

# Solving puzzles in deformed JT gravity: phase transitions and non-perturbative effects

Clifford V. Johnson<sup>a</sup> and Felipe Rosso<sup>a,b</sup>

<sup>a</sup>*Department of Physics and Astronomy, University of Southern California,  
Los Angeles, CA 90089, U.S.A.*

<sup>b</sup>*Kavli Institute for Theoretical Physics, University of California,  
Santa Barbara, CA 93106, U.S.A.*

*E-mail:* [johnson1@usc.edu](mailto:johnson1@usc.edu), [felipero@usc.edu](mailto:felipero@usc.edu)

**ABSTRACT:** Recent work has shown that certain deformations of the scalar potential in Jackiw-Teitelboim gravity can be written as double-scaled matrix models. However, some of the deformations exhibit an apparent breakdown of unitarity in the form of a negative spectral density at disc order. We show here that the source of the problem is the presence of a multi-valued solution of the leading order matrix model string equation. While for a class of deformations we fix the problem by identifying a first order phase transition, for others we show that the theory is both perturbatively and non-perturbatively inconsistent. Aspects of the phase structure of the deformations are mapped out, using methods known to supply a non-perturbative definition of undeformed JT gravity. Some features are in qualitative agreement with a semi-classical analysis of the phase structure of two-dimensional black holes in these deformed theories.

**KEYWORDS:** 2D Gravity, Matrix Models, Models of Quantum Gravity

**ARXIV EPRINT:** [2011.06026](https://arxiv.org/abs/2011.06026)

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Summary of results	3
<b>2</b>	<b>Matrix models</b>	<b>5</b>
2.1	Double-scaling limit	5
2.2	Leading perturbative analysis	8
2.3	Non-perturbative completion	9
<b>3</b>	<b>Deformed JT gravity: model A</b>	<b>11</b>
3.1	Semi-classical approximation	12
3.2	Perturbative genus expansion	15
3.3	Non-perturbative effects	18
<b>4</b>	<b>Deformed JT gravity: model B</b>	<b>20</b>
4.1	Semi-classical approximation	21
4.2	Perturbative genus expansion	22
4.3	Non-perturbative effects	23
<b>5</b>	<b>Final remarks</b>	<b>27</b>
<b>A</b>	<b>Inverting the Abel transform</b>	<b>29</b>
<b>B</b>	<b>Further details of the non-perturbative definition</b>	<b>29</b>
<b>C</b>	<b>Deriving a Bäcklund transformation</b>	<b>30</b>

---

## 1 Introduction

Low dimensional gravity theories have proven to be extremely fruitful playgrounds for exploring important questions in quantum gravity that are substantially more difficult to address in higher dimensions. One of the simplest examples is Jackiw-Teitelboim (JT) gravity [1, 2], a two-dimensional theory including a scalar (or “dilaton”) with the following Euclidean action:

$$I_{\text{JT}}[g_{\mu\nu}, \phi] = -S_0 \chi - \frac{1}{2} \int d^2x \sqrt{g} \phi (R + 2) + \text{boundary terms}, \quad (1.1)$$

where  $\chi$  is the Euler characteristic of the spacetime and  $S_0$  a constant. This simple theory captures the low energy sector of both near-extremal black holes and branes [3–8] and the SYK model [9–12] (see ref. [13] for a review). Most importantly, the full quantum theory,

where the spacetime is allowed to change topology, has been recently solved to all orders in perturbation theory in the topological expansion [14]. This was accomplished by using a variety of powerful mathematical results [15–17] to show that the full topological expansion of the Euclidean partition function:

$$Z(\beta) = \int_{-\infty}^{+\infty} \rho(E) e^{-\beta E} dE, \tag{1.2}$$

is reproduced by a double-scaled [18–20] random Hermitian matrix model.

Given this success, it is reasonable to explore more complicated two-dimensional dilation theories, obtained by deforming the JT gravity action (1.1). Studies of such deformations were recently initiated in refs. [21, 22], by considering the following class of theories:

$$I_{\text{dJT}}[g_{\mu\nu}, \phi] = I_{\text{JT}}[g_{\mu\nu}, \phi] - \frac{1}{2} \int d^2x \sqrt{g} U(\phi), \quad \text{where} \quad U(\phi) = 2 \sum_{i=1}^r \lambda_i e^{-2\pi(1-\alpha_i)\phi}. \tag{1.3}$$

For this to be a deformation of JT gravity,  $U(\phi)$  is required to vanish at  $\phi \rightarrow +\infty$  and so  $\alpha_i \leq 1$ . Further constraining  $\alpha_i \in (0, 1/2)$ , the Euclidean partition function of this theory was computed in refs. [21, 22], where the exponential terms in  $U(\phi)$  were interpreted as inserting sharp defects in the ordinary JT path integral. In a similar fashion to the work of ref. [14], the topological expansion of the partition function was shown to agree with that of a double scaled Hermitian matrix model.<sup>1</sup> The matrix model expansion is defined from the leading genus spectral density  $\rho_0(E)$ , obtained from the gravitational disc partition function after an inverse Laplace transform. All higher genus contributions are then obtained from  $\rho_0(E)$  via recursion relations derived from the matrix model.

While for generic deformations  $U(\phi)$  this procedure works nicely, it was noted in refs. [21, 22] that in certain cases the leading genus spectral density  $\rho_0(E)$  provided by the disc partition function is not positive definite. Given that  $\rho_0(E)$  is the seed of the recursion relation, starting from a sick spectral density results in an ill-defined perturbative expansion. This puzzle was left unresolved in refs. [21, 22], but it was speculated that either non-perturbative effects or phase transitions might resolve the issue.

This paper was motivated by the desire to better understand these deformations and determine whether the problems can be resolved using recently developed techniques that provide an alternative formulation of the matrix model description of JT gravity in terms of certain combinations of minimal string models [24, 25]. As emphasized in refs. [25–27], this formulation is particularly well-suited to enable the extraction of non-perturbative physics. Fully specifying the matrix model is equivalent to determining a function  $u(x)$  for  $x \in \mathbb{R}$ , which is the potential of an associated quantum mechanics problem with  $\hbar = e^{-S_0}$ . While the function  $u(x)$  satisfies a complicated non-linear ordinary differential equation (called the “string equation”), its leading piece as  $\hbar \rightarrow 0$ ,  $u_0(x)$ , satisfies a simple algebraic constraint.

*A central result of our work is the observation that the non-positive spectral density  $\rho_0(E)$  arising in some JT gravity deformations is only a symptom of the sickness of*

---

<sup>1</sup>The matching between the gravity and matrix model computations was made very explicit for arbitrary asymptotic boundaries and genus in ref. [21], using earlier results in refs. [17, 23].

the models, and the actual virus is a multi-valued potential  $u_0(x)$ . Sometimes the problem can be eradicated by identifying a phase transition, and at other times it cannot. In such cases, we give evidence that the non-perturbative physics (at least using the definitions employed here) does not exist. Furthermore, we explore the physics of the deformations in the semi-classical gravity approximation, the leading  $\hbar$  behavior of the matrix model, and the full non-perturbative definition, contrasting the three approaches and identifying several phenomena. We now summarize our core results in more detail, briefly explaining the salient features.

### 1.1 Summary of results

We start in section 2 by introducing the matrix model technology used throughout this work. An important issue we address is the non-perturbative formulation. The Hermitian matrix model definition [14] of JT gravity is afflicted by non-perturbative instabilities long known [28, 29] to be present in certain double-scaled Hermitian matrix models. In ref. [25] an alternative matrix model definition of JT gravity was provided which is free from such instabilities and reproduces the perturbative physics of ref. [14]. Here we present an extension of ref. [25]’s framework that is able to naturally incorporate the deformed JT gravity physics and serve as a non-perturbative definition. It reduces to the non-perturbatively stable definition when the deformations are removed and will therefore act as a sharp tool for determining if some deformations are non-perturbatively ill-defined (at least within a framework that works for ordinary JT gravity). The matrix model origins of the string equation are discussed.

Section 3 discusses a particular example of a deformation of JT gravity given by:

$$U(\phi) = 2\lambda \left( e^{-2\pi(1-\alpha_1)\phi} - e^{-2\pi(1-\alpha_2)\phi} \right), \tag{1.4}$$

which satisfies  $U(0) = 0$ . To start, we compute the Euclidean partition function in the semi-classical approximation and uncover an interesting variety of phase transitions of the two-dimensional black holes of the theory (see figures 2 and 3). We continue by studying the full Euclidean partition function to leading order in genus (disc topology), computed in refs. [21, 22]. The leading genus spectral density obtained from the results in ref. [22] is given by:

$$\rho_0(E) = \frac{\sinh(2\pi\sqrt{E})}{4\pi^2\hbar} + \lambda \frac{\cosh(2\pi\alpha_1\sqrt{E}) - \cosh(2\pi\alpha_2\sqrt{E})}{2\pi\hbar\sqrt{E}}, \tag{1.5}$$

where  $\hbar \equiv e^{-S_0}$  and the threshold energy is  $E_0 = 0$ . Expanding this expression for small energies, one finds  $\rho_0(E) \propto \sqrt{E} + \mathcal{O}(E^{3/2})$ , where the proportionality factor in the leading term is negative for  $\lambda > \lambda_c$ , see (3.15).

To understand the origin of this issue from the matrix model perspective, the model is reinterpreted as a particular combination [24] of minimal models, which enables a formulation to all orders in  $\hbar$  and beyond [25]. The central quantity in this formalism is the string equation, that determines a potential  $u(x)$ . The leading genus string equation associated to this model is given by:

$$\mathcal{R}_0[u_0, x] = \frac{\sqrt{u_0}}{2\pi} I_1(2\pi\sqrt{u_0}) + \lambda(I_0(2\pi\alpha_1\sqrt{u_0}) - I_0(2\pi\alpha_2\sqrt{u_0})) + x = 0, \tag{1.6}$$

(where  $I_n$  is the  $n$ th modified Bessel function), which provides an implicit definition of the leading genus potential  $u_0(x) \equiv \lim_{\hbar \rightarrow 0} u(x)$ . We first show that  $\rho_0(E)$  in (1.5) goes negative whenever the corresponding potential  $u_0(x)$  defined through equation (1.6) becomes multi-valued, see figure 4. From this perspective, the issue is easily solved by picking the (unique) solution to the implicit equation (1.6) that yields a single valued function  $u_0(x)$  in the region  $x < 0$ . This modifies the leading spectral density from (1.5) to:

$$\rho_0(E) = \frac{1}{2\pi\hbar} \int_{E_0}^E \frac{du_0}{\sqrt{E-u_0}} (\partial_{u_0} \mathcal{R}_0), \tag{1.7}$$

where  $E_0 \geq 0$  is determined from the largest solution to  $\mathcal{R}_0[E_0, 0] = 0$ . While for deformations for which  $E_0 = 0$  the integral (1.7) reduces to (1.5), when  $E_0 > 0$  the correct positive definite answer is instead obtained from the integral expression. The transition of  $E_0$  between the two regimes is not analytic (see figure 5) and corresponds to a phase transition which qualitatively matches with results obtained from the semi-classical analysis.<sup>2</sup>

At this point we must comment on the resolution of this first puzzle and its relation to the calculations in ref. [21]. In that work the disc partition function was shown to be given by (1.7) *irrespective* of the value of  $U(0)$ . As a result, the gravitational computations in ref. [21] already have the correct spectral density when  $U(0) = 0$ , given by equation (1.7) instead of equation (1.5). However, the issue *is* relevant from the perspective of ref. [22], where the gravitational calculation with  $U(0) = 0$  unequivocally leads to the poorly defined result in (1.5). It would therefore be interesting to understand how the derivation of (1.5) in ref. [22] must be modified in order to yield the correct result.

Section 3 concludes with a computation of the full spectral density  $\rho(E)$  including both higher genus and non-perturbative contributions (see figures 6 and 7). Non-perturbative effects generate corrections to  $\rho_0(E)$  in (1.7) analogous to those previously computed for ordinary JT gravity in refs. [25, 27], with the difference that in this case  $E_0$  can be non-zero.

Section 4 repeats the analysis for an example of a different class of deformation of JT gravity. Here,  $U(0) \neq 0$ , and:

$$U(\phi) = 2\lambda e^{-2\pi(1-\alpha)\phi}. \tag{1.8}$$

After computing the Euclidean partition function in the semi-classical approximation, the full partition function to leading order in genus [21, 22] is studied. The spectral density in this case is given by<sup>3</sup>

$$\rho_0(E) = \frac{1}{2\pi\hbar} \int_{E_0}^E \frac{du_0}{\sqrt{E-u_0}} (\partial_{u_0} \mathcal{R}_0), \quad \mathcal{R}_0 = \frac{\sqrt{u_0}}{2\pi} I_1(2\pi\sqrt{u_0}) + \lambda I_0(2\pi\alpha\sqrt{u_0}) + x, \tag{1.9}$$

with  $E_0$  again determined from the largest solution to  $\mathcal{R}_0[E_0, 0] = 0$ . While the integral expression is the same answer as given in (1.7) for  $U(0) = 0$ , when  $U(0) \neq 0$  the result of

---

<sup>2</sup>We should note that the semi-classical and matrix model computations apply to different  $\alpha_i$  regimes,  $\alpha_i \sim 1$  and  $\alpha_i \in (0, 1/2)$  respectively. While certain quantitative aspects do not match exactly (like the exact location and order of the transition), the qualitative agreement between the two approaches is quite remarkable.

<sup>3</sup>We are using the same conventions as in ref. [21]. This is obtained from ref. [22] by changing the integration variable in eq. (8.12) to  $u_0 = (1-s^2)E + s^2E_0$ .

the integral is not necessarily positive (see figure 9 for some examples). As in section 3, the issue originates in a multi-valued solution  $u_0(x)$  to the leading genus string equation (1.9), but in this case in the region  $x < 0$  (see figure 9). This makes the issue substantially different and ultimately results in a breakdown of the perturbative expansion of the matrix model. We also show that certain models which might appear to be well defined (since they have  $\rho_0(E)$  positive definite) still suffer from this perturbative breakdown.

To determine whether the model can be made sense of even when the perturbative expansion breaks down, the system is studied non-perturbatively. We find that the non-perturbative definition also breaks down precisely when  $u_0(x)$  becomes multi-valued in the region  $x < 0$ . Hence, we argue that deformations of JT gravity for which  $u_0(x)$  presents such multi-valuedness issues are both perturbatively and non-perturbatively unstable.

The paper concludes in section 5 with a brief discussion and some thoughts about future directions. Several appendices contain details and results used in the main text.

## 2 Matrix models

This section introduces some of the random matrix model tools used throughout this work. The introduction will be rather brief (for longer treatments see e.g., refs. [30–32]) but will help establish some notation, and contextualize key aspects of the perturbative and non-perturbative formulations.

### 2.1 Double-scaling limit

Both random Hermitian matrix models and random complex matrix models will be relevant here, but it is traditional to begin with the former. The model is defined through a probability measure  $dH e^{-N \text{Tr} V(H)}$  with  $H$  an  $N \times N$  Hermitian matrix and  $V(H)$  a polynomial potential generalizing Wigner’s prototype Gaussian case. Expectation values of matrix observables  $\mathcal{O}$  are computed as:

$$\langle \mathcal{O} \rangle \equiv \frac{1}{\mathcal{Z}} \int dH \mathcal{O} e^{-\frac{N}{\gamma} \text{Tr} V(H)} \quad \text{where} \quad \mathcal{Z} \equiv \int dH e^{-\frac{N}{\gamma} \text{Tr} V(H)}. \quad (2.1)$$

One of the central observables of the theory is the spectral density

$$\rho(E) \equiv \frac{1}{N} \left\langle \sum_{i=1}^N \delta(E - \alpha_i) \right\rangle, \quad (2.2)$$

where  $\alpha_i \in \mathbb{R}$  are the eigenvalues of  $H$ .<sup>4</sup> In the  $N \rightarrow \infty$  limit  $\rho(E)$  becomes a smooth density function, beginning at some classical threshold energy  $E_0$ .

