

Black hole/string transition for the small Schwarzschild black hole of $AdS_5 \times S^5$ and critical unitary matrix models

L. Álvarez-Gaumé^{1,a}, P. Basu^{2,b}, M. Mariño^{1,c}, S.R. Wadia^{2,d}

¹ CERN, 1211, Geneva, Switzerland

² Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai-400005, India

Received: 24 August 2006 /

Published online: 17 November 2006 – © Springer-Verlag / Società Italiana di Fisica 2006

Abstract. In this paper we discuss the black hole–string transition of the small Schwarzschild black hole of $AdS_5 \times S^5$ using the AdS/CFT correspondence at finite temperature. The finite temperature gauge theory effective action, at weak *and* strong coupling, can be expressed entirely in terms of constant Polyakov lines which are $SU(N)$ matrices. In showing this we have taken into account that there are no Nambu–Goldstone modes associated with the fact that the 10-dimensional black hole solution sits at a point in S^5 . We show that the phase of the gauge theory in which the eigenvalue spectrum has a gap corresponds to supergravity saddle points in the bulk theory. We identify the third order $N = \infty$ phase transition with the black hole–string transition. This singularity can be resolved using a double scaling limit in the transition region where the large N expansion is organized in terms of powers of $N^{-2/3}$. The $N = \infty$ transition now becomes a smooth crossover in terms of a renormalized string coupling constant, reflecting the physics of large but finite N . Multiply wound Polyakov lines condense in the crossover region. We also discuss the implications of our results for the resolution of the singularity of the lorentzian section of the small Schwarzschild black hole.

1 Introduction and synopsis

The problem of the fate of small Schwarzschild black holes is important to understand, in a quantum theory of gravity. In a unitary theory this problem is the same as the formation of a small black hole. An understanding of this phenomenon has bearing on the problem of spacelike singularities in quantum gravity and also (to some extent) on the information puzzle in black hole physics. It would also teach us something about non-perturbative string physics.

In the past Susskind [1], Horowitz and Polchinski (SHP) [2] and others [3–5] have discussed this, in the framework of string theory, as a black hole–string transition or more appropriately a crossover. Their proposal is that this crossover is parametrically smooth and it simply amounts to a change of description of the same quantum state in terms of degrees of freedom appropriate to the strength of the string coupling. The entropy and mass of the state change at most by $o(1)$. By matching the entropy formulas for black holes and perturbative string states, they arrived at a crude estimate of the small but non-zero string coupling at the crossover. The SHP description is difficult to make more precise because a formulation of

string theory in the crossover regime is not yet explicitly known.

There are many studies on the black hole–string transition and the nature of the black hole singularity in the case of two- and three-dimensional black holes [10–15]. Small extremal supersymmetric black holes have been discussed in string theory with enormous success [6–9]. In particular the α' corrections to the entropy of the supersymmetric string sized black holes has been matched to the microscopic counting.

In this paper we discuss the black hole–string crossover for the small 10-dimensional Schwarzschild black hole in the framework of the AdS/CFT correspondence. In [16], building on the work of [17–22], a simplified model for the thermal history of small and big black holes in AdS_5 (which were originally discussed by Hawking and Page [23]) was discussed in detail. In particular, the large N Gross–Witten–Wadia (GWW) transition [24–26] was identified with the SHP transition for the small AdS_5 black hole.

However, it turns out that the small black hole in AdS_5 , which is uniformly spread over S^5 , has a Gregory–Laflamme instability¹. When the horizon radius $r_h \sim$

^a e-mail: Luis.Alvarez-Gaume@cern.ch

^b e-mail: pallab@theory.tifr.res.in

^c e-mail: Marcos.Marino.Beiras@cern.ch

^d e-mail: wadia@theory.tifr.res.in

¹ This point was brought to our attention by Aharony and Minwalla. Understanding the Gregory–Laflamme transition from a boundary perspective is an important open issue. In the current work we will not try to address this and only assume the existence of such a transition.

$0.5R$ [27] the $l = 1$ perturbation is unstable. The final configuration this instability leads to, as r_h decreases and the horizon becomes less and less uniform over S^5 , is most likely to be the 10-dimensional Schwarzschild black hole. This small 10-dimensional Schwarzschild black hole does not have any further instability of Gregory–Laflamme type. This black hole also happens to be a solution with asymptotic $AdS_5 \times S^5$ geometry for $l_s \ll r_h \ll R$ (16).

When the horizon of this black hole approaches the string scale l_s , we expect the supergravity (geometric) description to break down and be replaced by a description in terms of degrees of freedom more appropriate at this scale. Presently we have no idea how to discuss this crossover in the bulk IIB string theory. Hence we will discuss this transition and its smoothening in the framework of a general finite temperature effective action of the dual $SU(N)$ gauge theory on $S^3 \times S^1$. In fact it is fair to say that in the crossover region we are really using the gauge theory as a definition of the non-perturbative string theory.

