all correlators was derived in Mertens et al. (2017) realizing that the Schwarzian theory is a limit of Liouville field theory, further elaborated in Mertens (2018); Lam et al. (2018). These correlators were later reproduced using a variety of approaches by Blommaert et al. (2018); Iliesiu et al. (2019) using the BF approach and by Kitaev and Suh (2019), Yang (2019), Suh (2020) using the particle in a magnetic field perspective on JT gravity, proposed by Kitaev in 2016.

#### 3.3.1 Path integral representation

Computing correlators involves taking functional derivatives with respect to matter sources as

$$\langle \mathcal{O}(\tau_1) \dots \mathcal{O}(\tau_n) \rangle = \frac{\delta}{\delta \phi_r(\tau_1)} \dots \frac{\delta}{\delta \phi_r(\tau_n)} Z(\beta, \phi_r) \Big|_{\phi_r \rightarrow 0}, \quad (3.15)$$

with generating functional

$$Z(\beta, \phi_r) = e^{S_0} \int [Df] e^{-I_{\text{Sch}}[f]} e^{\frac{1}{2} \int d\tau_1 d\tau_2 \left( \frac{f'(\tau_1) f'(\tau_2)}{\frac{\beta^2}{\pi^2} \sin^2 \frac{\pi}{\beta} [f(\tau_1) - f(\tau_2)]} \right)^\Delta} \phi_r(\tau_1) \phi_r(\tau_2). \quad (3.16)$$

The quantum gravity effects are encoded in the fluctuations of the Schwarzian mode.

**Two-point function** Let us begin with the two-point function (since one-point functions vanish):

$$\langle \mathcal{O}(\tau_1) \mathcal{O}(\tau_2) \rangle = e^{S_0} \int [Df] e^{C \int_0^\beta \{ \tan \frac{\pi}{\beta} f, \tau \}} \left( \frac{f'(\tau_1) f'(\tau_2)}{\frac{\beta^2}{\pi^2} \sin^2 \frac{\pi}{\beta} [f(\tau_1) - f(\tau_2)]} \right)^\Delta. \quad (3.17)$$

The building block here is the conformal two-point function, while the coupling to gravity is accounted for by the reparametrization mode. For example, when we freeze gravity by sending  $C \rightarrow \infty$ , we can set  $f(\tau) = \tau$  (the gravitational saddle), and we reduce to the conformal two-point function

$$\langle \mathcal{O}(\tau_1) \mathcal{O}(\tau_2) \rangle \approx Z(\beta) \left( \frac{\pi^2}{\beta^2 \sin^2 \frac{\pi}{\beta} \tau_{12}} \right)^\Delta. \quad (3.18)$$

We will see how this semi-classical answer emerges from the exact expression later, and how corrections are suppressed as long as  $\beta, |\tau_{12}| \ll C$ .



The previous discussion motivates introducing the following bilocal operator within the Schwarzian theory

$$G_{\Delta}(\tau_1, \tau_2) \equiv \left( \frac{f'(\tau_1)f'(\tau_2)}{\frac{\beta^2}{\pi^2} \sin^2 \frac{\pi}{\beta} [f(\tau_1) - f(\tau_2)]} \right)^{\Delta}. \quad (3.19)$$

This operator has an important property: it is invariant under the  $SL(2, \mathbb{R})$  transformations (2.82), and since by (2.86) the Hamiltonian is the quadratic Casimir, it hence commutes with the Hamiltonian. It will be useful later to denote the two-point function by the following diagram

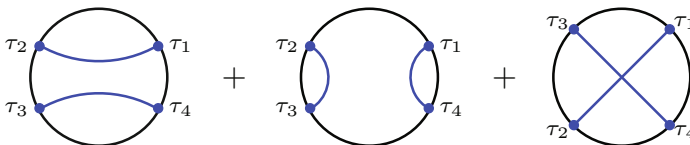
$$\langle G_{\Delta}(\tau_1, \tau_2) \rangle_{\text{Sch}} = \tau_2 \text{---} \bigcirc \text{---} \tau_1, \quad (3.20)$$

where we drew the boundary thermal circle, and the points denote the insertions labeled by the times  $\tau_1$  and  $\tau_2$ . The line going through the “interior” denotes the insertion of the bilocal operator  $G_{\Delta}(\tau_1, \tau_2)$ .

**Four-point function** We will also work out the four-point function in some detail. Assume for concreteness that  $\tau_1 < \tau_2 < \tau_3 < \tau_4$ . The boundary four-point function of a single free field is a sum of products of two-point functions in different channels since, turning off the reparametrization mode, the action in (3.14) is quadratic. After applying a reparametrization, each two-point function becomes the Schwarzian bilocal. The final answer is<sup>24</sup>

$$\begin{aligned} \langle \mathcal{O}(\tau_1) \mathcal{O}(\tau_2) \mathcal{O}(\tau_3) \mathcal{O}(\tau_4) \rangle &= \langle G_{\Delta}(\tau_1, \tau_2) G_{\Delta}(\tau_3, \tau_4) \rangle_{\text{Sch}} \\ &+ \langle G_{\Delta}(\tau_1, \tau_4) G_{\Delta}(\tau_2, \tau_3) \rangle_{\text{Sch}} + \langle G_{\Delta}(\tau_1, \tau_3) G_{\Delta}(\tau_2, \tau_4) \rangle_{\text{Sch}}. \end{aligned} \quad (3.21)$$

This gives the matter four-point function as a Schwarzian correlator of a product of bilocals. We describe the answer for each term in the next section. The four-point function can be represented in a diagrammatic way as the sum of these three terms:



<sup>24</sup> To simplify the discussion, we could also consider the presence of two free fields with scaling dimensions  $\Delta_1$  and  $\Delta_2$  dual to operators  $\mathcal{O}_1$  and  $\mathcal{O}_2$ . Then the four point function is given by  $\langle G_{\Delta_1} G_{\Delta_2} \rangle$ . In this case there is only a single channel depending on the position of the two insertions and there is no need to sum over three different diagrams.

The first two terms above are simple to compute: as mentioned above, the bilocal operator commutes with the Hamiltonian and states with the same energy propagate along the lines joining 1–4 and 2–3 in the first diagram, or 1–2 and 3–4 in the second diagram. This implies that in the energy eigenbasis the first two contributions are just products of two-point functions. The third diagram involves a new calculation due to the specific ordering of insertions and bilocals. As explained in Maldacena et al. (2016b), this correlator when coupled to the Schwarzian mode will describe gravitational shockwave scattering in the bulk.

**Generalizations** For free fields the generalization to higher-point functions is straightforward: they will involve higher correlators of the Schwarzian bilocal.

Another generalization we will not pursue here is to include self-interactions between matter fields in the bulk (besides their gravitational interactions). This involves first computing the interacting matter boundary correlator in rigid  $\text{AdS}_2$ . Then one applies a reparametrization and integrates over the Schwarzian mode. The conformal block expansion shows that it can always be expanded in a possibly infinite number of bilocal insertions. Even in the presence of bulk interactions, the one- and two-point functions are universal. Three-point functions are universal up to a single coupling factor: the OPE coefficient, but can likewise be written in terms of a product of three bilocal operators when coupling to the gravitational Schwarzian sector.

### 3.3.2 A first approach: Schwarzian perturbation theory

As a first approach of computing Schwarzian correlation functions of bilocal operators (3.19) explicitly, one can utilize perturbation theory as introduced before in Sect. 3.1.1. Schwarzian perturbation theory requires the propagator ( $u = 2\pi\tau/\beta$ ) (Maldacena and Stanford 2016; Maldacena et al. 2016b):

$$\langle \varepsilon(\tau)\varepsilon(0) \rangle = \frac{1}{2\pi C} \left( \frac{\beta}{2\pi} \right)^3 \left[ 1 - \frac{1}{2}(u - \pi)^2 + \frac{\pi^2}{6} + \frac{5}{2} \cos u + (\tau - \pi) \sin u \right], \quad (3.22)$$

and the expansion of the bilocal operator insertion as:

$$\left( \frac{F'_1 F'_2}{(F_1 - F_2)^2} \right)^\Delta = \frac{(1 + \varepsilon'_1)^\Delta (1 + \varepsilon'_2)^\Delta}{\left( \frac{\beta}{\pi} \sin \frac{\pi}{\beta} (\tau_{12} + \varepsilon_1 - \varepsilon_2) \right)^{2\Delta}}. \quad (3.23)$$

Any propagator carries a factor  $1/C$ , whereas any vertex gives a factor of  $C$ . The resulting perturbative expansion is then a power series in  $1/C$ ,<sup>25</sup> which gets complicated very quickly since new interaction vertices appear at higher orders as well. Nonetheless, interesting computations have been done, see e.g. Maldacena et al. (2016b), Haehl

<sup>25</sup> One convenient way of seeing this is to redefine  $\sqrt{C}\epsilon \rightarrow \epsilon$  to make the propagator independent of  $C$ . This makes all vertices (including those coming from expanding the operator as in (3.23)) contribute with negative integer powers of  $C$ .

and Rozali (2018), Maldacena et al. (2021), Cotler et al. (2020), Qi et al. (2019) for a selection of applications at lower orders in  $1/C$ .

Fortunately, the situation in JT gravity is even better: exact expressions are known to which we turn next, which can be compared to the perturbative  $1/C$  results as we will comment further on.

### 3.3.3 Exact two-point function

We now state the explicit answers of Schwarzian correlation functions, and review some of their immediate gravitational properties. We postpone an overview of their derivation until Sect. 3.7.

For the two-point function, the answer can be written in the following suggestive form as a double integral:

$$\langle \mathcal{O}(\tau_1) \mathcal{O}(\tau_2) \rangle = e^{S_0} \int_0^\infty \prod_{i=1,2} (dk_i^2 \sinh 2\pi k_i) e^{-\tau \frac{k_1^2}{2C} - (\beta - \tau) \frac{k_2^2}{2C}} \frac{\Gamma(\Delta \pm ik_1 \pm ik_2)}{8\pi^4 (2C)^{2\Delta} \Gamma(2\Delta)}. \quad (3.24)$$

where we defined  $\tau = |\tau_1 - \tau_2|$ . This is the unnormalized expression (where we did not divide by the partition function  $Z(\beta)$ ). We introduced the useful notation that whenever  $\pm$  appears inside a gamma function, it means one should take a product of that gamma function with both signs. For example  $\Gamma(x \pm y) = \Gamma(x + y)\Gamma(x - y)$  and hence the numerator in the right hand side of Eq. (3.27) involves a product of four gamma functions. As a simple check we can see that if  $\Delta \rightarrow 0$  then (3.24) gives back the Schwarzian partition function, using the identities  $\lim_{\Delta \rightarrow 0} \frac{\Gamma(\Delta \pm ix)}{\Gamma(2\Delta)} = 2\pi \delta(x)$  and  $\Gamma(\pm 2ik) = \frac{\pi}{2k \sinh 2\pi k}$ .

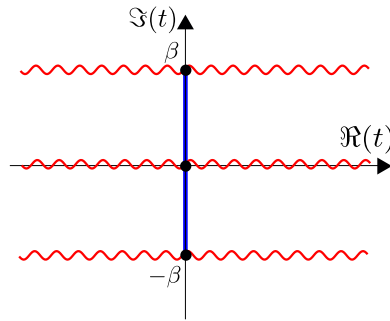
We can interpret this expression from the boundary holographic perspective. The quantum black hole in the bulk is supposed to be dual to a system with Hilbert space  $\mathcal{H}_{\text{BH}}$  defined above. Then  $\mathcal{O}$  should be an operator within this theory. Quantum mechanics gives an expansion

$$\langle \mathcal{O}(\tau_1) \mathcal{O}(\tau_2) \rangle = \text{Tr}_{\mathcal{H}_{\text{BH}}} \left[ e^{-\beta H} \mathcal{O}(\tau_1) \mathcal{O}(\tau_2) \right] \quad (3.25)$$

$$= \int \rho(E_1) dE_1 \rho(E_2) dE_2 e^{-\beta E_2 - \tau(E_1 - E_2)} |\mathcal{O}_{E_1, E_2}|^2, \quad (3.26)$$

where for systems with finite entropy  $\rho(E)$  is a sum of delta functions and  $\mathcal{O}_{E_1, E_2}$  are the operator matrix elements between energy eigenstates. Matter does not affect the partition function to leading order, so we should identify  $\rho(E) = \rho_{\text{JT}}(E)$  given in Eq. (3.10). Combining this with the expression (3.24), we can identify the integration variables with the intermediate energies as  $E \equiv \frac{k^2}{2C}$ . The remainder gives the off-diagonal matrix elements of the operator in the putative dual quantum mechanical system

$$|\mathcal{O}_{E_1, E_2}|^2 = 2e^{-S_0} \frac{\Gamma(\Delta \pm i\sqrt{2CE_1} \pm i\sqrt{2CE_2})}{(2C)^{2\Delta} \Gamma(2\Delta)}, \quad (3.27)$$



**Fig. 9** Complex time plane with Lorentzian time on the horizontal axis and Euclidean time on the vertical axis. Branch cuts of correlators are denoted by red wiggly lines

where the factor of  $e^{-S_0}$  implies that  $\mathcal{O}_{E_1, E_2} \sim \mathcal{O}(e^{-S_0/2})$  can be understood from the point of view of the eigenstate thermalization hypothesis (ETH) (Deutsch 1991; Srednicki 1994), as emphasized in Saad (2019), Jafferis et al. (2022a, b).<sup>26</sup>

We can extract from (3.26) the two-point function in the microcanonical ensemble, essentially by stripping off the  $E_2$ -integral, evaluated in a state with fixed energy  $E$ . The answer is

$$\langle E | \mathcal{O}(\tau_1) \mathcal{O}(\tau_2) | E \rangle = \int_0^\infty dE' \rho(E') e^{-\tau(E-E')} |\mathcal{O}_{E, E'}|^2. \quad (3.28)$$

We can verify that when the energy is of order  $C$ , this correlator becomes thermal with an effective temperature  $\beta(E) = \sqrt{2\pi^2 C/E}$ , consistent with ETH, see section 6.2 of Lam et al. (2018).

At  $\tau \rightarrow 0$  or  $\tau \rightarrow \beta$ , either the  $E_1$  or the  $E_2$  integral lose their exponential damping. The result is the UV divergence from the underlying matter CFT  $1/\tau^{2\Delta}$  or  $1/(\beta - \tau)^{2\Delta}$ . This pole gets smoothed out if one embeds JT within the microscopic SYK model.

The result (3.24) is written in Euclidean signature. It is possible to analytically continue this result to real time as follows. The singularity as  $\tau \rightarrow 0$ ,  $\beta$  is actually part of a branch cut when analytically continuing to the complex plane  $t$  where  $\Im(t) \equiv -\tau$ , with periodic cuts at  $\Im(t) = n\beta$ , as illustrated in Fig. 9. When continuing these expressions to real time  $\Re(t)$ , we have two answers depending on whether we approach the real axis from above or below.

These correspond to the two orderings of the operators  $\mathcal{O}(t_i)$ , and are the real-time Wightman correlators  $G_+(t) = \langle \mathcal{O}(t_1) \mathcal{O}(t_2) \rangle$  and  $G_-(t) = \langle \mathcal{O}(t_2) \mathcal{O}(t_1) \rangle$ , where the  $\pm$ -symbol denotes the continuation from the bottom (+) or top (−) of Fig. 9, and  $t = t_1 - t_2$ . These turn out to be complex conjugates of each other and  $G_\pm(t)$  is given by

$$e^{S_0} \int_0^\infty \prod_{i=1,2} (2k_i dk_i \sinh 2\pi k_i) e^{\mp i t \frac{k_1^2}{2C} - (\beta \mp i t) \frac{k_2^2}{2C}} e^{-\epsilon \frac{k_1^2}{2C} - \epsilon \frac{k_2^2}{2C}} \frac{\Gamma(\Delta \pm i k_1 \pm i k_2)}{8\pi^4 (2C)^{2\Delta} \Gamma(2\Delta)}, \quad (3.29)$$

<sup>26</sup> The phase of the matrix elements is arbitrary and would require knowing more complicated observables.

where  $\epsilon$  is a Euclidean regulator to make the integrals converge,  $t = t_1 - t_2$ , and the  $\pm$  symbol in the exponentials depends on which of the two correlators we consider (the Gamma-functions still involve a product over all signs). These are different since timelike separated operators generically do not commute.<sup>27</sup>

Finally, one can also separate the operators by  $\beta/2$  in Euclidean time and  $t$  in real time. This corresponds physically to having the operators acting on opposite halves of the thermofield double state (TFD). As visible in the above figure, the resulting real-time thermofield double correlation function is unique, reflecting the commutativity of operator insertions on opposite ends of the TFD state.

The semi-classical two-point function decays exponentially at large Lorentzian time. Indeed, under a Wick rotation  $\tau \rightarrow it$  of Eq. (3.18) we have  $\langle \mathcal{O}(t)\mathcal{O}(0) \rangle \sim e^{-\Delta \frac{2\pi t}{\beta}}$ . This is not so when quantum gravity is turned on. Equation (3.29) implies that at very late times  $t \gg C$  the correlator is dominated by small energies  $E_{1,2} \sim t^{-1}$ . In this limit,  $|\mathcal{O}_{0,0}|$  approaches a constant while  $\rho(E) \sim \sqrt{E}$ , implying that the correlators decay as

$$\langle \mathcal{O}(t)\mathcal{O}(0) \rangle \sim t^{-3}. \quad (3.30)$$

This behavior was derived before knowing the full answer for the correlator in Bagrets et al. (2017). Quantum gravity slows down the decay but does not stop it, leading to the information paradox raised in Maldacena (2003a). We will review later in Sect. 4 how the addition of spacetime wormholes resolves this problem.

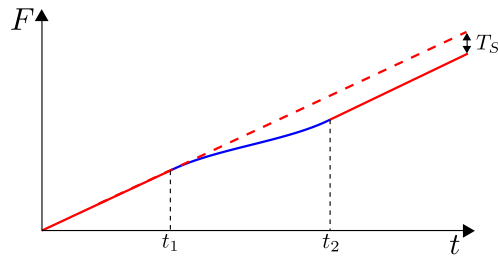
Let us finally make some comments on the large  $C$  (or semi-classical) content of these expressions, and the relation to boundary graviton expansions.

- Quantum gravity effects break conformal invariance, and only as we take the semi-classical limit  $C \rightarrow \infty$  with a light operator insertion  $\Delta \sim \mathcal{O}(1)$ , do we recover the conformal two-point function (3.18). This is true as long as both  $\beta, \tau \ll C$ . To see this, we first write  $E_1 = E_2 + \omega$ . Then the integral is dominated by configurations with large  $E_{1,2} \sim \mathcal{O}(C)$  and small  $\omega \sim \mathcal{O}(1)$ . Physically, this means we have a semi-classical black hole of large energy  $E_2$  that is perturbed by a small energy injection  $\omega$  caused by the operator insertion. Writing  $e^{-\beta E_2 - \tau(E_1 - E_2)} = e^{-\beta E_2 - \tau\omega}$ , in this limit the rest of the integrand becomes

$$\rho(E_1)\rho(E_2)|\mathcal{O}_{E_1,E_2}|^2 \sim e^{2\pi\sqrt{2CE_2}} \frac{e^{\pi \frac{\omega}{\sqrt{2E_2/C}}} \Gamma\left(\Delta \pm i \frac{\omega}{\sqrt{2E_2/C}}\right)}{(2E_2/C)^{\frac{1}{2}-\Delta} \Gamma(2\Delta)}, \quad (3.31)$$

where we omit energy-independent prefactors. The integral over  $E_2$  can be done by saddle point techniques combining the exponential in the RHS with the Boltzmann  $e^{-\beta E_2}$  factor, giving the saddle identification  $E_2 = 2\pi^2 C/\beta^2$  (which is again the JT black hole first law). This produces a factor of  $Z(\beta)$ . The rest of the RHS is

<sup>27</sup> We note that semi-classically both of these correlators have an infinitesimal imaginary part of opposite sign  $\pm i\epsilon$ . However, the full quantum expressions have finite imaginary parts and are hence different beyond infinitesimals.



**Fig. 10** Classical solution for the Schwarzian mode  $F(t)$  with sources inserted at time  $t_1$  and reabsorbed at  $t_2$ .  $T_S$  denotes the net time delay after this process

given by  $\beta^{1-2\Delta} e^{\frac{\beta\omega}{2}} \Gamma(\Delta \pm i\frac{\beta\omega}{2\pi})$ . The  $\omega$ -integral then finally reduces to the Fourier transform of Eq. (3.18).

- If the scaling dimension of the field is large in units of  $C$ ,  $\Delta \sim C$  as  $C \rightarrow \infty$ , the operator is called heavy. In this case classical gravitational backreaction is important: the saddle equation is sourced by the bilocal operator insertion since one can write the operator insertion into the action as:

$$S_{\text{eff}} = \int dt \left( -C \{F, t\} - \Delta \log \frac{F'_1 F'_2}{(F_1 - F_2)^2} \right). \quad (3.32)$$

Since  $\Delta \sim C$ , this contributes equally importantly as the Schwarzian action itself. The limit can be taken both directly in the expression (3.24), or using the new sourced saddle, with matching results Lam et al. (2018). An illustrative example is the case of zero temperature  $\beta^{-1} \rightarrow 0$ . In that case, the boundary clock starts at  $F(t) = t$ , transfers through a region of finite intermediate energy for  $t_1 < t < t_2$  between the ends of the bilocal operator, and ends up as  $F(t) = t - T_S$ , see Fig. 10. The quantity  $T_S > 0$  for  $\Delta > 0$  and is a net time delay, a Shapiro time delay, whose explicit expression depends on the time difference  $t_2 - t_1$  and the operator weight  $\Delta$ .

- Beyond the large  $C$  limits, the expression (3.24) can be expanded in powers of  $1/C$ , and compared to the perturbative analysis. The results match (Mertens 2021; Griguolo et al. 2021). Moreover, one can use this exact result to prove that the perturbative  $1/C$  expansion is asymptotic for generic real values of  $\Delta$ , and hence cannot be captured just by the perturbative boundary graviton treatment of Sect. 3.3.2.
- An exception occurs when  $\Delta \in -\mathbb{N}/2$ . In this case, the expression (3.24) would seem to vanish due to the poles of the Gamma-function in the denominator. What actually happens, is that the numerator also degenerates to a single  $k$ -integral. The resulting expressions are somewhat simpler structurally, but correspond to non-unitary matter insertions (Mertens 2021).<sup>28</sup> In this case, one can show that the  $1/C$  perturbative expansion is convergent, and all physics is captured by boundary graviton Feynman diagrams.

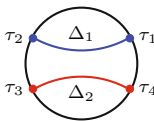
<sup>28</sup> These have a natural origin in terms of the BF gauge model description (as finite-dimensional non-unitary representations of  $\text{SL}(2, \mathbb{R})$ ), and in the Liouville CFT language (as degenerate Virasoro representations).

### 3.3.4 Exact four-point function

We now describe the results for the different types of four-point functions. To separate the channels that contribute, we consider a correlator between two pairs of operators  $\mathcal{O}_1$  and  $\mathcal{O}_2$  with scaling dimensions  $\Delta_{1,2}$ , defined by

$$\langle \mathcal{O}_{\Delta_1}(\tau_1) \mathcal{O}_{\Delta_1}(\tau_2) \mathcal{O}_{\Delta_2}(\tau_3) \mathcal{O}_{\Delta_2}(\tau_4) \rangle. \quad (3.33)$$

We begin by considering the case with  $\tau_1 < \tau_2 < \tau_3 < \tau_4$  such that the four-point function is equal to  $\langle G_{\Delta_1}(\tau_1, \tau_2) G_{\Delta_2}(\tau_3, \tau_4) \rangle_{\text{Sch}}$ . This corresponds to the diagram

$$\langle \mathcal{O}_{\Delta_1}(\tau_1) \mathcal{O}_{\Delta_1}(\tau_2) \mathcal{O}_{\Delta_2}(\tau_3) \mathcal{O}_{\Delta_2}(\tau_4) \rangle =$$


It is convenient to rewrite the correlator in terms of a fixed energy amplitude  $\mathcal{A}(k_1, \dots, k_4)$  as

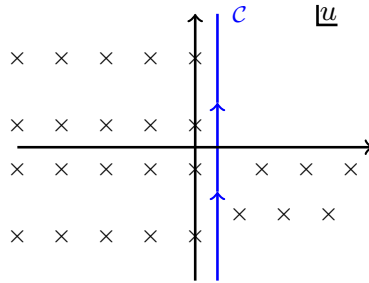
$$\begin{aligned} \langle \mathcal{O}_{\Delta_1}(\tau_1) \mathcal{O}_{\Delta_1}(\tau_2) \mathcal{O}_{\Delta_2}(\tau_3) \mathcal{O}_{\Delta_2}(\tau_4) \rangle &= \frac{e^{S_0}}{2} \int_0^\infty d\mu(k_1) d\mu(k_2) d\mu(k_3) d\mu(k_4) \\ &\times e^{-\frac{k_1^2}{2C}\tau_{21} - \frac{k_2^2}{2C}\tau_{32} - \frac{k_3^2}{2C}\tau_{43} - \frac{k_4^2}{2C}(\beta - \tau_{41})} \mathcal{A}(k_1, k_2, k_3, k_4), \end{aligned} \quad (3.34)$$

where  $d\mu(k) \equiv \frac{k \sinh 2\pi k}{\pi^2} dk$  and we can identify  $\frac{k^2}{2C}$  with the energies propagating between each insertion:  $k_1$  propagates between insertions 1–2,  $k_2$  between 2–3,  $k_3$  between 3–4 and  $k_4$  between 4–1. Moreover it is clear that under this identification  $d\mu(k)$  is proportional to the Schwarzian density of states. Then the quantity  $\mathcal{A}(k_1, k_2, k_3, k_4)$  can be interpreted as a fixed energy amplitude computing the correlator. This amplitude is (Mertens et al. 2017)

$$\mathcal{A}_{\text{uncross.}} = \frac{\Gamma(\Delta_1 \pm ik_1 \pm ik_2)}{(2C)^{2\Delta_1} \Gamma(2\Delta_1)} \frac{\Gamma(\Delta_2 \pm ik_3 \pm ik_4)}{(2C)^{2\Delta_2} \Gamma(2\Delta_2)} \frac{\pi^2 \delta(k_4 - k_2)}{k_4 \sinh 2\pi k_4}. \quad (3.35)$$

The delta function is a manifestation of the  $\text{SL}(2, \mathbb{R})$  invariance of the Schwarzian bilocal operator, and as expected the remaining factor is a product of two-point function amplitudes.

As a simple check, when  $\Delta_1 \rightarrow 0$  (3.35) the correlator reduces to the two-point function of the  $\mathcal{O}_{\Delta_2}$  operator, and viceversa. We can also verify it gives the correct semiclassical limit: defining  $\omega_{1,2} = (k_2^2 - k_{1,3}^2)/2C$ , then at large  $C$  the integral over  $k_2$  yields  $k_2 = 2\pi C/\beta$  via the saddle point method, while the integrals over  $\omega_{1,2}$  factorize into a product of conformal two-point functions. Finally, at very late times when quantum effects are important, the four-point function decays as  $\langle \mathcal{O}(t_1) \mathcal{O}(t_2) \mathcal{O}(t_3) \mathcal{O}(t_4) \rangle \sim t_{12}^{-3} t_{34}^{-3}$  when  $t_{12}, t_{34} \gg C$ , similarly to the two-point function. Again this late-time power law was derived before knowing the complete answer by Bagrets et al. (2017).



**Fig. 11** Complex  $u$ -plane relevant for computing the four-point function with crossing. The contour of integration is indicated in blue while the black crosses denote the poles of the integrand

We now move on to the four-point function including shockwave interaction, which has a much more interesting structure. This requires a computation of  $\langle G_{\Delta_1}(\tau_1, \tau_3) G_{\Delta_2}(\tau_2, \tau_4) \rangle_{\text{Sch}}$  for a configuration such that  $\tau_1 < \tau_2 < \tau_3 < \tau_4$  and therefore the  $\text{SL}(2, \mathbb{R})$  invariance of the bilocal does not help. This can be represented by

$$\langle \mathcal{O}_{\Delta_1}(\tau_1) \mathcal{O}_{\Delta_2}(\tau_2) \mathcal{O}_{\Delta_1}(\tau_3) \mathcal{O}_{\Delta_2}(\tau_4) \rangle =$$

The final result obtained in Mertens et al. (2017) is given by inserting in (3.34) the following amplitude:

$$\begin{aligned} \mathcal{A}_{\text{cross.}} = & \frac{\Gamma(\Delta_1 + ik_1 \pm ik_4) \Gamma(\Delta_1 - ik_3 \pm ik_2)}{(2C)^{2\Delta_1} \Gamma(2\Delta_1)} \frac{\Gamma(\Delta_2 - ik_1 \pm ik_2) \Gamma(\Delta_2 + ik_3 \pm ik_4)}{(2C)^{2\Delta_2} \Gamma(2\Delta_2)} \\ & \times \int_C \frac{du}{2\pi i} \frac{\Gamma(u \pm ik_4) \Gamma(u + ik_1 + ik_3 \pm ik_2) \Gamma(\Delta_1 - ik_1 - u) \Gamma(\Delta_2 - ik_3 - u)}{\Gamma(u + \Delta_1 + ik_1) \Gamma(u + \Delta_2 + ik_3)}. \end{aligned} \quad (3.36)$$

There is an extra auxiliary integral involved in the evaluation of the amplitude. The contour  $\mathcal{C}$  in the complex  $u$ -plane is defined in the following way, illustrated in Fig. 11. The integrand has four set of poles coming from  $\Gamma(u \pm ik_4)$  and  $\Gamma(u + ik_1 + ik_3 \pm ik_2)$  that extend towards negative real  $u$  and reach the  $\text{Re}(u) = 0$  axis. The contour  $\mathcal{C}$  is to the right of these poles. At the same time there are two set of poles coming from  $\Gamma(\Delta_1 - ik_1 - u)$  and  $\Gamma(\Delta_2 - ik_3 - u)$  that extend towards the right (we are assuming  $\Delta_{1,2} > 0$ ). The contour runs to the left of these poles.

With this choice, the integral can be evaluated into a linear combination of hypergeometric functions, although it will be useful to keep it in the integral form.

**A check:  $\Delta_1 \rightarrow 0$  limit**



To get some practice with this amplitude we take the limit where one of the scaling dimension vanishes  $\Delta_1 \rightarrow 0$ , so the four-point function should reduce to a two-point function with scaling dimension  $\Delta_2$ . To show this, first close the contour to the right. In the  $\Delta_1 \rightarrow 0$  limit only a single pole dominates given by  $u = \Delta_1 - ik_1$ . The residue around this pole contributes to the amplitude in the following way

$$\mathcal{A}_{\text{cross.}} = \frac{\Gamma(\Delta_1 \pm ik_1 \pm k_4)}{\Gamma(2\Delta_1)} \frac{\Gamma(\Delta_1 \pm ik_3 \pm ik_2)}{(2C)^{2\Delta_1} \Gamma(2\Delta_1)} \frac{\Gamma(\Delta_2 \pm ik_1 \pm ik_2)}{(2C)^{2\Delta_2} \Gamma(2\Delta_2)} \\ \times \frac{\Gamma(\Delta_2 + ik_3 \pm ik_4) \Gamma(\Delta_2 - \Delta_1 + ik_1 - ik_3)}{\Gamma(\Delta_2 + ik_1 \pm ik_2) \Gamma(\Delta_2 + \Delta_1 + ik_3 - ik_1)} + \dots, \quad (3.37)$$

where the dots denote other residues that are subleading. Using again the identity  $\lim_{\Delta \rightarrow 0} \frac{\Gamma(\Delta \pm ix)}{\Gamma(2\Delta)} = 2\pi\delta(x)$  and keeping track of the prefactors gives the expected answer

$$\lim_{\Delta_1 \rightarrow 0} \mathcal{A}_{\text{cross.}} = \frac{\Gamma(\Delta_2 \pm ik_2 \pm ik_4)}{(2C)^{2\Delta_2} \Gamma(2\Delta_2)} \frac{\pi^2 \delta(k_1 - k_4)}{k_4 \sinh 2\pi k_4} \frac{\pi^2 \delta(k_2 - k_3)}{k_2 \sinh 2\pi k_2}. \quad (3.38)$$

This has the two needed delta functions to completely remove the dependence of the four-point function on  $\tau_1$  and  $\tau_2$ , and leaves precisely the two-point amplitude.

—o—

Since this correlator is of central importance in the study of quantum chaos and its relation to shockwave scattering near the black hole horizon, we will analyze some properties in the next section.

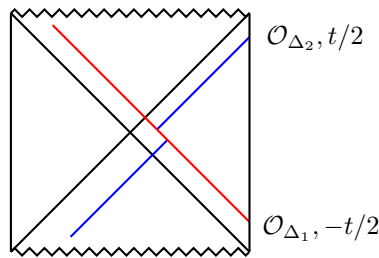
### 3.3.5 Application: quantum chaos

The exponential growth in time of double commutators is a signal of quantum chaos. This involves the computation of an out-of-time-ordered four-point function (OTOC). To compute it, we need to first analytically continue to Lorentzian signature  $\tau \rightarrow \epsilon + it$ , where  $t$  is Lorentzian time and  $\epsilon$  is a small Euclidean time, chosen partly to produce the desired operator ordering in real time. An out-of-time-ordered correlator is

$$\text{OTOC} \equiv \langle \mathcal{O}_{\Delta_1}(0) \mathcal{O}_{\Delta_2}(t) \mathcal{O}_{\Delta_1}(0) \mathcal{O}_{\Delta_2}(t) \rangle, \quad t > 0. \quad (3.39)$$

From the operator placing it is clear that this involves an analytic continuation of the more interesting four-point function in Euclidean space with crossed bilocal lines. In the energy basis, the OTOC can be written as

$$\langle \mathcal{O}_{\Delta_1}(t_1) \mathcal{O}_{\Delta_2}(t_2) \mathcal{O}_{\Delta_1}(t_3) \mathcal{O}_{\Delta_2}(t_4) \rangle = \frac{e^{S_0}}{2} \int_0^\infty d\mu(k_1) d\mu(k_2) d\mu(k_3) d\mu(k_4) \\ \times e^{-i \frac{k_1^2}{2C} t_{12} - i \frac{k_2^2}{2C} t_{23} - i \frac{k_3^2}{2C} t_{43} - \frac{k_4^2}{2C} (\beta - it_{41})} \mathcal{A}_{\text{cross.}}(k_1, k_2, k_3, k_4), \quad (3.40)$$



**Fig. 12** Penrose diagram of an AdS black hole. A particle is inserted at very early times  $-t/2$  (red) and another at very late times  $+t/2$  (blue). The scattering which happens near the horizon for large time separations is captured by the out-of-time ordered correlator (OTOC). Since the scattering happens at high energies it is described by a gravitational shockwave interaction

where the amplitude  $\mathcal{A}$  is the same as in (3.36). We also generalized the insertions to be at arbitrary four times with the understanding they are out of time ordered, meaning  $t_{1,3} < t_{2,4}$ .

In the context of semi-classical AdS/CFT, the OTOC is computed in gravity by a four-point function in a black hole background as shown in Fig. 12, see Shenker and Stanford (2015) for a detailed explanation. We have chosen  $t_1 = t_3 = -t/2$  and  $t_2 = t_4 = t/2$  to make it symmetric. In the semiclassical limit each insertion corresponds to a bulk particle denoted by red and blue for both types of operators.

As  $t$  grows, the particles meet closer to the black hole horizon and therefore are highly boosted. The OTOC is constructed by combining the scattering S-matrix of the two particles near the horizon with their respective in- and out-wavefunctions given by bulk-boundary propagators. At high energies, this scattering is described by a shockwave where one particle (red) creates a backreaction in the geometry that shifts the position of the other particle (blue) closer to the horizon. In higher dimensions this is a genuine bulk interaction, while in 2d the particles do not really interact in the bulk: the shockwave scattering is due to the dynamics of the boundary Schwarzian mode.