The central idea is to use random matrix models to compute universal properties of sums over random two-dimensional surfaces. Those surfaces can be seen [33, 34] to be present in the definition above by expanding the integral  $\mathcal{Z}$  in terms of Feynman diagrams (the quadratic term of  $V(H)$  gives a propagator, higher order terms give vertices), which are graph duals of tessellations of the surfaces into polygons. The power  $N^\chi$  is carried

---

<sup>4</sup>Denoted by  $\alpha_i$  since the more standard notation for them,  $\lambda_i$ , is used for the deformation parameters of JT gravity.

by a graph/polygonization with Euler characteristic  $\chi = 2 - 2g - b$  where  $g$  and  $b$  are the number of handles and boundaries, respectively. Universal physics from sum over surfaces (i.e., properties that do not depend on model-dependent details such as whether triangles or pentagons were used for the discretization) are picked out by taking the double-scaling limit [18–20]. This involves taking  $N \rightarrow \infty$ , while simultaneously tuning the polynomial coefficients in the potential  $V(H)$  to a critical point dominated by smooth surfaces that are very large compared to the number of constituent polygons. The  $1/N$  corrections yield the topological expansion. The critical point is characterized by the rate of vanishing of the leading spectral density  $\rho(E)$  at one of its endpoints (see e.g., ref. [35]). The generic behavior is Wigner’s semi-circle law:  $\rho \sim (E - E_0)^{\frac{1}{2}}$ . An additional  $(k - 1)$  zeros coincident at the endpoint  $E_0$  define the critical behavior corresponding to a distinct model [36], the  $k$ th model:  $\rho \sim (E - E_0)^{k - \frac{1}{2}}$ .<sup>5</sup>

Non-perturbative physics that lies beyond the topological  $1/N$  expansion is accessible too. A starting point for extracting it [37] is a family of polynomials  $P_n(\alpha) = \alpha^n +$  lower powers (where  $n = 1 \cdots N$ ), that are orthogonal with respect to the measure  $d\mu(\alpha) = d\alpha e^{-\frac{N}{\gamma} V(\alpha)}$ :

$$\int d\mu(\alpha) P_n(\alpha) P_m(\alpha) \propto \delta_{mn}, \tag{2.3}$$

and for which a recursion relation can be written (e.g. for even  $V(H)$ ) as:  $\alpha P_n(\alpha) = P_{n+1}(\alpha) + R_n P_{n-1}(\alpha)$ . All matrix model observables can be re-written in terms of the recursion coefficients  $R_n$  and identities satisfied by the defining matrix integral can be used to derive recursion relations for them. The  $P_n(\alpha)$  actually define a finite dimensional Hilbert space description of the system, built from  $|n\rangle \propto P_n(\alpha)$ , with  $\langle n|m\rangle = \delta_{nm}$ , within which observables, as discussed in equation (2.1), are associated with operators.

In the double scaling limit, the Hilbert space supplied by the  $P_n(\alpha)$  becomes infinite dimensional;  $|n\rangle \rightarrow |x\rangle$  with  $x \in \mathbb{R}$ , and an effective quantum mechanical system emerges from which matrix model observables can be computed [18, 28]. It is governed by a simple Schrödinger Hamiltonian:

$$\mathcal{H}[u] \equiv -\hbar^2 \frac{\partial^2}{\partial x^2} + u(\hat{x}), \tag{2.4}$$

with  $\hat{x}|x\rangle = x|x\rangle$ , and  $\hbar$  the scaling part of  $1/N$  in the limit. The potential  $u(x)$  arises as the scaling part of the  $R_n$  in the limit, and its functional form is determined by an ordinary differential equation called the “string equation” (to be discussed below) which is the scaling part of the aforementioned recursion relations for the  $R_n$ .

An important observable is the expectation value of a “macroscopic loop” (a single boundary) of length  $\beta$ , which in this formalism is:

$$Z(\beta) \equiv \langle \text{Tr} e^{-\beta H} \rangle = \int_{-\infty}^0 dx \langle x | e^{-\beta \mathcal{H}[u]} | x \rangle. \tag{2.5}$$

It is important to notice the integration range for  $x$ . The double scaling limit captures the universal physics to be found in the infinitesimal neighbourhood of the spectral density’s

---

<sup>5</sup>These were termed “multi-critical” points in the old language, but we will avoid that usage here. In the string theory context, where the random surfaces represent string worldsheets, these models are the  $(2k - 1, 2)$  “minimal string theories”, in the contemporary nomenclature.

endpoint (where the critical behaviour was specified earlier). Recall that the index  $n$  labeled the ordered  $N$  energies, and so at large  $N$  they are labeled by a continuous coordinate  $X = n/N \in [0, 1]$ . The coordinate  $x \in (-\infty, +\infty)$  describes the local physics of the  $X = 0$  endpoint, and in the present conventions they are related as  $X = 0 - x\delta^p$  (where  $p$  is some positive power and  $\delta \rightarrow 0$  as  $N \rightarrow \infty$ ). Crucially, the  $x < 0$  regime (integrated over in equation (2.5)) is the place where the surviving part of the spectral density classically has support, and  $x = 0$  is the location of the endpoint. Very importantly, note that *the full quantum mechanics is defined on the whole of the range of  $x$* . The physics from the  $x > 0$  regime controls non-perturbative aspects of the system.

The part of the spectral density that survives in the double scaling limit,  $\rho(E)$ , can be written in this language too. Starting with equation (2.5) and inserting a complete set of eigenstates of the Hamiltonian,  $\mathcal{H}|\psi_E\rangle = E|\psi_E\rangle$ :

$$Z(\beta) = \int_{E_0}^{\infty} dE \rho(E) e^{-\beta E}, \quad \rho(E) = \int_{-\infty}^0 dx |\psi_E(x)|^2, \quad (2.6)$$

where  $E_0$  is the lowest or “threshold” energy in the spectrum.

The macroscopic loop expectation value is denoted  $Z(\beta)$  in the above (as done in equation (1.2)) because it is of the same structural form as JT gravity, i.e. a sum over all surfaces with a single boundary of fixed length  $\beta$ . Evidently, the central object in this formalism is the potential  $u(x)$  that determines the Hamiltonian (2.4). For the appropriate choice of  $u(x)$ , equation (2.6) *will* be a definition of the JT gravity matrix model that is alternative to the recursion methods of ref. [14]. The core idea is that the full non-perturbative behavior can be accessed from  $u(x)$  obtained as the full solution to a differential equation, the string equation.

The all-orders perturbative expansion for  $u(x)$  is also nicely packaged in the string equation. However, it is important to note here that the string equation arising from Hermitian matrix models is non-perturbatively unstable (this will be discussed in section 2.2). Therefore, the Hermitian matrix model string equation that we now write, should only be used perturbatively in  $\hbar$  [18–20]:

$$\mathcal{R}[u, x] \equiv \sum_{k=0}^{\infty} t_k \tilde{R}_k[u] + x = 0. \quad (2.7)$$

The objects  $\tilde{R}_k[u]$  are  $k$ th order polynomials in  $u(x)$  and its derivatives called the Gel’fand-Dikii polynomials [38], normalized so that the coefficient of  $u^k$  is unity. They satisfy a simple recursion relation [38]:

$$\frac{2k+1}{2(k+1)} \tilde{R}'_{k+1} = u \tilde{R}'_k + \frac{1}{2} u' \tilde{R}_k - \frac{\hbar^2}{4} \tilde{R}_k''', \quad (2.8)$$

where  $\tilde{R}_0[u] = 1$  and primes are derivatives with respect to  $x$ . The coefficients  $t_k$  in equation (2.7) indicate how much the  $k$ th model contributes to the form of the potential  $u(x)$  and essentially define the double scaled model. To construct the matrix model equivalent to JT gravity, a natural strategy is to determine the combination of  $t_k$  that yields the leading perturbative  $\rho_0(E)$ , known from disc level computations in JT gravity. This determines the leading contribution to  $u(x)$ , denoted  $u_0(x)$ . This is described next.



## 2.2 Leading perturbative analysis

Writing a power series expansion for the potential  $u(x) = \sum_{n=0}^{\infty} u_n(x)\hbar^n$  and inserting it into the string equation (2.7), a perturbative analysis of the physics can be performed. The leading behavior (the classical limit in the quantum mechanics of equation (2.4)) of the potential  $u_0(x)$  satisfies the following simple algebraic equation:

$$\mathcal{R}_0[u_0, x] = \sum_{k=0}^{\infty} t_k u_0^k + x = 0 . \tag{2.9}$$

Subtleties regarding the implicit definition of  $u_0 = u_0(x)$  from this equation will play an important role in this paper. A simple and useful formula can be written for the loop expectation value (the putative JT gravity partition function)  $Z(\beta)$  defined in equation (2.5) at leading (disc) order in genus  $Z_0(\beta) = \lim_{\hbar \rightarrow 0} Z(\beta)$  in terms of  $u_0(x)$ . Using that the momentum  $\hat{p} = -i\hbar\partial_x$  and position operators  $\hat{x}$  commute to leading order in  $\hbar$  and inserting a complete set of momentum eigenstates  $|p\rangle$  we can solve the Gaussian integral in  $p$  and find:

$$Z_0(\beta) = \frac{1}{2\hbar\sqrt{\pi\beta}} \int_{-\infty}^0 dx e^{-\beta u_0(x)} , \tag{2.10}$$

where we have used  $\langle p|x\rangle = e^{ixp/\hbar}/\sqrt{2\pi\hbar}$ . Changing the integration variable to  $u_0$  and using (2.9), we can obtain the following formula for the spectral density  $\rho_0(E)$  after applying an inverse Laplace transform:

$$\rho_0(E) = \frac{1}{2\pi\hbar} \int_{E_0}^E \frac{du_0}{\sqrt{E-u_0}} (\partial_{u_0} \mathcal{R}_0) , \quad \text{where} \quad E_0 = u_0(0) . \tag{2.11}$$

This formula has played a central role in recent JT gravity explorations [24, 25]. In fact, in the mathematical literature this integral transform (when  $E_0 = 0$ ) is known as the Abel transform, and in such cases, it can be explicitly inverted. This is described in appendix A, the final result is:

$$E_0 = u(0) = 0 \quad \implies \quad \mathcal{R}_0[u_0, x] = 2\hbar \int_0^{u_0} \frac{dE}{\sqrt{u_0-E}} \rho_0(E) + x . \tag{2.12}$$

This very useful simple formula allows us to write the leading genus string equation from the spectral density  $\rho_0(E)$  and ultimately identify the coefficients  $t_k$  that define the model. For example, in the case of undeformed JT gravity the result is [24, 25]:

$$\rho_0(E) = \frac{\sinh(2\pi\sqrt{E})}{4\pi^2\hbar} \quad \implies \quad \mathcal{R}_0[u_0, x] = \frac{\sqrt{u_0}}{2\pi} I_1(2\pi\sqrt{u_0}) + x = 0 , \tag{2.13}$$

and explicitly in terms of the  $t_k$  this is:

$$t_k = \frac{\pi^{2(k-1)}}{2k!(k-1)!} . \tag{2.14}$$

As will become clear later, deformations of JT gravity will yield more general  $\rho_0(E)$ . This will imply different formulae for the  $t_k$ , and hence different matrix model definitions (i.e., new combinations of the underlying minimal models). The resulting  $u_0(x)$  can then be used as the seed for higher order corrections, as part of the boundary conditions of the string equation that supplies non-perturbative physics. The latter will be described next.

### 2.3 Non-perturbative completion

A complete definition of JT gravity should include a non-perturbative sector, and in this approach this means seeking contributions to  $u(x)$  that cannot be captured in an expansion in  $\hbar$ . As already mentioned, the Hermitian matrix model string equation (2.7) captures perturbation theory, but fails to include consistent non-perturbative physics. This issue has a long history, and can be traced to the fact that the even  $k$  minimal models defined by the Hermitian matrix model are individually non-perturbatively unstable [28, 29]. Since the Hermitian matrix model proposed by ref. [14] includes (as reviewed in the previous section) the contribution of the  $k$  even models, they inherit the instability.<sup>6</sup>

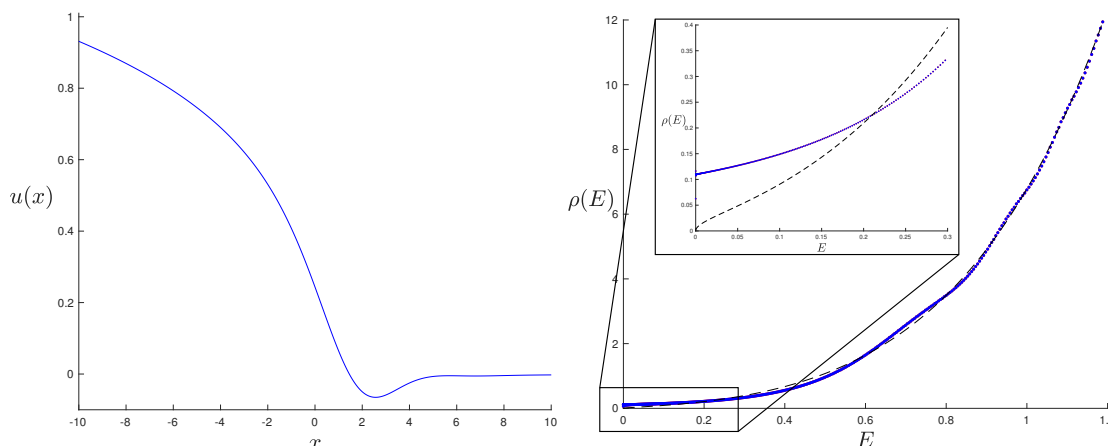
Overall, this means that the non-perturbative definition of the double scaled Hermitian matrix model is unstable and must be modified. The function  $u(x)$  must satisfy a different differential equation. In this work, the differential equation used can be obtained from two different perspectives: from an ensemble of *complex* (instead of Hermitian) random matrices or from a scaling argument. Our proposal is a slight generalization of the non-perturbative definition of JT gravity recently given in ref. [25], that incorporates cases for which  $E_0 \neq 0$ .

A different family of non-perturbatively stable minimal models can be obtained from an ensemble of complex matrices  $M$  with probability measure  $dM dM^\dagger e^{-\frac{N}{\gamma} \text{Tr} V(MM^\dagger)}$ . The double scaling limit of this system was obtained in [39–41], where much of the formalism described in section 2.1 for Hermitian matrix models also applies (for a thorough exposition, see ref. [40]). The crucial difference is given by the fact that the potential depends on the combination  $MM^\dagger \geq 0$ , and so its eigenvalues  $\alpha_i$  are non-negative. The spectral density is therefore non-zero only on the positive real line. It is instructive to think of this as structurally the same problem provided by the Hermitian matrix model, but with a “wall” at zero that prevents the energies from going below zero (presently we will extend the discussion to  $E_0 \neq 0$ ). The procedure for tuning to a critical point is analogous to what was done before. In terms of the endpoint of the distribution  $\rho(E)$ , for the  $k$ th model there are again  $(k - 1)$  additional zeros at the endpoint, but also the wall is present, which produces a key new feature. In terms of the variable  $x$ , the regime  $x < 0$  contains the physics of what happens all the way up to the wall, while the  $x > 0$  sector yields the non-perturbative (stabilising) physics supplied by the wall. It is the presence of this latter regime that makes complex matrix models radically differ from Hermitian matrix models, which contain no wall.<sup>7</sup>

---

<sup>6</sup>A suggestion for how to avoid this issue was given in ref. [27], obtained by thinking of the JT definition as a limit of a (stable)  $k$  odd model, with the even models as perturbations of it. However, it is not clear if this will work. Moreover, the overall picture developed in this paper suggests that the needed solution to the differential equation simply does not exist. See the remark at the end of section 4.3.

<sup>7</sup>In fact, it is easy to see that if one insists on using Hermitian matrix models, one can simply place a wall into such a model by hand, and reproduce the physics under discussion, as discussed in e.g., ref. [42]. Alternatively since Hermitian matrix models are not really any better motivated for studying sums over two-dimensional surfaces than complex matrix models, one might as well start with complex matrix models which supply the wall “naturally”.



**Figure 1.** The non-perturbative potential (left) and spectral density  $\rho(E)$  (right) for JT gravity computed in ref. [27] using equation (2.15), truncated to  $k = 7$ . The dashed line is the disc level result for  $\rho_0(E)$  given in equation (2.13). The inset shows the region near the origin, displaying a non-zero  $\rho(E = 0)$ . Here,  $\hbar$  is set to unity.