At large but finite N , since S^3 is compact, the partition function and all correlation functions are smooth functions of the temperature and other chemical potentials. There is no phase transition. However, in order to make a connection with a theory of gravity, which has an infinite number of degrees of freedom, we have to take the $N \rightarrow \infty$ limit and study the saddle point expansion in powers of $\frac{1}{N}$. It is this procedure that leads to non-analytic behavior. It turns out that by taking into account exact results in the $\frac{1}{N}$ expansion it is possible to resolve this singularity and recover a smooth crossover in a suitable double scaling limit.

In the specific problem at hand, it turns out that in the transition region the large N expansion is organized in powers of $N^{-2/3}$. In the bulk theory, assuming AdS/CFT, this would naively mean a string coupling expansion in powers of $g_s^{2/3}$. However in a double scaling limit, a renormalized string coupling $\tilde{g} = N^{\frac{2}{3}}(\beta_c - \beta)$ once again organizes the coupling constant expansion in integral powers. The free energy and correlators are smooth functions of \tilde{g} .

The use of the AdS/CFT correspondence for studying the black hole–string crossover requires that there is a description of small Schwarzschild black holes as solutions of type IIB string theory in $AdS_5 \times S^5$. Fortunately, Horowitz and Hubeny [28] have studied this problem with a positive conclusion. This result enables us to use the boundary gauge theory to address the crossover of the small Schwarzschild black hole into a state described in terms of ‘stringy’ degrees of freedom. Even so the gauge theory is very hard to deal with as we have to solve it in the $\frac{1}{N}$ expansion for large but finite values of the ’t Hooft coupling λ .

However, there is a window of opportunity to do some precise calculations, because it can be shown that the effective action of the gauge theory at finite temperature can be expressed entirely in terms of the Polyakov loop which does not depend on points on S^3 . This is a single $N \times N$ unitary matrix, albeit with a complicated interaction among the winding modes $\text{Tr } U^n$. This circumstance, that the order parameter U in the gauge theory is a constant on S^3 , matches well on the supergravity side with the fact

that all the zero angular momentum black hole solutions are also invariant under the $SO(4)$ symmetry of S^3 . The black hole may be localized in S^5 , but it does not depend on the co-ordinates of S^3 . The coefficients of the effective action depend upon the temperature, the ’t Hooft coupling λ and the VEVs of the scalar fields. Since the 10-dimensional black hole sits at a point in S^5 , one may be concerned about the spontaneous breaking of $SO(6)$ R -symmetry and corresponding Nambu–Goldstone modes. We will conclude, using a supergravity analysis, that the symmetry is not spontaneously broken. Instead we have to introduce collective coordinates for treating the zero modes associated with this symmetry.

The general unitary matrix model can be analyzed due to technical progress we have made in discussing the general multi-trace unitary matrix model. We prove an identity that enables us to express and study the critical properties of a general multi-trace unitary matrix model in terms of the critical properties of a general single trace matrix model.

As is well known, the single trace unitary matrix model at $N = \infty$ has a third order GWW transition, which occurs when the density of eigenvalues of the unitary matrix develop a gap on the unit circle. The vanishing of the density at a point on the circle leads to a relation among the coupling constants of the matrix model which defines a surface in the space of couplings (parameters of the effective action). The behavior of the matrix model in the neighborhood of this surface (call it the critical surface) is characterized by universal properties which are entirely determined by the way the gap in the eigenvalue density opens: $\rho(\theta) \sim (\pi - \theta)^{2m}$, where m is a positive integer. In our problem, there is only one tunable parameter, namely the temperature. Hence we will mainly focus only on the lowest, $m = 1$, critical point and present the relevant operator that opens the gap. We also discuss the possible relevance of higher order multi-critical points.

Using the properties of the $\frac{1}{N}$ expansion near and away from the critical surface, we will argue that the small black hole (or for that matter any saddle point of supergravity around which a well defined closed string perturbation expansion exists) corresponds to the phase of the matrix model where the density of eigenvalues on the unit circle has a gap. The small black hole therefore corresponds to the gapped phase of the unitary matrix model.

We make a reasonable physical assumption based on the proposal of SHP: that the thermal history of the unstable saddle point corresponding to the small black hole eventually intersects the critical surface at a critical temperature T_c , which is $o(1/l_s)$. T_c is smaller than the Hagedorn temperature. Once the thermal history crosses the critical surface, it would eventually meet the $AdS_5 \times S^5$ critical point corresponding to a uniform eigenvalue distribution. (Such a history was already discussed in the context of a simplified model in [16].) It is natural to identify the crossover across the critical surface in the gauge theory as the bulk black hole–string crossover.