The OTOC gives a signature of quantum chaos for the quantum system describing the black hole. One way to see this is to notice that the OTOC appears when computing the expectation value of  $[\mathcal{O}_{\Delta_1}(t), \mathcal{O}_{\Delta_2}(0)]^2$ . Systems with quantum chaos display an exponential growth in time of this quantity for times bigger than the dissipation time (order of the temperature for the black hole) and smaller than the scrambling time (logarithmic in  $N$ , when the double commutator saturates).<sup>29</sup>

Next we will show explicitly how all these elements of the semi-classical picture emerge from the exact OTOC expressions.

**Semiclassical limit of OTOC** To identify the semi-classical behavior of the OTOC, we take the large  $C$  limit. In this limit all  $k \sim \mathcal{O}(C)$ , such that all intermediate energies are of order  $C$ , but energy differences are order one. As shown in Lam et al. (2018), in this limit only the pole at  $u = ik_4$  is relevant. Its residue gives a contribution to the

<sup>29</sup> It is important for this interpretation that the four-point function involves pairwise equal insertions. The more general case is also interesting (Turiaci 2019) but is not captured by gravity in a simple way.

OTOC amplitude given by

$$\mathcal{A} = \frac{\Gamma(\Delta_1 + ik_4 \pm ik_1)\Gamma(\Delta_1 - ik_3 \pm ik_2)}{(2C)^{2\Delta_1}\Gamma(2\Delta_1)} \frac{\Gamma(\Delta_2 - ik_1 \pm ik_2)\Gamma(\Delta_2 + ik_4 \pm ik_3)}{(2C)^{2\Delta_2}\Gamma(2\Delta_2)} \\ \times \Gamma(-2ik_4)\Gamma(-ik_4 \pm ik_2 + ik_1 + ik_3) + \dots \quad (3.41)$$

Following the procedure applied for the two-point function we parametrize the energies as

$$\frac{k_1^2}{2C} = E + \omega_1, \quad \frac{k_2^2}{2C} = E - \omega_3 - \omega_4, \quad \frac{k_3^2}{2C} = E - \omega_4, \quad \frac{k_4^2}{2C} = E, \quad (3.42)$$

where  $E$  is taken to be of order  $C$  and  $\omega \sim \mathcal{O}(1)$ . The first observation is that the  $E$ -dependence simplifies and gives  $\text{OTOC} = \int dE e^{2\pi\sqrt{2CE} - \beta E} \times \mathcal{O}(1)$ , where the order one factor is a function of the  $\omega$ 's and  $E/C$ . Therefore, just like for the two-point function, we can perform the  $E$ -integral at large  $C$  by saddle point techniques, relating  $E$  to the temperature  $E = 2\pi^2 C / \beta^2$ . We can now write the order one term as

$$\frac{\text{OTOC}}{Z(\beta)} = \int \prod_{i=1}^4 \frac{d\omega_i}{2\pi} e^{\sum_{i=1}^4 i\omega_i t_i} e^{\frac{\beta}{4}(\omega_1 - \omega_3 - 2\omega_4)} \left(\frac{2\pi}{\beta}\right)^{2\Delta_1 + 2\Delta_2 - 3} \\ \times \frac{\Gamma(\Delta_1 - i\frac{\beta\omega_1}{2\pi})\Gamma(\Delta_1 - i\frac{\beta\omega_3}{2\pi})\Gamma(\Delta_2 + i\frac{\beta\omega_2}{2\pi})\Gamma(\Delta_2 + i\frac{\beta\omega_4}{2\pi})}{\Gamma(2\Delta_1)\Gamma(2\Delta_2)} \\ \times \left(\frac{\beta}{4\pi C}\right)^{-i\frac{\beta(\omega_1 + \omega_3)}{2\pi}} \Gamma\left(i\frac{\beta(\omega_1 + \omega_3)}{2\pi}\right) 2\pi\delta\left(\sum_{i=1}^4 \omega_i\right). \quad (3.43)$$

This expression has a nice holographic interpretation. The OTOC involves a scattering of two particles and  $\omega$  measures their energy according to an asymptotic observer. The factors in the first and second line of Eq. (3.43) are bulk-to-boundary propagators and the particles interact through an S-matrix given by the final line of Eq. (3.43).

The physical interpretation becomes more transparent when written in a basis of wavefunctions with fixed Kruskal energy. In this basis the S-matrix should be the Dray–'t Hooft shockwave S-matrix  $\mathcal{S} = e^{i\frac{\beta}{4\pi C}p_-q_+}$ , where  $p_-$  and  $q_+$  are the Kruskal momenta of the two particles being scattered. Applying the change of basis between the Dray–'t Hooft S-matrix from Kruskal to Schwarzschild energies, which can be found in Shenker and Stanford (2015), gives the expression:

$$\mathcal{S}(\omega_1, \omega_2, \omega_3, \omega_4) = \int_0^\infty \frac{dq_+}{q_+} \int_0^\infty \frac{dp_-}{p_-} q_+^{\frac{i\beta(\omega_1 + \omega_3)}{2\pi}} p_-^{-\frac{i\beta(\omega_2 + \omega_4)}{2\pi}} e^{i\frac{\beta}{4\pi C}p_-q_+} \\ = \left(\frac{\beta}{4\pi C}\right)^{-i\frac{\beta(\omega_1 + \omega_3)}{2\pi}} e^{-\frac{\beta}{4}(\omega_1 + \omega_3)} \Gamma\left(i\frac{\beta(\omega_1 + \omega_3)}{2\pi}\right) 2\pi\delta\left(\sum_{i=1}^4 \omega_i\right). \quad (3.44)$$

The remainder of the calculation is to match the other factors in (3.43) with bulk-boundary propagators. The details of the calculation are in Lam et al. (2018).

After rewriting Eq. (3.43) in the Kruskal basis, we can do the integrals and obtain a closed expression for the semiclassical OTOC. This was done in Maldacena et al. (2016b) without the full OTOC answer and by working directly in the semi-classical limit (which can be matched to Eq. (3.43)). The answer is:

$$\frac{\text{OTOC}}{Z(\beta)} = \frac{U(2\Delta_1, 1 + 2\Delta_1 - 2\Delta_2, 1/z)}{z^{2\Delta_1} (\frac{\beta}{\pi} \sin \frac{\pi i t_{13}}{\beta})^{2\Delta_1} (\frac{\beta}{\pi} \sin \frac{\pi i t_{24}}{\beta})^{2\Delta_2}}, \quad (3.45)$$

where  $U(a, b, x)$  is the Tricomi confluent hypergeometric function and we normalize by the partition function  $Z(\beta)$ . We have also defined

$$z = \frac{i\beta}{16\pi C} \frac{e^{\pi(t_2+t_4-t_1-t_3)/\beta}}{\sinh \frac{\pi t_{13}}{\beta} \sinh \frac{\pi t_{24}}{\beta}}. \quad (3.46)$$

This answer also matches with specific Virasoro vacuum blocks computed in Chen et al. (2017), for reasons explained in Lam et al. (2018). We see from the expression for  $z$  that the correlator is trivial at large  $C$  unless we scale times such that  $e^t \sim C/\beta$ . This can also be seen by noting that the effective Dray-'t Hooft scattering phase is actually  $\mathcal{S}_{\text{eff}} = e^{i \frac{\beta e^t}{4\pi C} p-q+}$ , which goes to 1 unless  $e^t \sim C/\beta$ .

**Behavior of OTOC** We assume  $C$  is large and look at how the OTOC depends on time. It is convenient to choose the operators evenly spread along the thermal circle:

$$t_1 = -i\frac{\beta}{2}, \quad t_2 = t - i\frac{\beta}{4}, \quad t_3 = 0, \quad t_4 = t + i\frac{\beta}{4}, \quad \Rightarrow \quad z = \frac{\beta}{16\pi C} e^{\frac{2\pi}{\beta} t}. \quad (3.47)$$

This regularizes divergences that appear at coincident points by Euclidean damping factors. The following quantity will play an important role

$$t_{\text{sc}} = \frac{\beta}{2\pi} \log \frac{2\pi C}{\beta}, \quad (3.48)$$

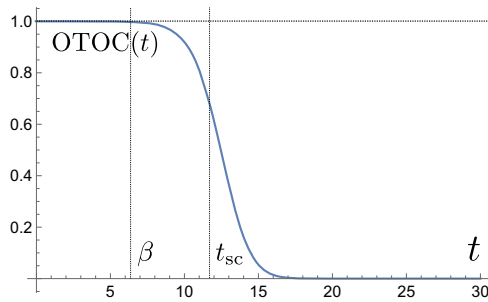
and is called the scrambling time of the black hole.

With these choices, we show a plot of the normalized OTOC( $t$ ) (3.45), where we additionally divide out the product of two-point functions, such that it starts at 1 for  $t \rightarrow -\infty$ . We marked the thermal time  $\beta$  and scrambling time  $t_{\text{sc}}$  for this configuration. This is illustrated in Fig. 13. The following qualitative features can be distinguished.

- We can first consider early times with small  $z$ , or equivalently  $t \ll t_{\text{sc}}$ . Using the expansion  $z^{-a} U(a, b, 1/z) \approx 1 - a(1+a-b)z$  for small  $z$ , we get

$$\frac{\text{OTOC}(t)}{Z(\beta)} \sim (\pi/\beta)^{2\Delta_1+2\Delta_2} \left( 1 - \frac{\Delta_1 \Delta_2 \beta}{4\pi C} e^{\frac{2\pi}{\beta} t} + \dots \right). \quad (3.49)$$

This is valid for times smaller than the thermal scale, where the OTOC becomes the product of two-point functions (the “1” in the above expansion), and is also



**Fig. 13** Plot of the OTOC as a function of time. We chose the parameters as  $\Delta_1 = \Delta_2 = 1$ ,  $\beta = 2\pi$  and  $8C = 10^6$ . The vertical lines separate different physical behavior. The first corresponds to time shorter than the dissipation time  $\sim \beta$ . Then the chaos regime kicks in at times  $\beta \ll t \ll t_{sc}$ . Finally after the scrambling time  $t_{sc}$ , the OTOC decays exponentially. For very late times  $t \gg C$  (not shown) quantum gravity effects become important

valid at times  $t \gtrsim \beta$  where it displays exponential deviations from this: this is chaotic Lyapunov behavior with maximal chaos exponent  $\frac{2\pi}{\beta}$ . This interpretation arises from the fact that

$$\langle [\mathcal{O}_1(t), \mathcal{O}_2(0)] \rangle \sim e^{\lambda t}, \quad \lambda = \frac{2\pi}{\beta}, \quad (3.50)$$

indicating the chaotic spread of an initial perturbation with Lyapunov exponent  $\lambda$ . The value for black holes saturates the chaos bound of Maldacena et al. (2016a), and black holes are therefore maximally chaotic.

- The correlator transitions to a different behavior at the scrambling time  $t \gg t_{sc}$  defined such that  $z$  is order one and the OTOC is small. In this regime  $U(a, b, 1/z) \approx \frac{\Gamma(1-b)}{\Gamma(1+a-b)} + z^{b-1} \frac{\Gamma(b-1)}{\Gamma(a)}$  when  $z \gg 1$  and therefore, assuming for concreteness that  $\Delta_1 > \Delta_2$ , the OTOC decays exponentially as

$$\text{OTOC}(t) \sim e^{-2\Delta_2 \frac{2\pi}{\beta} t}. \quad (3.51)$$

This is quasinormal mode decay. The reason is that the kinematical effect from the bulk-boundary propagation at these times dominates over the effect of the shockwave scattering.

- The scrambling time is logarithmic in  $C$ . A final transition happens when  $t \gg C$  (not shown on the figure above). At such late times quantum gravity effects are important, invalidating (3.45), and we need to go back to the exact expression. The low energy behavior of Eq. (3.36) and the density of states near the edge determines a power-law decay  $\text{OTOC}(t) \sim 1/t^6$ . This behavior has a quantum gravity origin and cannot be understood classically.



The quantity on the right hand side depends on six variables. It is the  $6j$ -symbol of  $SL(2, \mathbb{R})$ , and its explicit expression is

$$\begin{aligned} & \left\{ \begin{array}{ccc} \Delta_1 & k_1 & k_2 \\ \Delta_2 & k_3 & k_4 \end{array} \right\} \\ &= \sqrt{\frac{\Gamma(\Delta_1 + ik_1 \pm ik_4)\Gamma(\Delta_2 - ik_1 \pm ik_2)\Gamma(\Delta_1 - ik_3 \pm ik_2)\Gamma(\Delta_2 + ik_3 \pm ik_4)}{\Gamma(\Delta_1 - ik_1 \pm ik_4)\Gamma(\Delta_2 + ik_1 \pm ik_2)\Gamma(\Delta_1 + ik_3 \pm ik_2)\Gamma(\Delta_2 - ik_3 \pm ik_4)}} \\ & \times \int_{-i\infty}^{i\infty} \frac{du}{2\pi i} \frac{\Gamma(u \pm ik_4)\Gamma(u + ik_1 + ik_3 \pm ik_2)\Gamma(\Delta_1 - ik_1 - u)\Gamma(\Delta_2 - ik_3 - u)}{\Gamma(u + \Delta_1 + ik_1)\Gamma(u + \Delta_2 + ik_3)}. \end{aligned} \quad (3.55)$$

Multiplying the vertex factors with the  $6j$ -symbol involved in the crossing of lines for the case of the four-point function gives

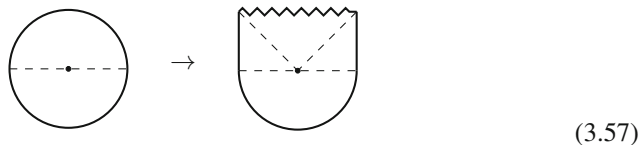
$$\mathcal{A}_{\text{cross.}} = \gamma_{\Delta_1}(k_1, k_4)\gamma_{\Delta_2}(k_4, k_3)\gamma_{\Delta_1}(k_3, k_2)\gamma_{\Delta_2}(k_2, k_1) \left\{ \begin{array}{ccc} \Delta_1 & k_1 & k_2 \\ \Delta_3 & k_3 & k_4 \end{array} \right\}. \quad (3.56)$$

which is precisely the amplitude quoted in (3.36). The final step consists in multiplying the whole correlator by  $e^{S_0}/2$ . Note that we are computing correlators without normalizing by the partition function  $Z(\beta)$ .

Some further examples of these rules, including higher-point OTOCs, can be found in Lam et al. (2018).

### 3.5 Pure states and end-of-the-world (EOW) branes

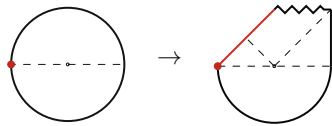
Most of this section studies JT gravity in the hyperbolic disk (which corresponds to the Euclidean thermal path integral of the boundary theory). Half of the hyperbolic disk can be used to prepare a state: the thermofield double state. This state is dual to the Einstein–Rosen bridge obtained from the maximally extended black hole solution:



This state is represented by  $|\text{TFD}\rangle = \sum_n e^{-\beta E_n/2} |E_n\rangle_L \otimes |E_n\rangle_R$ . After integrating out the left boundary, this gives a thermal density matrix for the right system.

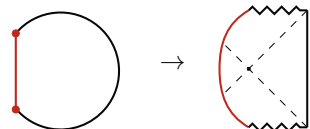
To instead produce a pure state on the right, one can project the left state onto a chosen state  $|B\rangle_L$  such that we get  $|\Psi\rangle_R = {}_L\langle B|\text{TFD}\rangle$ . The construction of such states was done explicitly in the SYK model. The projection on specific states can

be modeled by a heavy operator with scaling dimension of order  $C$ . The diagram proposed in Kourkoulou and Maldacena (2017) is


(3.58)

On the left we show the Euclidean preparation with the projection operator inserted in red. On the right is the diagram after continuing to Lorentzian signature, and shows how the projection creates a high-energy shockwave in the bulk that acts as a so-called EOW brane behind the black hole horizon. We could also consider the Penrose diagram fully in Lorentzian time: the EOW brane emerges from the past singularity, reaches the left boundary and then falls to the future singularity.

More general projections that act non-locally near the left boundary can also be considered. These are dual to states described geometrically by


(3.59)

The EOW brane moves behind the horizon, but does not reach the left boundary. In Euclidean signature, this can be obtained by a boundary curve made of two pieces: a holographic boundary of renormalized length  $\beta$  and a geodesic (corresponding to the EOW brane) of renormalized length  $\ell$ . The Euclidean action describing this set-up is

$$I = I_{\text{JT}} + m \int_{\text{brane}} ds, \quad (3.60)$$

where the second term is integrated over the EOW brane worldline and  $m$  is the tension. This term effectively imposes the boundary conditions  $\partial_n \Phi = m$  and  $K = 0$  along the EOW worldline, with  $\partial_n$  the normal derivative. The path integral in Euclidean signature over the disk with a partly holographic boundary of renormalized length  $\beta$  and an EOW brane with geodesic length  $\ell$  is then given by Harlow and Jafferis (2018); Yang (2019):

$$Z(\beta; \ell) = \int_{\text{disk}} \mathcal{D}\Phi = 2e^{-\ell/2} \int_0^\infty d\mu(k) e^{-\beta \frac{k^2}{2C}} K_{2ik}(4e^{-\ell/2}). \quad (3.61)$$

As a check of this formula, one can verify that gluing two such disks along the geodesic, adding a matter propagator  $e^{-\Delta\ell}$  and integrating over the geodesic length  $\int d\ell Z(\tau; \ell) Z(\beta - \tau; \ell) e^{-\Delta\ell}$  reproduces the two-point function (3.24). Equation



(3.61) can also be interpreted as the wavefunction for creating an ER bridge of length  $\ell$  in a TFD state of inverse temperature  $\beta$ .

There have been several applications in JT gravity involving EOW branes. The first was already mentioned (Kourkoulou and Maldacena 2017). One can also consider intermediate situations of partially entangled states which can be modeled by (not EOW) branes gluing to patches of  $\text{AdS}_2$  (Goel et al. 2019). More recently, they were used in Penington et al. (2019) to model pure state black hole evaporation in a simple model of a system with a Page curve. Embedding of dynamical EOW branes in a random matrix framework appeared in e.g. Gao et al. (2022); Blommaert and Usatyuk (2022).

### 3.6 Other operator insertions in JT gravity

The bilocal operator studied so far is not the only operator one can insert into the gravitational path integral. Here we go through other possibilities.

In fact, the simplest *local* operator insertion that is  $\text{PSL}(2, \mathbb{R})$  invariant is the local energy operator  $E(t) = -C \{F, t\}$ . Its correlators are relatively straightforward to determine and are piecewise constant up to contact terms. Within the diagrammatic formalism, they represent the energy  $k^2/2C$  in the different sectors of the diagram, depending on where we insert them in the Schwarzian path integral. To derive the expressions, one can write down a Ward identity relating a correlator with such an insertion to one without, see Stanford and Witten (2017); Mertens et al. (2017); Mertens (2019) for details from various perspectives on this result.

Mixing up these operator insertions with the bilocals, reinforces the physical interpretation of the bilocal operator itself in holography: in the sector between the endpoints of the bilocal, an energy has been injected into the gravitational bulk, only to be removed again at the second endpoint of the bilocal operator. A single endpoint of these operators can hence be compared to the classical energy pulses we studied in Sect. 2.

A structurally more interesting class of operators is the so-called defects in the JT bulk to which we turn next (Mertens and Turiaci 2019). We have previously shown that the thermal disk partition function is described by writing  $F(\tau) = \tan \frac{\pi}{\beta} f(\tau)$ , where  $f(\tau) \in \frac{G}{H} \equiv \frac{\text{Diff } S^1}{\text{PSL}(2, \mathbb{R})}$ . This integration space is special since it is symplectic, but it is also precisely the coadjoint orbit of the identity element of the Virasoro group (Witten 1988; Balog et al. 1998). This is no coincidence. In fact, there is a host of Schwarzian models that can be found by considering all of the different Virasoro orbits, with different choices of the preserved subgroup  $H$ . All of them can be interpreted gravitationally. Let us give an overview. In all cases, we write the Schwarzian action as:

$$S_H[f] = -C \int_0^\beta d\tau \{F \circ_H f(\tau), \tau\}. \quad (3.62)$$

The different models are at a crude level distinguished by a monodromy matrix  $M \in \mathrm{PSL}(2, \mathbb{R})$ , defined by the periodicity relation  $F(\tau + \beta) = M \cdot F(\tau) = \frac{aF(\tau)+b}{cF(\tau)+d}$ , and we label the orbits by the conjugacy class of this matrix  $M$ . The stabilizer  $H$  is then defined as the subgroup of all  $h \in \mathrm{PSL}(2, \mathbb{R})$  satisfying  $[h, M] = 0$ . We can distinguish the following different models<sup>30</sup>

- **Elliptic**  $H = U(1)_\theta$ . Elliptic orbits are parametrized by

$$F \circ_\theta f = \tan \frac{\pi}{\beta} \theta f, \quad M = \begin{pmatrix} \cos(\pi\theta) & \sin(\pi\theta) \\ -\sin(\pi\theta) & \cos(\pi\theta) \end{pmatrix} \in \mathrm{PSL}(2, \mathbb{R}). \quad (3.63)$$

In the bulk, one can choose conformal gauge and continue this time reparametrization into the bulk as in (2.32) by setting  $U(u) = F(u)$  and  $V(v) = F(v)$ . This choice is arbitrary, but coordinate-invariant results we obtain from it have meaning. This leads to the metric:

$$ds^2 = 4 \left( \frac{\pi\theta}{\beta} \right)^2 \frac{d\tau^2 + dz^2}{\sinh^2 \frac{2\pi}{\beta} \theta z}. \quad (3.64)$$

This geometry can be readily shown to have a conical singularity of periodicity  $2\pi\theta$ .

Within the path integral, one can show that the result is again one-loop exact:

$$\begin{aligned} Z(\beta, \theta) &= \beta \left( \text{triangle} \right) 2\pi\theta = \int_{\mathrm{Diff} S^1/U(1)_\theta} [\mathcal{D}f] e^C \int_0^\beta d\tau \{ \tan \frac{\pi}{\beta} \theta f(\tau), \tau \} \\ &= \int_0^{+\infty} dk \frac{\cosh(2\pi\theta k)}{\pi} e^{-\beta \frac{k^2}{2C}} = \frac{1}{2\pi} \left( \frac{2\pi C}{\beta} \right)^{1/2} e^{\frac{2\pi^2 C}{\beta} \theta^2}. \end{aligned} \quad (3.65)$$

The calculation only has a single zero-mode in this case, since the only isometry of the metric are rotations around the defect.

The action functional for these models precisely matches with the energy functional in Virasoro coadjoint orbit theory. In particular, bounds of this functional in the different orbits have been analyzed (Balog et al. 1998), which can then be used to show that for  $\theta > 1$  negative modes exist and the path integral evaluation is formal. Nonetheless, the answer can be derived by other means (e.g., from limits of Virasoro characters (Mertens and Turiaci 2019), or by complexifying the unstable

<sup>30</sup> Some fineprint is due. Firstly, there is no agreed-upon naming of the different orbits, so care must be taken when comparing different references. Secondly, the difference between orbits and conjugacy classes is mainly that we need to keep track of the winding number of the elliptic conjugacy class. This means the parameter  $\theta \in \mathbb{R}$  instead of  $\theta \in [0, 1)$ . Finally, we only consider orbits that have a constant representative. The reason is that the other orbits, by definition, have no solution to the equation  $\frac{d}{d\tau} \{F, \tau\} = 0$ , but this is precisely the saddle equation of all of these models. Such orbits would hence be highly quantum, and we are not aware of any use of these orbits (in any context).

fluctuations), see Alekseev and Shatashvili (2021) for some further speculation on an interpretation in terms of a change of integration cycle.<sup>31</sup>

The following observation made in Mertens and Turiaci (2019) will be important in Sect. 4.5.2. The presence of a defect at the point  $x_c$  of the hyperbolic disk means the scalar curvature has a delta-function singularity  $R(x) = -2 + 4\pi(1-\theta)\delta^2(x-x_c)$ . This is equivalent to the insertion in the JT gravity path integral of an operator  $e^{\frac{1}{4GN}(1-\theta)\Phi(x_c)}$  at the location of the defect.

- **Exceptional Elliptic**  $H = \mathrm{PSL}^n(2, \mathbb{R})$ . When  $\theta \in \mathbb{N}$ , the stabilizer is enhanced:

$$F \circ_n f = \tan \frac{\pi}{\beta} n f, \quad M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{PSL}(2, \mathbb{R}). \quad (3.68)$$

The conical singularity is  $2\pi n$ , and this represents a replicated version of the original disk. There are three zero-modes (just like for the  $n = 1$  case where there is no conical deficit), and we obtain the one-loop exact amplitude:

$$\begin{aligned} Z(\beta, n) &\equiv \int_{\mathrm{Diff} S^1 / \mathrm{PSL}^n(2, \mathbb{R})} [\mathcal{D}f] e^{C \int_0^\beta d\tau \{ \tan \frac{\pi}{\beta} n f(\tau), \tau \}} \\ &= \int_0^{+\infty} dk \frac{k \sinh(2\pi n k)}{2\pi^2} e^{-\beta \frac{k^2}{2C}} = \frac{n}{4\pi^2} \left( \frac{2\pi C}{\beta} \right)^{3/2} e^{\frac{2\pi^2 C}{\beta} n^2}. \end{aligned} \quad (3.69)$$

- **Hyperbolic**  $H = U(1)_\lambda$ . The hyperbolic orbit can be viewed as an analytic continuation of the elliptic case where in essence  $\theta \rightarrow i\lambda$ :

$$F \circ_\lambda f = \tanh \frac{\pi}{\beta} \lambda f, \quad M = \begin{pmatrix} \cosh(\pi \lambda) & \sinh(\pi \lambda) \\ \sinh(\pi \lambda) & \cosh(\pi \lambda) \end{pmatrix} \in \mathrm{PSL}(2, \mathbb{R}). \quad (3.70)$$

<sup>31</sup> We present some details on the boundedness of the Schwarzian action. We first write the Schwarzian action as  $-C \int_0^\beta d\tau \{F, \tau\} = -C \int_0^\beta d\tau \{f, \tau\} - \frac{2\pi^2 C}{\beta^2} \theta^2 \int_0^\beta d\tau f'^2$ . One can show the following inequalities for the Schwarzian action of  $f \in \mathrm{Diff} S^1$ :

$$-C \frac{2\pi^2}{\beta^2} \int_0^\beta d\tau (1 - f'^2) \leq -C \int_0^\beta d\tau \{f, \tau\} = \frac{C}{2} \int_0^\beta d\tau \left( \frac{f''}{f'} \right)^2. \quad (3.66)$$

The rightmost equality shows that the Euclidean action is unbounded from above by making  $f'$  have strong fluctuations. The inequality on the left can then be used directly to derive the following lower bound on the action:

$$-C \int_0^\beta d\tau \{F, \tau\} \geq -\frac{2\pi^2 \theta^2 C}{\beta} + \frac{2\pi^2 C}{\beta^2} (1 - \theta^2) \int_0^\beta d\tau (f' - 1)^2 \quad (3.67)$$

When  $\theta^2 < 1$ , the second term on the RHS is positive for all  $f$ , and the action is bounded below by  $-\frac{2\pi^2 \theta^2 C}{\beta}$ . On the other hand, for  $\theta^2 > 1$ , one can readily find examples that lower the action without bound. The result is that for  $\theta > 1$ , the path integral has negative modes and the one-loop computation is formal.

The geometry is of the form:

$$ds^2 = 4 \left( \frac{\pi \lambda}{\beta} \right)^2 \frac{d\tau^2 + dz^2}{\sin^2 \frac{2\pi}{\beta} \lambda z}, \quad (3.71)$$

and describes instead a macroscopic tube, with a minimal length circle (the neck of a wormhole) at  $z = \frac{\beta}{4\lambda}$ . This neck is a geodesic with a circumference  $b \equiv 2\pi\lambda$ . Once again, the answer is one-loop exact:

$$\begin{aligned} Z(\beta, \lambda) &= \beta \left( \text{tube diagram} \right) b = \int_{\text{Diff } S^1/U(1)_\lambda} [\mathcal{D}f] e^{C \int_0^\beta d\tau \{ \tanh \frac{\pi}{\beta} \lambda f(\tau), \tau \}} \\ &= \int_0^{+\infty} dk \frac{\cos(2\pi\lambda k)}{\pi} e^{-\beta \frac{k^2}{2C}} = \frac{1}{2\pi} \left( \frac{2\pi C}{\beta} \right)^{1/2} e^{-\frac{C}{2\beta} b^2}. \end{aligned} \quad (3.72)$$

• **Parabolic**  $H = U(1)_0$ . Finally, we have the parabolic orbit:

$$F \circ_0 f = f, \quad M = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \text{PSL}(2, \mathbb{R}), \quad (3.73)$$

which is found as a limiting case of both elliptic and hyperbolic defects. It corresponds geometrically to an infinite cusp-singularity, identifiable as thermal  $\text{AdS}_2$  in Poincaré coordinates, with one-loop exact amplitude:

$$Z(\beta, 0) = \beta \left( \text{cusp diagram} \right) = \int_0^{+\infty} dk \frac{1}{\pi} e^{-\beta \frac{k^2}{2C}} = \frac{1}{2\pi} \left( \frac{2\pi C}{\beta} \right)^{1/2}. \quad (3.74)$$

One can mix in these defect insertions with bilocal operator insertions by replacing the usual JT spectral density  $d\mu(k_i)$  in the sector of the diagram of interest by those found in (3.65), (3.69), (3.72), or (3.74). Some example expressions were written down in Mertens and Turiaci (2019).

A useful diagrammatic representation is to include a cross in the relevant sector of the diagram, labeled by the orbit parameter. An example for the OTOC with a hyperbolic defect  $\lambda$  in one of the four sectors is the diagram:



$$(3.75)$$

### 3.7 Outline of derivations in the literature

In this subsection we briefly outline and summarize the main approaches available to derive the JT amplitudes discussed in this section, and refer to the original papers for more details. Each of the studied approaches has benefits and downsides, which we describe, and it is worthwhile to have some understanding of all of them.

#### 3.7.1 Free particle approach

Using the redefinition  $F' = e^\varphi$ , which is always possible since  $F' \geq 0$  for a suitable time reparametrization, one immediately sees that the thermal Schwarzian action can be rewritten into a free boson action:

$$S = -C \int_0^\beta dt \{F, t\} = \frac{C}{2} \int_0^\beta dt (\partial_\tau \varphi)^2. \quad (3.76)$$

Implementing the (formal) periodicity requirement  $F(\tau + \beta) = F(\tau) + \infty$  leads to the constraint  $\int_0^\beta dt e^\varphi = +\infty$ , which can be implemented after regularization using a Lagrange multiplier, and which leads to an exponential potential added to the otherwise free boson system (3.76).

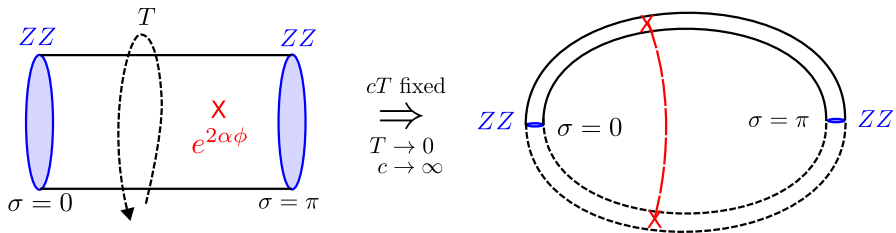
Bilocal operator insertions become more complicated non-local operators since

$$\left( \frac{F'_1 F'_2}{(F_1 - F_2)^2} \right)^\Delta = \frac{e^{\Delta \varphi_1} e^{\Delta \varphi_2}}{\left( \int_{\tau_1}^{\tau_2} d\tau e^\varphi \right)^{2\Delta}}. \quad (3.77)$$

Correlators can then be computed by using the identity  $A^{-p} = \frac{1}{\Gamma(p)} \int_0^\infty d\alpha \alpha^{p-1} e^{-\alpha A}$ , which leads to an exponential potential with piecewise constant prefactor (Bagrets et al. 2016, 2017), with the prefactor changing each time a bilocal endpoint is encountered.

As a closely related approach, one can instead use the black hole time coordinate  $f' = e^\varphi$  and rewrite the Schwarzian action again as a particle in exponential potential(s). This is however less useful when including operator insertions, and the full expression for the correlators has not been obtained from this approach, except in some special cases.

Strong points	Weak points
→ Fast and intuitive, since it relates to the 1d free boson.	→ Subtle in terms of boundary conditions and handling of $\mathrm{PSL}(2, \mathbb{R})$ gauge group → More difficult for more complicated correlators, not useful to derive final answers



**Fig. 14** 2d Liouville CFT between vacuum branes and its double-scaled Schwarzian limit. Operator insertions can be added and are depicted in red

### 3.7.2 Limit of Liouville theory

Schwarzian correlation functions can be found by taking a suitable double-scaling limit of Virasoro CFT (Mertens et al. 2017). Let us sketch how that works. Virasoro CFT at large central charge can be dynamically realized by the 2d Liouville CFT, described by the classical action:

$$S_L = \frac{1}{4\pi b^2} \int_0^T d\tau \int_0^\pi d\sigma \left[ (\partial\phi)^2 + 4\mu e^{2\phi} \right], \quad (3.78)$$

on a cylindrical surface with  $\sigma \in (0, \pi)$  and Euclidean time direction  $\tau$  of period  $T$ . The central charge is  $c = 1 + 6(b + b^{-1})^2$ . In 2d Liouville CFT, one can consider branes for the worldsheet to end on. In particular, generalizing Cardy's construction of boundary states to irrational CFT, considering a pair of identity or ZZ-branes (Zamolodchikov and Zamolodchikov 2001) at the ends of the cylinder worldsheet at  $\sigma = 0, \pi$  gives an amplitude that is the Virasoro vacuum character:

$$\chi_0(q) = \frac{q^{\frac{1-c}{24}} (1-q)}{\eta(iT/2\pi)}, \quad q = e^{-T}. \quad (3.79)$$

A convenient trick in dealing with identity branes is to use the mirror doubling trick with periodic boundary conditions around the doubled space. The procedure is now to study this amplitude in the limit where the cylinder becomes long and narrow, in conjunction with the classical large  $c$  limit ( $b \rightarrow 0$ ). This means that in this limit, the cylinder becomes a narrow (doubled) circular tube of length  $2\pi$  that degenerates. This is precisely the thermal circle where the Schwarzian model lives! The entire set-up is given in Fig. 14.

One can use this procedure at the level of the Liouville path integral directly to find both the Schwarzian action and path integral measure (Mertens et al. 2017; Mertens 2018) using an older field redefinition of Gervais and Neveu (1982a, b, 1983a, b). In the process, one encounters the 2d ancestor of the Schwarzian model as the so-called Alekseev–Shatashvili geometric action for coadjoint orbits of the Virasoro group (Alekseev and Shatashvili 1989, 1990).<sup>32</sup>

<sup>32</sup> This structural relation with Virasoro CFT is useful also in its relation with 3d gravity, see e.g. Cotler and Jensen (2019).

We can furthermore consider Liouville primary vertex operator insertions  $e^{2\alpha\phi(u,v)}$  on the cylinder worldsheet in the above double-scaling limit. This operator pairs up with its mirror image to form a single bilocal line (3.19), as drawn above. This calculation leads precisely to the JT disk boundary correlators as shown in detail in Mertens et al. (2017); Mertens (2018), Lam et al. (2018).

Diagrams with crossing bilocal lines are found by incorporating the Schwarzian double-scaling limit of the 2d braiding  $R$  matrix of Virasoro conformal blocks. Intuitively, if we have two uncrossed bilocal lines of the type drawn above, we can generate the crossed diagram if we swap either of the endpoints of both bilocals. This process acts on a chiral half of the Liouville CFT calculation and can be achieved by using the braiding  $R$  matrix of the relevant modular tensor category.