The string equation that results from taking the double-scaling limit of these models is [40, 42, 43]:

$$(u - E_0)\mathcal{R}^2 - \frac{\hbar^2}{2}\mathcal{R}\mathcal{R}'' + \frac{\hbar^2}{4}(\mathcal{R}')^2 = 0, \tag{2.15}$$

where  $\mathcal{R}$  is defined as in (2.7). Note that the Hermitian matrix model string equation  $\mathcal{R} = 0$  is a solution of this equation, which results in the potential  $u(x)$  agreeing to all orders in perturbation theory when expanding (2.15) in the  $x < 0$  regime, as is appropriate for JT gravity. The proposal of this paper is that for the deformations of JT gravity (as described in section 1) the equation (2.15) gives the appropriate non-perturbative definition. For  $E_0 \neq 0$  it generalizes the previous proposal of ref. [25], used in ref. [27] to explicitly compute the ordinary JT gravity non-perturbative potential  $u(x)$  and spectral density  $\rho(E)$  shown in figure 1.<sup>8</sup>

Let us comment on the  $E_0$  contribution in (2.15), that is the novel piece with respect to the proposal in ref. [25]. The leading order behavior  $u_0(x)$  obtained from (2.15) as  $\hbar \rightarrow 0$  is given by

$$(u_0 - E_0)\mathcal{R}_0^2[u_0, x] = 0 \quad \implies \quad u_0(x) \equiv E_0 \quad \text{or} \quad \mathcal{R}_0[u_0, x] = 0. \tag{2.16}$$

If we want to have a single-valued and non-trivial function  $u_0(x)$  that is defined in the whole real line  $x \in \mathbb{R}$ , none of the regimes is enough by itself, meaning there must be a transition between them. Since leading genus observables are determined by  $u_0(x)$  in the region  $x < 0$  (2.10), we should take  $u_0(x)$  defined from  $\mathcal{R}_0[u_0, x] = 0$  when  $x < 0$ . In this way, leading genus observables computed from (2.15) agree with those obtained from the original Hermitian matrix model string equation. The full solution for  $u_0(x)$  is then

---

<sup>8</sup>Note that the individual  $k$ th models of this subsection were (in minimal string language) identified in ref. [44] as the  $(4k, 2)$  type 0A minimal strings. Refs. [26, 27] combine them together in a *different* manner than done in ref. [25] in order to non-perturbatively define various JT supergravity theories.

given by

$$u_0(x) = \begin{cases} \mathcal{R}_0[u_0, x] = 0, & x < 0, \\ E_0, & x > 0. \end{cases} \quad (2.17)$$

In order to have a continuous solution, we must fix  $E_0$  as  $\mathcal{R}_0[u_0(0) = E_0, x = 0] = 0$ , which gives the following condition

$$\mathcal{R}_0[E_0, 0] = \sum_{k=0}^{\infty} t_k E_0^k = 0. \quad (2.18)$$

The threshold value  $E_0$  is the energy at the endpoint of the spectral density, where the wall is, at  $x = 0$ . The origin of the shifted equation (2.15), which first appeared in ref. [43], can be understood as a simple modification of the complex matrix model by shifting the matrix combination  $MM^\dagger \rightarrow (MM^\dagger + E_0)$ , so that the eigenvalues lie above the threshold  $E_0$ . The region  $x > 0$  is *beyond* the wall, and so classically  $u_0(x)$  should stay frozen at  $E_0$ , the value it reached at  $x = 0$ . The boundary conditions used to solve the full string equation (2.15) are obtained from the asymptotic expansions for  $x \rightarrow \pm\infty$  starting from each of the two regimes in (2.17).

There is a different way of motivating the full differential equation (2.15), following the work in refs. [40, 42]. Instead of going through the double scaling limit of complex matrices, one seeks an equation that has the same perturbative physics as given by the Hermitian matrix model (already known to agree with JT gravity perturbatively) but then assumes that it preserves the following two established properties of  $u(x)$  not just in perturbation theory but *non-perturbatively*:

$$\begin{aligned} 1. \text{KdV flow :} & \quad (k+1) \frac{\partial u}{\partial t_k} = \frac{\partial \tilde{R}_{k+1}[u]}{\partial x}, \\ 2. \text{Scaling :} & \quad u(sx; \{s^{2k+1}t_k\}, s^{-2}E_0) = s^{-2}u(x; \{t_k\}, E_0), \end{aligned} \quad (2.19)$$

where the first line refers to the explicit dependence of the potential  $u(x)$  on the coefficients  $\{t_k\}$ . Its evolution as a function of  $t_k$  is governed by the KdV flow equation [28, 45]. The second line is a simple scaling relation that specifies how  $u(x, \{t_k\}, E_0)$  changes with a rescaling of the parameters by  $s$ .<sup>9</sup> Appendix B reviews how to use these conditions, with (2.17), to arrive at the string equation (2.15).

All things considered, the non-perturbative definition given in this section is a well motivated proposal that allows for explicit computations of physical quantities in deformations of JT gravity.

### 3 Deformed JT gravity: model A

In this section we study a particular deformation of JT gravity in (1.3), where the dilaton potential is deformed according to

$$U(\phi) = 2\lambda \left( e^{-2\pi(1-\alpha_1)\phi} - e^{-2\pi(1-\alpha_2)\phi} \right), \quad (3.1)$$

---

<sup>9</sup>The simplest way to obtain this is by looking at the leading genus string equation  $\mathcal{R}_0[u_0, x] = 0$  in (2.9). By rescaling the parameters  $(x, t_k, E_0)$  by  $s$  as indicated, we obtain the scaling relation in (2.19).

where  $U(0) = 0$  and for simplicity we assume  $\alpha_1 < \alpha_2$ . In the following sections we analyze the Euclidean partition function using the semi-classical approximation as well as the matrix model description. Several interesting characteristics of this theory will be uncovered. Overall, it will emerge as both perturbatively and non-perturbatively well defined.

### 3.1 Semi-classical approximation

A standard semi-classical analysis of the deformed JT gravity theory involves computing the Euclidean partition function through a saddle point approximation:

$$Z(\beta) = \int Dg D\phi e^{-I_{\text{dJT}}[g_{\mu\nu}, \phi]} \simeq \sum_i e^{-I_{\text{dJT}}[g_{\mu\nu}^{(i)}, \phi^{(i)}]} + \dots, \quad (3.2)$$

where  $g_{\mu\nu}^{(i)}$  and  $\phi^{(i)}$  are on-shell solutions to the equation of motion with renormalized boundary length  $\beta$ . This procedure yields a good approximation when  $\alpha_i \sim 1$  (see appendix D of ref. [21]).

The first step in computing (3.2) is solving the equations of motion with a dilaton potential given by  $W(\phi) = 2\phi + U(\phi)$  and the appropriate boundary conditions. Luckily, this has been analysed a long time ago and the general solution can be written as [46–48]:

$$ds^2 = f(r)(dt/2)^2 + \frac{dr^2}{f(r)}, \quad f(r) = \int_{\phi_h}^r d\phi W(\phi), \quad \phi(r) = r, \quad (3.3)$$

where  $r \geq \phi_h$  and  $t \sim t + \beta$ .<sup>10</sup> Black hole solutions are labeled by  $\phi_h$ , which must satisfy the following constraint in order to have a well defined signature:

$$\int_{\phi_h}^{\phi} d\phi' W(\phi') \geq 0 \quad \text{for all } \phi \geq \phi_h, \quad (3.4)$$

which in particular implies  $W(\phi_h) \geq 0$ . Evaluating the on-shell action (3.2), a standard computation yields the following thermodynamic quantities [48]:

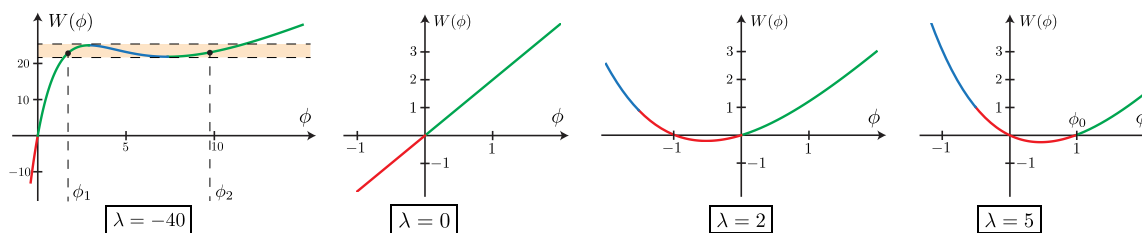
$$T = \frac{W(\phi_h)}{2\pi}, \quad S = S_0 + 2\pi\phi_h, \quad E(\phi_h) = \phi_h^2 - \int_{\phi_h}^{\infty} d\phi U(\phi). \quad (3.5)$$

From this we can compute the specific heat and find the following stability condition for black hole solutions

$$\frac{dE}{dT} = 4\pi \frac{W(\phi_h)}{W'(\phi_h)} \geq 0 \quad \implies \quad \text{stable black hole if } W'(\phi_h) > 0. \quad (3.6)$$

Let us now specialize to the dilaton potential in equation (3.1). The black hole solutions and their behavior can read off just by plotting the potential  $W(\phi) = 2\phi + U(\phi)$  for fixed  $(\alpha_1, \alpha_2) \sim 1$  and  $\lambda$ . Figure 2 shows sample characteristic behavior of the potential depending on the value of  $\lambda$ . Segments of the curve in red, blue and green indicate whether the black hole solution with  $\phi_h = \phi$  is nonexistent (violates (3.4)), unstable (negative specific heat) or stable (positive specific heat). The behavior can be quite different depending on the sign of  $\lambda$ , so each case is analysed separately.

<sup>10</sup>The extra factor of two in the time coordinate is such that the boundary conditions are satisfied with  $\beta$  matching with the conventions in ref. [21].



**Figure 2.** Deformed JT gravity dilaton potential  $W(\phi) = 2\phi + U(\phi)$  in equation (3.1) for the values  $(\alpha_1, \alpha_2) = (0.9, 0.95)$  and several values of  $\lambda$ . Segments of the curve in red, blue and green indicate whether the black hole solution with  $\phi_h = \phi$  is nonexistent (violates (3.4)), unstable (negative specific heat) or stable (positive specific heat).

**Positive  $\lambda$ :** since in this case the potential has  $W(\pm\infty) = +\infty$  there are two branches of black hole solutions, one stable and the other unstable. While at low enough temperatures there are only stable black holes, as we increase the temperature the unstable branch comes into play. In this regime, the favored black hole that dominates the partition function (3.5) is the one with lower free energy. A simple analysis using (3.5) (see also section 3 of ref. [48]) shows the stable black hole in the green branch always dominates the partition function over the unstable one in blue. As a result, the overall system is stable in the semi-classical approximation.

Note that there is an interesting behavior for the zero temperature black hole solution  $\phi_h = \phi_0$ , determined by the largest root of  $W(\phi_0) = 0$ . While for small enough  $\lambda$  we always have  $\phi_0 = 0$ , there is a transition at some finite value  $\lambda_c$  in which  $\phi_0$  becomes non-zero and positive. The critical value  $\lambda_c$  can be easily computed as

$$W'(\phi)|_{\phi=0} = 0 \quad \implies \quad \lambda_c = \frac{-1}{2\pi(\alpha_1 - \alpha_2)}. \quad (3.7)$$

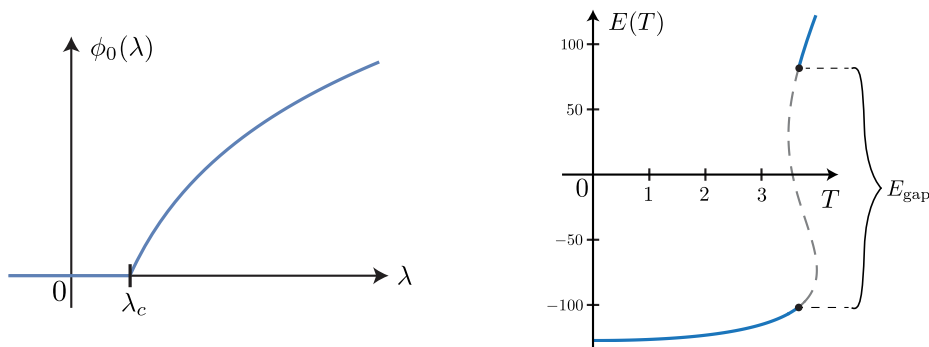
Solving for the zero temperature black hole solution  $\phi_0 = \phi_0(\lambda)$  in a series expansion around  $\lambda_c$  we find the following non-analytic behavior:

$$\phi_0(\lambda) = \begin{cases} 0, & \lambda \leq \lambda_c, \\ \frac{(1/\pi)}{(2 - \alpha_1 - \alpha_2)} \left( \frac{\lambda - \lambda_c}{\lambda_c} \right) + \mathcal{O}(\lambda - \lambda_c)^2, & \lambda \geq \lambda_c. \end{cases} \quad (3.8)$$

The full function  $\phi_0(\lambda)$  can be plotted numerically by solving  $W(\phi_0) = 0$ , so that we get the left diagram in figure 3. This corresponds to a phase transition of the zero temperature black hole solution. The order of the transition is obtained from the free energy  $F = E - TS$ , which at  $T = 0$  agrees with the energy (3.5):

$$F_0(\lambda) = E_0(\lambda) = \phi_0(\lambda)^2 - \frac{\lambda}{\pi} \left( \frac{e^{-2\pi(1-\alpha_1)\phi_0(\lambda)}}{1 - \alpha_1} - \frac{e^{-2\pi(1-\alpha_2)\phi_0(\lambda)}}{1 - \alpha_2} \right), \quad (3.9)$$

with  $\phi_0(\lambda)$  obtained from  $W(\phi_0) = 0$ . From the series expansion (3.8) around  $\lambda = \lambda_c$  we



**Figure 3.** The left diagram shows the phase transition at  $\lambda = \lambda_c$  in equation (3.7), where we plot  $\phi_h = \phi_0(\lambda)$  obtained by numerically solving  $W(\phi_0) = 0$ . In the right diagram is plotted the black hole energy as a function of the temperature for  $\lambda = -40$ , taking into account the phase transition (left diagram in figure 2). We observe a gap in the spectrum generated by the phase transition at  $\lambda$  below  $\lambda'_c$  in (3.11).

find that:<sup>11</sup>

$$E_0(\lambda) = E_0(\lambda_c) + \frac{(1/2\pi^2)}{(1-\alpha_1)(1-\alpha_2)} \left( \frac{\lambda - \lambda_c}{\lambda_c} \right) - \begin{cases} 0 & , \lambda \leq \lambda_c, \\ \frac{(1/3\pi^2)}{(2-\alpha_1-\alpha_2)^2} \left( \frac{\lambda - \lambda_c}{\lambda_c} \right)^3 + \dots & , \lambda \geq \lambda_c. \end{cases} \quad (3.10)$$

Since  $E_0'''(\lambda_c^-) \neq E_0'''(\lambda_c^+)$ , the semi-classical transition of the zero temperature black hole solution is third order.

**Negative  $\lambda$ :** in this case an interesting structure arises for negative enough values of  $\lambda$ . The potential develops a local maximum and minimum, see leftmost diagram in figure 2. This happens for  $\lambda < \lambda'_c$  given by:

$$\lambda'_c = \frac{1}{2\pi} \left[ (1 - \alpha_1) \left( \frac{1 - \alpha_1}{1 - \alpha_2} \right)^{\frac{2(1-\alpha_1)}{\alpha_1-\alpha_2}} - (1 - \alpha_2) \left( \frac{1 - \alpha_1}{1 - \alpha_2} \right)^{\frac{2(1-\alpha_2)}{\alpha_1-\alpha_2}} \right]^{-1}. \quad (3.11)$$

It is straightforward to check that when  $\lambda = \lambda'_c$  the potential has an inflection point

$$W'(\phi_c)|_{\lambda=\lambda'_c} = W''(\phi_c)|_{\lambda=\lambda'_c} = 0, \quad \text{where} \quad \phi_c = \frac{-1}{\pi(\alpha_1 - \alpha_2)} \ln \left[ \frac{1 - \alpha_1}{1 - \alpha_2} \right], \quad (3.12)$$

meaning that for  $\lambda < \lambda'_c$  a local maximum and minimum develops.