At the crossover, the $o(1)$ part of the gauge theory partition function (which depends on the renormalized string

coupling) can be exactly calculated in a double scaling limit. This is a universal result in the sense that it does not depend on the location of the critical point on the critical surface but depends only on deviations which are normal to the critical surface. If we parametrize this by t , the free energy $-F(t)$ solves the differential equation $\frac{\partial^2 F}{\partial t^2} = -f^2(t)$ where $f(t)$ satisfies the Painlevé II equation. The exact analytic form of $F(t)$ is not known, but $F(t)$ is a smooth function in the domain $(-\infty, \infty)^2$. All the operators $\rho_k = \frac{\text{Tr} U^k}{N}$ condense in the crossover region. In fact $\left\langle N^{\frac{2}{3}} (\rho_k - \rho_k^{\text{ug}}) \right\rangle = C_k \frac{d}{dt} F$, where $C_k = \frac{(-1)^k}{k}$ and ρ_k^{ug} represents the expectation value of ρ_k in the ungapped phase.

The smooth crossover of the euclidean black hole possibly has implications for the resolution of the singularity of the lorentzian black hole, because within the AdS/CFT correspondence we should be able to address all physical questions of the bulk theory in the corresponding gauge theory. In particular, we should be able to address phenomena both outside and inside the black hole horizon.

The plan of this paper is as follows. Section 2 discusses the SHP transition. Section 3 discusses the small 10-dimensional black hole in $AdS_5 \times S^5$. Section 4 discusses the finite temperature gauge theory and the effective action in terms of the unitary matrix model. Section 5 presents the multi-trace partition function as the calculable integral transform of the single trace unitary matrix model. Section 6 discusses critical behavior in the unitary matrix model. Section 7 discusses the saddle point equations of the matrix model. Section 8 discusses the double scaled partition function. Section 9 discusses the introduction of chemical potentials and higher critical points. Section 10 discusses the applications of the critical matrix model to the small 10-dimensional black hole. Section 11 discusses the lorentzian black hole.

2 Black hole–string transition

In this section we review the black hole–string crossover. Consider the 10-dimensional Schwarzschild black hole. As long as its horizon radius $r_h \gg l_s$ (l_s is the string length), the supergravity description is valid, and we can trust the lowest order effective action in l_s . When $r_h \sim l_s$, this description breaks down and one learns to derive an effective action valid to all orders in l_s or devises other methods to deal with the problem. Let us assume that the description in all orders of l_s is available; then presumably the geometrical description is still valid in principle, and one can indeed discuss the notion of a string size horizon with radius $r_h \sim l_s$ [6, 7, 9]. It is reasonable to expect that in such a description the qualitative fact that the mass decreases with the horizon radius and increasing temperature, is still valid. These facts are obviously valid to lowest order in l_s , because $r_h = 2G_N M$ and $T_h = (G_N M)^{1/7}$. Here G_N is Newton’s coupling and M is

the mass of the black hole. For definitiveness let us fix the mass and the entropy of the black hole. Then r_h and T_h vary with the gravitational coupling G_N . Now, since $g_s^2 = G_N l_s^{-8}$, we can say that r_h and T_h vary with g_s , and hence a crossover at $r_h \sim l_s$ occurs at a specific value of the string coupling.

When $r_h \lesssim l_s$ the above description of the state has to be replaced by a description in terms of microscopic degrees of freedom relevant to the scale l_s . Even in this description it is reasonable to assume that the temperature of the state varies as we change the string coupling. The assumption of Susskind–Horowitz–Polchinski is that the mass of the state would change by at most $o(1)$ in the string coupling.

From the above discussion it is clear that the black hole–string crossover occurs in a regime where the curvature of the black hole is $o(1)$ in string units, so as to render the supergravity description invalid. It is also clear that besides l_s related effects, the string coupling is non-zero and its effects have to be taken into account. Presently our understanding of string theory is not good enough for us to make a precise and quantitative discussion of the crossover. Hence we will discuss the problem using the AdS/CFT correspondence. In order to do this we need to be able to embed the small black hole in $AdS_5 \times S^5$. This has been discussed by Horowitz and Hubeny [28]. We briefly review their work in the next section.