Furthermore, defect insertions can be found by further enriching the story by adding Verlinde loops wrapped along the  $\tau$ -direction of the cylinder (Mertens and Turiaci 2019). This is a generalization of studying different brane types (beyond ZZ) at either end of the worldsheet.

Finally, local 2d CFT stress tensor insertions reduce to insertions of the energy operator  $E(t)$  in the Schwarzian limit.

Let us show the procedure in more detail for the partition function. We can evaluate (3.79) in the dual “closed” channel as

$$\chi_0(q) = \int_0^\infty dP S_0^P \chi_P(\tilde{q}), \quad \chi_P(\tilde{q}) = \frac{\tilde{q}^{P^2}}{\eta(i2\pi/T)}, \quad \tilde{q} = e^{-4\pi^2/T}, \quad (3.80)$$

with the Virasoro modular S-matrix element:

$$S_0^P \sim \sinh(2\pi bP) \sinh\left(\frac{2\pi P}{b}\right). \quad (3.81)$$

Parametrizing  $P = bk$ , we take the Schwarzian double-scaling limit (Mertens et al. 2017), where we let the central charge go to infinity  $c \rightarrow \infty$  (or  $b \rightarrow 0$ ) and simultaneously let  $T \rightarrow 0$  keeping the combination  $Tc = \frac{24\pi^2}{\beta}$  fixed where  $\beta$  will be the inverse temperature of the resulting 1d model. We get:

$$\chi_0(q) \rightarrow 2\pi b^2 \int dk k \sinh 2\pi k e^{-\beta k^2}, \quad (3.82)$$

which is the JT disk partition function, up to an immaterial prefactor that can be absorbed into  $S_0$  again. Notice that descendants scale out as is suitable for a classical limit, and that the momentum  $k$  comes from the “closed” channel and hence flows along the boundary circle.

<i>Strong points</i>	<i>Weak points</i>
→ Explains presence of 2d CFT structure in JT and Schwarzian. → One can use and exploit the past knowledge of 2d Liouville results.	→ Obscures the bulk JT disk picture, since the auxiliary Liouville CFT lives on a different space than JT.

### 3.7.3 Boundary particle approach

A different approach uses a direct rewriting of the GHY boundary term of the gravity path integral using the 2d Gauss–Bonnet theorem into a worldline path integral as (Kitaev and Suh 2019; Yang 2019; Suh 2020):

$$\sim \int [DX] e^{-\int_0^\beta d\tau \frac{1}{2} \frac{T'^2 + Z'^2}{Z^2} + q \frac{T'}{Z}}, \quad q = \frac{a}{2\epsilon} \rightarrow +\infty, \quad (3.83)$$

which is a non-relativistic particle in 2d hyperbolic space  $(T, Z)$  in an imaginary infinite magnetic field. One can reach the same conclusion in a Hamiltonian formulation of the model.

This non-relativistic system has been solved by Comtet and Houston (1985, 1987), with the spectral measure leading to the JT disk spectral density in the  $q \rightarrow +\infty$  limit:

$$\rho(k) = \frac{1}{4\pi^2} \frac{k \sinh 2\pi k}{\cosh 2\pi q + \cosh 2\pi k} \xrightarrow{q \rightarrow +\infty} \sim k \sinh 2\pi k, \quad (3.84)$$

with an infinite but  $k$ -independent prefactor in the Schwarzian limit.

Using the known propagator of the particle in the magnetic field in the same  $q \rightarrow +\infty$  limit, one can derive the following suggestive expression for JT boundary correlators for the simplest case of ordered operators (Yang 2019):

$$\begin{aligned} & \langle \mathcal{O}(\tau_1) \dots \mathcal{O}(\tau_n) \rangle \\ &= \int \frac{\prod_{i=1}^n \frac{d^2 x_i}{Z_i^2}}{V(\text{SL}(2, \mathbb{R}))} G(\tau_{12}; x_1, x_2) \dots G(\tau_{n1}; x_n, x_1) \langle \mathcal{O}(x_1) \dots \mathcal{O}(x_n) \rangle_{\text{QFT}}, \end{aligned} \quad (3.85)$$

where the ingredients are the bulk propagators of the particle in magnetic field model  $G(\tau_{ij}; x_i, x_j) \equiv \langle x_i | e^{-\tau_{ij} H} | x_j \rangle$  in proper time  $\tau_{ij}$  (between points in the hyperbolic disk  $x_i$  and  $x_j$ ) and the “undressed” bulk matter correlator  $\langle \mathcal{O}(x_1) \dots \mathcal{O}(x_n) \rangle_{\text{QFT}}$  that can be determined in terms of Witten diagrams.

### 3.7.4 Two-dimensional gauge theory

We have previously mentioned the BF formulation of JT gravity on a closed manifold in Sect. 2.1.1. For a group  $G$  BF model on a manifold  $\mathcal{M}$  with a boundary  $\partial\mathcal{M}$ , we



<i>Strong points</i>	<i>Weak points</i>
→ Direct approach. → Makes the JT bulk explicitly visible, and allows for bulk interactions to be treated universally since we start with the generic undressed $\langle \mathcal{O}(x_1) \dots \mathcal{O}(x_n) \rangle_{\text{QFT}}$ .	→ Structural and symmetry properties less visible a priori in more complicated correlators.

can include the boundary term and condition (Mertens 2018):

$$I_{\text{BF}} = - \int_{\mathcal{M}} \text{Tr}(BF) + \frac{1}{2} \oint_{\partial\mathcal{M}} d\tau \text{Tr}(BA_\tau), \quad B = A_\tau|_{\partial\mathcal{M}}. \quad (3.86)$$

With this boundary term, we have a well-defined variational principle. Structurally, one can motivate this also by realizing one gets precisely this term when dimensionally reducing 3d Chern–Simons theory to the 2d BF model.

Path-integrating over  $B$  along an imaginary contour makes  $F$  pure gauge on the disk, and  $A_\mu = g^{-1} \partial_\mu g$ . The dynamics reduces to a pure boundary dynamics given by the action of a non-relativistic particle on a group manifold:

$$I_{\text{BF}} = \frac{1}{2} \oint_{\partial\mathcal{M}} d\tau \text{Tr}(g^{-1} \partial_\tau g)^2. \quad (3.87)$$

Notice the similarity of this argument with that for JT gravity of Sect. 2.3.4.

Interesting operator insertions in this model are boundary-anchored Wilson lines. Since they intersect the boundary twice, these correspond to bilocal operator insertions in the boundary model defined by the action (3.87). Correlation functions of any number of these bilocal operator insertions can be determined using techniques very similar to 2d Yang–Mills theory (Migdal 1975; Witten 1991a; Cordes et al. 1995). The result of the disk partition function, and that with a single Wilson line insertion, in this model are given by (Mertens 2018; Blommaert et al. 2018; Iliesiu et al. 2019):

$$Z(\beta) = \sum_R (\dim R)^2 e^{-\beta C_R}, \quad (3.88)$$

$$\left\langle G_{MM}^R(\tau, 0) \right\rangle = \sum_{R_1, R_2, m_1, m_2} \dim R_1 \dim R_2 e^{-C_{R_1} \tau} e^{-C_{R_2}(\beta - \tau)} \left( \begin{matrix} R_1 & R_2 & R \\ m_1 & m_2 & M \end{matrix} \right)^2, \quad (3.89)$$

where the  $R_i$  labels run through all unitary irreducible representations of the group  $G$  of interest, with  $\dim R_i$  and  $C_{R_i}$  its dimension and quadratic Casimir respectively. The  $m$ -label runs through the different states within each representation. Finally, the  $3j$ -symbols  $\left( \begin{matrix} R_1 & R_2 & R \\ m_1 & m_2 & M \end{matrix} \right)$  in the above example), and  $6j$ -symbols can appear for Wilson line endpoints and bulk crossings respectively (Mertens 2018; Blommaert et al. 2018; Iliesiu et al. 2019). One can directly see the analogy with the JT correlation functions

by setting  $\sum_R \rightarrow \int_0^{+\infty} dk$  and interpreting<sup>33</sup>

$$\dim R \rightarrow \frac{1}{\pi^2} k \sinh 2\pi k, \quad 3j \rightarrow \sqrt{\frac{\Gamma(\Delta \pm ik_1 \pm ik_2)}{(2C)^2 \Delta \Gamma(2\Delta)}}, \quad 6j \rightarrow \left\{ \begin{array}{ccc} \Delta_1 & k_1 & k_2 \\ \Delta_2 & k_3 & k_4 \end{array} \right\}. \quad (3.90)$$

These observations can be made more explicit in terms of a BF model closely related to the non-compact  $SL(2, \mathbb{R})$  group.<sup>34</sup> There are two proposals: one based on the positive semigroup  $SL^+(2, \mathbb{R})$  (Blommaert et al. 2018, 2019b), and one based on an analytic continuation of the universal cover of  $SL(2, \mathbb{R})$  (Iliesiu et al. 2019).

Defect insertions can also be studied and correspond to insertions of  $\text{Tr}_R e^{2\pi B}$ , which descend from “vertical” Wilson loops in irrep  $R$  in the Chern–Simons ancestor of the BF model. Finally, insertions of the quadratic Casimir itself correspond directly to energy operator insertions in the JT model.

It should be noted that one could also be interested in the case of compact group  $G$  in its own right, and add this sector to the gravitational JT sector. We present some details later on in Sect. 4.5.3. This describes more general models of 2d gravity that include conserved charges, of relevance to higher-dimensional black hole physics.

Both the approach of this section and that of Sect. 3.7.2 above have a structurally natural embedding of the different JT operator insertions, and from both perspectives, these operator insertions exhaust the interesting possibilities.

---

#### *Strong points*

→ Intuitive picture of natural operator insertions.  
 → Group theoretical structure explicit.  
 → Provides efficient generalizations to models that include other conserved quantum numbers.

---

#### *Weak points*

→ Subtleties in the group theoretic structure make this approach a priori difficult.  
 → Higher topology moduli space is Teichmüller space, instead of the moduli space of Riemann surfaces.

---

<sup>33</sup> One might worry about the  $(\dim R)^2$  in the disk amplitude (3.88). One way to understand how to reproduce the JT answer, is that gravity has asymptotic Brown–Henneaux boundary conditions, which mathematically lead to a coset model instead of a pure group model. For a coset model, the indices running in loops are somewhat constrained, removing one  $(\dim R)$  factor and leading directly to an expression similar as in JT gravity.

<sup>34</sup> Around (3.61) we noted one can get the two-point function by performing an integral over the boundary-to-boundary geodesic length  $\ell$ . That integral has a direct interpretation in terms of a group integral of three representation matrices.

## 4 Spacetime wormholes and random matrices

In this section, we review how to compute wormhole contributions to gravitational amplitudes, and how they further correct the previous answers towards results that are more in line with a finite entropy boundary system in holography. This section heavily draws on material from the seminal work by Saad et al. (2019).

### 4.1 Motivation: information loss and late time decay

First, let us take stock of where we are up to this point and how our JT amplitudes are integrated into a holographic framework of a microscopic UV-complete system.

Consider a discrete microscopic holographic model, i.e. a 0+1d quantum mechanical system with a bulk dual. Its boundary two-point correlator can be expanded in the energy eigenbasis as:

$$\mathrm{Tr} \left[ e^{-\beta H} \mathcal{O}(t) \mathcal{O}(0) \right] = \sum_{n,m} e^{-\beta E_n} |\langle n | \mathcal{O} | m \rangle|^2 e^{-it(E_m - E_n)}, \quad (4.1)$$

where this truly is a discrete sum over energy eigenstates labeled by  $n$  and  $m$ . Due to the energy phase factors, this expression oscillates erratically as a function of time  $t$ . At late times, the terms with  $n = m$  dominate the correlator, and we find the late-time mean non-zero value<sup>35</sup>

$$\mathrm{Tr} \left[ e^{-\beta H} \mathcal{O}(t) \mathcal{O}(0) \right] \approx \sum_n e^{-\beta E_n} |\langle n | \mathcal{O} | n \rangle|^2. \quad (4.2)$$

This late-time behavior of oscillations around a non-zero mean, is to be contrasted with our results in JT gravity so far. Indeed, semi-classically we have the late-time quasinormal mode exponential decay (as in Eq. (3.18)):

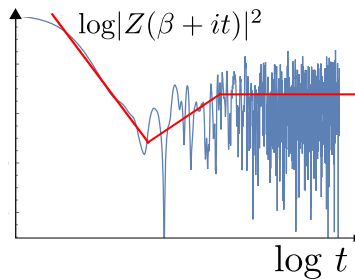
$$\frac{\langle \mathcal{O}(t) \mathcal{O}(0) \rangle_\beta}{Z(\beta)} = \left( \sinh \frac{\pi}{\beta} t \right)^{-2\Delta} \sim e^{-\frac{2\pi}{\beta} \Delta t}. \quad (4.3)$$

Quantum gravitational corrections discussed up to this point, would modify this into a power-law decay instead (3.30):

$$\frac{\langle \mathcal{O}(t) \mathcal{O}(0) \rangle_\beta}{Z(\beta)} \sim t^{-3}. \quad (4.4)$$

Both of these are in conflict with the behavior of a discrete finite entropy boundary system (4.1). This mismatch is one of the manifestations of information loss (Maldacena

<sup>35</sup> Physically, we can motivate this dominance of diagonal contributions by introducing a late-time averaging procedure to smoothen erratic fluctuations. E.g.  $\frac{1}{\Delta} \int_{T-\Delta/2}^{T+\Delta/2} e^{-it(E_m - E_n)} dt = \frac{2 \sin \frac{1}{2} \Delta(E_m - E_n)}{\Delta(E_m - E_n)} e^{-iT(E_m - E_n)}$ , which is suppressed as long as we average over multiple periods of the oscillating signal:  $\Delta \gg (E_m - E_n)^{-1}$ . Diagonal contributions ( $E_m = E_n$ ) are not suppressed by this procedure.



**Fig. 15** Spectral form factor for a matrix drawn from the GUE, with the averaged result in red

2003a), related to a continuum of energies in the energy spectrum due to the infinitely large near-horizon black hole region, or to the infinite redshift of near-horizon matter according to an asymptotic observer.

In order to address this, it is useful to first simplify a bit the observable to its core essence. For sufficiently simple operators  $\mathcal{O}$  that have zero one-point-function,<sup>36</sup> by ETH one expects the squared energy matrix elements  $|\langle n|\mathcal{O}|m\rangle|^2$  to be smooth functions of the energy (Deutsch 1991; Srednicki 1994). Hence the main qualitative features we want to study here solely come from the erratically oscillating phase factors in (4.1). So, as a zeroth-order approximation, we remove these energy matrix elements and study instead the spectral form factor:

$$Z(\beta + it)Z(\beta - it) = \sum_{n,m} e^{-(\beta+it)E_n} e^{-(\beta-it)E_m}, \quad (4.5)$$

which is the modulus squared of an analytically continued partition function. This quantity starts off at  $t = 0$  as  $|Z(\beta)|^2$  and oscillates erratically around a late-time mean  $Z(2\beta)$ , which is suppressed by  $e^S$ .<sup>37</sup> These high-frequency oscillations have a period on the scale of the energy difference between adjacent levels. This is for a system whose energy levels are sufficiently random, such as a chaotic system, with the hallmark example being random matrix ensembles. For integrable systems on the other hand, conspiracies between energy levels can create regular oscillations and recurrences.

The spectral form factor (4.5) for a chaotic system has a very characteristic shape as a function of time (Cotler et al. 2017). We sketch an example where we draw the spectrum from the Gaussian Unitary Ensemble (GUE) in Fig. 15. We can see an initial downward slope, followed by an erratic rising ramp and erratic oscillations around a final plateau. The proposal of Cotler et al. (2017) is that the black hole spectrum is chaotic and the level spacing statistics is captured by random matrix theory.

In order to see any indications of this underlying discreteness and chaotic behavior of a microscopic holographic system in its bulk gravitational description, we will have to include quantum gravity effects beyond those we have studied up to this point:

<sup>36</sup> This can be achieved either by subtracting the one-point functions explicitly, or by the presence of e.g. a  $\mathbb{Z}_2$ -symmetry mapping  $\mathcal{O} \rightarrow -\mathcal{O}$ .

<sup>37</sup> Poincaré recurrences are expected at much later times  $e^{e^S}$ , and will not play a role in this review.

gravitational configurations whose topology is different from the disk, i.e. spacetime wormholes. We first turn to describing how one computes such amplitudes gravitationally in JT gravity. We will come back to the spectral form factor and signs of discreteness for JT gravity in Sect. 4.4.2. Some aspects of this connection between wormholes and chaos generalize to black holes in any dimension through the double-cone wormhole of Saad et al. (2018).

## 4.2 Multiboundary higher genus amplitudes

In this section, we review the computation of the gravitational path integral on geometries with spacetime wormholes, and sum over topologies. We consider first the case of pure JT gravity studied by Saad et al. (2019, 2018).

As discussed at length in Sects. 2 and 3, computing the partition function using the gravitational path integral involves considering the situation with a Euclidean geometry with a boundary curve (in 2d) such that the proper length  $L$  and the boundary dilaton  $\Phi_\partial$  both diverge at a constant rate  $L, \Phi_\partial \rightarrow \infty$  with  $L/\Phi_\partial = \beta/C$ . The parameter  $C$  is fixed once and for all and sets the units in which the inverse temperature  $\beta$  is measured, see Sect. 2 for more details.

It will be convenient for the remainder of this section to generalize this set-up to the case of  $n$  boundaries, each described by their own inverse temperatures  $\beta_1, \dots, \beta_n$ . We will denote the gravity path integral with  $n$  boundaries by  $Z_{\text{grav}}(\beta_1, \dots, \beta_n)$ . The aim of this section is to analyze whether

$$Z_{\text{grav}}(\beta_1, \dots, \beta_n) \stackrel{?}{=} \text{Tr}_{\mathcal{H}_{\text{BH}}} \left( e^{-\beta_1 H} \right) \dots \text{Tr}_{\mathcal{H}_{\text{BH}}} \left( e^{-\beta_n H} \right) \quad (4.6)$$

is true, where  $H$  is a Hamiltonian acting on the putative black hole Hilbert space  $\mathcal{H}_{\text{BH}}$ . The purpose of this subsection is to compute the left hand side of this expression, while the relation on the right hand side will be analyzed in the next subsection.

To start the analysis, we repeat here the Euclidean action of pure JT gravity including the topological term:

$$I = -\frac{S_0}{4\pi} \left[ \int_{\mathcal{M}} \sqrt{g} R + 2 \oint_{\partial\mathcal{M}} \sqrt{h} K \right] - \frac{1}{16\pi G_N} \left[ \int_{\mathcal{M}} \sqrt{g} \Phi (R + 2) + 2 \oint_{\partial\mathcal{M}} \sqrt{h} \Phi (K - 1) \right], \quad (4.7)$$

where  $g$  is the metric on the two-dimensional space  $\mathcal{M}$  and  $h$  the induced metric on its boundary  $\partial\mathcal{M}$ . We focus on a connected surface here. Each time we add a handle to the Euclidean geometry, we increase the genus  $g$  of the surface. These handles are often called spacetime wormholes in this context.<sup>38</sup> Since the first line of the action  $I$  is proportional to the Euler characteristic  $\chi = 2 - 2g - n$ , with  $g$  the number of handles and  $n$  the number of boundaries, this allows us to expand the full path integral

<sup>38</sup> This nomenclature is to distinguish them from purely spatial wormholes that live at a fixed timeslice, such as the Einstein–Rosen bridge of an eternal black hole. Spacetime wormholes on the other hand extend in the full (Euclidean) spacetime. See e.g. Kundu (2022) for a recent review.

as a genus expansion

$$Z_{\text{grav,conn}}(\beta_1, \dots, \beta_n) = \sum_{g=0}^{\infty} e^{S_0(2-2g-n)} Z_{g,n}(\beta_1, \dots, \beta_n). \quad (4.8)$$

Notice the suppressing prefactor at higher genus  $g$  for fixed  $n$ . In this expression  $Z_{g,n}$  denotes the contribution from geometries with fixed topology. For convenience we focus solely on connected contributions, which are the building blocks for the full answers.

Before directly computing  $Z_{g,n}$ , it will be useful first to consider the case of a surface of genus  $g$  with  $n$  geodesic boundaries of lengths  $b_i$ ,  $i = 1 \dots n$ , and no holographic (or Schwarzian) boundaries. A geodesic boundary by definition has zero extrinsic curvature trace:  $K = 0$ . The JT path integral on such a surface becomes

$$\int \frac{[Dg_{\mu\nu}][D\Phi]}{V(\text{Diff})} e^{-I_{\text{JT}}} = \int \frac{[Dg_{\mu\nu}]}{V(\text{Diff})} \delta(R+2), \quad (4.9)$$

where no boundary actions appear. Here we mod out by bulk diffeomorphisms. The RHS is suggestive: the JT path integral on such a manifold is a volume integral (since the integrand is “1”) over all hyperbolic ( $R = -2$ ) metrics mod diffeomorphisms on the given surface. This is the volume of the moduli space of Riemann surfaces. We denote this moduli space as  $\mathcal{M}_{g,n}(\vec{b})$ . The real dimension of this moduli space is  $\dim(\mathcal{M}_{g,n}(\vec{b})) = 6g - 6 + 2n$ .<sup>39</sup>

The measure in the integral over  $\mathcal{M}_{g,n}(\vec{b})$  requires a calculation of the one-loop determinant arising after integrating out the dilaton and restricting to hyperbolic metrics. Insightful discussions can be found in Sect. 3.2 of Saad et al. (2019) in a perturbative string language, and in Sect. 3 of Stanford and Witten (2019), using the relation between analytic and combinatoric torsion. The upshot is that the correct measure over  $\mathcal{M}_{g,n}(\vec{b})$  that appears in JT gravity is the one induced from the Weil–Petersson symplectic form as follows. We can decompose any hyperbolic surface with genus  $g$  into a set of  $2g - 2 + n$  pair-of-pants with  $3g - 3 + n$  connecting tubes with lengths and twists  $(\tilde{b}_i, \tau_i)$  for  $i = 1, \dots, 3g - 3 + n$ . The Weil–Petersson form in these length-twist or Fenchel–Nielsen coordinates, is

$$\omega_{g,n} = \sum_{i=1}^{3g-3+n} d\tilde{b}_i \wedge d\tau_i. \quad (4.10)$$

<sup>39</sup> This is explained above equation (4.10) but another way to derive this is as follows. We start with the two-sphere  $S^2$ . Increasing the genus by one can be done by cutting out two disks from a surface and gluing them together. This requires the specification of four real moduli for the locations of both disks, one real moduli for their relative size, and one real moduli for their relative angle. This gives  $6g$  real moduli overall. The  $-6$  comes from the  $SL(2, \mathbb{C})$  global conformal transformations on the starting  $S^2$ , producing the same surface. Finally, a boundary has three real moduli: two for its location, and one for its length. Here we keep the length fixed, so we are left with two real moduli per geodesic boundary. Note that one can get a puncture if we let  $b_i \rightarrow 0$ .

The volume form extracted from this Weil–Petersson form is  $\frac{\omega_{g,n}^{3g-3+n}}{(3g-3+n)!}$ . Even though this measure is very simple, it is not easy to determine the region of integration over the variables  $(\bar{b}_i, \tau_i)$  when modding by the group of so-called large diffeomorphisms, i.e. the mapping class group. These are by definition the set of diffeomorphisms that are not continuously connected to the identity. We have to do this to make sure geometries are not overcounted. We will come back to this later.

Now we are ready to present the computation of  $Z_{g,n}$ .

#### 4.2.1 The disk $Z_{0,1}(\beta)$

There are two cases we need to work out separately. The first is the disk with  $g = 0$  and  $n = 1$  (and hence  $\chi = 1$ ), which we already computed in Sect. 3.1.1:

$$Z_{0,1}(\beta) = Z_{\text{Sch,disk}}(\beta) = \frac{C^{3/2}}{(2\pi)^{1/2} \beta^{3/2}} e^{\frac{2\pi^2 C}{\beta}}. \quad (4.11)$$

This gives a contribution of order  $e^{S_0}$  to the partition function, since  $\chi = 1$ . We can use this to extract the leading order density of states:

$$\rho_0(E) = \frac{C}{2\pi^2} \sinh\left(2\pi\sqrt{2CE}\right). \quad (4.12)$$

where we have defined the  $g = 0$  density  $\rho_0(E)$  with the overall factor  $e^{S_0}$  removed (compare with (3.10)). When summing over topologies this is only the first term in an  $e^{-2S_0}$  expansion.

#### 4.2.2 The cylinder $Z_{0,2}(\beta_1, \beta_2)$

The second case we need to consider separately is the genus zero two-boundary amplitude. The quantity  $Z_{\text{grav}}(\beta_1, \beta_2)$  gets two contributions. The first comes from two disconnected disks  $e^{2S_0} Z_{0,1}(\beta_1) Z_{0,1}(\beta_2)$ . More interesting is the contribution from the connected geometry without handles, a cylinder (or annulus) between the two boundaries,  $Z_{0,2}(\beta_1, \beta_2)$ . Its Euler characteristic is  $\chi = 0$ , and for large  $S_0$  it is at first sight strongly suppressed compared to the disconnected two disks. This conclusion, however, is too quick and can be invalidated in certain parametric regimes as we will explain later on in Sect. 4.4.2. This geometry is so important that it got a new name: the double trumpet.

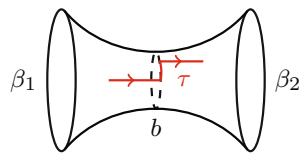
To compute this quantity we first integrate out the dilaton, imposing that the metric is hyperbolic. We can write down the two-parameter family of hyperbolic metrics on the double trumpet as:

$$ds^2 = dr^2 + \cosh^2 r [bdx + \tau\delta(r)dr]^2, \quad x \sim x + 1. \quad (4.13)$$

where  $b > 0$  and  $0 < \tau < b$ . For this metric,  $r$  is the radial distance and the boundaries are at large negative  $r$  (left boundary) or large positive  $r$  (right boundary).

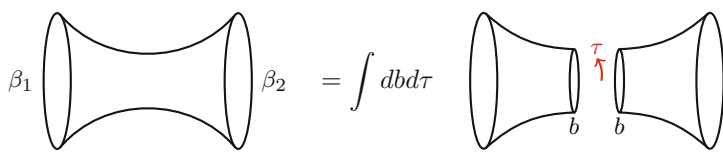
The precise radial cutoff is picked to match the boundary conditions. The  $x$ -coordinate is discontinuous; one can transfer to a continuous coordinate  $y$  defined as  $dy = bdx + \tau\delta(r)dr$ .

In the center at  $r = 0$  there is a geodesic with length  $b$ . Finally the parameter  $\tau$  denotes a possible twist when crossing the middle geodesic. Given  $(b, \tau)$ , the geometry can be depicted as:



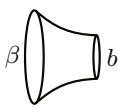
(4.14)

We first compute the path integral on this cylinder with fixed  $(b, \tau)$ , and afterward integrate over these parameters. This can be done by separating the cylinder into two single trumpets: one for  $r < 0$  and another for  $r > 0$ . Each trumpet is a hyperbolic cylinder bounding a holographic boundary and a geodesic. This procedure can be depicted as



(4.15)

On each single trumpet, the path integral is given by the contribution of the Schwarzian mode on the holographic boundary.<sup>40</sup> The boundary mode of the single trumpet is not the same as the disk, it corresponds to an integral over the hyperbolic orbit  $\text{Diff}(S^1)/U(1)$ . This orbit is labeled by a continuous parameter  $b$ . We quote the result (3.72) from Sect. 3:



$$Z_T(\beta, b) = \frac{\sqrt{C}}{\sqrt{2\pi}\sqrt{\beta}} e^{-\frac{Cb^2}{2\beta}}.$$

(4.16)

The final step is to glue the two trumpets into the complete cylinder, which requires knowing the correct measure on the moduli space of hyperbolic surfaces as derived earlier. Locality guarantees that the same measure and volume form (4.10) should be used here when gluing along the pair of geodesics: the volume form is thus  $dbd\tau$ . Since the integrand is independent of the twist, we can integrate over it producing a

<sup>40</sup> On the inner geodesic boundary there are no dynamical degrees of freedom since the extrinsic curvature is being fixed to zero.



factor of  $\int_0^b d\tau = b$ , giving a total measure of  $bdb$ .<sup>41</sup> The final answer is:

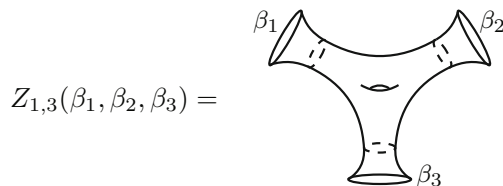
$$Z_{0,2}(\beta_1, \beta_2) = \int_0^\infty bdb Z_T(\beta_1, b) Z_T(\beta_2, b) = \frac{1}{2\pi} \frac{\sqrt{\beta_1 \beta_2}}{\beta_1 + \beta_2}. \quad (4.17)$$

We note that this is a completely off-shell calculation: there is no saddle point value determining  $b$  and stabilizing the wormhole, as can be seen from (4.16) where the exponential factor pushes  $b$  to zero.<sup>42</sup> A general argument for this lack of classical solution can be found in Maldacena et al. (1999) and Qi et al. (2019). This is not necessarily a problem, but it does complicate things when trying to generalize to higher-dimensional physics, where we do not have an exact quantum solution. Related to this, we make two further comments:

- Fixing  $b$  does allow for a saddle to exist. This saddle is hence the solution of a constrained extremization problem, and is called a constrained instanton. With this constraint, wormhole saddle solutions have been found explicitly in higher dimensions (Cotler and Jensen 2021, 2022).
- Analytically continuing  $\beta \rightarrow \beta + it$ , we can find a saddle at  $\beta = 0$ , a genuine saddle if we also transfer to a microcanonical version of Eq. (4.17) (Saad et al. 2018). This geometry has been constructed explicitly and looks like a double cone of two cones joined at the tip, and where Lorentzian time is periodically identified with period  $t$ . Stability of this saddle has been investigated in Mahajan et al. (2021). This wormhole solution can actually be written for black holes in any number of dimensions.

#### 4.2.3 The general case $Z_{g,n}(\beta_1, \dots, \beta_n)$

We can now describe a formula valid for all other  $Z_{g,n}$  as long as  $(g, n) \neq (0, 1), (0, 2)$ . To derive it, first notice that in hyperbolic geometry we can always find geodesics homologous to each holographic boundary, such that these  $n$  geodesics bound an inner hyperbolic surface in the bulk. For example, for three boundaries at genus one:



<sup>41</sup> If we utilize Teichmüller space instead of  $\mathcal{M}_{g,n}(\vec{b})$ , the integral over twists would give a divergent but  $b$ -independent factor. Teichmüller space is the universal cover of  $\mathcal{M}_{g,n}(\vec{b})$ , where we “undo” the modding by the mapping class group of large diffeomorphisms. Another way of representing Teichmüller space is by adding a Moore-Seiberg graph to a given Riemann surface. Using the local gluing procedure in the BF gauge theoretical TQFT framework, leads precisely to the Teichmüller space gluing integrals (Blommaert et al. 2019b).

<sup>42</sup> Although one could think of  $b = 0$  as a saddle on the boundary of the moduli space (Witten 2021).

Each of the three boundaries has an associated geodesic, drawn as a dashed line, bounding a single trumpet. The interior remainder is a hyperbolic surface with  $g$  handles (in the example just one) bounded by geodesics. The hyperbolic surface now is described by two types of moduli: the one associated to the inner surface  $\mathcal{M}_{g,n}(\vec{b})$ , and  $(b_1, \tau_1, \dots, b_n, \tau_n)$  that determine how the “outer” geodesics are glued to the boundary trumpets. Using this decomposition, after integrating over the boundary twists  $\tau_1, \dots, \tau_n$ , we find the following expression (Saad et al. 2019):

$$Z_{g,n}(\beta_1, \dots, \beta_n) = \int_0^\infty \left[ \prod_{i=1}^n b_i db_i Z_T(\beta_i, b_i) \right] V_{g,n}(b_1, \dots, b_n). \quad (4.18)$$

The integrals on the right hand side are over the lengths of the geodesics homologous to each boundary and include the single trumpet contributions. The last factor  $V_{g,n}(b_1, \dots, b_n)$  is the volume of the interior moduli space  $V_{g,n}(b_1, \dots, b_n) = \text{Vol}(\mathcal{M}_{g,n}(\vec{b}))$ , or the Weil–Petersson volumes.

The last ingredient we need is to provide a way to compute efficiently these Weil–Petersson volumes, so that the general expression for  $Z_{g,n}$  can become useful. There are two approaches to this problem, both developed by Mirzakhani (2007a, b). The first consists in using the following identity:

$$V_{g,n}(b_1, \dots, b_n) = \int_{\overline{\mathcal{M}}_{g,n}} e^{\omega_{g,n} + \frac{1}{2} \sum_i b_i^2 \psi_i}, \quad (4.19)$$

where  $\overline{\mathcal{M}}_{g,n}$  is the Deligne–Mumford compactification of the moduli space of hyperbolic surfaces of genus  $g$  and  $n$  punctures (which corresponds to a geodesic boundary with  $b \rightarrow 0$ ) and  $\omega_{g,n}$  is the Weil–Petersson form on the moduli space of genus  $g$  and  $n$  punctures.  $\psi_i$  is defined as follows: let  $\mathcal{L}_i$  be the cotangent space to the  $i$ th puncture. As we move along the moduli space,  $\mathcal{L}_i$  varies as the fiber of a complex line bundle over moduli space. The psi-class is the first Chern class of this line bundle  $\psi_i = c_1(\mathcal{L}_i)$ . A simple recursion formula to compute this type of integrals over psi-classes and Weil–Petersson form (or more generally kappa-classes) was conjectured by Witten (1991b) and proven by Kontsevich (1992). This point of view was exploited in the context of JT gravity mainly by Okuyama and Sakai (2020a, b).

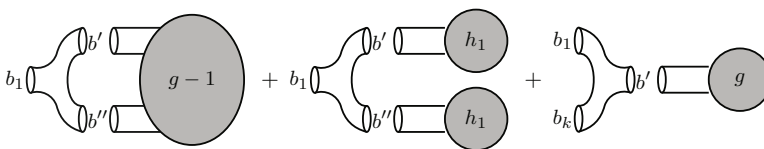
Instead, we will follow the approach originally taken by SSS which has a more geometric origin. First of all, even though the volume form is very simple, taking into account the mapping class group to determine the region of integration in the  $(\vec{b}, \vec{\tau})$  space is a very complicated problem. In Mirzakhani (2007a) this was circumvented by using the following trick.<sup>43</sup> First we pick a (geodesic) boundary, labeled as 1 below, and start shooting geodesics perpendicular to the boundary, along the entire circumference of the boundary. There are three options: (A) the curve intersects itself at some point or returns to the same geodesic boundary, and we truncate it as soon as either happens; (B) the geodesic reaches a different boundary geodesic; (C) the geodesic continues forever

<sup>43</sup> A useful presentation of the argument aimed at physicists can be found in Appendix D of (Stanford and Witten 2019).

in the interior without self-intersection and without reaching any boundary. Now we can write the length of the geodesic  $b_1$  as a sum over the length of segments where geodesics of type A, B and C start. It was proven by Birman and Series that the set C has measure zero so we can ignore its contribution to the length. When the hyperbolic surface is decorated by a geodesic of either type A or B, we can single out a pair of pants (i.e. a three-holed sphere surface) that in case A bounds the geodesic 1 and two internal ones; or in case B the geodesic 1, another boundary geodesic and an internal one. This decomposes the surface of genus  $g$  and  $n$  boundaries into either one or two surfaces with either lower genus or a smaller number of boundaries. Using the fact that there is only a single hyperbolic surface of genus zero with three given geodesic lengths:  $V_{0,3}(b_1, b_2, b_3) = 1$ , and a short calculation in hyperbolic geometry reviewed in Stanford and Witten (2019), one arrives at Mirzakhani's recursion relation:

$$\begin{aligned} b_1 V_{g,n}(b_1, B) &= \frac{1}{2} \int_0^\infty b' db' b'' db'' D(b_1, b', b'') V_{g-1}(b', b'', B), \\ &+ \frac{1}{2} \int_0^\infty b' db' b'' db'' D(b_1, b', b'') \sum_{\text{stable}} V_{h_1}(b', B_1) V_{h_2}(b'', B_2), \\ &+ \sum_{k=2}^n \int_0^\infty b' db' (b_1 - T(b_1, b', b_k)) V_g(b', B/b_k), \end{aligned} \quad (4.20)$$

where  $B = \{b_2, \dots, b_n\}$ , and we defined  $T(b_1, b_2, b_3) = \log \frac{\cosh \frac{b_3}{2} + \cosh \frac{b_1+b_2}{2}}{\cosh \frac{b_2}{2} + \cosh \frac{b_1-b_2}{2}}$  and  $D(b_1, b_2, b_3) = b_1 - T(b_1, b_2, b_3) - T(b_1, b_3, b_2)$ . These functions arise from computing within each pair of pants the length along  $b_1$  of segments of types A or type B in hyperbolic geometry. The three lines in (4.20) correspond graphically to the decomposition of surfaces:



where the grey blobs denote the remainder hyperbolic Riemann surface in this decomposition. The first line in (4.20) corresponds to case A, cutting the pair of pants involves only one boundary and does not divide the surface, but lowers the genus to  $g - 1$ . The second line in (4.20) corresponds also to case A, when removing the pair of pants divides the surface in two with genus  $h_1$  and  $h_2$  and boundaries  $(b', B_1)$  and  $(b'', B_2)$ , such that  $h_1 + h_2 = g$  and  $B_1 \cup B_2 = B$ .<sup>44</sup> Finally, the third line in (4.20) corresponds to case B where the pair of pants to be removed involves two boundaries  $b_1$  and  $b_k$ . Then  $B/b_k$  denotes the set  $B$  with the element  $b_k$  removed.