The behavior of the dilation potential  $W(\phi)$  in this regime induces a phase transition of a different kind from the one discussed for positive  $\lambda$ . In the canonical ensemble (at fixed temperature  $T = W(\phi_h)/2\pi$ ), we identify a shaded region in the left diagram in figure 2 where there is a finite range of temperatures with three black hole solutions. The dominant black hole in this regime is the one with the lowest free energy  $F = E - TS$ . From the equations in (3.5), a simple analysis shows that seeking the dominant black hole produces

<sup>11</sup>To get this expression we need to go to quadratic order in  $\phi_0(\lambda)$ . Solving for  $\lambda = \lambda(\phi_0)$  from  $W(\phi_0) = 0$  we use that  $\phi_0(\lambda) = (1/\lambda'(0))(\lambda - \lambda_c) - (\lambda''(0)/2\lambda'(0)^3)(\lambda - \lambda_c)^2 + \dots$ .

a discontinuous jump from  $\phi_1$  to  $\phi_2$ , as indicated in the leftmost component of figure 2. In the right diagram of figure 3 is plotted the energy as a function of the temperature (see eqs. (3.5)), taking into account the leap in  $\phi_h$ . At the critical temperature we observe a discontinuous jump in the energy. This generates a gap  $E_{\text{gap}}$  for which there are no dominant black hole solutions in a finite range of energies. This kind of first order phase transition (where three black holes compete) was first observed for charged AdS black holes in ref. [49]. Along with other transitions, they have been recently explored in the current JT gravity context in ref. [48].

### 3.2 Perturbative genus expansion

While the semi-classical analysis of the previous section showed a rich and interesting structure, it was obtained from the crude saddle point approximation to the Euclidean path integral (3.2). To compute  $Z(\beta)$  more carefully we use the results in refs. [21, 22], where it was shown the full partition function can be computed from a double scaled matrix model. In this subsection we focus on the leading genus contribution,  $Z_0(\beta)$ , given by the disc topology:

$$Z(\beta) = \int Dg D\phi e^{-I_{\text{dJT}}[g_{\mu\nu}, \phi]} = \int_{\text{Disc}} Dg D\phi e^{-I_{\text{dJT}}[g_{\mu\nu}, \phi]} + \dots = Z_0(\beta) + \dots, \quad (3.13)$$

where we are neglecting additional higher genus and non-perturbative contributions that will be studied in the next subsection. The leading genus disc spectral density (1.5) derived in ref. [22], that we rewrite here for convenience, is given by:

$$\rho_0(E) = \frac{\sinh(2\pi\sqrt{E})}{4\pi^2\hbar} + \lambda \frac{\cosh(2\pi\alpha_1\sqrt{E}) - \cosh(2\pi\alpha_2\sqrt{E})}{2\pi\hbar\sqrt{E}}. \quad (3.14)$$

This formula has a problem, since  $\rho_0(E)$  is not always positive definite. This can be seen by expanding for small energies, where we find:<sup>12</sup>

$$\rho_0(E) = \left( \frac{\lambda_c - \lambda}{2\pi\hbar\lambda_c} \right) \sqrt{E} + \mathcal{O}(E^{3/2}), \quad \text{where} \quad \lambda_c = \frac{-1}{2\pi^2(\alpha_1^2 - \alpha_2^2)}. \quad (3.15)$$

Here  $\lambda_c > 0$  since we are assuming  $\alpha_1 < \alpha_2$ . The spectral density goes negative for  $\lambda > \lambda_c$  and small enough energies.<sup>13</sup>

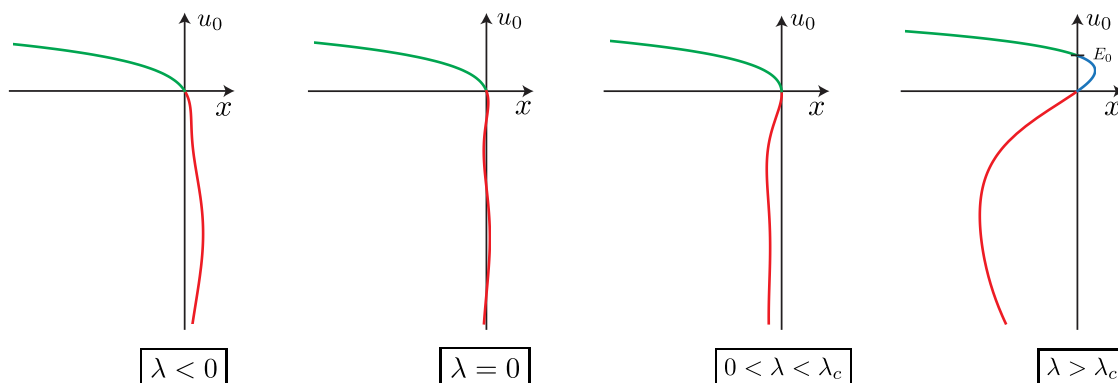
This certainly does not fit within the matrix model analysis and framework outlined in section 2.2. In particular, the definition of  $\rho(E)$  in equation (2.2) is manifestly positive, and large  $N$  and double-scaling limits cannot change that, as is clear from equation (2.6). To understand the origin of the issue from the matrix model perspective, consider the behavior of  $u_0(x)$ , defined from the leading genus string equation. Using equation (2.12) we can easily compute  $\mathcal{R}_0$  and find

$$\mathcal{R}_0[u_0, x] = \frac{\sqrt{u_0}}{2\pi} I_1(2\pi\sqrt{u_0}) + \lambda(I_0(2\pi\alpha_1\sqrt{u_0}) - I_0(2\pi\alpha_2\sqrt{u_0})) + x = 0. \quad (3.16)$$

<sup>12</sup>We have suggestively defined  $\lambda_c$ , which is different from the semi-classical definition in (3.7). We comment on the relation between the two at the end of this subsection.

<sup>13</sup>As already explained in the introduction, this issue does not arise from the perspective of ref. [21].





**Figure 4.** Examples of the curve  $x = x(u_0)$  obtained from (3.16) in the  $(x, u_0)$  plane for several values of  $\lambda$  and  $(\alpha_1, \alpha_2) = (1/4, 1/3)$ . The green curve is the unique choice that yields a single valued function defined in the range  $x < 0$ . While for  $\lambda \leq \lambda_c$  the green curve always satisfy  $u(0) = 0$ , for  $\lambda > \lambda_c$  we observe  $u(0) = E_0 > 0$ .

To have a well defined single valued function  $u_0(x)$  on the real line  $x \in \mathbb{R}$ , the implicit function theorem imposes the following constraint

$$u_0 = u_0(x) \text{ for } x \in \mathbb{R} \quad \implies \quad \partial_{u_0} \mathcal{R}_0 \neq 0. \tag{3.17}$$

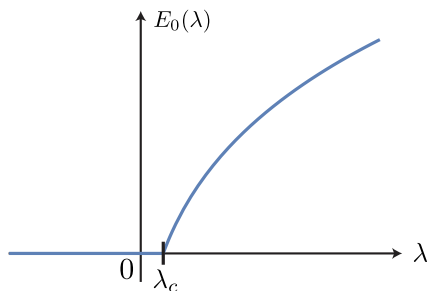
From (3.16) we can easily check this condition is never satisfied, not even for undeformed JT gravity (i.e.,  $\lambda = 0$ ). This might seem strange, since ordinary JT gravity has a well defined disc spectral density (2.13). However it is important to note that the leading genus partition function  $Z_0(\beta)$  (2.10) is built from the regime  $x < 0$ . Therefore we only need  $\mathcal{R}_0[u_0, x] = 0$  to define a unique single-valued function  $u_0(x)$  in the range  $x < 0$ .

To understand the situation more clearly it is useful to plot the curve  $x = x(u_0)$  defined through equation (3.16). The undeformed JT gravity case is the curve shown in the  $\lambda = 0$  diagram in figure 4. While for  $u_0 < 0$  there are an infinite number of oscillations around  $x = 0$  (marked in red), for  $u_0 > 0$  there is a single curve marked in green going from  $u_0 = 0$  to  $u_0 = +\infty$ . This green curve defines the unique single valued function  $u_0(x)$  in  $x < 0$ . As can be seen from the other plots in figure 4, the behavior for  $\lambda < \lambda_c$  is quite similar: the upper (green) branch can be used to define a unique single-valued function defined in the range  $x < 0$ . Using that in all these cases  $u_0(0) = E_0 = 0$ , the integral (2.11) can be solved explicitly, recovering the starting result (3.14), which is positive for such  $\lambda$ .

Consider now  $\lambda > \lambda_c$ . From the rightmost diagram in figure 4 it is clear that there is an interesting feature: the value of  $u_0(0)$  is no longer zero. The shift is caused by the fact that exactly at  $\lambda = \lambda_c$  defined in (3.15), the derivative of the leading genus string equation vanishes at  $u_0 = 0$

$$(\partial_{u_0} \mathcal{R}_0)|_{(u_0, \lambda) = (0, \lambda_c)} = 0. \tag{3.18}$$

Consider the curve starting at  $u_0(0) = 0$  in the rightmost diagram in figure 4, defining the multi-valued function  $u_0(x)$  (i.e., including the blue section of the curve). Using it to compute  $\rho_0(E)$  from the integral formula (2.11) yields the non-positive result in equation (3.14). On the other hand, if we proceed and define  $u_0(x)$  by picking the unique curve



**Figure 5.** Threshold energy  $E_0$  as a function of  $\lambda$  for  $(\alpha_1, \alpha_2) = (1/4, 1/3)$ , obtained by numerically solving (3.19). For  $\lambda = \lambda_c$  in (3.15) we observe a first order phase transition.

that results in a single valued function  $u_0(x)$  for  $x < 0$ , i.e. the green portion of the diagram, the  $\rho_0(E)$  obtained from integral (2.11) is explicitly positive (since  $\partial_{u_0} \mathcal{R}_0 \geq 0$  in this case). The primary new feature is the value of  $u_0(x)$  at the origin, the threshold energy  $E_0 = u_0(0)$ . For  $\lambda > \lambda_c$ , instead of having  $E_0 = 0$  it is a non-zero value obtained from the largest solution to the leading genus equation at  $x = 0$

$$\mathcal{R}_0[u_0(0) = E_0, x = 0] = \frac{\sqrt{E_0}}{2\pi} I_1(2\pi\sqrt{E_0}) + \lambda(I_0(2\pi\alpha_1\sqrt{E_0}) - I_0(2\pi\alpha_2\sqrt{E_0})) = 0. \quad (3.19)$$

This equation can be easily solved numerically to yield figure 5. The behavior of  $E_0$  as a function of  $\lambda$  is non-analytic at  $\lambda_c$  and corresponds to a phase transition. We can determine the order of the transition using (3.19) to expand  $E_0(\lambda)$  on either side of  $\lambda_c$ , so that:

$$E_0(\lambda) = \begin{cases} 0 & , \quad \lambda \leq \lambda_c, \\ \frac{(2/\pi)^2}{(2 - \alpha_1^2 - \alpha_2^2)} \left( \frac{\lambda - \lambda_c}{\lambda_c} \right) + \mathcal{O}(\lambda - \lambda_c)^2 & , \quad \lambda \geq \lambda_c. \end{cases} \quad (3.20)$$

Since  $E_0'(\lambda_c^+) \neq E_0'(\lambda_c^-)$  this is a first order phase transition. This phase transition solves the negativity of the leading genus spectral density, which for  $\lambda > \lambda_c$  is given by (2.11) with  $u(0) = E_0 \neq 0$  instead of the problematic formula (3.14).

It is interesting to compare these leading order matrix model results with those obtained in the previous subsection from the semi-classical gravity analysis. The most important similarity is given by the fact that in both cases we observe a phase transition of the threshold energy  $E_0$  for some positive value of  $\lambda$ . By comparing figures 2 and 4 we also note some resemblance in the mechanisms triggering the transition. There are however important quantitative differences, since the exact location of the transition  $\lambda_c$  in (3.7) and (3.15) does not agree. Moreover, while the matrix model transition is first order, the semi-classical analysis yields a third order transition. When doing this comparison, keep in mind there is no overlap between the  $\alpha_i$  regimes of validity of each computation. Moreover, the saddle-point approximation of the path integral (3.2) does not take into account quantum corrections that might be important at low energies. All things considered, it is quite pleasing (and perhaps surprising) that the semi-classical analysis captures the transition of the vacuum solution at all.

On the other hand, the semi-classical gravity analysis yielded a phase transition among black hole solutions that is present at negative  $\lambda$  (below  $\lambda'_c$  in (3.11)). It suggested the existence of a gap  $E_{\text{gap}}$  opening up in the available energy of black holes. There is no sign of such a gap in the spectrum in the matrix model analysis at disc level, but what about higher orders in  $\hbar$  in the perturbative expansion?

A simple way of showing perturbative corrections are not able to account for  $E_{\text{gap}}$  is by computing the first genus correction to the disc spectral density  $\rho_0(E)$  in (3.14). The partition function at this order is given in equation (9.19) of ref. [22]. Carefully applying an inverse Laplace transform (dealing with divergences in the same way as in footnote 9 of ref. [21]) we find the corrections to spectral density are of the form  $\hbar c_0/E^{3/2} + \hbar c_1/E^{5/2}$  for some constants  $c_0$  and  $c_1$ . Contributions of this form are not able to explain the energy gap  $E_{\text{gap}}$ , which should arise from a low energy behavior given by:

$$\rho(E) \sim \sqrt{E - E_{\text{gap}}} = \sqrt{E} - \frac{E_{\text{gap}}}{2\sqrt{E}} - \frac{E_{\text{gap}}^2}{8E^{3/2}} + \mathcal{O}(1/E^{5/2}) . \quad (3.21)$$

Neither the leading genus spectral density  $\rho_0(E)$  in (3.14) nor its first correction contribute with the term  $\sim 1/\sqrt{E}$  that is necessary to resum the series and obtain the gap. Overall, the energy gap  $E_{\text{gap}}$  identified in the semi-classical analysis is not observed when computing the Euclidean partition function including contributions of arbitrary genus.

### 3.3 Non-perturbative effects

So far, the Euclidean partition function of the theory has been studied in one or other approximation scheme: either the semi-classical gravity limit (3.2) or the perturbative expansion of the matrix model (3.13). Using our double-scaled matrix model definition presented in section 2.3, we now compute  $Z(\beta)$  *exactly*, including all higher order genus and non-perturbative effects.

The first step is to identify the coefficients  $t_k(\lambda)$  in (2.7) that define the matrix model. These can be obtained by expanding the leading genus string equation (3.16) in a power series in  $u_0$ , giving:

$$\mathcal{R}_0[u_0, x] = \sum_{k=1}^{\infty} t_k(\lambda) u_0^k + x, \quad \text{where} \quad t_k(\lambda) = \frac{\pi^{2(k-1)}}{2k!^2} (k + 2\pi^2 \lambda (\alpha_1^{2k} - \alpha_2^{2k})) . \quad (3.22)$$

Using this, the combination  $\mathcal{R}[u, x]$  in (2.7) can be constructed, and hence the non-linear ordinary differential equation (2.15) that determines the potential  $u(x)$ . The boundary condition is the full  $\hbar = 0$  solution for  $u_0(x)$ , which is given by the  $u_0(x)$  discussed in the previous section for  $x < 0$ , and for  $x \geq 0$  by  $u_0(x) = E_0$ . Since all the  $t_k(\lambda)$  are non-zero, the differential equation is of infinite order, but (as demonstrated in the case of  $\lambda = 0$  in ref. [27]) there is a meaningful sense in which it can be truncated to finite order, by including all  $t_k$  contributions up to a maximum,  $t_{k_{\text{max}}}$ :

$$\mathcal{R} = \sum_{k=1}^{\infty} t_k(\lambda) \tilde{R}_k[u] + x \quad \xrightarrow{\text{truncation}} \quad \mathcal{R}^{(k_{\text{max}})} = \sum_{k=1}^{k_{\text{max}}} t_k(\lambda) \tilde{R}_k[u] + x . \quad (3.23)$$

This is a controlled truncation, as can be seen from the fact that the  $t_k$  rapidly decrease in relative size as  $k$  increases. The truncation gives a good estimate for  $E_0$ ,  $\lambda_c$ , and moreover gives an energy spectrum that matches well to the exact one. Adding higher  $k$  improves the truncation accuracy. It was found that  $k_{\max} = 7$  gave good accuracy for the range of energies where non-perturbative effects are important. This required a solution of a highly non-linear 14th order differential equation, which turns out to be possible.<sup>14</sup>

After finding an accurate enough numerical solution for  $u(x)$ , the spectrum of the Schrödinger Hamiltonian  $\mathcal{H}[u]$  of equation (2.5) can be computed (again, using numerical techniques), and ultimately the full spectral density  $\rho(E)$  computed using (2.6).<sup>15</sup> Picking the values  $(\alpha_1, \alpha_2) = (1/4, 1/3)$ , the next step is to analyse each sign of  $\lambda$  separately.