<sup>44</sup> The sum is restricted to only stable surfaces. This means surfaces on which a hyperbolic metric exists; this excludes for example genus zero with less than three boundaries or genus one with no boundary.

In spite of not having an explicit closed expression for the Weil–Petersson volumes,<sup>45</sup> one can immediately derive the following important property:  $V_{g,n}(b_1, \dots, b_n)$  is a symmetric multivariate polynomial of degree  $3g - 3 + n$  in the  $b_i^2$ . This is most easily seen from (4.19).

Using the recursion relation above, one can compute any Weil–Petersson volume and combine it with the exterior trumpets to compute any desired  $Z_{g,n}$ . This completes the exact gravitational solution of JT gravity in a topological (or genus) expansion, i.e. by summing over all spacetime wormhole configurations.

As a final comment, going back to the intersection numbers approach of Eq. (4.19) and inserting that directly in (4.18), one can obtain a formula for  $Z_{g,n}$  as an integral over moduli space with punctures:

$$Z_{g,n}(\beta_1, \dots, \beta_n) = \frac{\sqrt{\beta_1 \dots \beta_n}}{(2\pi C)^{n/2}} \int_{\mathcal{M}_{g,n}} \frac{e^{\omega_{g,n}}}{\prod_{i=1}^n (1 - \frac{\beta_i}{C} \psi_i)}. \quad (4.21)$$

This expression can be understood as an application of localization to JT gravity, as explained in Eberhardt (2022). This approach generalizes nicely to chiral gravity in three dimensions.

To illustrate this rather abstract treatment, we will work out explicitly the example of the one-wormhole correction to the disk partition function  $Z_{1,1}(\beta)$ .

**Example:**  $Z_{1,1}(\beta)$ :

The only option to decompose the surface with one boundary and one handle is:

$$b \times \text{(disk with handle)} = \frac{1}{2} \int_0^\infty b' db' D(b, b', b') \times \text{(pair of pants with identified boundaries)},$$

where in the right hand side we have a single pair of pants with the two geodesics of length  $b'$  identified (indicated by the dashed line). The equation for the volume becomes

$$V_{1,1}(b) = \frac{1}{2b} \int_0^\infty b' db' D(b, b', b') = \frac{b^2 + 4\pi^2}{48}. \quad (4.22)$$

This integral is actually divergent. The standard way of dealing with this is to compute instead  $\partial_b(bV_{1,1}(b))$ , which involves  $\partial_b D(b, b', b')$  for which the integral converges. After integrating the result over  $b$ , the integration constant can be fixed by demanding  $V_{1,1}(b)$  has a smooth  $b \rightarrow 0$  limit.

<sup>45</sup> An exception is the case  $g = 0$ . An explicit formula for  $V_{0,n}(b_1, \dots, b_n)$  was recently derived in Mertens and Turiaci (2021), equation (7.54).

This case does not quite work the same as in the generic case of Eq. (4.20). Firstly, there is no volume factor appearing on the right hand side of the recursion and also one less integral since when gluing two ends of a single pair of pants those lengths are identified. Note that if we define  $V_{0,2}(b_1, b_2) = \frac{1}{b_1} \delta(b_1 - b_2)$  then this could be written as a special case of Eq. (4.20). This case is special for a second reason: the torus with one hole has a  $\mathbb{Z}_2$  symmetry and the volume given above is taking this symmetry into account.

Using (4.22) we can plug it in (4.18) and obtain the one-wormhole correction to the partition function

$$Z_{1,1}(\beta) = \text{[Diagram of a torus with one hole]} = \int_0^\infty b db Z_T(\beta, b) V_{1,1}(b) = \frac{\beta^{3/2}}{24\sqrt{2\pi}C^{3/2}} + \frac{\pi^{3/2}\sqrt{\beta}}{12\sqrt{2}\sqrt{C}}. \quad (4.23)$$

—o—

The above example already displays some interesting and generic features. Firstly, the amplitude with one wormhole is a polynomial in  $\sqrt{\beta}$ . This is in contrast to the disk which has an essential singularity at  $\sqrt{\beta} = 0$ . The lack of an exponential term in  $Z_{1,1}(\beta)$  is related to the fact that there is no classical solution of a disk with one wormhole. Much like the double trumpet, all of these configurations have no saddle point in the JT path integral. The reason was already pointed out around (2.41): the equation of motion for the metric leads to a Killing vector which does not exist on higher genus hyperbolic surfaces. Secondly, this correction is of order one and therefore it is suppressed by  $e^{-S_0}$  compared to the disk with no wormhole. Nevertheless this is too naive. Adding the disk and one-wormhole contributions, we find schematically for small temperatures  $Z \sim e^{S_0} \beta^{-3/2} + e^{-S_0} \beta^{3/2}$ , and therefore the one-wormhole contribution becomes dominant when  $\beta \sim \mathcal{O}(e^{2S_0/3})$ . However, at such low temperatures we cannot even trust the perturbative expansion in  $e^{-S_0}$  to begin with.<sup>46</sup>

### 4.3 JT and random matrices

In this subsection, we turn towards the connection between JT gravity and random matrix models in the double-scaling limit. This connection can be seen in the intersection number approach to computing Weil–Petersson volumes (4.19), since those can be obtained in a double-scaled matrix model through the Witten–Kontsevich theorem (Witten 1991b; Kontsevich 1992). A discussion along these lines can be found in Dijkgraaf and Witten (2018). Instead we will follow SSS (Saad et al. 2019), and use

<sup>46</sup> In general, the large  $\beta$  behavior is dominated by the large  $b$  behavior of the volumes. When one of the boundaries, say the first, has a very large length  $b_1 \gg 1$  one can show  $V_{g,n}(\vec{b}) \sim \frac{b_1^{6g-6+2n}}{(24)^g g! 2^3 g^{-3+n} (3g-3+n)!}$  (see Appendix A of Maxfield and Turiaci 2021) and therefore in the case of one boundary  $Z_{g,1}(\beta) \sim \frac{C^{3/2}}{\sqrt{2\pi}\beta^{3/2}} \frac{(\beta/C)^3 g}{(24)^g g!}$  at large  $\beta$ . We see that at higher genus, the divergence at low temperatures becomes stronger. This can be interpreted as a perturbative  $e^{-S_0}$  correction to the zero-point energy.

Mirzakhani's recursion relation and compare it with the loop equations of the matrix model.

### 4.3.1 SSS duality

We come back to the original question of whether a relation such as (4.6) is true in pure JT gravity. From the calculations done in the previous section, it is immediately clear that such a relation cannot be correct since, (1) there is no indication (in perturbation theory at least) that the density of states is discrete in any way which is in tension with (finite-volume) holography, and (2) the connected two-boundary path integral due to the connected wormhole amplitude does not factorize:

$$Z_{\text{grav}}(\beta_1, \beta_2) \neq Z_{\text{grav}}(\beta_1)Z_{\text{grav}}(\beta_2). \quad (4.24)$$

This is in tension with (4.6), which would predict a factorized answer for a given  $H$  and  $\mathcal{H}_{\text{BH}}$ . This is the factorization puzzle for holography raised first by Witten and Yau (1999). Both problems can be resolved by relaxing (4.6) in the following way. If pure JT gravity is not dual to a single quantum system with a discrete spectrum, can it be dual to an ensemble average of quantum systems? The results of SSS give a positive answer to this question. This resolves the tension (1) since the continuous spectrum can arise after averaging, and also resolves tension (2) since the non-factorization is due to the statistics of the ensemble.

We now describe the holographic dual of JT gravity. Take the black hole Hilbert space  $\mathcal{H}_{\text{BH}}$  to be of dimension  $L$ , with  $L$  a large integer which will be later related to  $e^{S_0}$ . A quantum theory is described by an  $L \times L$  Hermitian Hamiltonian matrix  $H$ , acting on  $\mathcal{H}_{\text{BH}}$ . The JT gravitational path integral is then equal to an ensemble average over theories, in this case an integral over all  $L \times L \sim e^{S_0} \times e^{S_0}$  Hermitian matrices:

$$Z_{\text{grav}}(\beta_1, \dots, \beta_n) = \int dH P(H) \text{Tr} \left( e^{-\beta_1 H} \right) \dots \text{Tr} \left( e^{-\beta_n H} \right), \quad (4.25)$$

where  $P(H) \propto e^{-L\text{Tr}V(H)}$  is the probability distribution over theory space (in this case over the Hamiltonian matrix itself), and it is normalized such that  $\int dH P(H) = 1$ . The quantity  $V(H)$  is the potential, which is taken to be a sum (possibly infinite) of positive powers of  $H$ . Hence the claim is that JT gravity is a matrix integral (Saad et al. 2019)!

Matrix models and techniques to solve them have a long history that we will not do justice here. We refer the reader to Mehta (2004), and to Eynard et al. (2015) for a pedagogical review of more recent techniques. Shorter accounts can also be found in reviews on Liouville gravity and non-critical string models (Ginsparg and Moore 1993; Di Francesco et al. 1995; Nakayama 2004), a topic that also has a deep link with matrix models, which turns out not to be a coincidence as we will point out later on in Sect. 5.10.

Let us continue with our main story. On the left hand side of Eq. (4.25), the JT gravity path integral can be expanded as a sum over topologies, so the first question is: what is the meaning of this expansion on the matrix model side?

Matrix integrals have natural insertions given by the resolvent  $R(E)$ , the “partition sum”  $Z(\beta)$  and the spectral density  $\rho(E)$ :

$$R(E) \equiv \text{Tr} \frac{1}{E - H}, \quad Z(\beta) \equiv \text{Tr} \left( e^{-\beta H} \right), \quad \rho(E) \equiv \sum_{m=1}^L \delta(E - E_m) \quad (4.26)$$

where we denoted the eigenvalues of  $H$  as  $E_m$ . There are simple relations between these different quantities as:

$$\int_0^\infty e^{\beta E} Z(\beta) = R(E), \quad Z(\beta) = \int dE \rho(E) e^{-\beta E}, \quad (4.27)$$

$$R(E + i\epsilon) - R(E - i\epsilon) = -2\pi i \rho(E), \quad (4.28)$$

so it suffices to study one of them. We then define correlation functions on the matrix integral side with  $n$  of these insertions:

$$R_{\text{MM}}(E_1, \dots, E_n) = \langle R(E_1) \dots R(E_n) \rangle, \quad (4.29)$$

$$Z_{\text{MM}}(\beta_1, \dots, \beta_n) = \langle Z(\beta_1) \dots Z(\beta_n) \rangle, \quad (4.30)$$

$$\rho_{\text{MM}}(E_1, \dots, E_n) = \langle \rho(E_1) \dots \rho(E_n) \rangle, \quad (4.31)$$

where  $\langle X \rangle$  denotes the ensemble average  $\langle X \rangle = \int dH P(H) X$ . Correlation functions of resolvents in the matrix integral are known to have a  $1/L$  expansion in the large  $L$  limit, of the form

$$R_{\text{MM}}(E_1, \dots, E_n)_{\text{conn.}} \simeq \sum_{g=0}^{\infty} \frac{R_{g,n}(E_1, \dots, E_n)}{L^{2g+n-2}}, \quad (4.32)$$

for resolvents and similarly for partition functions. The left hand side of this equation denotes the connected piece only of the  $n$ -resolvent correlator. The series on the right hand side is only asymptotic and there are non-perturbative corrections, but we can take it as a definition of the coefficients  $R_{g,n}$ . This expansion arises from a perturbation theory in 't Hooft double line diagrams ('t Hooft 1974; Brezin et al. 1978). In this context,  $n$  is the number of insertions while  $g$  is the genus of the surface needed to embed the particular double-line diagram into. This means the large  $L$  expansion of the matrix model should be identified with the large  $e^{S_0}$  expansion of JT gravity, and the coefficients should match term by term.

To make the identification (4.25) precise, hence requires a precise relation between  $L$  and  $S_0$ , and the probability distribution  $P(H)$  through the matrix potential  $V(H)$ . We address these two points in the next subsections and complete the derivation of Eq. (4.25).

### 4.3.2 Double-scaling limit of matrix integrals

Consider the spectral density of the matrix model:  $\rho(E) = \sum_{i=1}^L \delta(E - E_i)$ . Even if each realization for  $H$  has a discrete spectrum for finite  $L$ , after averaging over  $H$  and taking the large  $L$  limit, the resulting density of states will be continuous.

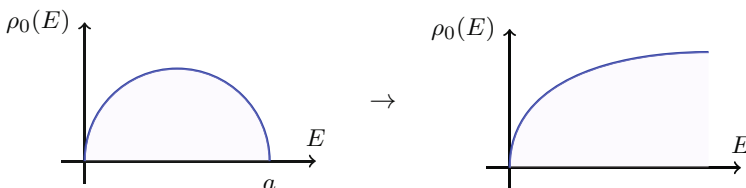
In the simplest models, the leading density of states at large  $L$  has support on a single interval  $E \in [a_-, a_+]$  with  $a_-$  and  $a_+ > a_-$  real numbers.<sup>47</sup> In these one-cut cases, there is a simple relation between the function  $V(H)$  and  $\rho_0(E)$ , see Eynard et al. (2015) for details. The prototypical example where this applies is a Gaussian model with leading density of states

$$V(H) = \frac{8}{a^2} \left(H - \frac{a}{2}\right)^2 \quad \rightarrow \quad \rho_0(E) \equiv \langle \rho(E) \rangle_{L \rightarrow \infty} = \frac{8L}{\pi a^{3/2}} \sqrt{\frac{E(a-E)}{a}}, \quad (4.33)$$

where the brackets mean we average over  $H$  with the measure determined by  $V(H)$ . This is the famous Wigner semicircle distribution. We chose the potential such that  $a_- = 0$  and  $a_+ = a > 0$  and the overall normalization of the density of states is such that its integral gives  $L$ .

To apply this to the JT gravity case, we should find a potential  $V(H)$  such that its density of states to leading order at large  $L$  matches with JT gravity on the disk. This is not possible, since JT gravity has an infinite support over  $E \in [0, +\infty)$ . The resolution of this puzzle is that JT gravity is dual not to an ordinary matrix integral, but to a double-scaled matrix integral.

Let us see first how the double scaling limit works in the case of the Gaussian matrix model (4.33). We wish to obtain a distribution that has support on the entire positive real axis. So we should take the limit of  $a \rightarrow \infty$ . This runs into an issue: since we have a finite number of eigenvalues  $L$  to be distributed on the whole real line, the spectral density  $\rho_0(E)$  goes to zero. To resolve this, we need to send the number of eigenvalues to infinity  $L \rightarrow \infty$  at the same time, such that we keep the following quantity  $8L/\pi a^{3/2} \equiv e^{S_0}$  fixed. This gives the double-scaled density of states  $\rho_0(E) \approx e^{S_0} \sqrt{E}$ , which is smooth and has support on the positive energy axis, although it is not normalized: the number of eigenvalues is infinite. Much like the original Gaussian matrix model (4.33), the resulting double-scaled matrix model is exactly solvable and is called the Airy model. Schematically, the operation we realize on the leading density of states is



<sup>47</sup> More general models can have several disjoint cuts.



This operation can be translated as due to a particular fine-tuning of the matrix potential.

This double-scaling limit affects the topological expansion of the matrix integral in a simple way. Instead of weighing different topologies by  $L$ , now the expansion is in the new parameter  $e^{S_0}$  which is large but fixed as  $L, a \rightarrow \infty$ . In terms of resolvents, for example, we can write:

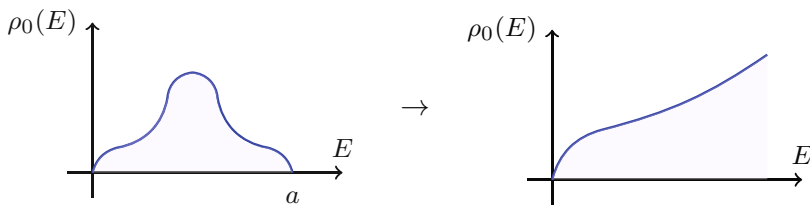
$$R_{\text{MM}}(E_1, \dots, E_n)_{\text{conn.}} \simeq \sum_{g=0}^{\infty} \frac{R_{g,n}(E_1, \dots, E_n)}{(e^{S_0})^{2g+n-2}}. \quad (4.34)$$

It requires a separate calculation to justify that once  $L$  is scaled in the right way as  $a \rightarrow \infty$ , with  $E$  fixed, all the terms in the topological expansion have a well-defined limit (4.34). The quantity  $e^{S_0}$  controls the scale of the eigenvalue density, and  $e^{-S_0}$  hence controls the average eigenvalue spacing.

We now have the tools to describe the holographic dual of JT gravity. Since standard matrix integrals in the large  $L$  limit only have finite support, we begin by “regularizing” the disk density of states, and constructing a matrix potential  $V(H)$ , such that to leading order in  $1/L$  e.g.

$$\rho_0(E) \equiv \langle \rho(E) \rangle_{L \rightarrow \infty} = e^{S_0} \frac{C}{2\pi^2} \sinh \left( 2\pi \sqrt{2CE \frac{a-E}{a}} \right), \quad (4.35)$$

where  $e^{S_0}$  is a function of  $L$  and  $a$  that can be determined by fixing normalization. We do not need to know this relation explicitly, since we are next going to take  $a \rightarrow \infty$  and  $L \rightarrow \infty$ , such that  $e^{S_0}$  is fixed. Note that there are many ways to “regularize” the JT disk density of states. Schematically, this gives us:



The topological expansion for  $Z_{\text{MM}}(\beta_1, \dots, \beta_n)$  is now in terms of  $e^{S_0}$  instead of  $L$ . It is this double-scaled parameter  $S_0$  that we identify with the prefactor of the topological term in the JT gravity action (4.7), that weighs different topologies in gravity by their Euler characteristic  $\chi$ .

The non-trivial statement we will describe next is that, once the matrix potential is tuned to give the disk JT gravity density of states in the large  $e^{S_0}$  limit, all other subleading terms in the topological expansion will automatically match between the matrix model and the gravity calculation we described in Sect. 4.2.

### 4.3.3 Derivation of the duality

A very direct way to derive the equivalence between JT gravity and a double-scaled matrix integral is through the so-called loop equations or topological recursion relations of the matrix model, which provide a recursive relation that determines all  $R_{g,n}$  or equivalently  $Z_{g,n}$ , in the  $1/L$  or  $e^{-S_0}$  expansion. The loop equations themselves were studied a long time ago (Migdal 1975), but they were put in its most useful form by Eynard (2004). The strategy will be to relate this recursion relation to the geometrical one used to compute Weil–Petersson volumes (4.20).

To state the topological recursion relation in the matrix model, it is convenient to introduce the following quantities related to the resolvent correlators:

$$W_{g,n}(z_1, \dots, z_n) \equiv (-2)^n z_1 \dots z_n R_{g,n}(-z_1^2, \dots, -z_n^2)_{\text{conn.}}, \quad E = -z^2. \quad (4.36)$$

Just as in the gravitational genus expansion, the matrix model topological expansion requires previous knowledge of two special cases  $W_{0,1}$  and  $W_{0,2}$ . These are related to the leading  $\langle \rho(E) \rangle_{L \rightarrow \infty}$ , which is determined by the potential, and to the density-density correlator  $\langle \rho(E) \rho(E') \rangle_{L \rightarrow \infty}$ , which is universal (independent of the matrix potential). They are explicitly given by

$$W_{0,1}(z) = 2zy(z), \quad W_{0,2}(z_1, z_2) = \frac{1}{(z_1 - z_2)^2}. \quad (4.37)$$

The function  $y(z) = -i\pi\rho_0(-z^2)$  determines the spectral curve, given by the locus  $(-E(z) = z^2, y(z)) \subset \mathbb{C}^2$ . Once these two quantities are determined, all other corrections to the resolvent correlators are determined by:

$$W_{g,n}(z_1, Z) = \text{Res}_{z \rightarrow 0} \left\{ \frac{1}{z_1^2 - z^2} \frac{1}{4y(z)} \left[ W_{g-1, n+1}(z, -z, Z) + \sum_{\text{stable}} W_{h_1, 1+|Z_1|}(z, Z_1) W_{h_2, 1+|Z_2|}(-z, Z_2) \right] \right\}, \quad (4.38)$$

where  $Z = \{z_2, \dots, z_n\}$ . The sum in the second line is over sets  $Z_{1,2}$  such that  $Z_1 \cup Z_2 = Z$  and  $g = h_1 + h_2$ , and the sum is again over stable curves. This form of the loop equations was derived in Eynard (2004), and is written for the case of a spectrum with support on the positive real line  $[0, +\infty)$ . For more general spectral support  $[a_-, a_+]$ , the expression is the same, but with the residue taken at both endpoints.

We now compare this with JT gravity, for which the spectral curve is chosen as

$$y(z) = \frac{C \sin(2\pi\sqrt{2C}z)}{2\pi}. \quad (4.39)$$

First of all,  $W_{0,1}$  and  $W_{0,2}$  are related to (4.11) and (4.17) respectively by a Laplace transform that turns the resolvents into thermal partition functions. More generally,

$W_{g,n}$  is related to  $Z_{g,n}$  by the  $n$ -fold Laplace transform:

$$W_{g,n}(z_1, \dots, z_n) = \left( \prod_{i=1}^n 2z_i \int_0^\infty d\beta_i e^{-\beta_i z_i^2} \right) Z_{g,n}(\beta_1, \dots, \beta_n). \quad (4.40)$$

Expressing  $Z_{g,n \neq (0,1), (0,2)}$  in terms of the Weil–Petersson volumes in JT gravity (4.18), one finds:

$$W_{g,n}(z_1, \dots, z_n) = (2C)^{n/2} \left( \prod_{i=1}^n \int_0^\infty b_i db_i e^{-\sqrt{2C} b_i z_i} \right) V_{g,n}(b_1, \dots, b_n). \quad (4.41)$$

Since the volumes are polynomials in  $b_i^2$ , these integrals are convergent for  $z \neq 0$ . Now comes the non-trivial part of the proof. Eynard and Orantin (2007) started from Mirzakhani’s recursion relation (4.20) and applied to it the Laplace transform (4.41) to derive a recursion relation for  $W_{g,n}$ . After a lengthy calculation, they obtained precisely the recursion (4.38) with the spectral curve  $y(z) = \frac{C \sin(2\pi \sqrt{2C} z)}{2\pi}$  used in JT gravity. This result completes the derivation that JT gravity is equivalent, order by order in the topological expansion, to a double-scaled matrix integral.

#### Example: $Z_{1,1}(\beta)$ revisited:

We repeat now the example worked out in Sect. 4.2.3 but from the matrix model side. To do this we use the matrix model loop equation in the form (4.38) to obtain the genus one contribution to the resolvent

$$W_{1,1}(z_1) = \text{Res}_{z \rightarrow 0} \left[ \frac{1}{z_1^2 - z^2} \frac{1}{4y(z)} W_{0,2}(z, -z) \right] = \frac{4\pi^2 C z_1^2 + 3}{48\sqrt{2C}^{3/2} z_1^4}. \quad (4.42)$$

The matrix model prediction for the genus one correction to  $Z(\beta)$  is then given by solving for  $Z_{1,1}(\beta)$  in the relation (4.40). The answer is

$$Z_{1,1}(\beta) = \frac{\beta^{3/2}}{24\sqrt{2\pi} C^{3/2}} + \frac{\pi^{3/2} \sqrt{\beta}}{12\sqrt{2} \sqrt{C}}. \quad (4.43)$$

This is precisely the same as the gravitational answer found in Sect. 4.2.3.

—o—

## 4.4 Non-perturbative effects in topological expansion

By the above matching of recursion relations, the SSS duality is proven order-by-order in the topological expansion in powers of  $e^{-S_0}$ . Since in gravity  $S_0 \sim 1/G_N$ , these are non-perturbative effects in  $G_N$ . From the matrix integral side, since the size of

the matrix is  $L \sim e^{S_0}$ , these are perturbative effects in the (double-scaled) 't Hooft expansion.

We know this cannot be the whole story: from either perspective, the genus expansion is divergent since each term grows for large genus as  $(2g)!$ . This indicates the need for doubly non-perturbative effects of order  $e^{-\#e^{S_0}}$ , as originally suggested by Shenker (1990) in the context of the non-critical string. In that context, this issue has a beautiful resolution: these doubly non-perturbative effects can be accounted for by adding D-branes to string theory.

In JT gravity, the interpretation of these doubly non-perturbative corrections is more mysterious. When seen as a limit of the non-critical string (as we review in Sect. 5.10), the spacetime of JT gravity is identified with the string worldsheet. Therefore, D-branes create holes in the JT gravity spacetime. However, as usual in D-brane physics (Polchinski 1994), we need to sum over configurations where the worldsheet (= spacetime here) ends on a single D-brane an arbitrary number of times, so a universe with an arbitrary number of boundaries. This obscures an intuitive bulk picture, but we can still write down the equations. A D-brane insertion in the matrix integral (in the non-critical string of FZZT type) is a determinant insertion:

$$\det(E - H) = e^{\text{Tr Log}(E - H)}. \quad (4.44)$$

Here we wrote it as the exponential of a new object. The insertion  $\text{Tr Log}(E - H) = \int \frac{d\beta}{\beta} e^{\beta E} Z(\beta)$  in the matrix integral can be viewed as an unmarked fixed energy boundary.<sup>48</sup> The exponentiation of this insertion then shows that we consider an arbitrary number of these unmarked boundaries, accounting for their indistinguishability by  $1/n!$  in the Taylor expansion. Computations can be done for such insertions, and we refer to (Saad et al. 2019).

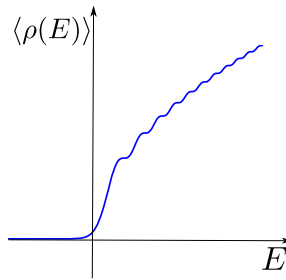
In this subsection, we will review how the doubly non-perturbative effects work from the matrix integral perspective, and leave their interpretation in bulk physics open (see Saad et al. 2019 for some discussion). For concreteness, we will analyze these corrections in the context of computing the black hole density of states and spectral form factor.

#### 4.4.1 Density of states

Consider the exact computation of  $\langle \rho(E) \rangle$  in the random matrix framework for large  $e^{S_0}$ . As previously explained, perturbative corrections are suppressed by powers of  $e^{-S_0}$  and non-perturbative corrections are naively even further suppressed as  $e^{-\#e^{S_0}}$ . However, it is known that counter-intuitively the non-perturbative corrections can actually be more important than the perturbative ones, essentially due to the factor  $\#$  in the exponential being imaginary.

In the large  $S_0$  regime, the full spectral density of the JT matrix model (for  $E \lesssim 1/C$ ) is sketched in Fig. 16. There are several ways to find it. One approach is that of Johnson (2020), which uses numerical techniques. This requires some random matrix

<sup>48</sup> It is unmarked due to the dividing by  $\beta$  as familiar in worldsheet string theory. The resolvent  $R(E)$  is marked since it does not have this  $1/\beta$  present as in (4.27).



**Fig. 16** Sketch of full spectral density of JT gravity including non-perturbative effects

machinery from the '90s that we will not explain here. Another approach, only valid at large  $S_0$ , is to approximate the JT answer using only contributions from topological disks and cylinders.

This “disk and cylinder” approximation was developed first for determinant insertions in Maldacena et al. (2004), but can be translated to density insertions.<sup>49</sup> This approach does not work at small  $E$ , but one can match it there onto the exactly solvable Airy model mentioned earlier.

Let us discuss the qualitative features of the spectral density  $\langle \rho(E) \rangle$ .

**Corrections to forbidden region  $E < 0$**  We begin by analyzing the classically forbidden region  $E < 0$ . It is easy to show that any perturbative corrections will never give a non-zero answer in this region:  $\rho_{g,1}(E < 0) = 0$  for all  $g$ . However, there is a non-perturbative correction to the density of states, computable using the “disk and cylinder” method, in the classically forbidden region  $E < 0$ :

$$\langle \rho(E < 0) \rangle = \frac{e^{-2e^{S_0} \int_0^{|E|} dx y(\sqrt{x})}}{8\pi |E|}, \quad (4.45)$$

which is a very small exponential tail going all the way to  $E \rightarrow -\infty$ .<sup>50</sup> The function in the exponential  $V_{\text{eff}}(E) \equiv 2e^{S_0} \int_0^{|E|} dx y(\sqrt{x})$  is the effective potential felt by an eigenvalue placed at that energy  $E < 0$ . The one-eigenvalue instanton corresponds to the locations where this effective potential is stationary (Shenker 1990).<sup>51</sup>

If the effective potential would be positive in the forbidden region, this would have a simple interpretation, giving a doubly exponentially suppressed in  $e^{-e^{S_0}}$  probability to find a black hole state with an energy below the classically allowed region. Nevertheless this is not so simple for JT gravity, since the effective potential is

<sup>49</sup> In Saad et al. (2019), the authors went from single determinant (or brane) insertions to resolvents and then evaluated their discontinuity (4.28) to get the density in the large  $S_0$  regime. In Blommaert et al. (2021b), density operator insertions were directly determined from brane pair insertions. The results are the same.

<sup>50</sup> For  $E \approx 0$ , this expression is no longer useable and one has to resort to other means, as mentioned above.

<sup>51</sup> See Alexandrov et al. (2003) for a string theory perspective as a ZZ brane contribution.

$V_{\text{eff}}(E) = \frac{e^{S_0}}{4\pi^3} [\sin(2\pi\sqrt{2C|E|}) - 2\pi\sqrt{2C|E|} \cos(2\pi\sqrt{2C|E|})]$  for  $E < 0$ . This means that the matrix model dual to JT gravity is non-perturbatively ill-defined. An option proposed in Saad et al. (2019) is to choose a steepest-descent contour for the matrix eigenvalues that runs along the real axis only for  $E > -1/8C$  and at  $E = -1/8C$  goes to the complex plane in a convergent direction. A different option is to change its random matrix completion to remove the forbidden region even non-perturbatively. This approach has been developed in a series of papers, nicely reviewed in Johnson (2022).

**Corrections to allowed region  $E > 0$**  The non-perturbative corrections in the allowed region  $E > 0$ , computed using the “disk and cylinder” method, give instead

$$\langle \rho(E > 0) \rangle = e^{S_0} \tilde{\rho}_0(E) - \frac{1}{4\pi E} \cos\left(2\pi e^{S_0} \int_0^E dE' \tilde{\rho}_0(E')\right), \quad (4.46)$$

where the first term is the disk answer with the  $e^{S_0}$  written explicitly.<sup>52</sup> This correction is not small, but is instead extremely rapidly oscillating  $e^{i\#e^{S_0}}$  (For example, it is larger than the leading one-wormhole correction to the disk  $Z_{1,1}$  which is of order  $e^{-S_0}$ ). These rapid oscillations with frequency  $\sim e^{S_0}$ , of the order of the average separation between matrix eigenvalues, are an indication of the underlying discreteness of the system.

The rapidly oscillating corrections are subleading after performing an average over an energy window, for example when computing the thermal partition function. In this sense, they can be ignored compared with the perturbative corrections in the topological expansion. Nevertheless, as emphasized in Saad et al. (2019), these oscillations can become very important in some observables, such as the very late time behavior of the spectral form factor, which we analyze next.

#### 4.4.2 Late time decay of spectral form factor

As our second example, let us look at the late-time behavior of the spectral form factor in JT gravity. Within the random matrix ensemble, the spectral form factor is the average  $\langle Z(\beta + it)Z(\beta - it) \rangle$ . First we double Laplace transform to write

$$\langle Z(\beta + it)Z(\beta - it) \rangle = \int dE dE' \langle \rho(E)\rho(E') \rangle e^{-\beta(E+E')} e^{-it(E-E')}. \quad (4.47)$$

The pair density correlator  $\langle \rho(E)\rho(E') \rangle$  in random matrix theory is of monumental importance. At large  $e^{S_0}$ , for small energy separation  $|E - E'| \ll 1$ , and sufficiently far away from spectral edges such that the spectral density  $\rho_0(E)$  is not wildly varying,

<sup>52</sup> From the non-critical string perspective, (4.46) can be thought of as coming from FZZT-brane corrections to the the partition function.

it is given by the expression:

$$\langle \rho(E)\rho(E') \rangle \approx \rho_0(E)\rho_0(E') - \frac{\sin^2(\pi\rho_0(E)(E-E'))}{\pi^2(E-E')^2} + \rho_0(E)\delta(E-E'). \quad (4.48)$$

This equation holds for *any* random matrix potential  $V(H)$  in the unitary universality class, which is of relevance to oriented bosonic JT gravity.<sup>53</sup> This is a direct manifestation of random matrix universality, governing universal level statistics at small energy separations.

The RHS contains the famous sine kernel  $\frac{\sin^2(\pi\rho_0(E)(E-E'))}{\pi^2(E-E')^2}$ , implementing level repulsion in a chaotic system. Indeed, as  $E \rightarrow E'$ , the first two terms in (4.48) cancel out as  $\sim (E-E')^2$ , indicating energy levels do not wish to be on top of each other.

It is useful to first mention the scaling in  $e^{S_0}$  of the different terms, hidden within  $\rho_0(E) \sim e^{S_0}$ . The first, factorized, term scales as  $e^{2S_0}$  and seems to vastly dominate the other terms in the regime of interest  $e^{S_0} \gg 1$ : the sine kernel oscillates rapidly but does not scale in amplitude with  $e^{S_0}$ , and the contact term scales as  $e^{S_0}$ . However, this is no longer true if the energy separation is extremely tiny:  $E-E' \sim e^{-S_0}$ . In that case, the sine kernel can compete against the factorized contribution. Since small energy separation means late times, this is precisely what happens at very late times as we will now explain.

It will be convenient to explicitly extract the average value of the sine kernel by rewriting  $\langle \rho(E)\rho(E') \rangle$  (4.48) as:

$$\rho_0(E)\rho_0(E') - \frac{1}{2\pi^2(E-E')^2} + \frac{\cos(2\pi\rho_0(E)(E-E'))}{2\pi^2(E-E')^2} + \rho_0(E)\delta(E-E'). \quad (4.49)$$

We go through all four terms of Eq. (4.49) and look at their late-time behavior and gravitational interpretation.

- The first term  $\rho_0(E)\rho_0(E')$  is just the factorized contribution we studied before, and arises gravitationally from two disconnected disks  $\sim e^{2S_0}$ . It leads to the expression:

$$Z(\beta + it) = \frac{e^{S_0}}{4\pi^2} \left( \frac{2\pi C}{\beta + it} \right)^{3/2} e^{\frac{2\pi^2 C}{\beta + it}} \rightarrow \langle |Z(\beta + it)|^2 \rangle \rightarrow \frac{e^{2S_0}}{2\pi} \frac{C^3}{t^3}. \quad (4.50)$$

The spectral form factor decays with a power-law expression as studied extensively for the more complicated boundary two-point functions in Sect. 3. This is the so-called slope.

<sup>53</sup> See Stanford and Witten (2019) for a thorough treatment of the relation between the other ensembles and gravity models.

- Inserting the second term  $-\frac{1}{2\pi^2(E-E')^2}$  of Eq. (4.49) into (4.47), and using the Fourier transform identity

$$-\int_{-\infty}^{+\infty} dx \frac{1}{2\pi^2 x^2} e^{itx} = \frac{1}{2\pi} |t|, \quad (4.51)$$

this leads to a linear growth in time  $\langle |Z(\beta + it)|^2 \rangle \supset \frac{1}{4\pi\beta} t$ . The term  $-\frac{1}{2\pi^2(E-E')^2}$  has no  $e^{S_0}$  scaling, and should gravitationally then be sought for in the connected piece of  $Z_{0,2}$ , where  $\chi = 0$ . We already computed this in (4.17). Setting  $\beta_1 = \beta + it$  and  $\beta_2 = \beta - it$ , we get at late times:

$$Z_{0,2}(\beta_1, \beta_2) = \frac{1}{2\pi} \frac{\sqrt{\beta_1 \beta_2}}{\beta_1 + \beta_2} = \frac{1}{2\pi} \frac{\sqrt{\beta^2 + t^2}}{2\beta} \rightarrow \frac{1}{4\pi\beta} t, \quad (4.52)$$

indeed gravitationally accounting for the linear growth in late time  $t$  with precisely the correct prefactor to match the random matrix result. This is the so-called ramp. Hence we can summarize that the ramp is gravitationally explained by the double trumpet amplitude.<sup>54</sup> In a microcanonical ensemble version of the spectral form factor, there is a solution of this form, the double cone we mentioned above. This contribution has a linear growth in time regardless of spacetime dimension, hinting at a random matrix universality of the black hole spectrum in all dimensions.

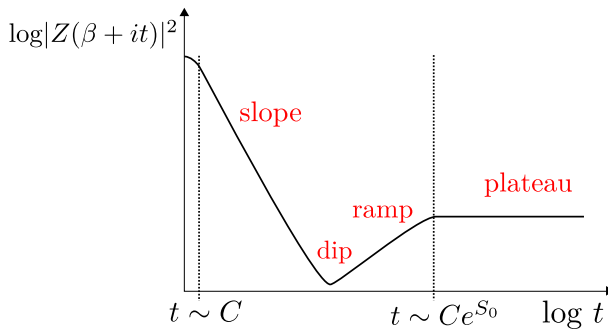
One can wonder about higher genus corrections  $Z_{g,2}$  at late times. One can readily show starting with (4.37) and (4.38) that only  $W_{0,2}$  is diverging at small energy differences, and for which hence the enhancement can compensate the suppression by  $e^{-2gS_0}$  at late times.

The decreasing slope ( $\sim \frac{e^{2S_0}}{t^3}$ ) and rising ramp ( $\sim t$ ) intersect at a late time  $t_d = (2\beta C^3)^{1/4} e^{S_0/2}$ . This is called the dip time  $t_d$  and its precise location depends on details of the dynamics of the model. Note that its scaling with the entropy  $t_d \sim e^{S_0/2}$  is more robust since it only depends on the topologies responsible for the slope and ramp.

- The third term of Eq. (4.49) is wildly oscillating:  $\frac{\cos(2\pi\rho_0(E)(E-E'))}{2\pi^2(E-E')^2}$  due to  $\rho_0(E) \sim e^{S_0} \gg 1$ , and non-perturbative in  $e^{-S_0}$  (due to  $\exp(i\#e^{S_0})$ ). As mentioned before, this is actually doubly non-perturbative in  $G_N \sim 1/S_0$ . Its Fourier transform yields a linear downward piece, starting at the time  $t \sim Ce^{S_0}$ , and perfectly canceling the slope linear growth. This location is smoothed out due to the remaining Laplace transform in (4.47). This causes the spectral form factor at times  $\sim Ce^{S_0}$  to flatten out and reach the plateau.
- Finally, the height of the plateau is determined by the contact term  $\rho_0(E)\delta(E-E')$  in (4.49), which when inserted into (4.47) leads to the final non-zero value  $Z(2\beta)$  as required for the late-time average as discussed in Sect. 4.1.

<sup>54</sup> An alternative proposal was made in Germani (2022), but we disagree with their approach.





**Fig. 17** Averaged behavior of the spectral form factor from random matrix theory

The full late-time behavior of the spectral form factor is conveniently summarized in a log-log plot, and is qualitatively of the shape shown in Fig. 17.

These random matrix results resolve the tension between the perpetual late-time decay of the semi-classical correlators and unitarity, as mentioned before in Sect. 4.1.

We understand the origin of the slope and ramp from gravity: it is an exchange in dominance from the disconnected pair of disks to the leading wormhole contribution. The gravity origin of the plateau is more mysterious, although there is recent progress indicating it might be reproduced by resumming a subset of the terms that appear in the topological expansion (Blommaert et al. 2023; Saad et al. 2022).

This is the shape predicted for a discrete system for which some sort of late-time averaging has been done to remove the erratic oscillations (such as the red curve drawn in the figure in Sect. 4.1). It is still an open question at the time of writing to understand the gravitational origin of the erratic oscillations, and fully make contact with an unaveraged discrete system. We will present some directions on this in Sect. 5.2 later on.

## 4.5 Generalization

We finish this section with a brief outline of some key generalizations on the duality between two-dimensional gravitational models and matrix integrals.

### 4.5.1 Other ensembles

The SSS duality has on one side an integral over Hermitian matrices, interpreted as a disordered Hamiltonian of a holographic theory, and on the other JT gravity on orientable manifolds with arbitrary topology. This involves the unitary ensemble, for which the canonical example is the Gaussian Unitary Ensemble (GUE). There are a total of ten types of random matrix ensembles characterized by Altland and Zirnbauer (1997). The simplest generalizations are the GOE and GSE ensembles which are dual to JT gravity on unorientable manifolds. For all the ten ensembles, it is possible to consider gravitational duals as modifications of JT gravity, either by adding topological

theories in the bulk, or by working with  $\mathcal{N} = 1$  supersymmetric JT gravity. The details of all dualities can be found in Stanford and Witten (2019).

As an example consider  $\mathcal{N} = 1$  supersymmetric JT gravity. Its disk density of states is known (Stanford and Witten 2017; Mertens et al. 2017)

$$\rho_{\mathcal{N}=1, \text{disk}}(E) = e^{S_0} \frac{2C}{\pi} \frac{\cosh\left(\frac{2\pi\sqrt{2CE}}{\sqrt{CE}}\right)}{\sqrt{CE}}, \quad (4.53)$$

and in the topological expansion one also has to sum over spin structures. Whereas bosonic JT gravity has the universal square root spectral edge  $\rho \sim \sqrt{E}$ , (4.53) has instead the universal  $\rho \sim 1/\sqrt{E}$  small  $E$  behavior. The dual is a Hermitian matrix integral over the  $L \times L$  supercharge matrix  $Q$  such that the Hamiltonian is  $H = Q^2 \geq 0$ .<sup>55</sup> The eigenvalues of  $Q$  have support over the whole real axis in the double-scaling limit and have no edge. The  $\rho \sim 1/\sqrt{E}$  behavior is just coming from transforming the  $Q$  eigenvalue density to the  $H$  spectral density.

There is a topological theory in the two-dimensional bulk that weighs odd and even spin structures differently, and we can add this theory to  $\mathcal{N} = 1$  JT gravity. This has the same disk density of states as (4.53) but a different topological expansion. The Hilbert space of the dual theory separates in two equal sized blocks (fermionic and bosonic, the  $(-1)^F$  is a symmetry) and the supercharge acts as  $Q = \begin{pmatrix} 0 & M^\dagger \\ M & 0 \end{pmatrix}$ , with  $M$  an arbitrary complex matrix. The ensemble average is done over  $M$  and the Hamiltonian is  $H = Q^2$ . The universal edge of this ensemble is  $\rho \sim 1/\sqrt{E}$ , consistent with (4.53). This realizes another ensemble in the Altland–Zirnbauer classification.

It is possible to define theories of pure JT gravity with  $\mathcal{N} = 2$  or  $\mathcal{N} = 4$  supersymmetry. These theories of gravity also appear when describing near-extremal black holes in string theory, in either asymptotically flat space or higher-dimensional AdS. The matrix ensemble dual to pure  $\mathcal{N} = 2$  JT gravity was derived in Turiaci and Witten (2023) and surprisingly only involves the Altland–Zirnbauer ensembles as building blocks.

#### 4.5.2 2d dilaton gravity with general potential

We can consider other theories of pure 2d dilaton gravity on orientable manifolds. At the two-derivative level, we reviewed in Sect. 2.1 that such theories of gravity are parametrized by a single function, the dilaton potential  $U(\Phi)$ . On the other side of the duality, matrix integrals that we consider are parametrized as well by a single function: the spectral density (or equivalently the matrix potential  $V(H)$ , before double scaling).

It is natural to expect a duality, similar to the JT case, between pure 2d dilaton gravity and a matrix integral by finding a relation between the dilaton potential and the matrix potential. This duality was proven in Maxfield and Turiaci (2021) and independently in Witten (2020), and a precise relation was found between these two functions.

<sup>55</sup> Notice that the Hilbert space does not separate into bosonic and fermionic subsectors: the  $(-1)^F$  symmetry is anomalous.

The idea is the following. Starting from JT gravity, each time a defect with opening angle  $2\pi\theta$  is added to the geometry at point  $x$  in the two-dimensional spacetime, it is equivalent to inserting an operator  $e^{-2\pi(1-\theta)\Phi(x)}$  (Mertens and Turiaci 2019). Integrating over the position of the defect is equivalent to inserting the integrated operator  $\int d^2x \sqrt{g} e^{-2\pi(1-\theta)\Phi(x)}$ . Finally, we can insert an arbitrary number of these defects and sum over their number, introducing a defect ‘fugacity’  $\lambda$ . The effect of this operator insertion in the JT path integral is very simple: the summation over the number of defects exponentiates the operator and corresponds to a deformation of the JT action that can be encoded in a shift of the dilaton potential from  $U_{\text{JT}}(\Phi) = 2\Phi$  to

$$U(\Phi) = 2\Phi + W(\Phi), \quad W(\Phi) = \int_0^1 d\theta \lambda(\theta) e^{-2\pi(1-\theta)\Phi}, \quad (4.54)$$

where we allowed for an arbitrary distribution of deficit angles and weights (For example a discrete set of defect flavors would correspond to  $\lambda(\theta) = \sum_i \lambda_i \delta(\theta - \theta_i)$ ). We can now compute the gravitational path integral for theories with such dilaton potentials by summing over defects in the JT path integral. This requires a generalization of the Weil–Petersson volumes to include the possibility of defects, geodesic boundaries and handles. These volumes are well-understood for the case that all  $\theta < 1/2$  (sharp defects) and that allowed (Maxfield and Turiaci 2021) to prove the equivalence with a random matrix integral. The density of states in this case, even at the disk level, requires a non-trivial resummation over defects and the answer is

$$\rho_0(E; W) = \frac{e^{S_0}}{2\pi} \int \frac{dy}{2\pi i} e^{2\pi y} \tanh^{-1} \left( \frac{\sqrt{E - E_0}}{\sqrt{y^2 - 2W(y) - E_0}} \right), \quad (4.55)$$

where now the support is  $E \in (E_0, +\infty)$  and  $E_0$  is the largest solution of the equation  $\int dy e^{2\pi y} (y - \sqrt{y^2 - E_0 - 2W(y)}) = 0$ . The integration over the auxiliary variable  $y$  is on a contour that runs along the imaginary axis such that all singularities of the integrand are placed to the left of it.

The expression above for  $\rho_0(E; W)$  and the duality of 2d dilaton gravity with matrix models can only be proven when the defect distribution  $\lambda(\theta)$  has support for  $\theta \in (0, 1/2)$ . It was proposed in Turiaci et al. (2021) by comparing this result with a limit of the minimal string that the expression above is valid for all  $\theta \in (0, 1)$ , and justified in Eberhardt and Turiaci (2023). This matches the answer originally found in Maxfield and Turiaci (2021), Witten (2020) for sharp defects, but the dependence with the deficit angle can be highly non-analytic.

Since the correction to the dilaton potential decays exponentially at large  $\Phi$ , and since  $\Phi = r$  classically as discussed in Sect. 2.3.2, all these theories are asymptotically AdS. It would be interesting to extend this approach to other spaces.

This argument has been extended to  $\mathcal{N} = 1$  supersymmetric dilaton gravity theories in Rosso and Turiaci (2022).

### 4.5.3 JT gravity coupled to gauge fields

Having considered a general class of 2d dilaton gravity models, the next step is to generalize the 2d black hole/random matrix duality to the case where matter is present. A simple realization of this is to couple JT gravity to 2d gauge fields (Davison et al. 2017; Mertens and Turiaci 2019) (another option previously discussed is to add also fermions using supersymmetry). These models of gravity coupled to gauge fields are also dual to double-scaled matrix models (Iliesiu 2019; Kapec et al. 2020) in a way we now describe.

Let us begin with the gravity side. For simplicity, we consider JT gravity coupled to a group  $G$  BF theory (similar considerations can be made for 2d Yang–Mills). We work in the grand canonical ensemble. The total Euclidean action is (3.86):

$$I = I_{\text{JT}} - \int_{\mathcal{M}} \text{Tr } BF + \frac{K}{2} \oint_{\partial\mathcal{M}} d\tau \text{Tr} \left[ (A_\tau + \mu)^2 \right], \quad (4.56)$$

supplemented with the chemical potential  $\mu$  (an element of the Lie algebra) and  $K$  is a dimensionful coupling constant (sometimes called the compressibility Davison et al. 2017). We impose the following mixed boundary conditions<sup>56</sup>

$$B = K(A_\tau + \mu)|_{\partial\mathcal{M}}. \quad (4.57)$$

The exact solution of 2d Yang–Mills theory is known explicitly (Witten 1991a, 1992), from which the BF amplitudes can be determined almost immediately: some amplitudes were written down in Eqs. (3.88) and (3.89) in Sect. 3.7.4. We first need to obtain the disk partition function. Integrating out the  $B$  field (along an imaginary direction) results in an integral over flat connections  $A = g dg^{-1}$ . We mod out by bulk (small) gauge transformations, but the resulting action depends on the boundary value  $g(\tau)$  giving the particle-on-a-group action<sup>57</sup>:

$$I = I_{\text{JT}} + \frac{K}{2} \int_0^\beta d\tau \text{Tr} \left[ \left( g \partial_\tau g^{-1} + \mu \right)^2 \right], \quad (4.58)$$

where  $g(\tau + \beta) = g(\tau)$ . This theory is one-loop exact (Picken 1989), and the partition function is given by

$$Z_{\text{Disk, BF}} = \frac{1}{\text{vol}(G)} \sum_R \dim(R) \chi_R(e^{\beta\mu}) e^{-\beta \frac{C_2(R)}{2K}}, \quad (4.59)$$

<sup>56</sup> In higher dimensions, working in the grand canonical ensemble implies keeping fixed the holonomy of the gauge potential at the boundary. The 2d case is special in this regard and requires a boundary condition mixing  $B$  and  $A$  in this fashion.

<sup>57</sup> The field transformation from  $A$  to  $g$  has a redundancy under right multiplication  $g(\tau) \rightarrow g(\tau)U$  with  $U \in G$ . This results in an integration space  $\text{Loop}(G)/G$ . This is not true for left multiplication which acts as a true global symmetry. This redundancy is the gauge theory version of that discussed in footnote 11 earlier.

where  $R$  are the irreducible representations of  $G$ ,  $C_2(R)$  the quadratic Casimir, and  $\chi_R(V)$  the character in representation  $R$  of group element  $V \in G$ . Combining this result with the JT gravity sector we obtain the total disk partition function:

$$Z_{\text{Disk, JT+BF}}(\beta, \mu) = \sum_R \int_{\frac{C_2(R)}{2K}}^{\infty} dE \, \rho_R(E) \chi_R(e^{\beta\mu}) e^{-\beta E}, \quad (4.60)$$

with a density of states per representation given by (Mertens and Turiaci 2019)

$$\rho_R(E) = \frac{\dim(R)}{\text{Vol}(G)} \frac{e^{S_0 C}}{2\pi^2} \sinh \left( 2\pi \sqrt{2C \left( E - \frac{C_2(R)}{2K} \right)} \right). \quad (4.61)$$

This is precisely the Schwarzian spectrum with a representation-dependent gap proportional to  $C_2(R)$ .

The calculation can be extended to higher genus surfaces in the following way. Compute first the JT and BF path integral on single trumpets, where we additionally fix the gauge holonomy of  $A$  along the geodesic boundaries. Next, we need to perform the path integral over the interior surface with geodesic boundaries with fixed length and holonomies. Similarly to JT, there are no boundary terms on these gluing cycles and the result is the volume of the moduli space of flat  $G$  connections. These are easy to compute locally (since there is no issue analogous to the mapping class group) by cut-and-glue TQFT techniques (Witten 1991a, 1992). The final step is to glue on the single trumpets by integrating over geodesic lengths and holonomies.

Now we move on to the holographic dual which is a random matrix model. In holography, when a group  $G$  gauge field is present in the bulk, a global symmetry  $G$  should commute with the boundary Hamiltonian. This implies that the boundary Hilbert space  $\mathcal{H}_{\text{BH}} = \oplus_R \mathcal{H}_{\text{BH}}^R$  breaks up into separate subspaces in different representations  $R$  of  $G$ . We can define separate Hamiltonians  $H_R$  acting on each subspace labeled by the representation  $R$ . The dual of JT gravity coupled to a group  $G$  gauge field, derived in Iliesiu (2019), Kapec et al. (2020), corresponds to an ensemble average over Hamiltonians, where each  $H_R$  is taken from a matrix ensemble described by a disk density of states  $\rho_R(E)$  given by (4.61). The Hamiltonians acting on different representation subsectors are statistically independent. It is a non-trivial calculation to verify that this matrix integral reproduces the higher genus bulk calculations described in the previous paragraph.

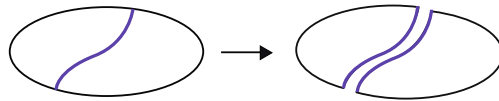
This story can be easily generalized to a more general 2d dilaton gravity theory coupled to gauge fields, by applying the defect expansion of Sect. 4.5.2 to JT gravity with gauge fields. As far as we know, this has not been studied in the literature, but it is straightforward.

#### 4.5.4 JT gravity coupled to matter

One further extension is to describe matter boundary correlation functions of the type studied in Sect. 3.3 including wormholes. Increasingly detailed investigations of these were performed in Saad et al. (2018), Blommaert et al. (2019a), Saad (2019), see also

Blommaert (2020); Iliesiu et al. (2022). A concrete random matrix dual of JT gravity with matter was proposed in Jafferis et al. (2022b, a).

We focus here on the boundary two-point function. The algorithm is the following. We first cut open the disk along the bilocal line:



In a second step, we can add all wormhole corrections in all possible ways to the resulting cut geometry:



It was shown in Saad (2019) that this procedure is identical to first starting with a given higher genus hyperbolic surface, and then summing over all topologically distinct bilocal lines one can draw on the surface. The result is quite elegant, and upon summing over genus combines into

$$\langle \mathcal{O}(\tau_1) \mathcal{O}(\tau_2) \rangle = \int dE_1 dE_2 \langle \rho(E_1, E_2) \rangle e^{-\beta E_2 - \tau(E_1 - E_2)} |\mathcal{O}_{E_1, E_2}|^2, \quad (4.62)$$

where, comparing to (3.26), we only replace the product of densities  $\rho_0(E_1)\rho_0(E_2)$  at the disk level, by the pair density correlator  $\langle \rho(E_1, E_2) \rangle$  of the JT random matrix ensemble described in Sect. 4.3. This matches with inserting simple operators  $\mathcal{O}(\tau)$  with given matrix elements  $\langle E_1 | \mathcal{O} | E_2 \rangle$  between  $H$ -eigenvalues  $E_{1,2}$ , directly into the JT random matrix integral. The decomposition of the pair density correlator (4.49) in random matrix theory, mirrors the gravitational amplitudes. In particular, the first two diagrams above correspond to the disconnected contribution to  $\langle \rho(E_1, E_2) \rangle$ . The last diagram contains the connected contributions, which at lowest genus leads to a linearly increasing ramp for the boundary two-point function at late times (Blommaert et al. 2019a; Saad 2019). Investigations beyond the two-point function can be found in Blommaert (2020), Stanford et al. (2022).

There are still some open questions surrounding these matter correlation functions. One open problem is the role of “self-intersecting” bilocal lines, whose intersection is topologically supported by wrapping around the wormhole. These were not included in the above prescription. At late times, we expect these are subdominant, but it is an open question how to think of them from the random matrix ensemble, and even whether they should be included in the first place. Another problem is that when considering dynamical matter in the bulk, one can have matter loops encircling wormhole necks. It is easy to show that the resulting negative Casimir energy causes a divergence due to the small wormhole UV region  $b \rightarrow 0$ , essentially the tachyon divergence familiar from string theory.<sup>58</sup> The recent work by Jafferis et al. (2022b, a) studies a way to regularize this divergence by deforming the model directly at the matrix level.

<sup>58</sup> Details of this calculation can be found in section 2.2 of Jafferis et al. (2022a).

A final comment is that the matrix integral language does not a priori know about the dynamical ingredients in correlation functions of these local boundary operators (the vertex factors (3.53) and the crossing kernel (3.54)), and these have to be supplied externally. In this sense, the matrix integral is merely a machine that produces the correct combinatorics of higher genus contributions in a non-perturbative framework.

## 5 Applications and future directions

In this last section, we provide several applications and generalizations of the results we reviewed up to this point on the exact solution of JT gravity. This section is not aimed to be entirely self-contained, we instead refer to the literature for more details.

### 5.1 The entropy of Hawking radiation

Since the discovery of Hawking radiation (Hawking 1975), it was realized that semi-classical black hole evaporation is inconsistent with quantum mechanical unitarity (Hawking 1976). Holography however implies that black holes behave as a quantum system, and it is important to resolve this tension and determine the fate of spacetime inside and around the horizon. This tension manifests in two ways when naively computing the gravitational path integral semi-classically.

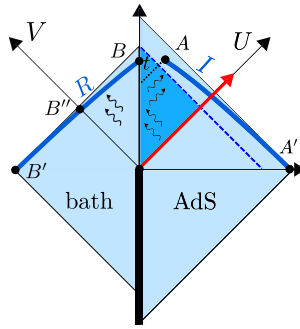
The first is that semi-classical gravity does not produce a discrete black hole spectrum, manifested e.g. in the late-time decay behavior described in Sect. 4.1. This is inconsistent with the microscopic interpretation of the Bekenstein–Hawking entropy. We already discussed in Sect. 4.4.2 how it can be improved via the inclusion of space-time wormholes. Obtaining an entirely discrete spectrum and seeing the erratic wiggles discussed in Sect. 4.1 is not yet well-understood. We come back to this in Sect. 5.2 further on.

The second is that the radiation entropy of an evaporating black hole grows forever, see  $S_{\text{rad}}(t)$  plotted in the figure below Eq. (2.104) for the JT computation. If the evaporation is unitary and the initial state is pure, the black hole microstates can only purify as much as their own entropy. This was recently resolved as well by the inclusion of spacetime wormholes that appear when computing the entropy using the replica trick, sometimes called replica wormholes (Penington et al. 2019; Almheiri et al. 2020).

The result of this discussion is the following proposal for a formula that computes the radiation entanglement entropy in a gravitating system (Penington 2020; Almheiri et al. 2019; Penington et al. 2019; Almheiri et al. 2020):

$$S = \min_X \left\{ \text{ext}_X \left[ \frac{\text{Area}(X)}{4G_N} + S_{\text{semicl}}(\Sigma_X) \right] \right\}. \quad (5.1)$$

The meaning of the right hand side is the following: first extremize the generalized entropy  $S_{\text{gen}}(X) = \frac{\text{Area}(X)}{4G_N} + S_{\text{semicl}}(\Sigma_X)$  over a codimension-two surface  $X$  such that it defines a codimension-one surface  $\Sigma_X$ , whose boundary is  $X$ , on the region whose entropy we are computing. After finding all extrema  $X$  for  $S_{\text{gen}}$ , the one with the mini-



**Fig. 18** Evaporating  $\text{AdS}_2$  black hole coupled to a flat space bath to the left

imum value is selected. Finally, the quantity  $S_{\text{semicl}}(\Sigma_X)$  is the radiation entanglement entropy computed in the semiclassical state. The regions in  $\Sigma_X$  disconnected from the semi-classical radiation region are called entanglement islands.

The above island formula resolves the issue with the entropy of Hawking radiation: when the semiclassical radiation entropy  $S_{\text{semicl}}(\Sigma_X)$  grows larger than the Bekenstein–Hawking entropy  $\frac{\text{Area}(X)}{4G_N}$ , the entropy is computed by choosing  $X$  to be a surface slightly in the interior of the black hole. In the end, this means that one essentially follows the *minimum* of both curves plotted earlier below Eq. (2.104) in Sect. 2.4.1. We note that the procedure of swapping between extrema as a function of time is a non-perturbative (in  $G_N$ ) tunneling process in quantum gravity.

In the coming three subsections, we first perform this computation in JT gravity, followed by reviews on the role of replica wormholes in understanding these developments, and in particular deriving formula (5.1).

### 5.1.1 Entanglement islands in JT gravity

Having simple solvable black hole systems as JT gravity was fundamental for uncovering the entropy formula (5.1) to begin with. We described a setup in JT gravity of evaporating black holes in Sect. 2.4.1 that we retake here. We supplement that set-up by adding an asymptotically flat region, acting as a bath collecting the radiation. On this flat region, we use the metric  $ds^2 = -dudv$  where  $u = t + z$  and  $v = t - z$  in terms of the proper coordinates defined using the near-boundary asymptotics in Sect. 2.3. Use of this time coordinate  $t$  allows a continuous gluing to the bulk region. Here we will sketch the computations that lead to a non-trivial entanglement island and resolve the tension with quantum unitarity. We follow a simplified combination of Almheiri et al. (2019), Goto et al. (2021).

We consider the (naive, semi-classical) radiation region  $R$  consisting of the flat space bath spacelike interval  $R = [B', B]$ , where we will imagine taking  $B$  on the boundary (at measured time  $t$ ) for simplicity, although our equations are more general. The bath is decoupled for  $t < 0$ , and the interface is made transparent for  $t \geq 0$ . Next to this, we allow for a non-trivial spacelike island  $I = [A, A']$  ranging from the Poincaré horizon  $A'$  to some point  $A$  that could be behind the black hole horizon, as illustrated in Fig. 18.



Our goal is to find the location  $A = X$  keeping fixed all other quantities by following the above prescription (5.1). In this enlarged set-up, we can continue with the expressions obtained in Sect. 2.4.1. We can read the stress tensor vevs around (2.100) as those corresponding to the vacuum state defined by the coordinate  $(U = F(u), v)$  after the pulse, where  $F(u)$  is the explicit solution (2.102).<sup>59</sup>

First let us consider the case of no island ( $A \rightarrow A'$ ). Then we only need to consider the entanglement entropy of the radiation region  $R$ . The emitted Hawking radiation is only captured there from  $U > 0$ , and we can consider the entanglement entropy of the interval  $[B'', B]$ . There is an additive reference (and uninteresting) entanglement entropy in the state where the bath and system would remain decoupled forever. If one were to remove this, we would obtain the same result as before (2.104).

If a non-trivial island endpoint  $A$  can be found, then the radiation region consists of both the naive one  $R$  and the island region  $I$ . As the black hole evaporates, the union  $R \cup I$  increases in size, and the entanglement entropy decreases with time, yielding a result consistent with unitarity.

Let us work this out more explicitly. The computation is more straightforward when trying to solve the closely related problem of finding a quantum extremal surface at point  $A$ , where we extremize the same quantity as before (5.1), except that we take the QFT entanglement entropy in the complementary interval  $[A, B]$  instead, and extremize for  $A$ . It is straightforward to show that in the case of no ingoing flux :  $T_{vv} := 0$ , the dilaton solution (2.44) can be written explicitly in terms of the solution  $F(t)$  of Eq. (2.98) as (Goto et al. 2021):

$$\Phi(U_A, v_A) = \frac{a}{2} \left( \frac{F''(v_A)}{F'(v_A)} + \frac{2F'(v_A)}{U_A - V_A(v_A)} \right). \quad (5.2)$$

For the metric  $ds^2 = -\Omega^{-2} dz^+ dz^-$ , and for the vacuum state defined in the  $(z^+, z^-)$  coordinates, the curved spacetime 2d CFT entanglement entropy for an interval between points  $(z_A^+, z_A^-)$  and  $(z_B^+, z_B^-)$  is given by (Fiola et al. 1994)

$$S_{\text{CFT}} = \frac{c}{6} \log \frac{|(z_A^+ - z_B^+)(z_A^- - z_B^-)|}{\epsilon_A \epsilon_B \Omega(z_A) \Omega(z_B)}, \quad (5.3)$$

where the UV cutoffs  $\epsilon_{A,B}$  are measured in the inertial coordinates at the endpoints. In our case, the coordinates  $(U, v)$  are those with respect to which we choose the vacuum. The somewhat peculiar feature here is that one endpoint ( $A$ ) is in the AdS bulk, whereas the other endpoint ( $B$ ) is in the glued flat region. Writing the bulk AdS metric as

$$ds^2 = \frac{-4dUdV}{(U-V)^2} = \frac{-4F'(v)dUdv}{(U-V)^2} \Rightarrow \Omega(z_A) = \frac{1}{2}(U_A - V_A) \frac{1}{\sqrt{F'(v_A)}}, \quad (5.4)$$

<sup>59</sup> To compare notation with (Almheiri et al. 2019; Goto et al. 2021), our  $(u, v)$  are their  $(y^+, y^-)$ . Our  $(U, V)$  are their  $(x^+, x^-)$ .

we can read off  $\Omega(z_A)$  in the bulk point  $A$  as written. Analogously, in the flat region, we write the metric as:

$$ds^2 = -dudv = -\frac{1}{F'(u)}dUdv, \Rightarrow \Omega(z_B) = \sqrt{F'(u_B)}. \quad (5.5)$$

With this information, we evaluate the QFT entanglement entropy (5.3) as

$$S_{\text{CFT}} = S_{\text{semic}} = \frac{c}{6} \log \frac{2(U_A - U_B)(v_B - v_A)\sqrt{F'(v_A)}}{\epsilon_A \epsilon_B (U_A - V_A)\sqrt{F'(u_B)}}. \quad (5.6)$$

Adding this to the gravity contribution (5.2), we write the entropy functional  $S_{\text{gen}}(X)$  as:

$$S_0 + \frac{a/2}{4G_N} \left( \frac{F''(v_A)}{F'(v_A)} + \frac{2F'(v_A)}{U_A - V_A(v_A)} \right) + \frac{c}{6} \log \frac{2(U_A - U_B)(v_B - v_A)\sqrt{F'(v_A)}}{\epsilon_A \epsilon_B (U_A - V_A)\sqrt{F'(u_B)}}. \quad (5.7)$$

The UV cutoffs  $\epsilon_{A,B}$  yield additive terms that can be reabsorbed into a renormalization of  $G_N$ . Now, for fixed  $B$ -coordinates, this expression is to be extremized for the island endpoints  $U_A$  and  $v_A$ . Upon doing this, the following important conclusions are found:

- At early times  $t < t_{\text{Page}}$ , the minimal  $S_{\text{gen}}(X)$  is found to be the no-island situation where  $A \rightarrow A'$ , located on the initial Poincaré horizon. This leads to zero contribution from the BH entropy in the entropy formula, and the rising Hawking semi-classical radiation entropy (2.104) is retrieved:

$$S \approx S_{\text{rad}}(t). \quad (5.8)$$

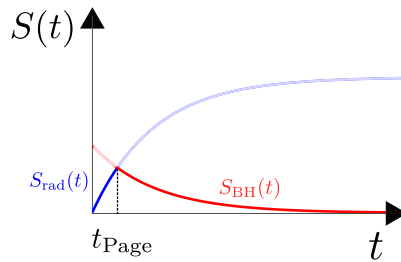
- At late times  $t > t_{\text{Page}}$ , a non-trivial island endpoint  $A$  minimizes the entropy functional. Plugging the resulting location back into the entropy formula, one finds

$$S \approx S_{\text{BH}}(t), \quad (5.9)$$

and the black hole entropy decreases in time, following roughly the quasi-statically decreasing Bekenstein–Hawking entropy. The resulting Page curve is then just the minimum of these two entropies, as illustrated in Fig. 19.

- The Page time itself gets its dimensions from  $C$ , and is inversely proportional to the (assumed large) matter central charge  $c$ , as  $t_{\text{Page}} \sim C/c$ .
- The location  $U_A > F(t \rightarrow +\infty)$ , putting the island endpoint behind the black hole horizon.<sup>60</sup>

<sup>60</sup> For the eternal black hole, the island endpoint is located outside of the black hole horizon (Almheiri et al. 2019c).



**Fig. 19** Page curve of evaporating black hole using the island prescription

- For the other lightcone component, one can write (Penington 2020; Almheiri et al. 2019):

$$v_A \sim t - \frac{1}{2\pi T(t)} \log \left( 16 \frac{S(t) - S_0}{c} \right), \quad (5.10)$$

where  $T(t) = \frac{1}{\pi} \sqrt{\frac{E}{2C}} e^{-\frac{c}{48\pi C} t} \theta(t)$  is the quasi-static temperature determined from the energy profile (2.101) and the energy-temperature relation (2.59). The island  $v_A$ -coordinate (5.10) trails behind a scrambling time behind the time  $t$  when we measure the entropy at the boundary. Both this and the previous feature were illustrated in the Penrose diagram above.

The apparent conflict with unitarity is hence resolved by taking into account entanglement islands at times greater than the Page time.

### 5.1.2 Replica wormholes

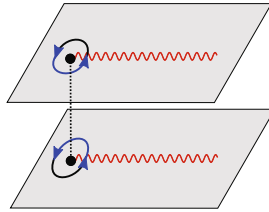
Next we review ways to motivate and derive the island formula (5.1).

A first route is to consider the case where the matter CFT sector is on its own strongly coupled and holographic. This is a version of double holography studied in Almheiri et al. (2020). The quantum extremal surface or island formula (5.1) reduces to the classical Ryu–Takayanagi formula in this higher-dimensional space. The 2d JT model is then described by a dynamical metric on the “Planck brane” (or Randall–Sundrum brane (Randall and Sundrum 1999; Karch and Randall 2001)).

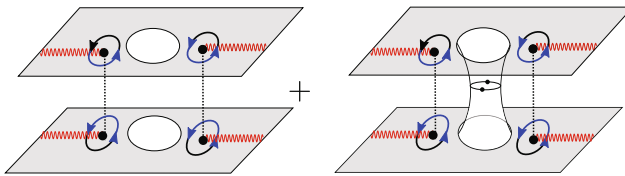
There is a more generic derivation of the island rule (5.1) based on Euclidean replica trick techniques. This approach was investigated by two groups Almheiri et al. (2020); Penington et al. (2019). Here we provide a short summary of the first one, referred to as the “east coast model”, the other one follows in the next subsection.