**Positive  $\lambda$ :** the potential  $u(x)$  was found, and the spectrum computed to enable the construction of  $\rho(E)$  for positive  $\lambda$  in the  $k_{\max} = 7$  truncation. In the left diagram of figure 6 is plotted the potential  $u(x)$  for several values of  $\lambda$ , where the phase transition (3.20) in  $E_0$  is manifest in the shift of the boundary condition  $u(+\infty) = E_0$ . Comparing with the ordinary JT gravity solution with  $\lambda = 0$  (blue curve), it can be appreciated how the potential  $u(x)$  is deformed.

On the right diagram of figure 6 is plotted the spectral density  $\rho(E)$  computed from each of the potentials. For  $\lambda = 0$  it is known that the threshold density value  $\rho(0)$  is non-zero due to non-perturbative effects [25, 27]. While for  $\lambda \in (0, \lambda_c)$  we also observe  $\rho(0)$  non-zero (the slight numerical jitter near the threshold energy can be safely ignored), it takes a smaller value as compared to that of ordinary JT gravity (blue and red curves). As  $\lambda$  increases past the transition value  $\lambda_c$ , the threshold energy  $E_0$  shifts to the positive values obtained from equation (3.19). The value of  $\rho(E_0)$  is still non-zero and increasing with  $\lambda$  (compare green and black curves).

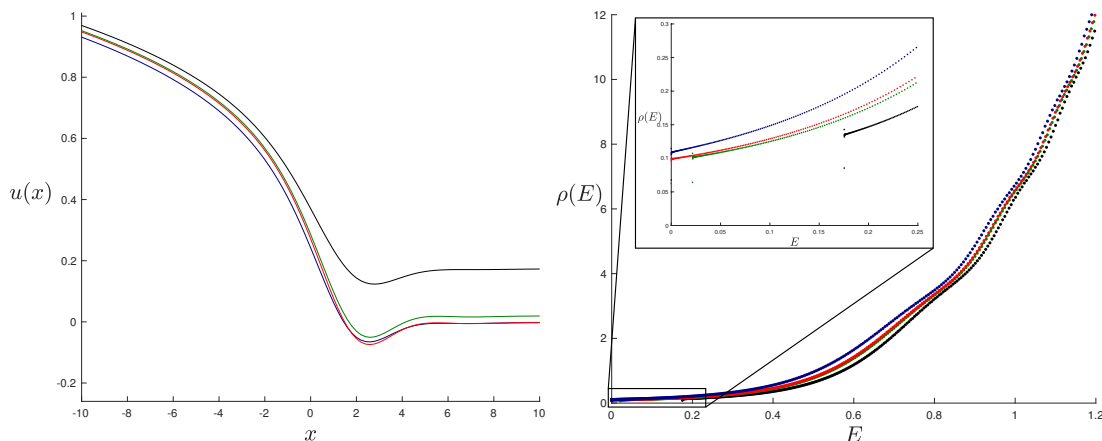
The behavior of  $\rho(E_0)$  is related to the value of the non-perturbative potential  $u(x)$  at  $x = 0$ . The spectral density  $\rho(E)$  is computed from (2.6) by integrating over the (modulus-squared) tail of the  $E = E_0$  wavefunction that penetrates into the  $x < 0$  region. How much of that tail penetrates depends upon the value of  $u(x) - E_0$  at  $x = 0$ . Here, note that  $u(x)$  is the *full* non-perturbative potential — its leading behavior  $u_0(x) - E_0$  of course vanishes at  $x = 0$  by construction. Comparing the diagrams in figure 6 we observe that  $\rho(E_0)$  increases whenever  $u(x) - E_0$  decreases, since  $|\psi(x, E_0)|^2$  (integrated to find  $\rho(E_0)$  in (2.6)) has a larger support in the region  $x < 0$ .

**Negative  $\lambda$ :** let us now consider negative  $\lambda$ , where there is no phase transition in the threshold energy  $E_0 = 0$ . On the diagrams in figure 7 we plot the non-perturbative

---

<sup>14</sup>In practice, a first derivative of the string equation (and then dividing by  $\mathcal{R}$ ) results in an equation that is structurally less complicated, so a 15th order equation was solved for  $k_{\max} = 7$ . The boundary conditions, obtained by fixing the appropriate number of derivatives  $\partial_x^n u(x) \equiv u^{(n)}(x)$  at  $x = \pm x_{\max}$ , are fixed to  $u^{(n)}(\pm x_{\max}) = u_0^{(n)}(\pm x_{\max})$ . See the appendix of ref. [27] for detailed tips on how to numerically solve the equations, and to solve the spectral problem for  $\mathcal{H}[u]$  once  $u(x)$  is found. Throughout, the choice  $\hbar = 1$  was made, although explorations were done with  $\hbar < 1$ , with qualitatively similar results.

<sup>15</sup>The same numerical techniques as the ones developed in refs. [25–27] were used, and so they will not be described here. See an appendix of ref. [27] for suggestions about solving the equations and extracting the spectrum of  $\mathcal{H}$ .



**Figure 6.** On the left are plots of the non-perturbative potential  $u(x)$  in the  $k_{\max} = 7$  truncation for several  $\lambda \geq 0$  and  $(\alpha_1, \alpha_2) = (1/4, 1/3)$ . While the blue curve corresponds to pure JT gravity, red and green correspond to  $\lambda = 9\lambda_c/10$  and  $\lambda = 11\lambda_c/10$  respectively, where  $\lambda_c \simeq 1.042$  (3.15). The black curve is for  $\lambda = 2\lambda_c$ . On the right (using the corresponding colours) are plotted the full spectral densities obtained from equation (2.6) after computing the spectrum of  $\mathcal{H}[u]$ .

numerical solutions to  $u(x)$  and  $\rho(E)$  for several values of  $\lambda$ . The quantity  $\rho(0)$ , the spectral density at the threshold energy  $E_0 = 0$ , increases for increasingly larger negative  $\lambda$ . From the discussion immediately above, this is in agreement with the value of the potential  $u(x)$  at the origin  $x = 0$ .

Overall, we conclude that the deformation of JT gravity according to (3.1) is both perturbatively and non-perturbatively well defined (using the non-perturbative scheme adopted in this paper). For  $\lambda$  positive the phase structure obtained from the semi-classical analysis of the two-dimensional black holes solutions is in qualitative agreement with the full partition function as captured by the matrix model. For  $\lambda$  negative (we explored out to  $\lambda = -100$  to ensure that the patterns of figure 7 persist) this does not appear to be the case, as the energy gap  $E_{\text{gap}}$  observed in the semi-classical approximation is invisible in the matrix model description, even after accounting for non-perturbative effects.<sup>16</sup>

#### 4 Deformed JT gravity: model B

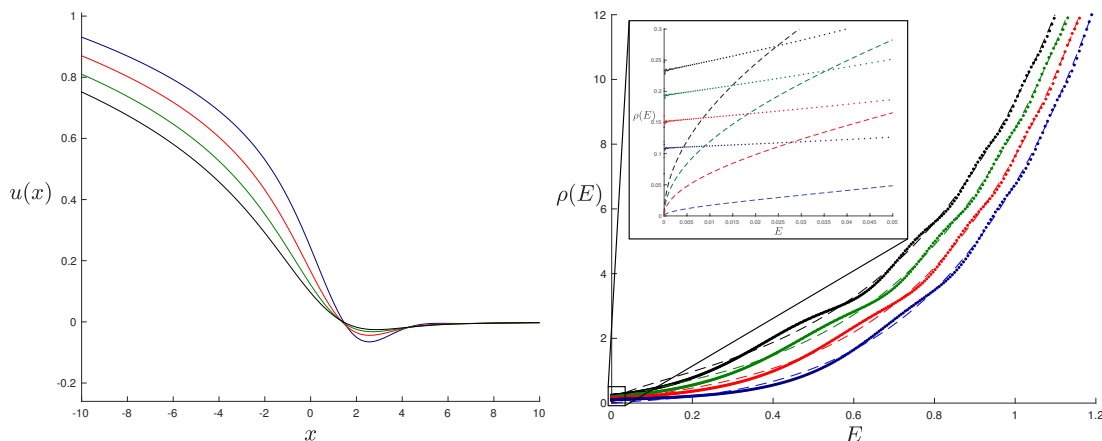
This section will repeat the previous the analysis of Euclidean partition function but for a different deformation of JT gravity, given by:

$$U(\phi) = 2\lambda e^{-2\pi(1-\alpha)\phi}, \tag{4.1}$$

where now  $U(0) \neq 0$ . We show that for certain values of  $\lambda$  this deformation is both perturbatively and non-perturbatively inconsistent, using our definition in section 2.3.

---

<sup>16</sup>The first version of this manuscript reported effects that were interpreted as the appearance of a non-perturbative  $E_{\text{gap}}$ . They were due to a subtle error in the code used to numerically solve the full string equation. This error has been corrected.



**Figure 7.** On the left are plots of the non-perturbative potential  $u(x)$  obtained from numerically solving (2.15) for  $(\alpha_1, \alpha_2) = (1/4, 1/3)$  with  $\lambda = 0$  (blue),  $\lambda = -10/3$  (red),  $\lambda = -20/3$  (green) and  $\lambda = -10$  (black). On the right are the associated spectral densities obtained from equation (2.6) and plotted using the same colours, with the dashed lines corresponding to the disc spectral density  $\rho_0(E)$  in (3.14).

### 4.1 Semi-classical approximation

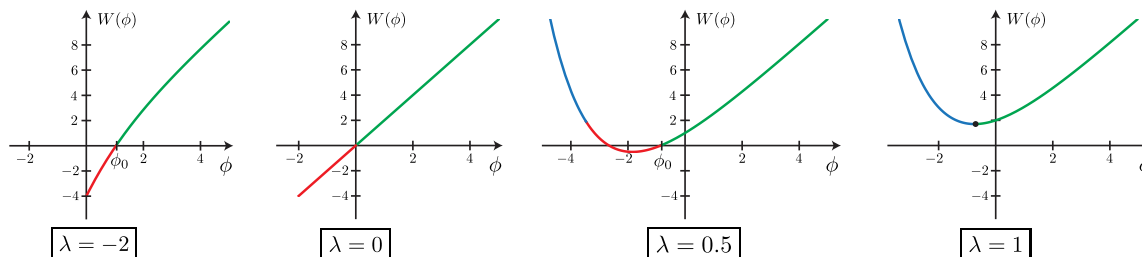
As done in section 3.1, the semi-classical approximation can be used to compute the Euclidean partition function (3.2) of the theory with potential  $W(\phi) = 2\phi + U(\phi)$ , but now with  $U(\phi)$  given in equation (4.1). The black hole solutions are written in equation (3.3) and depend on the single parameter  $\phi_h$ . The conditions derived on  $W(\phi)$  have the same interpretation as before, and the thermodynamic quantities are computed from equation (3.5).<sup>17</sup>

In figure 8 the potential is plotted for  $\alpha = 0.9 \sim 1$  and several values of  $\lambda$  in order to illustrate its key qualitative features. As before, segments of the curve in red, blue and green indicate whether the black hole solution with  $\phi_h = \phi$  is non-existent, unstable or stable. For  $\lambda$  negative there is a single branch of stable black hole solutions, with the zero temperature solution determined by  $\phi_h = \phi_0(\lambda) > 0$ . The value of  $\phi_0(\lambda)$  changes smoothly with  $\lambda$  and there are no finite temperature phase transitions. For positive  $\lambda$  there are two separate branches with stable and unstable black holes. At  $\lambda = \lambda_c > 0$  there is a transition to a phase for which there is no zero temperature black hole solution. The critical value associated to this transition can be easily computed and is given by

$$W(\phi_0^{(c)})|_{\lambda=\lambda_c} = W'(\phi_0^{(c)})|_{\lambda=\lambda_c} = 0, \quad \lambda_c = \frac{1}{2\pi e(1-\alpha)}, \quad \phi_0^{(c)} = \frac{-1}{2\pi(1-\alpha)}, \tag{4.2}$$

where  $e$  is Euler’s number. Note that this is a different kind of transition than the one analysed around equation (3.7). While in that case the transition was between two different zero temperature regimes, in this case there is a transition between a system with a  $T = 0$  regime and one without.

<sup>17</sup>A semi-classical analysis of this theory was recently presented in appendix D of ref. [21]. There is a slight difference in the proportionally factor in  $U(\phi)$  that does not modify the qualitative behavior.



**Figure 8.** Deformed JT gravity dilaton potential  $W(\phi) = 2\phi + U(\phi)$  in (4.1) for  $\alpha = 0.9 \sim 1$  and several values of  $\lambda$ . Black hole solutions exist for  $\phi = \phi_h$  indicated in green (stable) and blue (unstable). For  $\lambda > \lambda_c$  in (4.2) there is a transition to a regime in which there are no zero temperature black hole solutions.

## 4.2 Perturbative genus expansion

Consider now the matrix model computation of the partition function at leading (disc) order, as in equation (3.13). The spectral density computed in refs. [21, 22] is obtained from the integral (1.9), where  $E_0$  is obtained from  $\mathcal{R}_0[E_0, 0] = 0$  with

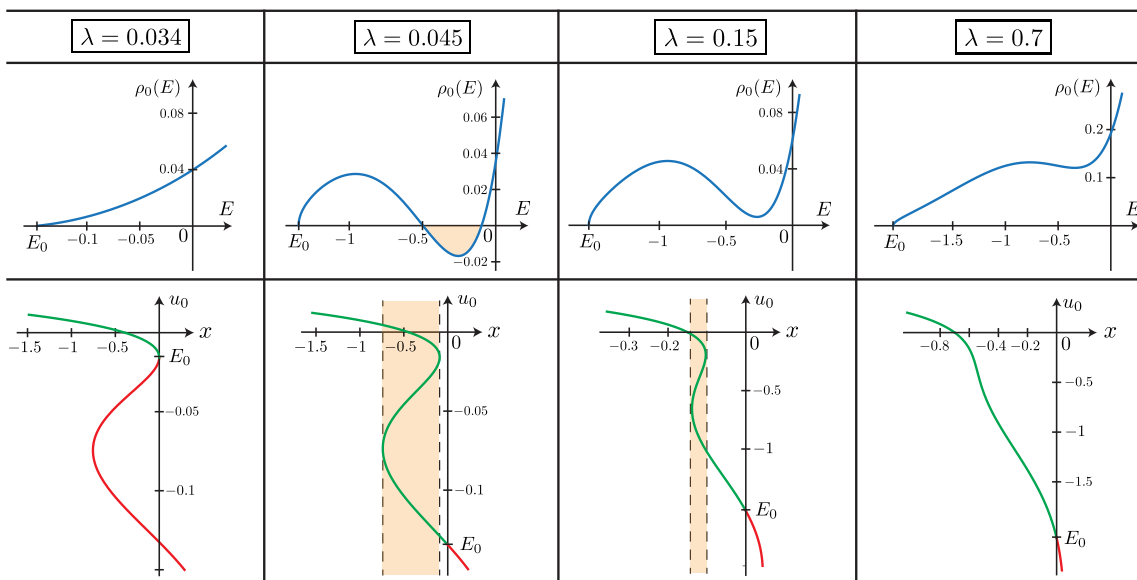
$$\mathcal{R}_0[u_0, x] = \frac{\sqrt{u_0}}{2\pi} I_1(2\pi\sqrt{u_0}) + \lambda I_0(2\pi\alpha\sqrt{u_0}) + x = 0. \quad (4.3)$$

The top row of figure 9 displays the spectral density  $\rho_0(E)$  (extracted numerically from (1.9) with (4.3)) for  $\alpha = 1/4$  and several values of  $\lambda$ . Strikingly,  $\rho_0(E)$  develops negative regions for a finite window of  $\lambda$  approximately given by  $\lambda \in (0.035, 0.122)$ . As discussed before, this clearly signals a problem, and again its origins lie in the behavior of  $u_0(x)$  as a solution to equation (4.3).

The place to look at is the implicit definition of  $u_0(x)$  from the string equation (4.3). In the second row of figure 9 is plotted the curve  $x = x(u_0)$  for the same values of  $\alpha$  and  $\lambda$  as before. From the first two diagrams starting from the left, we note that after  $\lambda \simeq 0.034$ , the threshold energy  $E_0$  jumps discontinuously as a function of  $\lambda$ . This zeroth order phase transition generates a multi-valued potential  $u_0(x)$  in the crucial region  $x < 0$  as can be seen at e.g.,  $\lambda = 0.045$ . This ultimately results (for the reasons discussed after equation (3.16)) in a negative spectral density. Amusingly, while a phase transition in  $E_0$  solved the issues of  $\rho_0(E)$  in the previous section, in this case it is the origin of the problem.

Since the potential  $u_0(x)$  is multivalued in the region  $x < 0$ , the issue is more severe than the one discussed in the previous section. This is because leading order observables in the matrix model are computed from the behavior of  $u_0(x)$  in  $x < 0$ . Recall for instance, the disc partition function  $Z_0(\beta)$ , given in terms of  $u_0(x)$  in equation (2.10). That expression makes no sense if the function  $u_0(x)$  is multivalued in the region  $x < 0$ , meaning the  $\hbar \rightarrow 0$  limit that defines  $Z_0(\beta)$  does not exist. Since any single trace observable (like the spectral density) can be computed from  $Z_0(\beta)$ , it means none of these observables have a well defined leading genus behavior. This results in the breakdown of the  $\hbar$  perturbative expansion of the matrix model.<sup>18</sup> This type of failure is not extremely unusual, for instance observables

<sup>18</sup>Moreover, from the point of view of the associated quantum mechanical problem, having a multi-valued potential for  $u_0(x)$ , in the classical limit or beyond, is highly problematic.



**Figure 9.** The first row corresponds to the  $\rho_0(E)$  numerically computed from (1.9) for  $\alpha = 1/4$  and several values of  $\lambda$ . For  $\lambda \in (0.035, 0.122)$  the spectral density is non-positive. In the second row we plot the curve  $x = x(u_0)$  in the  $(x, u_0)$  plane obtained from (4.3) with  $\alpha = 1/4$  and the same values of  $\lambda$ . The function  $u_0(x)$  is multi-valued in the shaded region in  $x < 0$  for a (different) finite window of  $\lambda$ , approximately given by  $\lambda \in (0.035, 0.334)$ .

in a multi-cut Hermitian matrix model do not have the ordinary power series expansion in  $1/N$  [50].

It must be emphasized that the multi-valuedness of  $u_0(x)$  for  $x < 0$  is the *root cause* of the physical problems here. Negativity of  $\rho_0(E)$  is only a symptom. A more careful comparison of the diagrams in figure 9 illustrates this. Since for  $\lambda = 0.15$  the leading genus spectral density is positive, one might think the matrix model perturbative expansion in this case is sensible. However, a closer inspection shows that  $u_0(x)$  for  $\lambda = 0.15$  is still multi-valued in the region  $x < 0$ , meaning the leading genus spectral density obtained from equation (2.10) is not defined at all. This shows the negativity of the spectral density is not a robust indicator when assessing the health of the matrix model.

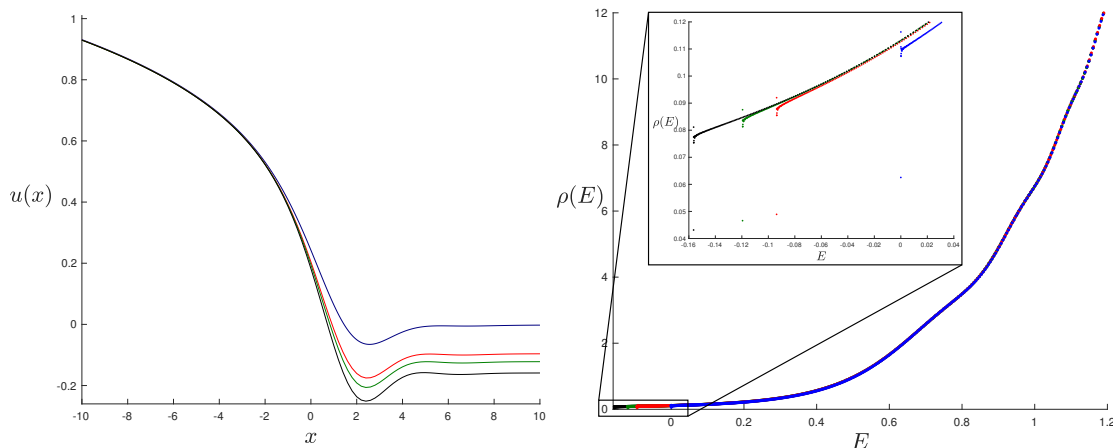
### 4.3 Non-perturbative effects

The breakdown of the perturbative expansion calls for an appeal to non-perturbative physics. Here again, the definition of section 2.3 is employed. To do so, the coefficients  $t_k$  appearing in  $\mathcal{R}[u, x]$  in equation (2.7) must be obtained. These can be identified by expanding the disc level equation (4.3), giving:

$$\mathcal{R}_0[u_0, x] = \sum_{k=1}^{\infty} t_k(\lambda) u_0^k + \lambda + x, \quad \text{where} \quad t_k(\lambda) = \frac{\pi^{2(k-1)}}{2k!^2} (k + 2\pi^2 \lambda \alpha^{2k}). \quad (4.4)$$

Using these coefficients we can write  $\mathcal{R}[u, x]$  and the non-perturbative differential equation (2.15) for  $u(x)$ . As discussed in the case of section 3.3, explicit solution can be obtained *via* a truncation to some highest  $k = k_{\max}$  as in equation (3.23). It is important





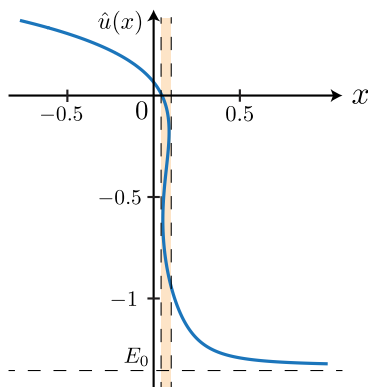
**Figure 10.** On the left are numerical solutions for the potential  $u(x)$  obtained by solving the string equation (2.15) for  $\alpha = 1/4$  in the  $k_{\max} = 7$  truncation and  $\lambda = 0$  (blue),  $\lambda = 0.03$  (red),  $\lambda = 0.033199$  (green) and  $\lambda = 0.034798$  (black). All of these values are before the zeroth order transition of  $E_0$  that occurs at  $\lambda_c \simeq 0.03480$ . On the right are the associated spectral densities obtained from (2.6) and plotted using the same colors.

to pick a truncation that captures the phase transition of  $E_0$  shown in figure 9. Only odd truncations  $k_{\max}$  are useful in this regard, with  $k_{\max} = 5$  the minimum truncation with the necessary structure, given the properties of the  $t_k(\lambda)$ . To obtain a more accurate spectral density for a larger range of energies, the truncation  $k_{\max} = 7$  is used. The curve  $x = x(u_0)$  obtained from  $\mathcal{R}_0^{(7)}[u_0, x] = 0$  is analogous to the ones shown in figure 9, with the transition occurring at  $\lambda_c \simeq 0.03480$ . For  $k_{\max} = 7$ , the differential equation is 14th order but as before (see footnote 14) some persistence can yield highly accurate numerical solutions for  $u(x)$  (errors smaller than  $10^{-4}$ ).

**Before the transition:** first is the case of  $\lambda < \lambda_c$ , where  $\lambda_c$  is the critical value triggering the transition in  $E_0$ . Similarly to the previous section solutions can be found for the non-perturbative potential  $u(x)$  and its associated spectral density  $\rho(E)$  computed. The results are shown in figure 10 for several values of  $\lambda < \lambda_c$ . There is no phase transition for  $E_0$ , and its value changes continuously with  $\lambda$ , as can be seen from the behavior at  $u(+\infty) = E_0$ . The spectral density  $\rho(E)$  does not show any particularly novel behavior for  $\lambda < \lambda_c$  (in contrast to some of the phenomena seen in the previous section).

**Across and beyond the transition:** next is a study of the non-perturbative behavior of the model as  $\lambda$  crosses the transition value  $\lambda_c$  where  $E_0$  has a finite jump. In the  $k_{\max} = 7$  truncation the range of  $\lambda$  for which the perturbative expansion breaks down is given by  $\lambda \in (0.03480, 0.36733)$ , with the left edge corresponding to the zeroth order phase transition in  $E_0$ .

As soon as  $\lambda$  crosses the transition the string equation (2.15) for  $u(x)$  becomes extremely unstable. While for  $\lambda$  arbitrarily close to  $\lambda_c$  before the transition, well behaved solutions can be constructed, the numerics becomes unpredictable as soon as  $\lambda$  goes be-



**Figure 11.** Numerical plots obtained by solving for the potential  $\hat{u}(x)$  in (4.6) corresponding to the t’Hooft limit with  $q = 0.1$  and  $\lambda = 0.038 > \lambda_c = 0.035$ . The asymptotic behavior of the solutions shows the distinct behavior at  $\hat{u}(+\infty) = E_0$ . There is a shaded region for which  $\hat{u}(x)$  is multi-valued.

yond  $\lambda_c$ .<sup>19</sup> Before jumping to the conclusion that the deformed JT gravity theory is non-perturbatively unstable beyond the transition, it is prudent to carefully consider other alternatives that might explain the issue. The first obvious possibility is a lack of numerical precision: the discontinuous jump in  $E_0$  might be making the differential equation harder to solve, even though a well behaved solution might in principle still exist. Deforming away from a neighboring solution already found (a common trick for numerically exploring difficult equations) does not work here since  $E_0$  jumps. To explore this, a toy model was constructed that has the same features as the deformed JT gravity theory with the difference that the jump in  $E_0$  can be controlled and made small. The model has the following leading genus string equation:

$$\mathcal{R}_0^{\text{toy}}[u_0, x] = u_0^2(u_0 + a)(u_0^2 + u_0 + 1) + b + x = 0, \quad (4.5)$$

where  $b \in \mathbb{R}$  and  $a > 0$ . While for  $b = 0$  this model has  $u_0(0) = E_0 = 0$ , a zeroth order phase transition is triggered when  $b > 0$ , where the threshold energy jumps to  $E_0 \simeq a$  and  $u_0(x)$  becomes multi-valued in the region  $x < 0$ . The plot of  $u_0(x)$  for this toy model is analogous to the first two plots in the second row of figure 9. By tuning the parameter  $a$ , the discontinuity in  $E_0$  across the phase transition can be made as small as needed. The full non-perturbative string equation (2.15) associated to this toy model is written after identifying the coefficients  $t_k(a, b)$  from (4.5). Attempting to numerically solve the resulting equation for  $b > 0$ , we find the instabilities are still present, no matter how small the parameter  $a$  is taken to be. This shows the instability issue is not addressed by reducing the magnitude of the jump in  $E_0$ .

There is an analytical argument that has been previously used in refs. [44, 51] to argue in favor of the existence of solutions to the differential equation (2.15) (and variants of it),

<sup>19</sup>Numerical instability can manifest itself in a number of ways. The clearest and most reliable sign of instability, seen here, is when inaccurate solutions are found (sometimes with sudden increase in run time for convergence to occur), and for which there is strong dependence in their shape to small changes in the size of the grid, and its granularity.

when the right hand side has the constant  $\hbar^2\Gamma^2$  instead of 0. It involves taking a t’Hooft limit as  $\hbar \rightarrow 0$  such that  $\Gamma$  becomes large while the combination  $q = \hbar\Gamma$  remains finite. In this limit, the string equation simplifies to the following algebraic constraint:

$$(\hat{u} - E_0) \left( \sum_{k=1}^{\infty} t_k(\lambda) \hat{u}^k + \lambda + x \right)^2 = q^2, \tag{4.6}$$

where  $\hat{u}(x) \equiv \lim_{(\hbar, 1/\Gamma) \rightarrow 0} u(x)$ . Smooth algebraic solutions of this equation for  $\hat{u}(x)$  suggest the existence of solutions  $u(x)$  for other  $\hbar$  and  $\Gamma$ . The algebraic constraint for  $\hat{u}(x)$  in (4.6) can be easily solved numerically and visualized, see figure 11 for a plot for a fixed  $q$  and  $\lambda > \lambda_c$ . The solution  $\hat{u}(x)$  automatically satisfies the appropriate boundary conditions, and for large  $q$  a smooth single-valued solution exists. However for small enough  $q$  the solution  $\hat{u}(x)$  is multi-valued within a certain range in  $x$  (which interestingly occurs for  $x > 0$ ).

The emergence of the multi-valuedness (for small enough  $q$ ) suggests that there is no path connecting possible smooth stable solutions of the full differential equation that may exist at finite  $\Gamma$  to the ones needed at  $\Gamma = 0$ .<sup>20</sup> More generally, it is difficult to see how a multivalued solution of an algebraic equation, our  $u_0(x)$ , (or even one smoothed out by the  $q$ -deformation above) can be completed into a solution of a boundary value problem for an ordinary differential equation once  $\hbar$  is turned on. The multi-valuedness overconstrains the information supplied by the boundary conditions. As a last remark, it is with noting that a multivalued classical potential  $u_0(x)$  also makes no sense from the point of view of the quantum mechanical problem on the real line  $x \in \mathbb{R}$ . At a given point  $x$  there would be an ambiguity as to the value of the force  $F_x = -\partial_x u_0(x)$ .

All the evidence therefore points towards a non-perturbative instability or non-existence of the deformed JT gravity theory for  $\lambda$  across the transition, when  $u_0(x)$  becomes multi-valued in the  $x < 0$  regime. Since  $u_0(x)$  is multi-valued in the finite window  $\lambda \in (0.03480, 0.36733)$ , this means we should be able to construct numerical solutions for  $\lambda$  on the other side of this window, i.e.  $\lambda > 0.36733$ . Figure 12 shows that this is indeed the case, displaying regular solutions to  $u(x)$  for  $\lambda$  beyond the multi-valued window of  $u_0(x)$ . The associated spectral densities are also well-behaved and shown in that figure. Overall, there is enough evidence to reach the following conclusion about the leading  $u_0(x)$  (defined, possibly piecewise, as a continuous function over the whole domain  $-\infty \leq x \leq +\infty$ ):

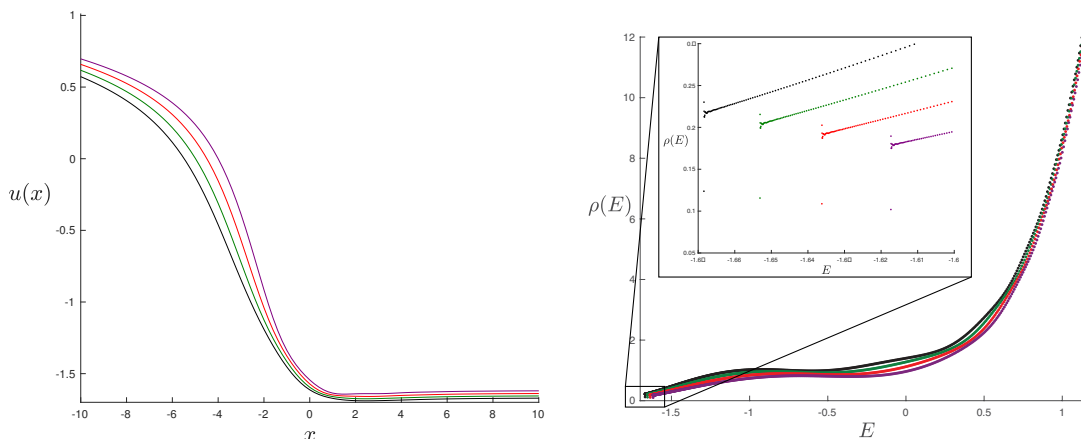
For  $\lambda$  with  $u_0(x)$  multivalued, the system is non-perturbatively unstable

This is the central result of this section.