In QFT, the  $n$ ’th Rényi entropy  $S_n \equiv \text{Tr}(\rho_R^n)$  of an interval is computed by considering  $n$  replicas of the original manifold, glued together cyclically across branch cuts, with branch points at all endpoints of the interval of interest. We depict the  $n = 2$  case for a single endpoint, where the color-coding denotes how one moves between

replicas as one crosses a branch cut:

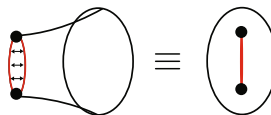


However, if there is a gravitational region in the space, we are instructed by holography (or quantum gravity more generally) to fill in the gravitational interior in all possible ways, and then sum over these possibilities with the Euclidean path integral. One of these possibilities is to connect all of the different replicas together into a single connected geometry, forming a replica wormhole. In this context, this is a saddle solution of the replicated equations of motion. For example, for the specific case of the two-sided eternal black hole in AdS, for which the radiation region contains an interval in both the left- and right flat bath region, we sketch the 2-replica configuration below:



The grey region is the non-gravitating region (the flat bath) in Euclidean signature. Locations on opposite sides of the eternal black hole (the two intervals) are in Euclidean signature separated by a rotation by 180 degrees, as drawn. The white region is a gravitating region for which we are instructed to sum over geometries, for which the two saddle topologies are drawn. For replica number  $n$ , the connected saddle is topologically suppressed as  $\sim e^{S_0(2-n)}$ , but can get enhanced at late times.

For a  $\mathbb{Z}_n$ -symmetric fully connected replica saddle configuration, one can represent the replica wormhole in two ways. As a single smooth surface as above; or by modding by the  $\mathbb{Z}_n$  symmetry as a surface of disk topology that has two twist points present (drawn in the figure on the right for  $n = 2$ ), around which one has a  $2\pi/n$  conical singularity:



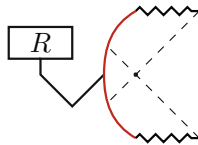
With these ideas, the following conclusions can be drawn:

- The island formula (5.1) follows directly from a generalization of the argument of Lewkowycz and Maldacena (2013) that was used to prove the Ryu–Takanayagi formula.
- The twist operator locations in the replica geometry are also part of the extremization procedure, and these locations are precisely the island endpoints in the gravitational bulk.
- The disconnected replica geometry dominates at early times (before the Page time), whereas the fully connected replica geometry dominates at late times (after the Page time).

This approach was first used for the information paradox for the eternal black hole in Almheiri et al. (2020), for which we drew the 2-replica geometries above, and generalized to the evaporating case in Goto et al. (2021).

### 5.1.3 Replica wormholes and EOW branes

We next review a very simple model proposed in Penington et al. (2019), commonly referred to as the “west coast model”, to study the role of wormholes in replica trick computations. It consists of a black hole in a pure state using the end-of-the-world (EOW) branes introduced in Sect. 3.5. To produce an information-like paradox, we can introduce a very large number  $k$  of internal states of the EOW brane, entangled with an auxiliary system  $R$ , schematically as:



The motivation of Penington et al. (2019) to consider this model was to interpret the EOW states as the interior radiation modes and the auxiliary system  $R$  as the early Hawking radiation. The state of the whole system and the radiation density matrix are therefore

$$|\Psi\rangle = \frac{1}{\sqrt{k}} \sum_{i=1}^k |\psi_i\rangle_B |i\rangle_R, \quad \rightarrow \quad \rho_R = \frac{1}{k} \sum_{i,j=1}^k |j\rangle\langle i|_R \langle\psi_i|\psi_j\rangle_B, \quad (5.11)$$

where the state of the black hole with an EOW brane in state  $i$  is  $|\psi_i\rangle_B$ . It is easy to see that when wormholes are not included, and if we pick an orthogonal basis of EOW states, gravity will give:

$$\langle\psi_i|\psi_j\rangle = \int_{j=i}^i \propto \delta_{ij}, \quad (5.12)$$

implying  $\rho_R = k^{-1} \sum_{i=1}^k |i\rangle\langle i|$  is maximally mixed with entropy  $\log(k)$ . This is the information paradox: as we increase  $k$  the entropy of the radiation can grow and become bigger than the black hole entropy  $S_{\text{BH}}$ . If the black hole is described by a quantum system with  $e^{S_{\text{BH}}}$  degrees of freedom, this is an inconsistency: for system B to purify R, the dimension of B should be at least as large as that of R.

Instead of considering entropies, it is instructive to study the purity of a state defined as  $\text{Tr}(\rho_R^2)$ . This quantity is one for a pure state and strictly smaller than one for a mixed state. Ignoring wormholes, the fact that the entropy of Hawking radiation grows indefinitely is related to the fact that the purity can decay without bound in time. Using the radiation density matrix (5.11) gives

$$\text{Tr}(\rho_R^2) = \frac{1}{k^2} \sum_{i,j=1}^k |\langle \psi_i | \psi_j \rangle|^2. \quad (5.13)$$

If we take  $\langle \psi_i | \psi_j \rangle \propto \delta_{ij}$  then the purity is  $\text{Tr}(\rho_R^2) = 1/k$  and decays indefinitely as  $k$  grows. Computing this quantity in gravity involves a computation of  $|\langle \psi_i | \psi_j \rangle|^2$ , which in the presence of wormholes gets the following contributions to leading order:

$$\begin{aligned} |\langle \psi_i | \psi_j \rangle|^2 &= \text{diagram 1} + \text{diagram 2}, \\ &= \frac{Z_1^2 \delta_{ij} + Z_2}{Z_1^2} = \delta_{ij} + \frac{Z_2}{Z_1^2}, \end{aligned} \quad (5.14)$$

where the path integral on a disk with one holographic boundary and one EOW brane is denoted by  $Z_1$  (only non-zero when  $i = j$ ) and discussed in Sect. 3.5, while the path integral on a surface with two holographic boundaries and two EOW branes is denoted by  $Z_2$  and is independent of the EOW brane flavor indices  $i, j$ . The denominator  $Z_1^2$  is only a choice of normalization of the states  $|\psi_i\rangle$ . There are two important features we need for this discussion. The first is that both  $Z_1 = \tilde{Z}_1 e^{S_0}$  with  $\tilde{Z}_1$  of order one,  $Z_2 = \tilde{Z}_2 e^{S_0}$ . The second is that while the disconnected diagram imposes  $j = i$  and therefore is given by  $Z_1^2 \delta_{ij}$ , the second diagram is non-zero for any  $i$  and  $j$ , giving an off-diagonal component to the inner product. We can use this result to compute the purity

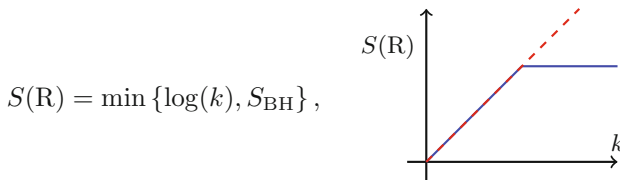
$$\text{Tr}(\rho_R^2) = \frac{k Z_1^2 + k^2 Z_2}{(k Z_1)^2} = \frac{1}{k} + \frac{Z_2}{Z_1^2}, \quad (5.15)$$

$$= \frac{1}{k} + \frac{1}{e^{S_0}} \frac{\tilde{Z}_2}{\tilde{Z}_1^2}, \quad (5.16)$$

with  $\tilde{Z}_2/\tilde{Z}_1^2$  an order one number computed in Penington et al. (2019). We now see how the wormhole term resolves this version of the information paradox: while the

first term  $1/k$  decays without bound as  $k \rightarrow \infty$  whenever the number of states is of order  $k \gtrsim e^{S_0}$  the purity will saturate to  $\text{Tr}(\rho_R^2) \sim e^{-S_0}$  thanks to the wormhole contribution.

The previous argument can be applied to the computation of Renyi entropies involving  $\text{Tr}(\rho_R^n)$  and used to extract the entanglement entropy of the radiation  $S(R) = -\text{Tr}(\rho_R \log \rho_R)$ . The result is:



which is precisely what one obtains from the “island rule” defined in Eq. (5.1), when applying that discussion to an eternal black hole [see e.g. Almheiri et al. (2019c)] and interpreting  $k$  loosely as time. In the right we show the analog of the Page curve for this model. This answer is accurate except when we are close to the transition where  $k \sim e^{S_{\text{BH}}}$ . That same work (Penington et al. 2019) also found the answer that precisely interpolates between the early and late behaviors in this simple toy model. This is a striking result: without gravity, if we want to compute the entropy of a quantum system  $R$  it is enough to know  $\rho_R$ . If the auxiliary system  $R$  is entangled with a system with gravity, such as the pure JT black hole, then the entanglement wedge of  $R$  includes the region behind the black hole horizon with an area term contributing  $S_{\text{BH}}$ .

Finally, another interesting feature of this calculation is related to the interplay between wormholes, the lack of factorization and the need for ensemble average. In terms of the matter density matrix, the derivation above does not seem to make sense since  $\langle \psi_i | \psi_j \rangle = \delta_{ij}$  while  $|\langle \psi_i | \psi_j \rangle|^2 = \delta_{ij} + Z_2/Z_1^2$ . The interpretation is that gravity is computing ensemble averages over these matrix elements, and wormholes capture the variance around the leading order delta function.

## 5.2 Factorization, discreteness, and ensemble averaging in gravity

The non-perturbative features of random matrix theory contain high frequency wiggles of temporal size  $\sim e^{-S_0}$  that are an indication of the underlying discreteness of the matrix integral. However, JT gravity’s matrix model description is double-scaled and moreover evaluated in a  $e^{-S_0}$  expansion, where the actual discreteness of the underlying system is washed out. But if JT gravity is truly a low-energy approximation of some microscopic model (such as SYK), then it should be possible to describe the discreteness of that UV model in the gravitational low-energy language. There have been several works attempting to reintroduce discreteness in JT gravity by introducing additional brane-like objects in the gravitational bulk, see e.g. Blommaert et al. (2021b), Johnson (2022), Blommaert et al. (2021a) for various approaches.

One of the important features of this link with a discrete spectrum is the so-called factorization problem mentioned in Sect. 4.3.1 before. The paradox goes as follows. Suppose we have an  $n$ -boundary amplitude  $\langle Z(\beta_1) \dots Z(\beta_n) \rangle$ . Then the boundary theories are disconnected and the amplitude should factorize. However, holography instructs us to sum over all possible geometric interiors that end on these  $n$  boundaries, disconnected *and* connected. How does gravity know in this set-up that it ultimately should factorize microscopically, since many individual geometries do not? One possibility is that there are actually additional saddles in the bulk path integral that do not appear in an averaged theory, but would appear for any specific microscopic version of it. A simple analog of the mechanism is to consider the spectral form factor (at  $\beta = 0$ ) of a discrete system:

$$\left( \sum_i e^{i E_i t} \right) \left( \sum_i e^{-i E_j t} \right) = \sum_{i=j} 1 + \sum_{i \neq j} e^{i(E_i - E_j)t}. \quad (5.17)$$

After averaging (e.g. in time), the off-diagonal terms on the RHS cancel out due to the heavy fluctuations; only the diagonal terms survive. But this diagonal contribution on its own does not factorize as the LHS does. Hence, the non-factorization is an artifact of the averaging procedure, but is not there in the truly microscopic model. For a simplified version of the SYK model, this mechanism was investigated in Saad et al. (2021b), Mukhametzhanov (2022), Saad et al. (2021a), where the new configurations were dubbed “half-wormholes”. A geometric interpretation of these is still lacking at the time of writing.

This discussion is related to the deep problem of how one should really think about quantized gravity when including wormholes. There have been several distinct indications (Harlow and Jafferis 2018; Penington et al. 2019; Almheiri et al. 2020; Giddings and Turiaci 2020; Stanford 2020; Engelhardt et al. 2021; Johnson 2021; Saad et al. 2021a) that the gravitational path integral is not equivalent to an ordinary quantum mechanical system (where we interpret the boundary circles as representing Euclidean time).

As a further example of this tension, we noted at several points that the higher genus sectors in the gravitational path integral do not contain classical saddle points in dilaton gravity. However standard canonical quantization starts by solving the classical equations of motion, which means that, in canonical quantization, it is seemingly impossible to see effects from any surfaces besides disconnected disks!

The new interpretation that we seem to be led to, is that gravity should instead be viewed as an ensemble average that happens at the very end of the computation. This is precisely the same situation as in disorder averages in spin glass systems in condensed matter physics. For example, for the free energy of the system, one should compute the quenched free energy  $\langle \log Z \rangle$ , and not the annealed free energy  $\log \langle Z \rangle$ , where the brackets denote the gravitational path integral. The situation is exactly the same in gravity (Engelhardt et al. 2021; Johnson 2021; Alishahiha et al. 2021).



### 5.3 Near-extremal black holes

In Sect. 2.1.2, we gave a brief overview of how JT gravity describes physics close to the horizon of a near-extremal black hole, at the classical gravity level. This is an old observation which was recently developed further thanks to an improved understanding of JT gravity. A partial collection of references that explore this connection is Nayak et al. (2018), Moitra et al. (2019b), Hadar (2019), Castro et al. (2018), Larsen and Zeng (2019), Kolekar and Narayan (2018), Moitra et al. (2019a), Sachdev (2019), Hong et al. (2019), Castro and Godet (2020), Charles and Larsen (2020), Larsen and Paranjape (2021), Narayan (2021), Castro and Verheijden (2021), Castro et al. (2021b), Castro et al. (2021a).

In Ghosh et al. (2020) and Iliesiu and Turiaci (2021), it was explained how the link between JT gravity and near-extremal black holes can be made also at the quantum level. This means that the leading low-temperature behavior of quantum gravity corrections from large fluctuations of a certain mode in the metric of the higher-dimensional black hole can be approximated by the quantum effects in JT gravity we analyzed in Sect. 3. This leads to interesting conclusions such as the fact that there are no classical extremal black holes in gravity and the Bekenstein–Hawking area law for extremal black holes is meaningless, except when the black hole is embedded in a supersymmetric theory and preserves enough supersymmetry (Heydemann et al. 2022; Boruch et al. 2022). A more detailed review on these results can be found in the upcoming review Iliesiu and Turiaci (2023).

### 5.4 Supersymmetric JT

JT gravity has since its start been generalized to include supersymmetry. Next to providing interesting solvable models of supergravity, whose solution can be done in parallel with the bosonic case, they are also of direct relevance to match with higher-dimensional supersymmetric black hole physics, as briefly mentioned in Sect. 2.1.2.

Here we summarize what has been computed in supersymmetric generalization of JT gravity at the time of writing this review.

The simplest case with  $\mathcal{N} = 1$  has one supercharge and describes on the boundary side the symmetry breaking pattern associated to  $\text{OSp}(2|1) \supset \text{SL}(2, \mathbb{R})$ . The super-Schwarzian action has been derived using a fluctuating boundary curve analysis in Forste and Golla (2017). At the level of the pure  $\mathcal{N} = 1$  JT gravity partition function, it has been computed for all genus and matched with random matrix models in Stanford and Witten (2019). In the presence of matter, at the disk level, the time-ordered correlators were computed in Mertens et al. (2017); Fan and Mertens (2022b) while the OTOCs are unknown. Defects have been classified in Fan and Mertens (2022b).

The next case has two supercharges,  $\mathcal{N} = 2$  JT gravity describing the symmetry breaking pattern of  $\text{SU}(1, 1|1) \supset \text{SL}(2, \mathbb{R}) \times U(1)_R$ . The super-Schwarzian analysis appears in Förste et al. (2018). The disk partition function was computed in Stanford and Witten (2017), and the density of states in Mertens et al. (2017). This theory can have fractional charge and also be supplemented by a theta angle, see Boruch et al. (2022) and Heydemann et al. (2023). The time-ordered correlators were computed

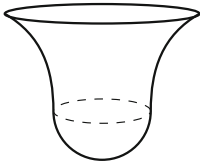
recently in Lin et al. (2023a, b). The OTOCs are unknown. The contribution from higher topologies and its matrix dual were recently proposed in Turiaci and Witten (2023).

The case with four supercharges,  $\mathcal{N} = 4$  JT gravity, describing the breaking pattern of  $\text{PSU}(1, 1|2) \supset \text{SL}(2, \mathbb{R}) \times \text{SU}(2)_R$ , was defined and the disk partition function computed in Heydeman et al. (2022). Matter correlators have not been computed and the sum over topologies is not known, although some preliminary calculations are in Turiaci and Witten (2023).

## 5.5 Two-dimensional cosmology

JT gravity with positive cosmological constant coupled to matter was considered in Maldacena et al. (2021), Cotler et al. (2020). The action takes the form (2.4) with a dilaton potential  $U(\Phi) = -2\Phi$ , supplemented by a  $\Phi_0$  term. This theory has asymptotically two-dimensional de Sitter solutions and it therefore models lower-dimensional cosmology.

There are several set-ups one can analyze in JT gravity with positive cosmological constant. One option is to follow the Hartle–Hawking no-boundary proposal (Hartle and Hawking 1983). In a frame where the metric is rigid  $\text{dS}_2$ , there is a conformal symmetry that is broken by the (asymptotics of the) dilaton which is slowly increasing with time (similar pattern of symmetry breaking as in inflation). The value of the dilaton acts as a clock, and we fix its value on the future boundary  $\Phi_b = \Phi|_{\partial}$ . The wavefunction of the universe depends on the proper length of the spatial universe which we call  $\ell$  (analogous to  $\beta/\epsilon$  in the AdS case), and it is computed by the geometry

$$\Psi(\ell, \Phi_b) \rightarrow \int [Dg][D\Phi] \quad \ell, \Phi_b \text{ large} \quad (5.18)$$


Considering  $\Phi_b$  and  $\ell$  large is an invariant way of specifying we want to compute the wavefunction at late times. The result can be computed by reducing JT gravity to a space-like Schwarzian mode living in the future boundary. The result is  $\Psi(\ell, \Phi_b) \propto \frac{\Phi_b^{3/2}}{\ell^{3/2}} e^{-i\Phi_b \ell + S_0 + i \frac{2\pi^2 \Phi_b}{\ell}}$ . In this context the parameter  $S_0$  can be interpreted as half of the dS entropy. An interpretation of this formula can be found in Maldacena et al. (2021), Cotler et al. (2020). When matter is present, it is possible to compute quantum gravity corrections to late time matter correlators using the Schwarzian action (Maldacena et al. 2021). Since dS correlators involve an integral over the wavefunction (instead of derivatives as in AdS) (Maldacena 2003b), this is considerably harder than the AdS case and the full answer for these correlators is not known.

For the case of pure JT gravity, the sum over non-trivial topologies when computing this wavefunction of the universe can be reproduced by a matrix model similar to SSS

(Maldacena et al. 2021; Cotler et al. 2020). The ultimate role of these cosmological wormholes is not understood. Finally, when matter is present there is a paradox due to the fact that the matter entropy can grow larger than the de Sitter entropy. A simplified version of a resolution of this paradox uses wormholes (Chen et al. 2021) not too different from the ones discussed in Sect. 5.1.

JT gravity with positive cosmological constant describes the physics near the horizon of a near-Nariai black hole in higher-dimensional dS. This limit is very different than the near-extremal near-horizon counterpart: it does not require charge and it is a solution of pure gravity. Four-dimensional black holes in dS have a maximal mass and when this is approached, a  $dS_2 \times S^2$  throat emerges close to the black hole and cosmological horizons.

Having discussed AdS and dS it is important to point out that some works also studied JT gravity in flat spacetime, see Dubovsky et al. (2017); Godet and Marteau (2021), and even proposed a matrix model dual (Kar et al. 2022).

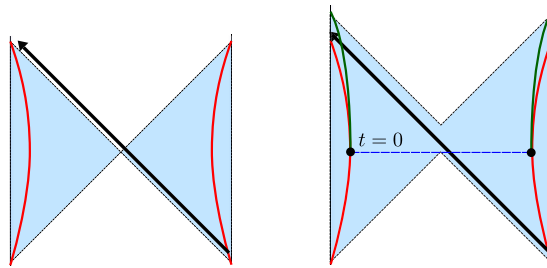
## 5.6 Traversable wormholes

The main ingredient of wormhole traversability in a holographic system is a specific coupling between two boundary systems that allows an initial state, localized to one boundary system, after a complicated evolution become a simple final state localized in the second boundary system. The holographic dual of such a process, is geometrically smooth travel through a traversable wormhole. We need to distinguish between two classes of such wormholes, the first containing black hole horizons and the second containing no horizons.

**Traversable ER wormholes** The thermofield double state in holography has a spatial wormhole, which is the AdS analog of the Einstein–Rosen (ER) bridge. This wormhole is famously not traversable by any bulk probe, but it is very close.

It was argued in Gao et al. (2017) that one can make this TFD wormhole traversable by introducing a double trace boundary interaction, leading to a schematic path integral insertion  $e^{ig O_L^{(0)} O_R^{(0)}}$  at a reference time  $t = 0$  for simple boundary operators  $O_{L,R}$  localized to either the left ( $L$ ) or right ( $R$ ) boundary, and with suitable sign of the coupling  $g$ . This results in a negative bulk energy insertion violating the null energy condition and making the wormhole traversable. This procedure was explored in detail in the NAdS<sub>2</sub> context in Maldacena et al. (2017). The physics of the resulting traversability can be conveniently captured by thinking about the wiggly boundary curve experiencing a kick whenever bulk matter is injected. In particular, after the interaction happens, the left- and right boundaries are pushed towards each other due to the negative energy injection, making a bulk signal able to traverse the wormhole and arrive at the other boundary.

In Fig. 20, the left panel illustrates the non-traversability of any signal through the Einstein–Rosen bridge. The right panel of the same figure illustrates that after negative energy injection through the double trace interaction (blue), the boundary is kicked inwards (green) and the wormhole becomes traversable.



**Fig. 20** Left: two-sided black hole as a non-traversable wormhole, or Einstein–Rosen bridge. Right: an interaction between the two sides can open up the wormhole

The bulk particle, in turn, also backreacts (not shown above) on the boundary curve, kicking the boundary curve towards the asymptotic boundary, and hence limits the amount of information that can be sent through the wormhole.

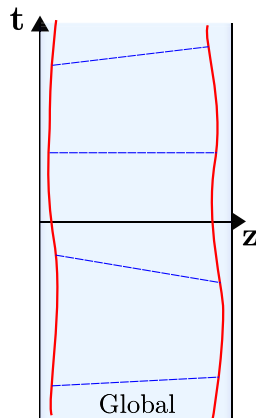
**Eternal traversable wormholes** There is a second class of traversable wormholes that are eternal and horizonless, first described in the JT framework in Maldacena and Qi (2018). These can be described in  $\text{NAdS}_2$  by understanding how to interpret both global boundaries as  $\text{NAdS}$  boundaries. The vacuum solution for the dilaton field  $\Phi$  in global coordinates is  $\Phi = a \cot 2z$ ,<sup>61</sup> where the right boundary is at  $z = 0$  and leads to  $\Phi \rightarrow +\infty$  as before. The left boundary on the other hand is at  $z = \pi$  and hence  $\Phi \rightarrow -\infty$ . This boundary hence does not satisfy the boundary conditions of Sect. 2.3.1.

In order to make sense of the two-boundary system in  $\text{NAdS}_2/\text{NCFT}_1$ , we need to have  $\Phi \rightarrow +\infty$  at both ends. To implement this, one has to violate the null energy condition of the matter sector, since by integrating the first equation of motion in (2.42) along the  $u$ -direction from boundary to boundary, we get  $-(\Phi_b^L + \Phi_b^R) = 8\pi G_N \int_{-\infty}^{+\infty} T_{x^+x^+} dx^+$ , where  $x^+$  is the affine coordinate along the null line  $dx^+ = e^{2\omega} du$ . The LHS is negative, but the RHS is positive if the null energy condition is satisfied. Violating the null energy condition can again be implemented by a left-right coupling of similar type as before:  $S_{\text{int}} = g \int dt \sum_{i=1}^N O_i^L(t) O_i^R(t)$ . At leading order in large  $N$  and small  $g$  with fixed  $Ng$ , the effect of this coupling can be captured in a double Schwarzian language in terms of left and right global time reparametrizations  $f_L$  and  $f_R$  of the respective boundaries. The global frame is shown in Fig. 21, with bulk negative energy interactions (blue) and two wiggly curves at the holographic boundaries.

The effective action describing this coupled system is

$$I = -C \int dt \left[ \left\{ \tan \frac{f_L}{2}, t \right\} + \left\{ \tan \frac{f_R}{2}, t \right\} - \frac{gN}{2^{2\Delta} C} \left( \frac{f'_L(t) f'_R(t)}{\cos^2 \frac{f_L(t) - f_R(t)}{2}} \right)^\Delta \right], \quad (5.19)$$

<sup>61</sup> We set  $\mu = -a$  (and  $b = 0$ ) in Eq. (2.43) and use the global frame (2.34). Note that because  $\mu < 0$ , the global frame has a negative energy  $-a/8\pi G_N$  compared to the Poincaré patch. This is very similar to the situation in 3d gravity.



**Fig. 21** Global AdS<sub>2</sub> as an eternal traversable wormhole

where there is only a single “diagonal” Möbius redundancy. For the analysis of this system, and an alternative derivation from the low-energy limit of two coupled SYK models, we refer to the original work by Maldacena and Qi (2018).

Physically, the resulting wormhole is horizonless. From the higher-dimensional perspective, it has been argued that one could obtain it by gluing together the throats of two near-extremal black holes by integrating out sufficiently energetic field modes. For further investigations into the presence of these kinds of solutions in our universe, we refer to Maldacena et al. (2018).

### 5.7 Finite cut-off and $\bar{T}\bar{T}$

The  $T\bar{T}$  deformation is a solvable irrelevant deformation of two-dimensional conformal field theories (Smirnov and Zamolodchikov 2017). It was proposed in McGough et al. (2018) that the AdS dual interpretation is to move the holographic boundary inside the bulk with Dirichlet boundary conditions. Even though this has passed many tests, it is still not fully understood.

In Iliesiu et al. (2020), Stanford and Yang (2020) the problem of computing the partition function of JT gravity in AdS with a finite-length boundary was studied, using different methods. Up to now we considered boundary conditions where both the dilaton  $\Phi_b$  and the boundary proper length  $L$  diverge, with a fixed ratio  $L/\Phi_b$  given in terms of the boundary temperature. This corresponds to sending the boundary to infinity in AdS<sub>2</sub>. The finite cutoff version of this calculation is to work with a fixed dilaton and proper length not parametrically large. The results found by the two groups are different, and it is still an open question to resolve it. The result in Iliesiu et al. (2020) is given by

$$Z(L, \Phi_b) = \frac{L\Phi_b^2}{L^2 + 4\pi^2} e^{-L\Phi_b} K_2\left(-\sqrt{\Phi_b^2(L^2 + 4\pi^2)}\right), \quad (5.20)$$

and was matched to a quantum-mechanical version of the  $T\bar{T}$  deformation introduced in Gross et al. (2020) applied to the Schwarzian theory.

A problem with this theory is that it has a non-unitary spectrum, an issue anticipated in McGough et al. (2018). The result in Stanford and Yang (2020) is more complicated, but has a more transparent geometric interpretation.

Further references include Griguolo et al. (2022) which applies resurgence techniques to the problem, and Ebert et al. (2022) which analyzes the connection with 3d gravity.

## 5.8 Volume of black hole interior and complexity

A further success of JT gravity is the explicit evaluation of the volume of the black hole interior, including quantum gravity corrections. The extremal volume of a codimension-one slice, reduces to the geodesic length in one spatial dimension. In the hyperbolic two-plane  $ds^2 = \frac{dZ^2 + dT^2}{Z^2}$ , geodesics are circle segments. For two bulk points  $P = (T_1, Z_1)$  and  $Q = (T_2, Z_2)$ , the length  $\ell$  along the circle geodesic connecting these two points is readily computed

$$\ell(P, Q) = \int_P^Q ds = 2 \operatorname{Arcsinh} \sqrt{\frac{(Z_1 - Z_2)^2 + (T_1 - T_2)^2}{4Z_1 Z_2}}. \quad (5.21)$$

In the case where both points are on the wiggly boundary curve, and hence  $Z = \epsilon T' = \epsilon F'$ , we get

$$\ell = \ln \frac{(F_1 - F_2)^2}{F'_1 F'_2} - \ln \epsilon^2. \quad (5.22)$$

If we finally plug in the black hole solution  $F(\tau) = \tan \frac{\pi}{\beta} \tau$ , move one endpoint to the other side of the TFD state by letting  $\tau_1 \rightarrow \tau_1 + \beta/2$ , and Wick-rotate as  $t \equiv i\tau_1 - i\tau_2$ , we obtain the renormalized real-time result:

$$\ell_{\text{ren}}(t) = \ln \cosh^2 \frac{\pi}{\beta} t \xrightarrow{t \rightarrow \infty} \frac{2\pi}{\beta} t, \quad (5.23)$$

which is linear in the time separation  $t$  for large times (Brown et al. 2019; Goto et al. 2019; Akhavan et al. 2018). This geodesic length is growing indefinitely due to the unbounded growth of the black hole interior as time evolves (Hartman and Maldacena 2013).

The beauty of the expression (5.22) is that we can go beyond classical gravity, and plug it into the Schwarzian path integral directly. The resulting path integral evaluation can be done exactly by relating it to the boundary two-point function using the trick:  $\frac{\partial}{\partial \Delta} G_{\Delta}(\tau_1, \tau_2)|_{\Delta=0} = -\ln \frac{(F_1 - F_2)^2}{F'_1 F'_2}$ . The resulting expression is written in Yang (2019) and somewhat surprisingly shows a continued linear growth. This means the quantum gravity fluctuations described in Sect. 3 do not cause a change in very late time behavior.<sup>62</sup>

<sup>62</sup> The same late time growth can be shown to hold in the  $\mathcal{N} = 1$  supersymmetric version of JT gravity as well (Fan and Mertens 2022b).

Going beyond the Schwarzian, one can include wormhole corrections to this computation. As previously, the calculation can be done by directly relating it to the boundary two-point function (Iliesiu et al. 2022). The result is a saturation of the growth of the black hole interior size at the timescale where non-perturbative random matrix effects kick in  $t \sim Ce^{S_0}$ . Within the complexity = volume conjecture (Susskind 2016), it was argued by L. Susskind and collaborators (Brown and Susskind 2018) that this late time rise and saturation is expected for the growth of computational complexity of the dual boundary quantum system.

## 5.9 Bulk correlators and observables

Within any theory of quantum gravity, defining bulk observables is a delicate business. Namely, when coordinate transformations (or diffeomorphisms) are treated as a gauge symmetry, only diff-invariant operators are physically observable. Defining such diff-invariant operators requires gravitationally dressing a bulk matter operator. There are uncountably many ways of doing this. Here we focus on one particular choice that leads to explicit results in JT gravity (Blommaert et al. 2019a; Mertens 2019). Using the wiggly boundary curve for any off-shell  $F(t)$ , we use a radar procedure to define a bulk point as  $(U = F(u), V = F(v))$  in terms of two proper times  $u$  and  $v$  on the boundary clock. This then defines an observable matter operator  $\mathcal{O}(U = F(u), V = F(v))$  that has implicit dependence on the gravitational degrees of freedom through  $F(t)$ .

For instance, for a massless scalar field in the bulk, the CFT bulk two-point function with Dirichlet boundary conditions at the AdS boundary equals:

$$\langle \phi(u_1, v_1) \phi(u_2, v_2) \rangle_{\text{CFT}} = -\frac{1}{4\pi} \ln \frac{(F(u_1) - F(u_2))(F(v_1) - F(v_2))}{(F(u_1) - F(v_2))(F(v_1) - F(u_2))}. \quad (5.24)$$

This object is then viewed in a second step as a gravitational operator to be inserted in the Schwarzian path integral. To actually evaluate it, we can make use of the following trick:

$$\ln \frac{(F(u_1) - F(u_2))(F(v_1) - F(v_2))}{(F(u_1) - F(v_2))(F(v_1) - F(u_2))} = \int_{v_1}^{u_1} dt_1 \int_{v_2}^{u_2} dt_2 \frac{F'(t_1)F'(t_2)}{(F(t_1) - F(t_2))^2}, \quad (5.25)$$

relating it to a double integral of the known boundary two-point function.

The technical trick (5.25) is actually the HKLL bulk reconstruction formula (Hamilton et al. 2006a, b) applied to a massless scalar in  $\text{AdS}_2$ . This prescription is hence elevated here into generic off-shell configurations  $F(t)$ .

One can Fourier transform these expressions to get the spectral occupation of the bulk matter modes, leading to quantum gravitational corrections to the Unruh heat bath (Mertens 2019; Blommaert et al. 2021c). These dressed bulk observables seem to correspond to measurements done by static (fiducial) observers in the quantum bulk.

The near-horizon region is very far from the holographic boundary and corresponds to an IR region. This means that correlators of these bulk observables experience strong

quantum gravitational fluctuations in the near-horizon Rindler region. Hence, insisting on using such observable operators leads to tension with assuming quantum field theory in curved spacetime close to the black hole horizon, one of the main assumptions in arguing for information loss during black hole evaporation. This comment is independent from other relations to the information paradox discussed in earlier sections, which relied on wormhole configurations.

## 5.10 Liouville gravity and minimal string

When considering solvable models of 2d gravity, next to JT gravity, there is an older model that attracted considerable attention in the 1980s and 1990s: that is Liouville gravity. An important recent result is that these models are not independent, but JT gravity can be embedded in Liouville gravity. This was first noticed for the spectrum in Saad et al. (2019), and subsequently developed in detail in Mertens and Turiaci (2021), Mertens (2021), Turiaci et al. (2021), Fan and Mertens (2022a), Suzuki and Takayanagi (2021), Hirano and Kuroki (2022), Artemev (2022).

Liouville gravity, or the non-critical string, is a model of two-dimensional quantum gravity, where a 2d CFT matter sector  $S_M[\chi]$  is coupled to gravity as

$$Z = \sum_{\text{topologies}} \int \frac{[Dg][D\chi]}{\text{Vol}(\text{Diff})} e^{-S_M[\chi, g] - \mu_0 \int_{\Sigma} d^2x \sqrt{g}}. \quad (5.26)$$

The conformal factor  $e^{2\omega}$  of the metric gets dynamics at the quantum level governed by the Liouville CFT  $S_L$  (3.78) (Polyakov 1981; Distler and Kawai 1989; David 1988). This gauge fixing leads to additional  $bc$  ghosts, and the total model is hence 2d Liouville + matter + ghost CFT, with the conformal anomaly constraint  $c_M + c_L + c_{\text{gh}} = 0$ , where  $c_L = 1 + 6Q^2 = 1 + 6(b + b^{-1})^2$ .

Within this string theoretic framework, the fixed length amplitude on the disk world-sheet, in suitable units, equals

$$Z(\beta) \sim \int_1^\infty dE e^{-\beta E} \rho_0(E), \quad \rho_0(E) = \sinh\left(\frac{1}{b^2} \text{arccosh} E\right). \quad (5.27)$$

We now interpret this in the same way as in JT gravity as a canonical partition function of a quantum black hole with temperature  $\beta^{-1}$ , but with a modified density of states. At low energies,  $\rho_0(E)$  limits to the JT density of states. To see this, we identify new kinematic variables as  $E = 1 + 2\pi^2 b^4 E_{\text{JT}}$  and  $\beta = \frac{\beta_{\text{JT}}}{2\pi^2 b^4}$ . Zooming in on the region close to the spectral cutoff by  $b \rightarrow 0$ , keeping  $E_{\text{JT}}$  and  $\beta_{\text{JT}}$  fixed, we reduce to the JT density of states  $\rho_0(E) \approx \sinh 2\pi \sqrt{E_{\text{JT}}}$ . At high energies, the Liouville gravity density of states (5.27) deviates from JT and rises with a power-law as  $\sim E^{1/b^2}$ .

This observation of embedding JT in the IR regime of Liouville gravity is no coincidence, and substantial more evidence can be gathered for it:

- Bulk and boundary vertex operators can be inserted on the disk, and their fixed length amplitudes computed. The results match in each case with the JT expressions in the IR kinematic regime (Mertens and Turiaci 2021).