It is worth remarking that the above statement/result has a more general validity. In the case of  $\lambda = 0$ , ordinary JT gravity, the non-perturbative string equation from a Hermitian matrix model definition is just  $\mathcal{R} = 0$ . There, the full  $u_0(x)$  is multivalued in

---

<sup>20</sup>Such a path can be found when there is no multi-valuedness, by for example using a solution-generating transformation [52, 53] that acts on solutions of the string equation to generate new functions that are solutions of the string equation with  $\Gamma$  changed by an integer. It is in fact a kind of “Bäcklund” transformation, known in the integrable systems literature for changing the soliton number of certain kinds of solutions to integrable systems, here the relevant integrable system is KdV. The transformation can be made explicit, and an extension of it for cases with non-zero  $E_0$  described and derived in appendix C.



**Figure 12.** On the left are numerical solutions for the potential  $u(x)$  obtained by solving the string equation (2.15) for  $\alpha = 1/4$  in the  $k_{\max} = 7$  truncation and  $\lambda$  beyond the window  $\lambda \in (0.03480, 0.36733)$  in which  $u_0(x)$  is multi-valued (see figure 9). Here  $\lambda = 4$  (purple),  $\lambda = 4.5$  (red),  $\lambda = 5$  (green) and  $\lambda = 5.5$  (black). On the right the associated spectral densities obtained from (2.6) and plotted using the same colors.

the  $x > 0$  regime, and the statement about multi-valuedness above is consistent with the fact that the string equation does not seem to have a stable solution. The non-perturbative definition of ref. [25] (and its extension to this paper) is such that the piece of  $u_0(x)$  in the  $x > 0$  regime is replaced by  $u_0 = E_0$ , removing the multi-valuedness, and the string equation (2.15), has a nice (unique) solution. It is interesting to speculate if there are alternative non-perturbative definitions that somehow repair the multi-valuedness in the  $x < 0$  regime when it occurs, maintain perturbation theory, and yield sensible solutions to some new differential equation.

## 5 Final remarks

In this work we have investigated and resolved some puzzles regarding the double scaled matrix models shown in refs. [21, 22] to describe certain deformations of JT gravity. In doing so, we have supplied a non-perturbative completion of the double scaled matrix model physics that extends the proposal for ordinary JT given in ref. [27]. Studying the deformations in detail, we have uncovered an interesting phase structure, in some special cases in qualitative agreement with the semi-classical analysis of the two-dimensional dilaton gravity. Let us mention a few open questions.

**Beyond sharp defects:** the matching between deformations of JT gravity and double scaled matrix models in refs. [21, 22] was performed for potentials  $U(\phi)$  in (1.3) with  $\alpha_i \in (0, 1/2)$ . This corresponds to inserting sharp defects to the JT path integral. The issue with extending this range further is due to the fact that the decomposition of the hyperbolic surfaces in the topological expansion as done in ref. [14] breaks down for  $\alpha_i \in [1/2, 1)$ . This technical difficulty has been very recently overcome in ref. [54].<sup>21</sup> Building on earlier studies

<sup>21</sup>We thank the authors for sharing a manuscript of their work before it was submitted to arXiv.

of deformations of minimal models [55, 56], the authors were able to compute the leading genus string equation for  $\alpha_i \in (0, 1)$ . For a single defect (i.e.  $r = 1$  in (1.3)) it can be written as [54]

$$\mathcal{R}_0[u_0, x] = \frac{\sqrt{u_0}}{2\pi} I_1(2\pi\sqrt{u_0}) + \sum_{n=1}^{\lfloor \frac{1}{1-\alpha} \rfloor} \frac{\lambda^n}{n!} \left[ \frac{2\pi(1-n(1-\alpha))}{\sqrt{u_0}} \right]^{n-1} I_{n-1}(2\pi(1-n(1-\alpha))\sqrt{u_0}) + x. \tag{5.1}$$

While for  $\alpha \in (0, 1/2]$  there is a single term in the summation and we recover (1.9) as given in refs. [21, 22], this equation provides a very interesting and non-trivial generalization in the range  $\alpha \in (1/2, 1)$ .

This string equation supplies the necessary ingredients to extend our matrix model analysis from section 4 to  $\alpha \in (0, 1)$ . In particular, it raises the following question: are the non-positive spectral density  $\rho_0(E)$  and multi-valued potential  $u_0(x)$  observed in figure 9 still present in the regime  $\alpha \in (1/2, 1)$ ? A simple analysis shows the answer is affirmative. For instance, taking  $\alpha = 3/4$  the threshold energy  $E_0$  obtained from the largest solution to  $\mathcal{R}_0[E_0, 0] = 0$  above, exhibits a zeroth order phase transition around  $\lambda \sim 0.17$ . In a completely analogous way as in figure 9, the potential  $u_0(x)$  becomes multi-valued in the region  $x < 0$  and  $\rho_0(E)$  non-positive. Overall, this suggests that while the disc string equation (5.1) becomes quite complicated in the extended range of  $\alpha$ , our analysis in section 4 does not change that much. That being said, studying the equation (5.1) in further detail is an interesting question we would like to explore in future work.

**Normalization of the dilaton potential:** as explained in ref. [22], the computation of the Euclidean partition function from the classical definition of the action in (1.3) entails some ambiguities. These have to do with different choices of operator orderings and renormalization procedures in the quantum theory. This results in an ambiguity regarding the normalization of the potential  $U(\phi)$ , already seen by comparing equations (1.6) and (D.2) in refs. [22] and [21] respectively. In this work we have used the normalization of ref. [22]. The choice of normalization turns out being important, specially for cases in which  $U(\phi)$  in (1.3) contains several terms, as it affects the behavior of  $U(\phi = 0)$ . In particular, it determines whether the phase transition at  $T = 0$  discussed around (3.7) is present or not in the semi-classical analysis. It would be interesting to better understand the correct normalization of the dilaton potential (see section 5.1 in ref. [54] for progress in this direction).

## Acknowledgments

This work is supported by the DOE grant DE-SC0011687 (USC) and NSF grant PHY-1748958 (KITP). CVJ and FR thank Henry Maxfield, Krzysztof Pilch, Joaquin Turiaci, Mykhaylo Usatyuk, Wayne Weng and Edward Witten for comments and conversations. CVJ thanks Amelia for her support.

## A Inverting the Abel transform

In this appendix we invert the Abel transform and derive the simple formula for  $\mathcal{R}_0[u_0, x]$  given in (2.12). Our starting point is (2.11) with  $E_0 = 0$

$$\rho_0(E) = \frac{1}{2\pi\hbar} \int_0^E \frac{du_0}{\sqrt{E-u_0}} (\partial_{u_0} \mathcal{R}_0) = \frac{(g * \mathcal{R}'_0)(E)}{2\pi\hbar}, \quad (\text{A.1})$$

where in the second equality we have defined  $g(y) = 1/\sqrt{y}$ , written the integral as a convolution and the prime is a derivative with respect to  $u_0$ . Applying the Laplace transform, defined in the usual way

$$\mathcal{L}\{f\}(\beta) \equiv \int_0^{+\infty} dE f(E) e^{-\beta E}, \quad (\text{A.2})$$

and using the convolution theorem we find

$$\mathcal{L}\{\rho_0\}(\beta) = \frac{\mathcal{L}\{g\}(\beta)}{2\pi\hbar} \mathcal{L}\{\mathcal{R}'_0\}(\beta) = \frac{1}{2\hbar\sqrt{\pi\beta}} \left[ \beta \mathcal{L}\{\mathcal{R}_0\}(\beta) - \mathcal{R}_0|_{u_0=0} \right], \quad (\text{A.3})$$

where in the second equality we used  $\mathcal{L}\{f'\}(\beta) = \beta \mathcal{L}\{f\}(\beta) - f(0)$  and solved the Laplace transform of  $g$ . From the leading genus string equation (2.9) we can evaluate the boundary term  $\mathcal{R}_0|_{u_0=0} = x$ , given that  $t_0$  must vanish so that  $u_0(0) = E_0 = 0$ . Solving for the Laplace transform of  $\mathcal{R}_0$  we find

$$\mathcal{L}\{\mathcal{R}_0\}(\beta) = 2\hbar\sqrt{\pi/\beta} \mathcal{L}\{\rho_0\}(\beta) + x/\beta = 2\hbar \mathcal{L}\{(g * \rho_0)\}(\beta) + x \mathcal{L}\{1\}(\beta), \quad (\text{A.4})$$

where we have again used the convolution theorem and used  $\mathcal{L}\{1\}(\beta) = 1/\beta$ . Applying the inverse Laplace transform we obtain the explicit formula for  $\mathcal{R}_0$  in terms of  $\rho_0(E)$  given in (2.12).

## B Further details of the non-perturbative definition

Based on the work in refs. [40, 42], in this appendix we motivate in a simple way the non-perturbative differential equation for  $u(x)$  in (2.15) using the assumptions in (2.19). Differentiating the scaling relation (2.19) with respect to  $s$ , setting  $s = 1$  and using the KdV flow equation we find

$$u + \frac{1}{2}u'x + \sum_{k=0}^{\infty} t_k \frac{(2k+1)}{2(k+1)} \tilde{R}'_{k+1}[u] = E_0 \frac{\partial u}{\partial E_0} \quad \implies \quad u\mathcal{R}' + \frac{1}{2}u'\mathcal{R} - \frac{\hbar^2}{4}\mathcal{R}''' = E_0 \frac{\partial u}{\partial E_0}, \quad (\text{B.1})$$

where a prime is a derivative with respect to  $x$ . In the second step we have used the recursion relation satisfied by  $\tilde{R}_k[u]$  (2.8) and we have written everything in terms of  $\mathcal{R}$  defined in (2.7).

Our additional condition is that whatever equation we obtain must perturbatively agree with the Hermitian matrix model, i.e., it must have the solution  $\mathcal{R} = 0$  in  $x < 0$  perturbation theory. Given that  $E_0$  scales the same way as  $u$ , one of the simplest options

is that  $\partial u/\partial E_0$  is proportional to  $\mathcal{R}'$ . Powers of  $\mathcal{R}$  and other powers of  $\mathcal{R}'$  are in principle possible, which would need to be combined with powers of  $x$  to achieve the correct scaling. However, requiring that the boundary condition for  $u(x)$  at large positive  $x$  must be fixed as  $u(x \rightarrow +\infty) = E_0$  (the perturbative continuity discussed in equation (2.17)), means that  $\partial u/\partial E_0$  must go to unity as  $x \rightarrow +\infty$ . Indeed, from the boundary condition  $u(x \rightarrow +\infty) = E_0$  we have  $\lim_{x \rightarrow +\infty} \mathcal{R}' = u'(x) \sum_{k=0}^{\infty} t_k k u(x)^{k-1} + 1 = 1$ , where we used  $\partial_x^n u(x)$  vanishes for  $n > 0$  and large positive  $x$ . No other combinations of  $\mathcal{R}$  and  $\mathcal{R}'$  (combined with powers of  $x$ , for scaling) can satisfy this additional constraint. Therefore,<sup>22</sup>

$$\frac{\partial u}{\partial E_0} = \mathcal{R}' . \tag{B.2}$$

Using this we can multiply both sides in (B.1) by  $\mathcal{R}$ , giving a total derivative, then integrate and obtain:

$$(u - E_0)\mathcal{R}^2 + \frac{\hbar^2}{2}\mathcal{R}\mathcal{R}'' + \frac{\hbar^2}{4}(\mathcal{R}')^2 = 0 , \tag{B.3}$$

where we have fixed the integration constant to zero, so that  $\mathcal{R} = 0$  (the Hermitian model string equation) is a solution. We have arrived at the non-perturbative string equation (2.15) for  $u(x)$ .

Some interesting features of the physics are apparent from this derivation (some of which have been studied in ref. [57]). For example, equation (B.2) shows that as  $E_0$  changes (as it will under some deformations) by some amount  $\delta E_0$ , the corresponding change to the function in the asymptotic region (because  $\mathcal{R}' \rightarrow 1$  there) is  $\delta u = \delta E_0$ , preserving the boundary condition as it should. More generally, equation (B.2) is a new non-perturbative flow equation (joining the KdV ones) for  $u$  as a function of  $E_0$ . Recall also that the parameter  $E_0$  (the end of the matrix model spectrum) defines “the wall”. It arises from the potential that defines the probability measure, now as a function of  $(MM^\dagger + E_0) \geq E_0$ . In the limit of  $E_0 \rightarrow -\infty$  the model is increasingly equivalent to an Hermitian matrix model (arbitrarily negative eigenvalues are now allowed), and indeed the dominant part of the equation (B.3) in that limit is simply  $E_0\mathcal{R}^2 = 0$ , which forces  $u(x)$  to solve the Hermitian matrix model equation  $\mathcal{R} = 0$ .

### C Deriving a Bäcklund transformation

In this appendix we derive a Bäcklund transformation that relates solutions to the differential equation (2.15) when the right-hand side has the constant  $\hbar^2\Gamma^2$  instead of zero. This is a slight generalization of the derivation in refs. [52, 53], where this transformation was derived for  $E_0 = 0$ .

Let us start by assuming we have  $u(x)$ , a solution to the differential equation (2.15). We can use this solution to define two functions  $v_\pm(x)$  according to

$$v_\pm = \frac{\hbar}{2} \left( \frac{\mathcal{R}'[u, x] \mp 2\Gamma}{\mathcal{R}[u, x]} \right) \implies X_\pm \equiv \frac{\hbar}{2} \mathcal{R}'[u, x] - v_\pm \mathcal{R}[u, x] \mp \hbar\Gamma = 0 . \tag{C.1}$$

---

<sup>22</sup>While there are other more complicated scale invariant terms that could be written down that vanish to all orders in perturbation theory (such as  $e^{-\hbar^2/\mathcal{R}}$ ), it is not clear if any of them can be produced from a matrix model, as is the case for the one given in equation (B.2).

Using  $X_{\pm} = 0$  together with the fact  $u(x)$  satisfies the differential equation (2.15) (with  $\hbar^2\Gamma^2$  on the right-hand side) we have the following trivial identity

$$0 = X_{\pm}(X_{\pm} \pm 2\hbar\Gamma) - \hbar\mathcal{R}X'_{\pm} = (u - E_0)\mathcal{R}^2 - \frac{\hbar^2}{2}\mathcal{R}\mathcal{R}'' + \frac{\hbar^2}{4}(\mathcal{R}')^2 - \hbar^2\Gamma^2, \quad (\text{C.2})$$

that we can use to solve for  $u(x)$  in terms of  $v_{\pm}(x)$

$$u(x) = v_{\pm}^2(x) + \hbar v'_{\pm}(x) + E_0. \quad (\text{C.3})$$

This is the Miura transformation, mapping between the KdV and mKdV hierarchies, generalized to non-zero  $E_0$ . The differential equations satisfied by the functions  $v_{\pm}(x)$  can be obtained combining (C.3) and (C.1), so that we find

$$\sum_{k=1}^{\infty} t_k \mathcal{S}_k[v_{\pm}] - v_{\pm}(x + t_0) = \hbar(\pm\Gamma - 1/2). \quad (\text{C.4})$$

where we have defined

$$\mathcal{S}_k[v_{\pm}] \equiv \frac{\hbar}{2} \tilde{R}'_k[v_{\pm}^2 + \hbar v'_{\pm} + E_0] - v_{\pm} \tilde{R}_k[v_{\pm}^2 + \hbar v'_{\pm} + E_0]. \quad (\text{C.5})$$

This is the string equation in the mKdV hierarchy (arising from unitary [58] and Hermitian multi-cut [59] matrix models) generalized to non-zero  $E_0$ .

Let us introduce an additional notation, by adding a subscript to  $u(x)$  and  $v_{\pm}(x)$  that indicates the parameter that appears on the right hand side of their respective differential equations. The solutions to (2.15) and (C.4) in this notation are given by

$$u(x) \equiv u_{\Gamma}(x), \quad v_{\pm}(x) \equiv v_{\pm\Gamma-1/2}(x). \quad (\text{C.6})$$

We now note the functional  $\mathcal{S}_k[v(x)]$  in (C.5) is odd under reflections, i.e.  $\mathcal{S}_k[-v(x)] = -\mathcal{S}_k[v(x)]$ . For instance, if we consider  $k = 1$  so that  $\tilde{R}_{k=1}[f(x)] = f(x)$  we have

$$\mathcal{S}_{k=1}[v(x)] = \frac{\hbar^2}{2} v''(x) - E_0 v(x) - v(x)^3 = -\mathcal{S}_{k=1}[-v(x)]. \quad (\text{C.7})$$

While this property was noted in refs. [52, 58] for  $E_0 = 0$  it is non-trivially extended to non-zero  $E_0$ . For any arbitrary value of  $k > 1$  it is straightforward to check it continues to hold. This turns out to be very useful, as it gives us another way of relating solutions to the differential equation in (C.4). Given any solution  $v_c(x)$  (with constant  $\hbar c$  on the right hand side), we can use the reflection property satisfied by  $\mathcal{S}_k$  to generate a solution with opposite constant, i.e.  $v_{-c}(x) = -v_c(x)$ .