- Describing  $S_M$  with a timelike Liouville field  $\chi$ , one can perform a field redefinition from the Liouville gravity variables  $(\phi, \chi)$  into the dilaton gravity variables  $(\omega, \Phi)$  by using  $\phi = b^{-1}\omega - b\pi\Phi$ , and  $\chi = b^{-1}\omega + b\pi\Phi$ . This transformation leads to a dilaton gravity model of the type (2.4) with dilaton potential

$$U(\Phi) \sim \sinh(2\pi b^2 \Phi), \quad (5.28)$$

which in the deep IR region, where  $\Phi = r$  is small, becomes the JT dilaton potential. Close to the boundary  $r \approx +\infty$ , the dilaton potential diverges leading to a curvature singularity at the holographic boundary.

- The amplitudes of this model are governed by an algebraic structure related to the  $U_q(\mathfrak{sl}(2, \mathbb{R}))$  quantum group. This can be explained directly at the Lagrangian level by the above dilaton potential (5.28) in the Poisson-sigma model framework of Sect. 2.1.1 (Fan and Mertens 2022a; Mertens 2022).
- For the particular case where the matter sector is the  $(p, q)$  minimal model, the resulting string model is called the  $(p, q)$  minimal string, which famously was conjectured to be equivalent to a random matrix integral, stemming from ideas around random triangulations of a 2d surface (Brezin and Kazakov 1990; Douglas and Shenker 1990; Gross and Migdal 1990). The fact that JT gravity is a matrix integral is hence embedded into this older statement.
- Multi-boundary and higher genus effects can be studied within the random matrix framework of the minimal string, leading to a deformation of the decomposition of amplitudes given in Sect. 4.2 (Mertens and Turiaci 2021).

## 5.11 Universe field theory and quantum chaos

The non-perturbative completion of JT gravity into a matrix integral is very useful, but it has its downsides. For instance, a matrix integral is not an ordinary quantum mechanical system, instead representing a maximally disordered average over models.

What one would want is an all-encompassing quantum mechanical framework that correctly describes multi-universes as quantum mechanical operator insertions, and allows for quantum processes that capture the emission and absorption of entire universes.<sup>63</sup> This is a third quantization of sorts, similar to string field theory. Such a quantum picture was already developed in the past by Coleman (1988); Giddings and Strominger (1988), and was reinvestigated in the current context in Marolf and Maxfield (2020), Giddings and Turiaci (2020), where basic principles were formulated.

A concrete description in which such multi-universe phenomena are captured in JT gravity is in terms of Kodaira–Spencer theory. This observation was made in Post et al. (2022), and we refer to that work for all the details of this correspondence. The Kodaira–Spencer model is defined on a 2d surface (Dijkgraaf and Vafa 2007) (the

<sup>63</sup> Notice that there is no tension with Sect. 5.2, since we are not working with a fixed number of boundaries, and do not require the interpretation of the boundaries as Euclidean time directions of this quantum mechanical system.

spectral curve of the associated matrix model)<sup>64</sup>:

$$S_{\text{KS}} = \int d^2z \left( \frac{1}{2} \partial \Phi \bar{\partial} \Phi - (\mathcal{J} - 2zy(z)) \bar{\partial} \Phi \right) + \frac{\lambda}{2} \oint_{z=0} \frac{dz}{2zy(z)} \Phi \mathcal{J}^2, \quad (5.29)$$

in terms of two fields  $\Phi$  (unrelated to the dilaton!) and  $\mathcal{J}$ , which can be thought of as  $\mathbb{Z}_2$ -twisted boson fields. The free theory has the propagators  $\langle \mathcal{J}(z_0) \Phi(z) \rangle = \frac{1}{z_0 - z} - \frac{1}{z_0 + z}$  and  $\langle \mathcal{J}(z_i) \mathcal{J}(z) \rangle = \frac{1}{(z_i - z)^2} + \frac{1}{(z_i + z)^2}$ .

The cubic interaction vertex in (5.29) is to be identified with the elementary three-holed sphere in hyperbolic geometry. The full hyperbolic surface is then created as a Feynman graph that coincides with the “skeleton graph” one draws on the surface. This theory is a string field theory model for which the perturbative string worldsheets are reinterpreted as spacetime universes, hence the name universe field theory.

### Example:

Let us consider the genus one correction to the disk amplitude, given in 1st order perturbation theory by bringing down a single instance of the interaction vertex (5.29) as

$$\langle \mathcal{J}(z_0) \rangle_1^c = \frac{\lambda}{2} \oint_{z=0} \frac{dz}{2zy(z)} \langle \mathcal{J}(z_0) \Phi(z) \rangle \langle \mathcal{J}(z) \mathcal{J}(z) \rangle, \quad (5.30)$$

corresponding to a one-loop Feynman diagram. Evaluating gives

$$\langle \mathcal{J}(z_0) \rangle_1^c = \frac{\lambda}{4} \text{Res}_{z=0} \frac{2}{z_0^2 - z^2} \frac{2\pi}{\sin 2\pi z} \frac{1}{4z^2} = \lambda \frac{3 + 2\pi^2 z_0^2}{24z_0^4}, \quad (5.31)$$

which is indeed the correct contribution  $W_{1,1}$ .

—o—

Feynman diagrammatics with the cubic vertex in (5.29) allows one to immediately write down recursive Dyson–Schwinger relations that can be shown to match with Eynard–Orantin’s topological recursion relations (Dijkgraaf and Vafa 2007).<sup>65</sup>

FZZT brane (and anti-brane) operators can be inserted into the model as  $e^{\Phi(z)}$  (and  $e^{-\Phi(z)}$ ), and can be utilized to get to the same non-perturbative physics as the matrix model completion of JT gravity reviewed in Sect. 4.4. The model with a fixed number of branes can be rewritten in terms of a Kontsevich-like matrix model in terms of flavor degrees of freedom. The resulting flavor matrix theory can then be shown to describe universal chaotic dynamics (Altland et al. 2022). This hence provides a direct connection between the universal ergodic behavior in chaotic systems, and the universe field theory of a gravity model.

<sup>64</sup> It starts its life as a 6d model on the topological closed string B-model that is reduced to 2d.

<sup>65</sup> A subtlety with this expression is that naively not all contractions are taken into account; the external  $\mathcal{J}$  lines only contract with the  $\Phi$ -leg of an internal vertex. This is related to the fact that the boson fields are chiral (Dijkgraaf and Vafa 2007). This is also present in (5.30) above.

**Acknowledgements** We would like to thank all our collaborators throughout the past years, without whom we would not have been able to write the current review. We thank Roberto Emparan, Juan Maldacena, Francesca Mariani and Subir Sachdev for comments on the draft. We would like to specially thank Herman Verlinde for introducing us to this subject. TM acknowledges financial support from Research Foundation Flanders (FWO Vlaanderen) and the European Research Council (Grant BHHQG-101040024). Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. GJT is supported by the Institute for Advanced Study and the National Science Foundation under Grant No. PHY-2207584, and by the Sivilian Fund. GJT thanks the University of Warsaw for hospitality during final stages of writing this review.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Achucarro A, Ortiz ME (1993) Relating black holes in two-dimensions and three-dimensions. *Phys Rev D* 48:3600–3605. <https://doi.org/10.1103/PhysRevD.48.3600>. arXiv:hep-th/9304068
- Akhavan A, Alishahiha M, Naseh A et al (2018) Complexity and behind the horizon cut off. *JHEP* 12:090. [https://doi.org/10.1007/JHEP12\(2018\)090](https://doi.org/10.1007/JHEP12(2018)090). arXiv:1810.12015 [hep-th]
- Alekseev A, Shatashvili SL (1989) Path integral quantization of the coadjoint orbits of the Virasoro group and 2D gravity. *Nucl Phys B* 323:719–733. [https://doi.org/10.1016/0550-3213\(89\)90130-2](https://doi.org/10.1016/0550-3213(89)90130-2)
- Alekseev A, Shatashvili SL (1990) From geometric quantization to conformal field theory. *Commun Math Phys* 128:197–212. <https://doi.org/10.1007/BF02097053>
- Alekseev A, Shatashvili SL (2021) Characters, coadjoint orbits and Duistermaat–Heckman integrals. *J Geom Phys* 170:104386. <https://doi.org/10.1016/j.geomphys.2021.104386>. arXiv:2004.03024 [hep-th]
- Alexandrov SY, Kazakov VA, Kutasov D (2003) Nonperturbative effects in matrix models and D-branes. *JHEP* 09:057. <https://doi.org/10.1088/1126-6708/2003/09/057>. arXiv:hep-th/0306177
- Alishahiha M, Faraji Astaneh A, Jafari G et al (2021) Free energy for deformed Jackiw–Teitelboim gravity. *Phys Rev D* 103(4):046005. <https://doi.org/10.1103/PhysRevD.103.046005>. arXiv:2010.02016 [hep-th]
- Almheiri A, Kang B (2016) Conformal symmetry breaking and thermodynamics of near-extremal black holes. *JHEP* 10:052. [https://doi.org/10.1007/JHEP10\(2016\)052](https://doi.org/10.1007/JHEP10(2016)052). arXiv:1606.04108 [hep-th]
- Almheiri A, Polchinski J (2015) Models of  $\text{AdS}_2$  backreaction and holography. *JHEP* 11:014. [https://doi.org/10.1007/JHEP11\(2015\)014](https://doi.org/10.1007/JHEP11(2015)014). arXiv:1402.6334 [hep-th]
- Almheiri A, Engelhardt N, Marolf D et al (2019) The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole. *JHEP* 12:063. [https://doi.org/10.1007/JHEP12\(2019\)063](https://doi.org/10.1007/JHEP12(2019)063). arXiv:1905.08762 [hep-th]
- Almheiri A, Hartman T, Maldacena J, et al (2020) Replica wormholes and the entropy of Hawking radiation. *JHEP* 05:013. [https://doi.org/10.1007/JHEP05\(2020\)013](https://doi.org/10.1007/JHEP05(2020)013). arXiv:1911.12333 [hep-th]
- Almheiri A, Mahajan R, Maldacena J (2019c) Islands outside the horizon. arXiv e-prints arXiv:1910.11077 [hep-th]
- Almheiri A, Mahajan R, Maldacena J et al (2020) The Page curve of Hawking radiation from semiclassical geometry. *JHEP* 03:149. [https://doi.org/10.1007/JHEP03\(2020\)149](https://doi.org/10.1007/JHEP03(2020)149). arXiv:1908.10996 [hep-th]
- Almheiri A, Hartman T, Maldacena J et al (2021) The entropy of Hawking radiation. *Rev Mod Phys* 93(3):035002. <https://doi.org/10.1103/RevModPhys.93.035002>. arXiv:2006.06872 [hep-th]
- Altland A, Zirnbauer MR (1997) Nonstandard symmetry classes in mesoscopic normal-superconducting hybrid structures. *Phys Rev B* 55:1142–1161. <https://doi.org/10.1103/PhysRevB.55.1142>. arXiv:cond-mat/9602137

- Altland A, Post B, Sonner J, et al (2022) Quantum chaos in 2D gravity. arXiv e-prints [arXiv:2204.07583](https://arxiv.org/abs/2204.07583) [hep-th]
- Artemev A (2022) Note on large- $p$  limit of  $(2,2p+1)$  minimal Liouville gravity and moduli space volumes. Nucl Phys B 981:115876. <https://doi.org/10.1016/j.nuclphysb.2022.115876>. arXiv:2203.06629 [hep-th]
- Bagrets D, Altland A, Kamenev A (2016) Sachdev-Ye-Kitaev model as Liouville quantum mechanics. Nucl Phys B 911:191–205. <https://doi.org/10.1016/j.nuclphysb.2016.08.002>, arXiv:1607.00694 [cond-mat.str-el]
- Bagrets D, Altland A, Kamenev A (2017) Power-law out of time order correlation functions in the SYK model. Nucl Phys B 921:727–752. <https://doi.org/10.1016/j.nuclphysb.2017.06.012>, arXiv:1702.08902 [cond-mat.str-el]
- Balog J, Feher L, Palla L (1998) Coadjoint orbits of the Virasoro algebra and the global Liouville equation. Int J Mod Phys A 13:315–362. <https://doi.org/10.1142/S0217751X98000147>. arXiv:hep-th/9703045
- Banks T, O’Loughlin M (1991) Two-dimensional quantum gravity in Minkowski space. Nucl Phys B 362:649–664. [https://doi.org/10.1016/0550-3213\(91\)90547-B](https://doi.org/10.1016/0550-3213(91)90547-B)
- Blommaert A (2020) Dissecting the ensemble in JT gravity. arXiv e-prints [arXiv:2006.13971](https://arxiv.org/abs/2006.13971) [hep-th]
- Blommaert A, Usatyuk M (2022) Microstructure in matrix elements. JHEP 09:070. [https://doi.org/10.1007/JHEP09\(2022\)070](https://doi.org/10.1007/JHEP09(2022)070). arXiv:2108.02210 [hep-th]
- Blommaert A, Mertens TG, Verschelde H (2018) The Schwarzian theory: a Wilson line perspective. JHEP 12:022. [https://doi.org/10.1007/JHEP12\(2018\)022](https://doi.org/10.1007/JHEP12(2018)022). arXiv:1806.07765 [hep-th]
- Blommaert A, Mertens TG, Verschelde H (2019) Clocks and Rods in Jackiw–Teitelboim quantum gravity. JHEP 09:060. [https://doi.org/10.1007/JHEP09\(2019\)060](https://doi.org/10.1007/JHEP09(2019)060). arXiv:1902.11194 [hep-th]
- Blommaert A, Mertens TG, Verschelde H (2019) Fine structure of Jackiw–Teitelboim quantum gravity. JHEP 09:066. [https://doi.org/10.1007/JHEP09\(2019\)066](https://doi.org/10.1007/JHEP09(2019)066). arXiv:1812.00918 [hep-th]
- Blommaert A, Iliesiu LV, Kruthoff J (2021a) Gravity factorized. arXiv e-prints [arXiv:2111.07863](https://arxiv.org/abs/2111.07863) [hep-th]
- Blommaert A, Mertens TG, Verschelde H (2021) Eigenbranes in Jackiw–Teitelboim gravity. JHEP 02:168. [https://doi.org/10.1007/JHEP02\(2021\)168](https://doi.org/10.1007/JHEP02(2021)168). arXiv:1911.11603 [hep-th]
- Blommaert A, Mertens TG, Verschelde H (2021) Unruh detectors and quantum chaos in JT gravity. JHEP 03:086. [https://doi.org/10.1007/JHEP03\(2021\)086](https://doi.org/10.1007/JHEP03(2021)086). arXiv:2005.13058 [hep-th]
- Blommaert A, Kruthoff J, Yao S (2023) An integrable road to a perturbative plateau. JHEP 04:048. [https://doi.org/10.1007/JHEP04\(2023\)048](https://doi.org/10.1007/JHEP04(2023)048). arXiv:2208.13795 [hep-th]
- Boruch J, Heydeman MT, Iliesiu LV, et al (2022) BPS and near-BPS black holes in  $AdS_5$  and their spectrum in  $\mathcal{N} = 4$  SYM. arXiv e-prints [arXiv:2203.01331](https://arxiv.org/abs/2203.01331) [hep-th]
- Brezin E, Kazakov VA (1990) Exactly solvable field theories of closed strings. Phys Lett B 236:144–150. [https://doi.org/10.1016/0370-2693\(90\)90818-Q](https://doi.org/10.1016/0370-2693(90)90818-Q)
- Brezin E, Itzykson C, Parisi G et al (1978) Planar diagrams. Commun Math Phys 59:35. <https://doi.org/10.1007/BF01614153>
- Brown AR, Susskind L (2018) Second law of quantum complexity. Phys Rev D 97(8):086015. <https://doi.org/10.1103/PhysRevD.97.086015>. arXiv:1701.01107 [hep-th]
- Brown AR, Gharibyan H, Lin HW et al (2019) Complexity of Jackiw–Teitelboim gravity. Phys Rev D 99(4):046016. <https://doi.org/10.1103/PhysRevD.99.046016>. arXiv:1810.08741 [hep-th]
- Callan CG Jr., Giddings SB, Harvey JA, et al (1992) Evanescent black holes. Phys Rev D 45(4):R1005. <https://doi.org/10.1103/PhysRevD.45.R1005>. arXiv:hep-th/9111056
- Callebaut N (2019) The gravitational dynamics of kinematic space. JHEP 02:153. [https://doi.org/10.1007/JHEP02\(2019\)153](https://doi.org/10.1007/JHEP02(2019)153). arXiv:1808.10431 [hep-th]
- Callebaut N, Verlinde H (2019) Entanglement dynamics in 2D CFT with boundary: entropic origin of JT gravity and Schwarzian QM. JHEP 05:045. [https://doi.org/10.1007/JHEP05\(2019\)045](https://doi.org/10.1007/JHEP05(2019)045). arXiv:1808.05583 [hep-th]
- Castro A, Godet V (2020) Breaking away from the near horizon of extreme Kerr. SciPost Phys 8:098. <https://doi.org/10.21468/SciPostPhys.8.6.089> arXiv:1906.09083 [hep-th]
- Castro A, Verheijden E (2021) Near- $AdS_2$  spectroscopy: classifying the spectrum of operators and interactions in  $N = 2$  4D supergravity. Universe 7(12):475. <https://doi.org/10.3390/universe7120475>. arXiv:2110.04208 [hep-th]
- Castro A, Larsen F, Papadimitriou I (2018) 5D rotating black holes and the  $nAdS_2/nCFT_1$  correspondence. JHEP 10:042. [https://doi.org/10.1007/JHEP10\(2018\)042](https://doi.org/10.1007/JHEP10(2018)042). arXiv:1807.06988 [hep-th]
- Castro A, Godet V, Simón J et al (2021a) Gravitational perturbations from NHEK to Kerr. JHEP 07:218. [https://doi.org/10.1007/JHEP07\(2021\)218](https://doi.org/10.1007/JHEP07(2021)218). arXiv:2102.08060 [hep-th]

- Castro A, Pedraza JF, Toldo C, et al (2021b) Rotating 5D black holes: interactions and deformations near extremality. *SciPost Phys* 11:102. <https://doi.org/10.21468/SciPostPhys.11.6.102>, [arXiv:2106.00649](https://arxiv.org/abs/2106.00649) [hep-th]
- Chamseddine AH, Wyler D (1989) Gauge theory of topological gravity in (1+1)-dimensions. *Phys Lett B* 228:75–78. [https://doi.org/10.1016/0370-2693\(89\)90528-5](https://doi.org/10.1016/0370-2693(89)90528-5)
- Charles AM, Larsen F (2020) A one-loop test of the near-AdS<sub>2</sub>/near-CFT<sub>1</sub> correspondence. *JHEP* 07(07):186. [https://doi.org/10.1007/JHEP07\(2020\)186](https://doi.org/10.1007/JHEP07(2020)186). [arXiv:1908.03575](https://arxiv.org/abs/1908.03575) [hep-th]
- Chen H, Fitzpatrick AL, Kaplan J et al (2017) Degenerate operators and the  $1/c$  expansion: Lorentzian resummations, high order computations, and super-virasoro blocks. *JHEP* 03:167. [https://doi.org/10.1007/JHEP03\(2017\)167](https://doi.org/10.1007/JHEP03(2017)167). [arXiv:1606.02659](https://arxiv.org/abs/1606.02659) [hep-th]
- Chen Y, Gorbenco V, Maldacena J (2021) Bra-ket wormholes in gravitationally prepared states. *JHEP* 02:009. [https://doi.org/10.1007/JHEP02\(2021\)009](https://doi.org/10.1007/JHEP02(2021)009). [arXiv:2007.16091](https://arxiv.org/abs/2007.16091) [hep-th]
- Chowdhury D, Georges A, Parcollet O, et al (2022) Sachdev-Ye-Kitaev models and beyond: Window into non-Fermi liquids. *Rev Mod Phys* 94(3):035004. <https://doi.org/10.1103/RevModPhys.94.035004>, [arXiv:2109.05037](https://arxiv.org/abs/2109.05037) [cond-mat.str-el]
- Christensen SM, Fulling SA (1977) Trace anomalies and the Hawking effect. *Phys Rev D* 15:2088–2104. <https://doi.org/10.1103/PhysRevD.15.2088>
- Coleman SR (1988) Black holes as red herrings: topological fluctuations and the loss of quantum coherence. *Nucl Phys B* 307:867–882. [https://doi.org/10.1016/0550-3213\(88\)90110-1](https://doi.org/10.1016/0550-3213(88)90110-1)
- Comtet A (1987) On the Landau levels on the hyperbolic plane. *Ann Phys* 173:185. [https://doi.org/10.1016/0003-4916\(87\)90098-4](https://doi.org/10.1016/0003-4916(87)90098-4)
- Comtet A, Houston PJ (1985) Effective action on the hyperbolic plane in a constant external field. *J Math Phys* 26:185. <https://doi.org/10.1063/1.526781>
- Cordes S, Moore GW, Ramgoolam S (1995) Lectures on 2-d Yang–Mills theory, equivariant cohomology and topological field theories. *Nucl Phys Proc Suppl* 41:184–244. [https://doi.org/10.1016/0920-5632\(95\)00434-B](https://doi.org/10.1016/0920-5632(95)00434-B), [arXiv:hep-th/9411210](https://arxiv.org/abs/hep-th/9411210) [hep-th]
- Cotler J, Jensen K (2019) A theory of reparameterizations for AdS<sub>3</sub> gravity. *JHEP* 02:079. [https://doi.org/10.1007/JHEP02\(2019\)079](https://doi.org/10.1007/JHEP02(2019)079). [arXiv:1808.03263](https://arxiv.org/abs/1808.03263) [hep-th]
- Cotler J, Jensen K (2021) Gravitational constrained instantons. *Phys Rev D* 104:081501. <https://doi.org/10.1103/PhysRevD.104.L081501>. [arXiv:2010.02241](https://arxiv.org/abs/2010.02241) [hep-th]
- Cotler J, Jensen K (2022) A precision test of averaging in AdS/CFT. *JHEP* 11:070. [https://doi.org/10.1007/JHEP11\(2022\)070](https://doi.org/10.1007/JHEP11(2022)070) [arXiv:2205.12968](https://arxiv.org/abs/2205.12968) [hep-th]
- Cotler J, Jensen K, Maloney A (2020) Low-dimensional de Sitter quantum gravity. *JHEP* 06:048. [https://doi.org/10.1007/JHEP06\(2020\)048](https://doi.org/10.1007/JHEP06(2020)048). [arXiv:1905.03780](https://arxiv.org/abs/1905.03780) [hep-th]
- Cotler JS, Gur-Ari G, Hanada M et al (2017) Black holes and random matrices. *JHEP* 05:118. [https://doi.org/10.1007/JHEP05\(2017\)118](https://doi.org/10.1007/JHEP05(2017)118). [arXiv:1611.04650](https://arxiv.org/abs/1611.04650) [hep-th]. [Erratum: *JHEP* 09, 002 (2018)]
- David F (1988) Conformal field theories coupled to 2-d gravity in the conformal gauge. *Mod Phys Lett A* 3:1651. <https://doi.org/10.1142/S0217732388001975>
- Davies PCW, Fulling SA, Unruh WG (1976) Energy momentum tensor near an evaporating black hole. *Phys Rev D* 13:2720–2723. <https://doi.org/10.1103/PhysRevD.13.2720>
- Davison RA, Fu W, Georges A, et al (2017) Thermoelectric transport in disordered metals without quasiparticles: the Sachdev-Ye-Kitaev models and holography. *Phys Rev B* 95(15):155131. <https://doi.org/10.1103/PhysRevB.95.155131>, [arXiv:1612.00849](https://arxiv.org/abs/1612.00849) [cond-mat.str-el]
- De Vuyst J, Mertens TG (2023) Operational islands and black hole dissipation in JT gravity. *JHEP* 01:027. [https://doi.org/10.1007/JHEP01\(2023\)027](https://doi.org/10.1007/JHEP01(2023)027). [arXiv:2207.03351](https://arxiv.org/abs/2207.03351) [hep-th]
- Deutsch JM (1991) Quantum statistical mechanics in a closed system. *Phys Rev A* 43:2046–2049. <https://doi.org/10.1103/PhysRevA.43.2046>
- Di Francesco P, Ginsparg PH, Zinn-Justin J (1995) 2-D gravity and random matrices. *Phys Rep* 254:1–133. [https://doi.org/10.1016/0370-1573\(94\)00084-G](https://doi.org/10.1016/0370-1573(94)00084-G). [arXiv:hep-th/9306153](https://arxiv.org/abs/hep-th/9306153)
- Dijkgraaf R, Vafa C (2007) Two dimensional Kodaira–Spencer theory and three dimensional Chern–Simons gravity. *arXiv e-prints* [arXiv:0711.1932](https://arxiv.org/abs/0711.1932) [hep-th]
- Dijkgraaf R, Witten E (2018) Developments in topological gravity. *Int J Mod Phys A* 33(30):1830029. <https://doi.org/10.1142/S0217751X18300296>. [arXiv:1804.03275](https://arxiv.org/abs/1804.03275) [hep-th]
- Distler J, Kawai H (1989) Conformal field theory and 2D quantum gravity. *Nucl Phys B* 321:509–527. [https://doi.org/10.1016/0550-3213\(89\)90354-4](https://doi.org/10.1016/0550-3213(89)90354-4)
- Douglas MR, Shenker SH (1990) Strings in less than one-dimension. *Nucl Phys B* 335:635. [https://doi.org/10.1016/0550-3213\(90\)90522-F](https://doi.org/10.1016/0550-3213(90)90522-F)

- Dubovsky S, Gorbenco V, Mirbabayi M (2017) Asymptotic fragility, near  $\text{AdS}_2$  holography and  $T\bar{T}$ . JHEP 09:136. [https://doi.org/10.1007/JHEP09\(2017\)136](https://doi.org/10.1007/JHEP09(2017)136). arXiv:1706.06604 [hep-th]
- Eberhardt L (2022) Off-shell partition functions in 3d gravity. arXiv e-prints arXiv:2204.09789 [hep-th]
- Eberhardt L, Turiaci GJ (2023) 2D dilaton gravity and the Weil–Petersson volumes with conical defects. arXiv e-prints arXiv:2304.14948 [hep-th]
- Ebert S, Ferko C, Sun HY, et al (2022)  $T\bar{T}$  in JT gravity and BF gauge theory. SciPost Phys 13:096. <https://doi.org/10.21468/SciPostPhys.13.4.096>, arXiv:2205.07817 [hep-th]
- Elitzur S, Moore GW, Schwimmer A et al (1989) Remarks on the canonical quantization of the Chern–Simons–Witten theory. Nucl Phys B 326:108–134. [https://doi.org/10.1016/0550-3213\(89\)90436-7](https://doi.org/10.1016/0550-3213(89)90436-7)
- Engelhardt N, Fischetti S, Maloney A (2021) Free energy from replica wormholes. Phys Rev D 103(4):046021. <https://doi.org/10.1103/PhysRevD.103.046021>. arXiv:2007.07444 [hep-th]
- Engelsöy J, Mertens TG, Verlinde H (2016) An investigation of  $\text{AdS}_2$  backreaction and holography. JHEP 07:139. [https://doi.org/10.1007/JHEP07\(2016\)139](https://doi.org/10.1007/JHEP07(2016)139). arXiv:1606.03438 [hep-th]
- Eynard B (2004) Topological expansion for the 1-Hermitian matrix model correlation functions. JHEP 11:031. <https://doi.org/10.1088/1126-6708/2004/11/031>, arXiv:hep-th/0407261 [hep-th]
- Eynard B, Orantin N (2007) Weil–Petersson volume of moduli spaces, Mirzakhani’s recursion and matrix models. arXiv e-prints arXiv:0705.3600 [math-ph]
- Eynard B, Kimura T, Ribault S (2015) Random matrices. arXiv e-prints arXiv:1510.04430 [math-ph]
- Fabbri A, Navarro-Salas J (2005) Modeling black hole evaporation. Imperial College Press, London
- Fan Y, Mertens TG (2022) From quantum groups to Liouville and dilaton quantum gravity. JHEP 05:092. [https://doi.org/10.1007/JHEP05\(2022\)092](https://doi.org/10.1007/JHEP05(2022)092). arXiv:2109.07770 [hep-th]
- Fan Y, Mertens TG (2022) Supergroup structure of Jackiw–Teitelboim supergravity. JHEP 08:002. [https://doi.org/10.1007/JHEP08\(2022\)002](https://doi.org/10.1007/JHEP08(2022)002). arXiv:2106.09353 [hep-th]
- Ferrari F (2021) Gauge theory formulation of hyperbolic gravity. JHEP 03:046. [https://doi.org/10.1007/JHEP03\(2021\)046](https://doi.org/10.1007/JHEP03(2021)046). arXiv:2011.02108 [hep-th]
- Fiola TM, Preskill J, Strominger A et al (1994) Black hole thermodynamics and information loss in two-dimensions. Phys Rev D 50:3987–4014. <https://doi.org/10.1103/PhysRevD.50.3987>. arXiv:hep-th/9403137
- Forste S, Golla I (2017) Nearly  $\text{AdS}_2$  sugra and the super-Schwarzian. Phys Lett B 771:157–161. <https://doi.org/10.1016/j.physletb.2017.05.039>. arXiv:1703.10969 [hep-th]
- Förste S, Kames-King J, Wiesner M (2018) Towards the holographic dual of  $N = 2$  SYK. JHEP 03:028. [https://doi.org/10.1007/JHEP03\(2018\)028](https://doi.org/10.1007/JHEP03(2018)028). arXiv:1712.07398 [hep-th]
- Fukuyama T, Kamimura K (1985) Gauge theory of two-dimensional gravity. Phys Lett B 160:259–262. [https://doi.org/10.1016/0370-2693\(85\)91322-X](https://doi.org/10.1016/0370-2693(85)91322-X)
- Gao P, Jafferis DL, Wall AC (2017) Traversable Wormholes via a double trace deformation. JHEP 12:151. [https://doi.org/10.1007/JHEP12\(2017\)151](https://doi.org/10.1007/JHEP12(2017)151). arXiv:1608.05687 [hep-th]
- Gao P, Jafferis DL, Kolchmeyer DK (2022) An effective matrix model for dynamical end of the world branes in Jackiw–Teitelboim gravity. JHEP 01:038. [https://doi.org/10.1007/JHEP01\(2022\)038](https://doi.org/10.1007/JHEP01(2022)038). arXiv:2104.01184 [hep-th]
- Germani C (2022) Retrieving black hole information from the main Lorentzian saddle point. Phys Rev D 106(6):066018. <https://doi.org/10.1103/PhysRevD.106.066018>. arXiv:2204.13046 [hep-th]
- Gervais JL, Neveu A (1982) Dual string spectrum in Polyakov’s quantization (II). Mode separation. Nucl Phys B 209:125–145. [https://doi.org/10.1016/0550-3213\(82\)90105-5](https://doi.org/10.1016/0550-3213(82)90105-5)
- Gervais JL, Neveu A (1982) The dual string spectrum in Polyakov’s quantization (I). Nucl Phys B 199:59. [https://doi.org/10.1016/0550-3213\(82\)90566-1](https://doi.org/10.1016/0550-3213(82)90566-1)
- Gervais JL, Neveu A (1983) A new quantum solution of Liouville field theory. Phys Lett B 123:86–88. [https://doi.org/10.1016/0370-2693\(83\)90964-4](https://doi.org/10.1016/0370-2693(83)90964-4)
- Gervais JL, Neveu A (1983) New quantum treatment of Liouville field theory. Nucl Phys B 224:329–348. [https://doi.org/10.1016/0550-3213\(83\)90008-1](https://doi.org/10.1016/0550-3213(83)90008-1)
- Ghosh A, Maxfield H, Turiaci GJ (2020) A universal Schwarzian sector in two-dimensional conformal field theories. JHEP 05:104. [https://doi.org/10.1007/JHEP05\(2020\)104](https://doi.org/10.1007/JHEP05(2020)104). arXiv:1912.07654 [hep-th]
- Gibbons GW, Hawking SW (1977) Action integrals and partition functions in quantum gravity. Phys Rev D 15:2752–2756. <https://doi.org/10.1103/PhysRevD.15.2752>
- Gibbons GW, Hawking SW, Perry MJ (1978) Path integrals and the indefiniteness of the gravitational action. Nucl Phys B 138:141–150. [https://doi.org/10.1016/0550-3213\(78\)90161-X](https://doi.org/10.1016/0550-3213(78)90161-X)
- Giddings SB, Strominger A (1988) Loss of incoherence and determination of coupling constants in quantum gravity. Nucl Phys B307:854–866. [https://doi.org/10.1016/0550-3213\(88\)90109-5](https://doi.org/10.1016/0550-3213(88)90109-5)



- Giddings SB, Turiaci GJ (2020) Wormhole calculus, replicas, and entropies. JHEP 09:194. [https://doi.org/10.1007/JHEP09\(2020\)194](https://doi.org/10.1007/JHEP09(2020)194). arXiv:2004.02900 [hep-th]
- Ginsparg PH, Moore GW (1993) Lectures on 2-D gravity and 2-D string theory. In: Proceedings, theoretical advanced study institute (TASI 92): from black holes and strings to particles: Boulder, USA, June 1–26, 1992, pp 277–469. arXiv:hep-th/9304011
- Giombi S, Maloney A, Yin X (2008) One-loop partition functions of 3D gravity. JHEP 08:007. <https://doi.org/10.1088/1126-6708/2008/08/007>. arXiv:0804.1773 [hep-th]
- Godet V, Marteau C (2021) From black holes to baby universes in CGHS gravity. JHEP 07:138. [https://doi.org/10.1007/JHEP07\(2021\)138](https://doi.org/10.1007/JHEP07(2021)138). arXiv:2103.13422 [hep-th]
- Goel A, Lam HT, Turiaci GJ et al (2019) Expanding the Black hole interior: partially entangled thermal states in SYK. JHEP 02:156. [https://doi.org/10.1007/JHEP02\(2019\)156](https://doi.org/10.1007/JHEP02(2019)156). arXiv:1807.03916 [hep-th]
- Goel A, Iliesiu LV, Kruthoff J et al (2021) Classifying boundary conditions in JT gravity: from energy-branes to  $\alpha$ -branes. JHEP 04:069. [https://doi.org/10.1007/JHEP04\(2021\)069](https://doi.org/10.1007/JHEP04(2021)069). arXiv:2010.12592 [hep-th]
- Goto K, Marrochio H, Myers RC et al (2019) Holographic complexity equals which action? JHEP 02:160. [https://doi.org/10.1007/JHEP02\(2019\)160](https://doi.org/10.1007/JHEP02(2019)160). arXiv:1901.00014 [hep-th]
- Goto K, Hartman T, Tajdini A (2021) Replica wormholes for an evaporating 2D black hole. JHEP 04:289. [https://doi.org/10.1007/JHEP04\(2021\)289](https://doi.org/10.1007/JHEP04(2021)289). arXiv:2011.09043 [hep-th]
- Griguolo L, Papalini J, Seminara D (2021) On the perturbative expansion of exact bi-local correlators in JT gravity. JHEP 05:140. [https://doi.org/10.1007/JHEP05\(2021\)140](https://doi.org/10.1007/JHEP05(2021)140). arXiv:2101.06252 [hep-th]
- Griguolo L, Panerai R, Papalini J et al (2022) Nonperturbative effects and resurgence in Jackiw–Teitelboim gravity at finite cutoff. Phys Rev D 105(4):046015. <https://doi.org/10.1103/PhysRevD.105.046015>. arXiv:2106.01375 [hep-th]
- Gross DJ, Migdal AA (1990) Nonperturbative two-dimensional quantum gravity. Phys Rev Lett 64:127. <https://doi.org/10.1103/PhysRevLett.64.127>
- Gross DJ, Kruthoff J, Rolph A, et al (2020)  $T\overline{T}$  in  $AdS_2$  and quantum mechanics. Phys Rev D 101:026011. <https://doi.org/10.1103/PhysRevD.101.026011>. arXiv:1907.04873 [hep-th]
- Grumiller D, Kummer W, Vassilevich DV (2002) Dilaton gravity in two-dimensions. Phys Rep 369:327–430. [https://doi.org/10.1016/S0370-1573\(02\)00267-3](https://doi.org/10.1016/S0370-1573(02)00267-3). arXiv:hep-th/0204253 [hep-th]
- Grumiller D, Ruzziconi R, Zwickel C (2022) Generalized dilaton gravity in 2d. SciPost Phys 12(1):032. <https://doi.org/10.21468/SciPostPhys.12.1.032>. arXiv:2109.03266 [hep-th]
- Hadari S (2019) Near-extremal black holes at late times, backreacted. JHEP 01:214. [https://doi.org/10.1007/JHEP01\(2019\)214](https://doi.org/10.1007/JHEP01(2019)214). arXiv:1811.01022 [hep-th]
- Haehl FM, Rozali M (2018) Fine grained chaos in  $AdS_2$  gravity. Phys Rev Lett 120(12):121601. <https://doi.org/10.1103/PhysRevLett.120.121601>. arXiv:1712.04963 [hep-th]
- Hamilton A, Kabat DN, Lifschytz G et al (2006) Holographic representation of local bulk operators. Phys Rev D 74:066009. <https://doi.org/10.1103/PhysRevD.74.066009>. arXiv:hep-th/0606141
- Hamilton A, Kabat DN, Lifschytz G et al (2006) Local bulk operators in  $AdS/CFT$ : a boundary view of horizons and locality. Phys Rev D 73:086003. <https://doi.org/10.1103/PhysRevD.73.086003>. arXiv:hep-th/0506118
- Harlow D, Jafferis D (2018) The factorization problem in Jackiw–Teitelboim gravity. arXiv e-prints arXiv:1804.01081 [hep-th]
- Hartle JB, Hawking SW (1983) Wave function of the universe. Phys Rev D 28:2960–2975. <https://doi.org/10.1103/PhysRevD.28.2960>
- Hartman T, Maldacena J (2013) Time evolution of entanglement entropy from black hole interiors. JHEP 05:014. [https://doi.org/10.1007/JHEP05\(2013\)014](https://doi.org/10.1007/JHEP05(2013)014). arXiv:1303.1080 [hep-th]
- Hawking SW (1975) Particle creation by black holes. Commun Math Phys 43:199–220. <https://doi.org/10.1007/BF02345020>. [Erratum: Commun. Math. Phys. 46, 206 (1976)]
- Hawking SW (1976) Breakdown of predictability in gravitational collapse. Phys Rev D 14:2460–2473. <https://doi.org/10.1103/PhysRevD.14.2460>
- Heydeman M, Iliesiu LV, Turiaci GJ et al (2022) The statistical mechanics of near-BPS black holes. J Phys A 55(1):014004. <https://doi.org/10.1088/1751-8121/ac3be9>. arXiv:2011.01953 [hep-th]
- Heydeman M, Turiaci GJ, Zhao W (2023) Phases of  $\mathcal{N} = 2$  Sachdev–Ye–Kitaev models. JHEP 01:098. [https://doi.org/10.1007/JHEP01\(2023\)098](https://doi.org/10.1007/JHEP01(2023)098). arXiv:2206.14900 [hep-th]
- Hirano S, Kuroki T (2022) Replica wormholes from Liouville theory. JHEP 01:094. [https://doi.org/10.1007/JHEP01\(2022\)094](https://doi.org/10.1007/JHEP01(2022)094). arXiv:2109.12539 [hep-th]
- Hong J, Larsen F, Liu JT (2019) The scales of black holes with  $nAdS_2$  geometry. JHEP 10:260. [https://doi.org/10.1007/JHEP10\(2019\)260](https://doi.org/10.1007/JHEP10(2019)260). arXiv:1907.08862 [hep-th]

- Horowitz GT (1989) Exactly soluble diffeomorphism invariant theories. *Commun Math Phys* 125:417. <https://doi.org/10.1007/BF01218410>
- Ikeda N (1994) Two-dimensional gravity and nonlinear gauge theory. *Ann Phys* 235:435–464. <https://doi.org/10.1006/aphy.1994.1104>. arXiv:hep-th/9312059
- Ikeda N, Izawa KI (1993) General form of Dilaton gravity and nonlinear gauge theory. *Prog Theor Phys* 90:237–246. <https://doi.org/10.1143/PTP.90.237>. arXiv:hep-th/9304012
- Iliesiu L, Turiaci GJ (2023) upcoming. *Rep Prog Phys*
- Iliesiu LV (2019) On 2D gauge theories in Jackiw–Teitelboim gravity. arXiv e-prints arXiv:1909.05253 [hep-th]
- Iliesiu LV, Turiaci GJ (2021) The statistical mechanics of near-extremal black holes. *JHEP* 05:145. [https://doi.org/10.1007/JHEP05\(2021\)145](https://doi.org/10.1007/JHEP05(2021)145). arXiv:2003.02860 [hep-th]
- Iliesiu LV, Pufu SS, Verlinde H et al (2019) An exact quantization of Jackiw–Teitelboim gravity. *JHEP* 11:091. [https://doi.org/10.1007/JHEP11\(2019\)091](https://doi.org/10.1007/JHEP11(2019)091). arXiv:1905.02726 [hep-th]
- Iliesiu LV, Kruthoff J, Turiaci GJ, et al (2020) JT gravity at finite cutoff. *SciPost Phys* 9:023. <https://doi.org/10.21468/SciPostPhys.9.2.023>. arXiv:2004.07242 [hep-th]
- Iliesiu LV, Mezei M, Sárosi G (2022) The volume of the black hole interior at late times. *JHEP* 07:073. [https://doi.org/10.1007/JHEP07\(2022\)073](https://doi.org/10.1007/JHEP07(2022)073). arXiv:2107.06286 [hep-th]
- Isler K, Trugenberger CA (1989) A gauge theory of two-dimensional quantum gravity. *Phys Rev Lett* 63:834. <https://doi.org/10.1103/PhysRevLett.63.834>
- Jackiw R (1985) Lower dimensional gravity. *Nucl Phys B* 252:343–356. [https://doi.org/10.1016/0550-3213\(85\)90448-1](https://doi.org/10.1016/0550-3213(85)90448-1)
- Jackiw R (1992) Gauge theories for gravity on a line. *Theor Math Phys* 92:979–987. <https://doi.org/10.1007/BF01017075>, [197(1992)], arXiv:hep-th/9206093 [hep-th]
- Jafferis DL, Kolchmeyer DK, Mukhametzhanov B, et al (2022a) JT gravity with matter, generalized ETH, and random matrices. arXiv e-prints arXiv:2209.02131 [hep-th]
- Jafferis DL, Kolchmeyer DK, Mukhametzhanov B, et al (2022b) Matrix models for eigenstate thermalization. arXiv e-prints arXiv:2209.02130 [hep-th]
- Jensen K (2016) Chaos in  $AdS_2$  holography. *Phys Rev Lett* 117(11):111601. <https://doi.org/10.1103/PhysRevLett.117.111601>. arXiv:1605.06098 [hep-th]
- Jensen K, Kachru S, Karch A et al (2011) Towards a holographic marginal Fermi liquid. *Phys Rev D* 84:126002. <https://doi.org/10.1103/PhysRevD.84.126002>. arXiv:1105.1772 [hep-th]
- Johnson CV (2020) Nonperturbative Jackiw–Teitelboim gravity. *Phys Rev D* 101(10):106023. <https://doi.org/10.1103/PhysRevD.101.106023>. arXiv:1912.03637 [hep-th]
- Johnson CV (2021) On the quenched free energy of JT gravity and supergravity. arXiv e-prints arXiv:2104.02733 [hep-th]
- Johnson CV (2022) The microstate physics of JT gravity and supergravity. arXiv e-prints arXiv:2201.11942 [hep-th]
- Joshi LK, Mukhopadhyay A, Soloviev A (2020) Time-dependent  $NAdS_2$  holography with applications. *Phys Rev D* 101(6):066001. <https://doi.org/10.1103/PhysRevD.101.066001>. arXiv:1901.08877 [hep-th]
- Kapec D, Mahajan R, Stanford D (2020) Matrix ensembles with global symmetries and 't Hooft anomalies from 2d gauge theory. *JHEP* 04:186. [https://doi.org/10.1007/JHEP04\(2020\)186](https://doi.org/10.1007/JHEP04(2020)186). arXiv:1912.12285 [hep-th]
- Kar A, Lamprou L, Marteau C, et al (2022) Celestial matrix model. *Phys Rev Lett* 129:201601. <https://doi.org/10.1103/PhysRevLett.129.201601>. arXiv:2205.02240 [hep-th]
- Karch A, Randall L (2001) Locally localized gravity. *JHEP* 05:008. <https://doi.org/10.1088/1126-6708/2001/05/008>. arXiv:hep-th/0011156
- Kitaev A (2014) 2015 Breakthrough prize fundamental physics symposium. <https://www.youtube.com/watch?v=OQ9qN8j7EZI>
- Kitaev A (2015a) A simple model of quantum holography (part 1). <http://online.kitp.ucsb.edu/online/entangled15/kitaev>
- Kitaev A (2015b) Hidden correlations in the Hawking radiation and thermal noise. <http://online.kitp.ucsb.edu/online/joint98/kitaev/>
- Kitaev A, Suh SJ (2019) Statistical mechanics of a two-dimensional black hole. *JHEP* 05:198. [https://doi.org/10.1007/JHEP05\(2019\)198](https://doi.org/10.1007/JHEP05(2019)198). arXiv:1808.07032 [hep-th]
- Kolekar KS, Narayan K (2018)  $AdS_2$  dilaton gravity from reductions of some nonrelativistic theories. *Phys Rev D* 98(4):046012. <https://doi.org/10.1103/PhysRevD.98.046012>. arXiv:1803.06827 [hep-th]



- Kontsevich M (1992) Intersection theory on the moduli space of curves and the matrix Airy function. *Commun Math Phys* 147:1–23. <https://doi.org/10.1007/BF02099526>
- Kourkoulou I, Maldacena J (2017) Pure states in the SYK model and nearly- $AdS_2$  gravity. arXiv e-prints [arXiv:1707.02325](https://arxiv.org/abs/1707.02325) [hep-th]
- Kundu A (2022) Wormholes and holography: an introduction. *Eur Phys J C* 82(5):447. <https://doi.org/10.1140/epjc/s10052-022-10376-z>. arXiv:2110.14958 [hep-th]
- Lam HT, Mertens TG, Turiaci GJ et al (2018) Shockwave S-matrix from Schwarzian quantum mechanics. *JHEP* 11:182. [https://doi.org/10.1007/JHEP11\(2018\)182](https://doi.org/10.1007/JHEP11(2018)182). arXiv:1804.09834 [hep-th]
- Larsen F, Paranjape S (2021) Thermodynamics of near BPS black holes in  $AdS_4$  and  $AdS_7$ . *JHEP* 10:198. [https://doi.org/10.1007/JHEP10\(2021\)198](https://doi.org/10.1007/JHEP10(2021)198). arXiv:2010.04359 [hep-th]
- Larsen F, Zeng Y (2019) Black hole spectroscopy and  $AdS_2$  holography. *JHEP* 04:164. [https://doi.org/10.1007/JHEP04\(2019\)164](https://doi.org/10.1007/JHEP04(2019)164). arXiv:1811.01288 [hep-th]
- Lemos JPS (1996) Thermodynamics of the two-dimensional black hole in the Teitelboim–Jackiw theory. *Phys Rev D* 54:6206–6212. <https://doi.org/10.1103/PhysRevD.54.6206>. arXiv:gr-qc/9608016
- Lemos JPS, Sa PM (1994) Nonsingular constant curvature two-dimensional black hole. *Mod Phys Lett A* 9:771–774. <https://doi.org/10.1142/S0217732394000587>. arXiv:gr-qc/9309023
- Lemos JPS, Sa PM (1994) The Black holes of a general two-dimensional dilaton gravity theory. *Phys Rev D* 49:2897–2908. <https://doi.org/10.1103/PhysRevD.49.2897>. arXiv:gr-qc/9311008. [Erratum: *Phys Rev D* 51, 5967–5968 (1995)]
- Lewkowycz A, Maldacena J (2013) Generalized gravitational entropy. *JHEP* 08:090. [https://doi.org/10.1007/JHEP08\(2013\)090](https://doi.org/10.1007/JHEP08(2013)090). arXiv:1304.4926 [hep-th]
- Lin HW, Maldacena J, Rozenberg L, et al (2023a) Holography for people with no time. *SciPost Phys* 14:150. <https://doi.org/10.21468/SciPostPhys.14.6.150>. arXiv:2207.00407 [hep-th]
- Lin HW, Maldacena J, Rozenberg L, et al (2023b) Looking at supersymmetric black holes for a very long time. *SciPost Phys.* 14:128. <https://doi.org/10.21468/SciPostPhys.14.5.128>. arXiv:2207.00408 [hep-th]
- Louis-Martinez D, Gegenberg J, Kunstatter G (1994) Exact Dirac quantization of all 2-D Dilaton gravity theories. *Phys Lett B* 321:193–198. [https://doi.org/10.1016/0370-2693\(94\)90463-4](https://doi.org/10.1016/0370-2693(94)90463-4). arXiv:gr-qc/9309018
- Mahajan R, Marolf D, Santos JE (2021) The double cone geometry is stable to brane nucleation. *JHEP* 09:156. [https://doi.org/10.1007/JHEP09\(2021\)156](https://doi.org/10.1007/JHEP09(2021)156). arXiv:2104.00022 [hep-th]
- Maldacena J, Qi XL (2018) Eternal traversable wormhole. arXiv e-prints [arXiv:1804.00491](https://arxiv.org/abs/1804.00491) [hep-th]
- Maldacena J, Stanford D (2016) Remarks on the Sachdev–Ye–Kitaev model. *Phys Rev D* 94(10):106002. <https://doi.org/10.1103/PhysRevD.94.106002>. arXiv:1604.07818 [hep-th]
- Maldacena J, Shenker SH, Stanford D (2016) A bound on chaos. *JHEP* 08:106. [https://doi.org/10.1007/JHEP08\(2016\)106](https://doi.org/10.1007/JHEP08(2016)106). arXiv:1503.01409 [hep-th]
- Maldacena J, Stanford D, Yang Z (2016) Conformal symmetry and its breaking in two dimensional nearly Anti-de-Sitter space. *PTEP* 12:12C104. <https://doi.org/10.1093/ptep/ptw124>. arXiv:1606.01857 [hep-th]
- Maldacena J, Stanford D, Yang Z (2017) Diving into traversable wormholes. *Fortsch Phys* 65(5):1700034. <https://doi.org/10.1002/prop.201700034>. arXiv:1704.05333 [hep-th]
- Maldacena J, Milekhin A, Popov F (2018) Traversable wormholes in four dimensions. arXiv e-prints [arXiv:1807.04726](https://arxiv.org/abs/1807.04726) [hep-th]
- Maldacena J, Turiaci GJ, Yang Z (2021) Two dimensional Nearly de Sitter gravity. *JHEP* 01:139. [https://doi.org/10.1007/JHEP01\(2021\)139](https://doi.org/10.1007/JHEP01(2021)139). arXiv:1904.01911 [hep-th]
- Maldacena JM (1999) The Large N limit of superconformal field theories and supergravity. *Int J Theor Phys* 38:1113–1133. <https://doi.org/10.1023/A:1026654312961>, <https://doi.org/10.4310/ATMP.1998.v2.n2.a1>, [Adv. Theor. Math. Phys.2,231(1998)], arXiv:hep-th/9711200 [hep-th]
- Maldacena JM (2003) Eternal black holes in anti-de Sitter. *JHEP* 04:021. <https://doi.org/10.1088/1126-6708/2003/04/021>. arXiv:hep-th/0106112
- Maldacena JM (2003) Non-Gaussian features of primordial fluctuations in single field inflationary models. *JHEP* 05:013. <https://doi.org/10.1088/1126-6708/2003/05/013>. arXiv:astro-ph/0210603
- Maldacena JM, Michelson J, Strominger A (1999) Anti-de Sitter fragmentation. *JHEP* 02:011. <https://doi.org/10.1088/1126-6708/1999/02/011>, arXiv:hep-th/9812073 [hep-th]
- Maldacena JM, Moore GW, Seiberg N, et al (2004) Exact vs. semiclassical target space of the minimal string. *JHEP* 10:020. <https://doi.org/10.1088/1126-6708/2004/10/020>, arXiv:hep-th/0408039

- Maloney A, Witten E (2010) Quantum gravity partition functions in three dimensions. JHEP 02:029. [https://doi.org/10.1007/JHEP02\(2010\)029](https://doi.org/10.1007/JHEP02(2010)029). arXiv:0712.0155 [hep-th]
- Marolf D, Maxfield H (2020) Transcending the ensemble: baby universes, spacetime wormholes, and the order and disorder of black hole information. JHEP 08:044. [https://doi.org/10.1007/JHEP08\(2020\)044](https://doi.org/10.1007/JHEP08(2020)044). arXiv:2002.08950 [hep-th]
- Maxfield H, Turiaci GJ (2021) The path integral of 3D gravity near extremality; or, JT gravity with defects as a matrix integral. JHEP 01:118. [https://doi.org/10.1007/JHEP01\(2021\)118](https://doi.org/10.1007/JHEP01(2021)118). arXiv:2006.11317 [hep-th]
- McGough L, Mezei M, Verlinde H (2018) Moving the CFT into the bulk with  $T\bar{T}$ . JHEP 04:010. [https://doi.org/10.1007/JHEP04\(2018\)010](https://doi.org/10.1007/JHEP04(2018)010). arXiv:1611.03470 [hep-th]
- Mehta ML (2004) Random matrices, 3rd edn. Elsevier Academic Press, Amsterdam
- Mertens TG (2018) The Schwarzian theory—origins. JHEP 05:036. [https://doi.org/10.1007/JHEP05\(2018\)036](https://doi.org/10.1007/JHEP05(2018)036). arXiv:1801.09605 [hep-th]
- Mertens TG (2019) Towards black hole evaporation in Jackiw–Teitelboim gravity. JHEP 07:097. [https://doi.org/10.1007/JHEP07\(2019\)097](https://doi.org/10.1007/JHEP07(2019)097). arXiv:1903.10485 [hep-th]
- Mertens TG (2021) Degenerate operators in JT and Liouville (super)gravity. JHEP 04:245. [https://doi.org/10.1007/JHEP04\(2021\)245](https://doi.org/10.1007/JHEP04(2021)245). arXiv:2007.00998 [hep-th]
- Mertens TG (2022) Quantum exponentials for the modular double and applications in gravity models. arXiv e-prints arXiv:2212.07696 [hep-th]
- Mertens TG, Turiaci GJ (2019) Defects in Jackiw–Teitelboim quantum gravity. JHEP 08:127. [https://doi.org/10.1007/JHEP08\(2019\)127](https://doi.org/10.1007/JHEP08(2019)127). arXiv:1904.05228 [hep-th]
- Mertens TG, Turiaci GJ (2021) Liouville quantum gravity—holography. JT and matrices. JHEP 01:073. [https://doi.org/10.1007/JHEP01\(2021\)073](https://doi.org/10.1007/JHEP01(2021)073). arXiv:2006.07072 [hep-th]
- Mertens TG, Turiaci GJ, Verlinde HL (2017) Solving the Schwarzian via the conformal bootstrap. JHEP 08:136. [https://doi.org/10.1007/JHEP08\(2017\)136](https://doi.org/10.1007/JHEP08(2017)136). arXiv:1705.08408 [hep-th]
- Migdal AA (1975) Recursion equations in gauge theories. Sov Phys JETP 42:413
- Mirzakhani M (2007) Simple geodesics and Weil–Petersson volumes of moduli spaces of bordered Riemann surfaces. Invent Math 167(1):179–222. <https://doi.org/10.1007/s00222-006-0013-2>
- Mirzakhani M (2007) Weil–Petersson volumes and intersection theory on the moduli space of curves. J Am Math Soc 20(01):1–24. <https://doi.org/10.1090/S0894-0347-06-00526-1>
- Moitra U, Sake SK, Trivedi SP, et al (2019a) Jackiw–Teitelboim gravity and rotating black holes. JHEP 11:047. [https://doi.org/10.1007/JHEP11\(2019\)047](https://doi.org/10.1007/JHEP11(2019)047). arXiv:1905.10378 [hep-th]
- Moitra U, Trivedi SP, Vishal V (2019b) Extremal and near-extremal black holes and near-CFT<sub>1</sub>. JHEP 07:055. [https://doi.org/10.1007/JHEP07\(2019\)055](https://doi.org/10.1007/JHEP07(2019)055). arXiv:1808.08239 [hep-th]
- Moitra U, Sake SK, Trivedi SP (2021) Jackiw–Teitelboim gravity in the second order formalism. JHEP 10:204. [https://doi.org/10.1007/JHEP10\(2021\)204](https://doi.org/10.1007/JHEP10(2021)204). arXiv:2101.00596 [hep-th]
- Mukhametzhanov B (2022) Half-wormholes in SYK with one time point. SciPost Phys 12(1):029. <https://doi.org/10.21468/SciPostPhys.12.1.029>, arXiv:2105.08207 [hep-th]
- Nakayama Y (2004) Liouville field theory: a decade after the revolution. Int J Mod Phys A 19:2771–2930. <https://doi.org/10.1142/S0217751X04019500>. arXiv:hep-th/0402009
- Narayan K (2021) Aspects of two-dimensional Dilaton gravity, dimensional reduction, and holography. Phys Rev D 104(2):026007. <https://doi.org/10.1103/PhysRevD.104.026007>. arXiv:2010.12955 [hep-th]
- Nayak P, Shukla A, Soni RM et al (2018) On the dynamics of near-extremal black holes. JHEP 09:048. [https://doi.org/10.1007/JHEP09\(2018\)048](https://doi.org/10.1007/JHEP09(2018)048). arXiv:1802.09547 [hep-th]
- Nojiri S, Odintsov SD (2001) Quantum Dilatonic gravity in (D = 2)-dimensions, (D = 4)-dimensions and (D = 5)-dimensions. Int J Mod Phys A 16:1015–1108. <https://doi.org/10.1142/S0217751X01002968>. arXiv:hep-th/0009202
- Okuyama K, Sakai K (2020) JT gravity, KdV equations and macroscopic loop operators. JHEP 01:156. [https://doi.org/10.1007/JHEP01\(2020\)156](https://doi.org/10.1007/JHEP01(2020)156). arXiv:1911.01659 [hep-th]
- Okuyama K, Sakai K (2020) Multi-boundary correlators in JT gravity. JHEP 08:126. [https://doi.org/10.1007/JHEP08\(2020\)126](https://doi.org/10.1007/JHEP08(2020)126). arXiv:2004.07555 [hep-th]
- Penington G (2020) Entanglement wedge reconstruction and the information paradox. JHEP 09:002. [https://doi.org/10.1007/JHEP09\(2020\)002](https://doi.org/10.1007/JHEP09(2020)002). arXiv:1905.08255 [hep-th]
- Penington G, Shenker SH, Stanford D, et al (2019) Replica wormholes and the black hole interior. arXiv e-prints arXiv:1911.11977 [hep-th]
- Picken R (1989) The propagator for quantum mechanics on a group manifold from an infinite-dimensional analogue of the Duistermaat–Heckman integration formula. J Phys A 22(13):2285

- Polchinski J (1994) Combinatorics of boundaries in string theory. *Phys Rev D* 50:R6041–R6045. <https://doi.org/10.1103/PhysRevD.50.R6041>. [arXiv:hep-th/9407031](https://arxiv.org/abs/hep-th/9407031)
- Polyakov AM (1981) Quantum geometry of bosonic strings. *Phys Lett B* 103:207–210. [https://doi.org/10.1016/0370-2693\(81\)90743-7](https://doi.org/10.1016/0370-2693(81)90743-7)
- Post B, van der Heijden J, Verlinde E (2022) A universe field theory for JT gravity. *JHEP* 05:118. [https://doi.org/10.1007/JHEP05\(2022\)118](https://doi.org/10.1007/JHEP05(2022)118). [arXiv:2201.08859](https://arxiv.org/abs/2201.08859) [hep-th]
- Qi YH, Sin SJ, Yoon J (2019) Quantum correction to chaos in Schwarzian theory. *JHEP* 11:035. [https://doi.org/10.1007/JHEP11\(2019\)035](https://doi.org/10.1007/JHEP11(2019)035). [arXiv:1906.00996](https://arxiv.org/abs/1906.00996) [hep-th]
- Randall L, Sundrum R (1999) A large mass hierarchy from a small extra dimension. *Phys Rev Lett* 83:3370–3373. <https://doi.org/10.1103/PhysRevLett.83.3370>. [arXiv:hep-ph/9905221](https://arxiv.org/abs/hep-ph/9905221)
- Rosenhaus V (2019) An introduction to the SYK model. *J Phys A* 52:323001. <https://doi.org/10.1088/1751-8121/ab2ce1>. [arXiv:1807.03334](https://arxiv.org/abs/1807.03334) [hep-th]
- Rosso F, Turiaci GJ (2022) Phase transitions for deformations of JT supergravity and matrix models. *JHEP* 02:187. [https://doi.org/10.1007/JHEP02\(2022\)187](https://doi.org/10.1007/JHEP02(2022)187). [arXiv:2111.09330](https://arxiv.org/abs/2111.09330) [hep-th]
- Ryu S, Takayanagi T (2006) Aspects of holographic entanglement entropy. *JHEP* 08:045. <https://doi.org/10.1088/1126-6708/2006/08/045>. [arXiv:hep-th/0605073](https://arxiv.org/abs/hep-th/0605073)
- Saad P (2019) Late time correlation functions, Baby Universes, and ETH in JT Gravity. *arXiv e-prints* [arXiv:1910.10311](https://arxiv.org/abs/1910.10311) [hep-th]
- Saad P, Shenker SH, Stanford D (2018) A semiclassical ramp in SYK and in gravity. *arXiv e-prints* [arXiv:1806.06840](https://arxiv.org/abs/1806.06840) [hep-th]
- Saad P, Shenker SH, Stanford D (2019) JT gravity as a matrix integral. *arXiv e-prints* [arXiv:1903.11115](https://arxiv.org/abs/1903.11115) [hep-th]
- Saad P, Shenker S, Yao S (2021a) Comments on wormholes and factorization. *arXiv e-prints* [arXiv:2107.13130](https://arxiv.org/abs/2107.13130) [hep-th]
- Saad P, Shenker SH, Stanford D, et al (2021b) Wormholes without averaging. *arXiv e-prints* [arXiv:2103.16754](https://arxiv.org/abs/2103.16754) [hep-th]
- Saad P, Stanford D, Yang Z, et al (2022) A convergent genus expansion for the plateau. *arXiv e-prints* [arXiv:2210.11565](https://arxiv.org/abs/2210.11565) [hep-th]
- Sachdev S (2010) Holographic metals and the fractionalized Fermi liquid. *Phys Rev Lett* 105:151602. <https://doi.org/10.1103/PhysRevLett.105.151602>. [arXiv:1006.3794](https://arxiv.org/abs/1006.3794) [hep-th]
- Sachdev S (2019) Universal low temperature theory of charged black holes with  $\text{AdS}_2$  horizons. *J Math Phys* 60(5):052303. <https://doi.org/10.1063/1.5092726>. [arXiv:1902.04078](https://arxiv.org/abs/1902.04078) [hep-th]
- Sachdev S, Ye J (1993) Gapless spin fluid ground state in a random, quantum Heisenberg magnet. *Phys Rev Lett* 70:3339. <https://doi.org/10.1103/PhysRevLett.70.3339>. [arXiv:cond-mat/9212030](https://arxiv.org/abs/cond-mat/9212030) [cond-mat]
- Sárosi G (2018)  $\text{AdS}_2$  holography and the SYK model. *PoS Modave2017:001*. <https://doi.org/10.22323/1.323.0001>. [arXiv:1711.08482](https://arxiv.org/abs/1711.08482) [hep-th]
- Schaller P, Strobl T (1994) Diffeomorphisms versus non abelian gauge transformations: an example of  $(1+1)$ -dimensional gravity. *Phys Lett B* 337:266–270. [https://doi.org/10.1016/0370-2693\(94\)90974-1](https://doi.org/10.1016/0370-2693(94)90974-1). [arXiv:hep-th/9401110](https://arxiv.org/abs/hep-th/9401110)
- Schaller P, Strobl T (1994) Poisson structure induced (topological) field theories. *Mod Phys Lett A* 9:3129–3136. <https://doi.org/10.1142/S0217732394002951>. [arXiv:hep-th/9405110](https://arxiv.org/abs/hep-th/9405110)
- Shenker SH (1990) The strength of nonperturbative effects in string theory. In: Cargèse study institute: random surfaces, quantum gravity and strings, pp 809–819
- Shenker SH, Stanford D (2015) Stringy effects in scrambling. *JHEP* 05:132. [https://doi.org/10.1007/JHEP05\(2015\)132](https://doi.org/10.1007/JHEP05(2015)132). [arXiv:1412.6087](https://arxiv.org/abs/1412.6087) [hep-th]
- Smirnov FA, Zamolodchikov AB (2017) On space of integrable quantum field theories. *Nucl Phys B* 915:363–383. <https://doi.org/10.1016/j.nuclphysb.2016.12.014>. [arXiv:1608.05499](https://arxiv.org/abs/1608.05499) [hep-th]
- Spradlin M, Strominger A (1999) Vacuum states for  $\text{AdS}(2)$  black holes. *JHEP* 11:021. <https://doi.org/10.1088/1126-6708/1999/11/021>. [arXiv:hep-th/9904143](https://arxiv.org/abs/hep-th/9904143)
- Srednicki M (1994) Chaos and Quantum Thermalization. *Phys Rev E* 50:888. <https://doi.org/10.1103/PhysRevE.50.888>. [arXiv:cond-mat/9403051](https://arxiv.org/abs/cond-mat/9403051)
- Stanford D (2020) More quantum noise from wormholes. *arXiv e-prints* [arXiv:2008.08570](https://arxiv.org/abs/2008.08570) [hep-th]
- Stanford D, Witten E (2017) Fermionic localization of the Schwarzian theory. *JHEP* 10:008. [https://doi.org/10.1007/JHEP10\(2017\)008](https://doi.org/10.1007/JHEP10(2017)008). [arXiv:1703.04612](https://arxiv.org/abs/1703.04612) [hep-th]
- Stanford D, Witten E (2019) JT gravity and the ensembles of random matrix theory. *arXiv e-prints* [arXiv:1907.03363](https://arxiv.org/abs/1907.03363) [hep-th]

- Stanford D, Yang Z (2020) Finite-cutoff JT gravity and self-avoiding loops. arXiv e-prints [arXiv:2004.08005](https://arxiv.org/abs/2004.08005) [hep-th]
- Stanford D, Yang Z, Yao S (2022) Subleading Weingartens. JHEP 02:200. [https://doi.org/10.1007/JHEP02\(2022\)200](https://doi.org/10.1007/JHEP02(2022)200). [arXiv:2107.10252](https://arxiv.org/abs/2107.10252) [hep-th]
- Strominger A, Vafa C (1996) Microscopic origin of the Bekenstein–Hawking entropy. Phys Lett B 379:99–104. [https://doi.org/10.1016/0370-2693\(96\)00345-0](https://doi.org/10.1016/0370-2693(96)00345-0). [arXiv:hep-th/9601029](https://arxiv.org/abs/hep-th/9601029)
- Suh SJ (2020) Dynamics of black holes in Jackiw–Teitelboim gravity. JHEP 03:093. [https://doi.org/10.1007/JHEP03\(2020\)093](https://doi.org/10.1007/JHEP03(2020)093). [arXiv:1912.00861](https://arxiv.org/abs/1912.00861) [hep-th]
- Susskind L (2016) Computational complexity and black hole horizons. Fortsch Phys 64:24–43. <https://doi.org/10.1002/prop.201500092>. [arXiv:1403.5695](https://arxiv.org/abs/1403.5695) [hep-th]. [Addendum: Fortsch Phys 64, 44–48 (2016)]
- Suzuki K, Takayanagi T (2021) JT gravity limit of Liouville CFT and matrix model. JHEP 11:137. [https://doi.org/10.1007/JHEP11\(2021\)137](https://doi.org/10.1007/JHEP11(2021)137). [arXiv:2108.12096](https://arxiv.org/abs/2108.12096) [hep-th]
- Teitelboim C (1983) Gravitation and Hamiltonian structure in two space-time dimensions. Phys Lett 126B:41–45. [https://doi.org/10.1016/0370-2693\(83\)90012-6](https://doi.org/10.1016/0370-2693(83)90012-6)
- ‘t Hooft G (1974) A planar diagram theory for strong interactions. Nucl Phys B 72:461. [https://doi.org/10.1016/0550-3213\(74\)90154-0](https://doi.org/10.1016/0550-3213(74)90154-0)
- Turiaci GJ (2019) An inelastic bound on chaos. JHEP 07:099. [https://doi.org/10.1007/JHEP07\(2019\)099](https://doi.org/10.1007/JHEP07(2019)099). [arXiv:1901.04360](https://arxiv.org/abs/1901.04360) [hep-th]
- Turiaci GJ, Witten E (2023)  $\mathcal{N} = 2$  JT Supergravity and matrix models. arXiv e-prints [arXiv:2305.19438](https://arxiv.org/abs/2305.19438) [hep-th]
- Turiaci GJ, Usatyuk M, Weng WW (2021) 2D dilaton-gravity, deformations of the minimal string, and matrix models. Class Quant Grav 38(20):204001. <https://doi.org/10.1088/1361-6382/ac25df>. [arXiv:2011.06038](https://arxiv.org/abs/2011.06038) [hep-th]
- Witten E (1988) Coadjoint orbits of the Virasoro Group. Commun Math Phys 114:1. <https://doi.org/10.1007/BF01218287>
- Witten E (1991) On quantum gauge theories in two dimensions. Commun Math Phys 141:153–209. <https://doi.org/10.1007/BF02100009>
- Witten E (1991) Two-dimensional gravity and intersection theory on moduli space. Surveys Diff Geom 1:243–310. <https://doi.org/10.4310/SDG.1990.v1.n1.a5>
- Witten E (1992) Two dimensional gauge theories revisited. J Geom Phys 9:303–368. [https://doi.org/10.1016/0393-0440\(92\)90034-X](https://doi.org/10.1016/0393-0440(92)90034-X). [arXiv:hep-th/9204083](https://arxiv.org/abs/hep-th/9204083)
- Witten E (2007) Three-dimensional gravity revisited. arXiv e-prints [arXiv:0706.3359](https://arxiv.org/abs/0706.3359) [hep-th]
- Witten E (2020) Matrix models and deformations of JT gravity. arXiv e-prints [arXiv:2006.13414](https://arxiv.org/abs/2006.13414) [hep-th]
- Witten E (2021) A note on complex spacetime metrics. arXiv e-prints [arXiv:2111.06514](https://arxiv.org/abs/2111.06514) [hep-th]
- Witten E, Yau ST (1999) Connectedness of the boundary in the AdS/CFT correspondence. Adv Theor Math Phys 3:1635–1655. <https://doi.org/10.4310/ATMP.1999.v3.n6.a1>. [arXiv:hep-th/9910245](https://arxiv.org/abs/hep-th/9910245)
- Yang Z (2019) The quantum gravity dynamics of near extremal black holes. JHEP 05:205. [https://doi.org/10.1007/JHEP05\(2019\)205](https://doi.org/10.1007/JHEP05(2019)205). [arXiv:1809.08647](https://arxiv.org/abs/1809.08647) [hep-th]
- York JW Jr. (1972) Role of conformal three geometry in the dynamics of gravitation. Phys Rev Lett 28:1082–1085. <https://doi.org/10.1103/PhysRevLett.28.1082>
- Zamolodchikov AB, Zamolodchikov AB (2001) Liouville field theory on a pseudosphere. arXiv e-prints [arXiv:hep-th/0101152](https://arxiv.org/abs/hep-th/0101152)
- Zurek WH (1982) Entropy evaporated by a black hole. Phys Rev Lett 49:1683–1686. <https://doi.org/10.1103/PhysRevLett.49.1683>