We now have all the ingredients necessary to derive the Bäcklund transformation for  $u_{\Gamma}(x)$ , shown below in (C.11). Let us prove it for  $(\Gamma - 1)$ , the other case being completely analogous. The starting point is (C.3), which in the notation in (C.6) becomes

$$u_{\Gamma}(x) = v_{\pm\Gamma-1/2}^2(x) + \hbar v'_{\pm\Gamma-1/2}(x) + E_0. \quad (\text{C.8})$$

Changing  $\Gamma \rightarrow \Gamma - 1$  and using  $v_{-c}(x) = -v_c(x)$  we can transform this to

$$u_{\Gamma-1}(x) = v_{\Gamma-1/2}^2(x) - \hbar v'_{\Gamma-1/2}(x) + E_0, \quad (\text{C.9})$$



which combined with (C.8) with the upper sign gives

$$u_{\Gamma-1}(x) = 2v_{\Gamma-1/2}^2(x) - u_{\Gamma}(x) + 2E_0 . \quad (\text{C.10})$$

The Bäcklund transformation is obtained after rewriting the first term in this expression using (C.1):

$$u_{\Gamma\pm 1}(x) = \frac{\hbar^2}{2} \left( \frac{\mathcal{R}'[u_{\Gamma}, x] \pm 2\Gamma}{\mathcal{R}[u_{\Gamma}, x]} \right)^2 - u_{\Gamma}(x) + 2E_0 . \quad (\text{C.11})$$

This is a useful transformation in a number of situations. For example, sometimes it is easier to solve the string equation at another value of  $\Gamma$  (because for example, it has a shallower potential well), and then afterwards apply the transformation to change  $\Gamma$ . This was sometimes done in ref. [27] as a convenience (e.g., solving for  $\Gamma = 1$  and then reducing to  $\Gamma = 0$  afterwards), although the current paper did not need to employ this method since the numerical methods used have since become more powerful for solving directly at  $\Gamma = 0$ . It is also possible to imagine using this transformation successively to reduce  $\Gamma$  from some high integer value down to zero, having started with it large, or even formally infinite.

Finally, note that in the 't Hooft limit of section 4.3 where the solution is algebraic, it is easy to see that there is no path to smaller  $\Gamma$  using these transformations. Since here  $u_{\Gamma}(x)$  is  $\hat{u}(x)$ , a solution to equation (4.6) valid in the assumed limit, the same limit should be used in the transformation. This means that the transformation becomes:

$$u_{\Gamma\pm 1}(x) = \frac{2q^2}{\mathcal{R}_0^2[\hat{u}, x]} - \hat{u}(x) + 2E_0 = \hat{u}(x), \quad (\text{C.12})$$

where in the last step equation (4.6) was used to write  $q^2/\mathcal{R}_0^2[\hat{u}, x] = \hat{u} - E_0$ . In other words, the 't Hooft limit solution  $\hat{u}(x)$  is invariant under the Bäcklund transformation.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

## References

- [1] R. Jackiw, *Lower Dimensional Gravity*, *Nucl. Phys. B* **252** (1985) 343 [[INSPIRE](#)].
- [2] C. Teitelboim, *Gravitation and Hamiltonian Structure in Two Space-Time Dimensions*, *Phys. Lett. B* **126** (1983) 41 [[INSPIRE](#)].
- [3] A. Achucarro and M.E. Ortiz, *Relating black holes in two-dimensions and three-dimensions*, *Phys. Rev. D* **48** (1993) 3600 [[hep-th/9304068](#)] [[INSPIRE](#)].
- [4] A. Fabbri, D.J. Navarro and J. Navarro-Salas, *Quantum evolution of near extremal Reissner-Nordstrom black holes*, *Nucl. Phys. B* **595** (2001) 381 [[hep-th/0006035](#)] [[INSPIRE](#)].
- [5] P. Nayak, A. Shukla, R.M. Soni, S.P. Trivedi and V. Vishal, *On the Dynamics of Near-Extremal Black Holes*, *JHEP* **09** (2018) 048 [[arXiv:1802.09547](#)] [[INSPIRE](#)].
- [6] K.S. Kolekar and K. Narayan, *AdS<sub>2</sub> dilaton gravity from reductions of some nonrelativistic theories*, *Phys. Rev. D* **98** (2018) 046012 [[arXiv:1803.06827](#)] [[INSPIRE](#)].

- [7] A. Ghosh, H. Maxfield and G.J. Turiaci, *A universal Schwarzian sector in two-dimensional conformal field theories*, *JHEP* **05** (2020) 104 [[arXiv:1912.07654](#)] [[INSPIRE](#)].
- [8] S. Sachdev, *Universal low temperature theory of charged black holes with  $AdS_2$  horizons*, *J. Math. Phys.* **60** (2019) 052303 [[arXiv:1902.04078](#)] [[INSPIRE](#)].
- [9] S. Sachdev and J. Ye, *Gapless spin fluid ground state in a random, quantum Heisenberg magnet*, *Phys. Rev. Lett.* **70** (1993) 3339 [[cond-mat/9212030](#)] [[INSPIRE](#)].
- [10] A. Kitaev, *A simple model of quantum holography*, talks at KITP, 7 April 2015 and 27 May 2015.
- [11] J. Maldacena and D. Stanford, *Remarks on the Sachdev-Ye-Kitaev model*, *Phys. Rev. D* **94** (2016) 106002 [[arXiv:1604.07818](#)] [[INSPIRE](#)].
- [12] K. Jensen, *Chaos in  $AdS_2$  Holography*, *Phys. Rev. Lett.* **117** (2016) 111601 [[arXiv:1605.06098](#)] [[INSPIRE](#)].
- [13] G. Sárosi,  *$AdS_2$  holography and the SYK model*, *PoS Modave2017* (2018) 001 [[arXiv:1711.08482](#)] [[INSPIRE](#)].
- [14] P. Saad, S.H. Shenker and D. Stanford, *JT gravity as a matrix integral*, [arXiv:1903.11115](#) [[INSPIRE](#)].
- [15] M. Mirzakhani, *Simple geodesics and Weil-Petersson volumes of moduli spaces of bordered Riemann surfaces*, *Invent. Math.* **167** (2006) 179.
- [16] B. Eynard, *Topological expansion for the 1-Hermitian matrix model correlation functions*, *JHEP* **11** (2004) 031 [[hep-th/0407261](#)] [[INSPIRE](#)].
- [17] B. Eynard and N. Orantin, *Invariants of algebraic curves and topological expansion*, *Commun. Num. Theor. Phys.* **1** (2007) 347 [[math-ph/0702045](#)] [[INSPIRE](#)].
- [18] D.J. Gross and A.A. Migdal, *A Nonperturbative Treatment of Two-dimensional Quantum Gravity*, *Nucl. Phys. B* **340** (1990) 333 [[INSPIRE](#)].
- [19] M.R. Douglas and S.H. Shenker, *Strings in Less Than One-Dimension*, *Nucl. Phys. B* **335** (1990) 635 [[INSPIRE](#)].
- [20] E. Brézin, V.A. Kazakov and A.B. Zamolodchikov, *Scaling Violation in a Field Theory of Closed Strings in One Physical Dimension*, *Nucl. Phys. B* **338** (1990) 673 [[INSPIRE](#)].
- [21] H. Maxfield and G.J. Turiaci, *The path integral of 3D gravity near extremality; or, JT gravity with defects as a matrix integral*, *JHEP* **01** (2021) 118 [[arXiv:2006.11317](#)] [[INSPIRE](#)].
- [22] E. Witten, *Matrix Models and Deformations of JT Gravity*, *Proc. Roy. Soc. Lond. A* **476** (2020) 20200582 [[arXiv:2006.13414](#)] [[INSPIRE](#)].
- [23] T.G. Mertens and G.J. Turiaci, *Liouville quantum gravity – holography, JT and matrices*, *JHEP* **01** (2021) 073 [[arXiv:2006.07072](#)] [[INSPIRE](#)].
- [24] K. Okuyama and K. Sakai, *JT gravity, KdV equations and macroscopic loop operators*, *JHEP* **01** (2020) 156 [[arXiv:1911.01659](#)] [[INSPIRE](#)].
- [25] C.V. Johnson, *Nonperturbative Jackiw-Teitelboim gravity*, *Phys. Rev. D* **101** (2020) 106023 [[arXiv:1912.03637](#)] [[INSPIRE](#)].
- [26] C.V. Johnson, *Jackiw-Teitelboim supergravity, minimal strings, and matrix models*, *Phys. Rev. D* **103** (2021) 046012 [[arXiv:2005.01893](#)] [[INSPIRE](#)].

- [27] C.V. Johnson, *Explorations of nonperturbative Jackiw-Teitelboim gravity and supergravity*, *Phys. Rev. D* **103** (2021) 046013 [[arXiv:2006.10959](#)] [[INSPIRE](#)].
- [28] T. Banks, M.R. Douglas, N. Seiberg and S.H. Shenker, *Microscopic and Macroscopic Loops in Nonperturbative Two-dimensional Gravity*, *Phys. Lett. B* **238** (1990) 279 [[INSPIRE](#)].
- [29] M.R. Douglas, N. Seiberg and S.H. Shenker, *Flow and Instability in Quantum Gravity*, *Phys. Lett. B* **244** (1990) 381 [[INSPIRE](#)].
- [30] P.H. Ginsparg and G.W. Moore, *Lectures on 2 – D gravity and 2 – D string theory*, in *Theoretical Advanced Study Institute (TASI 92): From Black Holes and Strings to Particles*, 10, 1993, [hep-th/9304011](#) [[INSPIRE](#)].
- [31] P. Di Francesco, P.H. Ginsparg and J. Zinn-Justin, *2 – D Gravity and random matrices*, *Phys. Rept.* **254** (1995) 1 [[hep-th/9306153](#)] [[INSPIRE](#)].
- [32] B. Eynard, T. Kimura and S. Ribault, *Random matrices*, [arXiv:1510.04430](#) [[INSPIRE](#)].
- [33] G. 't Hooft, *A Planar Diagram Theory for Strong Interactions*, *Nucl. Phys. B* **72** (1974) 461 [[INSPIRE](#)].
- [34] E. Brézin, C. Itzykson, G. Parisi and J.B. Zuber, *Planar Diagrams*, *Commun. Math. Phys.* **59** (1978) 35 [[INSPIRE](#)].
- [35] S. Dalley, C.V. Johnson and T.R. Morris, *Classification of critical hermitian matrix models*, *Mod. Phys. Lett. A* **6** (1991) 439 [[INSPIRE](#)].
- [36] V.A. Kazakov, *The Appearance of Matter Fields from Quantum Fluctuations of 2D Gravity*, *Mod. Phys. Lett. A* **4** (1989) 2125 [[INSPIRE](#)].
- [37] D. Bessis, C. Itzykson and J.B. Zuber, *Quantum field theory techniques in graphical enumeration*, *Adv. Appl. Math.* **1** (1980) 109 [[INSPIRE](#)].
- [38] I. Gelfand and L. Dikii, *Asymptotic behavior of the resolvent of Sturm-Liouville equations and the algebra of the Korteweg-De Vries equations*, *Russ. Math. Surveys* **30** (1975) 77.
- [39] T.R. Morris, *Checkered surfaces and complex matrices*, *Nucl. Phys. B* **356** (1991) 703 [[INSPIRE](#)].
- [40] S. Dalley, C.V. Johnson and T.R. Morris, *Multicritical complex matrix models and nonperturbative 2 – D quantum gravity*, *Nucl. Phys. B* **368** (1992) 625 [[INSPIRE](#)].
- [41] S. Dalley, C.V. Johnson and T.R. Morris, *Nonperturbative two-dimensional quantum gravity*, *Nucl. Phys. B* **368** (1992) 655 [[INSPIRE](#)].
- [42] S. Dalley, *On loop equations in KdV exactly solvable string theory*, *Mod. Phys. Lett. A* **7** (1992) 1263 [[hep-th/9111064](#)] [[INSPIRE](#)].
- [43] S. Dalley, C.V. Johnson and T.R. Morris, *Nonperturbative two-dimensional quantum gravity, again*, *Nucl. Phys. B Proc. Suppl.* **25** (1992) 87 [[hep-th/9108016](#)] [[INSPIRE](#)].
- [44] I.R. Klebanov, J.M. Maldacena and N. Seiberg, *Unitary and complex matrix models as 1 – D type 0 strings*, *Commun. Math. Phys.* **252** (2004) 275 [[hep-th/0309168](#)] [[INSPIRE](#)].
- [45] M.R. Douglas, *Strings in Less Than One-dimension and the Generalized  $K^-D^-V$  Hierarchies*, *Phys. Lett. B* **238** (1990) 176 [[INSPIRE](#)].
- [46] D. Louis-Martinez and G. Kunstatter, *On Birckhoff's theorem in 2 – D dilaton gravity*, *Phys. Rev. D* **49** (1994) 5227 [[INSPIRE](#)].

- [47] C.R. Nappi and A. Pasquinucci, *Thermodynamics of two-dimensional black holes*, *Mod. Phys. Lett. A* **7** (1992) 3337 [[gr-qc/9208002](#)] [[INSPIRE](#)].
- [48] E. Witten, *Deformations of JT Gravity and Phase Transitions*, [arXiv:2006.03494](#) [[INSPIRE](#)].
- [49] A. Chamblin, R. Emparan, C.V. Johnson and R.C. Myers, *Charged AdS black holes and catastrophic holography*, *Phys. Rev. D* **60** (1999) 064018 [[hep-th/9902170](#)] [[INSPIRE](#)].
- [50] G. Bonnet, F. David and B. Eynard, *Breakdown of universality in multicut matrix models*, *J. Phys. A* **33** (2000) 6739 [[cond-mat/0003324](#)] [[INSPIRE](#)].
- [51] R. Iyer, C.V. Johnson and J.S. Pennington, *Non-perturbative String Theory from Water Waves*, *J. Phys. A* **44** (2011) 375401 [[arXiv:1011.6354](#)] [[INSPIRE](#)].
- [52] S. Dalley, C.V. Johnson, T.R. Morris and A. Watterstam, *Unitary matrix models and 2 – D quantum gravity*, *Mod. Phys. Lett. A* **7** (1992) 2753 [[hep-th/9206060](#)] [[INSPIRE](#)].
- [53] J.E. Carlisle, C.V. Johnson and J.S. Pennington, *Backlund transformations, D-branes, and fluxes in minimal type 0 strings*, *J. Phys. A* **40** (2007) 12451 [[hep-th/0501006](#)] [[INSPIRE](#)].
- [54] G.J. Turiaci, M. Usatyuk and W.W. Weng, *Dilaton-gravity, deformations of the minimal string, and matrix models*, [arXiv:2011.06038](#) [[INSPIRE](#)].
- [55] G.W. Moore, N. Seiberg and M. Staudacher, *From loops to states in 2 – D quantum gravity*, *Nucl. Phys. B* **362** (1991) 665 [[INSPIRE](#)].
- [56] A.A. Belavin and A.B. Zamolodchikov, *On Correlation Numbers in 2D Minimal Gravity and Matrix Models*, *J. Phys. A* **42** (2009) 304004 [[arXiv:0811.0450](#)] [[INSPIRE](#)].
- [57] C.V. Johnson, T.R. Morris and P.L. White, *The Boundary cosmological constant in stable 2 – D quantum gravity*, *Phys. Lett. B* **292** (1992) 283 [[hep-th/9206066](#)] [[INSPIRE](#)].
- [58] V. Periwal and D. Shevitz, *Exactly Solvable Unitary Matrix Models: Multicritical Potentials and Correlations*, *Nucl. Phys. B* **344** (1990) 731 [[INSPIRE](#)].
- [59] C. Crnkovic and G.W. Moore, *Multicritical multicut matrix models*, *Phys. Lett. B* **257** (1991) 322 [[INSPIRE](#)].