3 Embedding the 10-dimensional Schwarzschild black hole in $AdS_5 \times S^5$

It is not difficult to argue that the small 10-dimensional Schwarzschild black hole exists as a solution of Einstein’s equation in $AdS_5 \times S^5$. A small patch of the $AdS_5 \times S^5$ space is locally identical to 10-dimensional euclidean space. Since the horizon radius of this black hole $r_h \ll R$, we can have a solution which is locally identical to a 10-dimensional Schwarzschild black hole in flat space-time. We would also require that the solution for large 10-dimensional radial distances asymptotically goes to $AdS_5 \times S^5$. This solution is not explicitly known, but it can be numerically constructed given the boundary conditions on the radial functions. The more non-trivial issue is concerning the fact that the type IIB theory also has a 5-form. In the absence of the black hole this form is the volume form of S^5 and carries N units of flux. It turns out that, in the presence of the small black hole, a consistent solution to the equations of motion is such that there is no energy flux into the black hole. Hence the small black hole remains small. In the above analysis one neglects the back reaction on the metric due to the fact that the black hole is small and the curvature near its horizon is large.

The solution is conveniently represented if we use a 10-dimensional radial coordinate system (fixed by the area of S^8) in $AdS_5 \times S^5$. One splits S^8 into S^3 and S^4 , corresponding to the rotational $SO(4)$ symmetry of AdS_5 and the remaining (unbroken) $SO(5)$ symmetry of S^5 . This is achieved by using the following coordinate transformation

² This universal formula also appeared in the discussion of the simplified model in [16]

in (16):

$$\begin{aligned} r &= \rho \sin \theta, \\ \chi &= \rho \cos \theta. \end{aligned} \quad (1)$$

In these coordinates, a flat patch within AdS is achieved in the limit $r, \xi \ll R$, where R is the radius of AdS_5 . The metric takes the form

$$ds^2 = -dt^2 + d\rho^2 + \rho^2 (d\theta^2 + \sin^2 \theta d\Omega_3^2 + \cos^2 \theta d\Omega_4^2). \quad (2)$$

(The angular term in parentheses is equivalent to $d\Omega_8^2$). Similarly the 5-form field strength takes the form

$$\begin{aligned} F &= -\rho^3 \sin^4 \theta dt \wedge d\rho \wedge d\Omega_3 \\ &\quad - \rho^4 \sin^3 \theta \cos \theta dt \wedge d\theta \wedge d\Omega_3 + r^4 \cos^5 \theta d\rho \wedge d\Omega_4 \\ &\quad - r^5 \sin \theta \cos^4(\theta) d\theta \wedge d\Omega_4. \end{aligned} \quad (3)$$

In this metric the Schwarzschild solution is given by

$$\begin{aligned} ds^2 &= -f(\rho) dt^2 + f^{-1}(\rho) d\rho^2 \\ &\quad + \rho^2 (d\theta^2 + \sin^2 \theta d\Omega_3^2 + \cos^2 \theta d\Omega_4^2) \\ F &= g_1(\rho, \theta) [-\rho^3 \sin^4 \theta dt \wedge d\rho \wedge d\Omega_3 \\ &\quad - r^5 \sin \theta \cos^4 \theta d\theta \wedge d\Omega_4] \\ &\quad + g_2(\rho, \theta) [\rho^4 \sin^3 \theta \cos \theta dt \wedge d\theta \wedge d\Omega_3 \\ &\quad + r^4 \cos^5 \theta d\rho \wedge d\Omega_4], \end{aligned} \quad (4)$$

where near the black hole horizon $f = 1 - \frac{r_+^7}{r^7}$. As $r \rightarrow \infty$, the functions $f(r), g_1(r, \theta), g_2(r, \theta)$ approach their corresponding values in $AdS_5 \times S^5$ geometry. The explicit solutions for these functions are not known, but their form can be determined by numerically integrating a set of coupled linear differential equations. These solutions have the desired property that imply that the small black hole remains small.

4 Finite temperature gauge theory, order parameter and effective action

We first present a general discussion of the order parameter of $SU(N)$ YM theory on the compact manifold S^3 . We consider the theory in the canonical ensemble, i.e. the euclidean time direction is periodically identified with a period of $\beta = \frac{1}{T}$. It was pointed out in [17, 46] that the Yang–Mills theory partition function on S^3 at a temperature T can be reduced to an integral over a unitary $SU(N)$ matrix U , which is the zero mode of Polyakov loop on the euclidean time circle. This analysis was done in the limit when the 't Hooft coupling $\lambda \rightarrow 0$. We have

$$Z(\lambda, T) = \int dU e^{S(U)}, \quad (6)$$

with

$$U = P \exp \left(i \int_0^\beta A_0 d\tau \right), \quad (7)$$

where $A_0(\tau)$ is the zero mode of the time component of the gauge field on S^3 . This follows from the fact that apart from A_0 all modes of the gauge theory on S^3 are massive. We will discuss the validity of the above expression in both the strong and weak (λ) coupling regimes. Hence we can use U as an order parameter. Gauge invariance requires that the effective action of U be expressed in terms of products of $\text{Tr } U^n$, with n an integer, since these are the only gauge invariant quantities that can be constructed from A_0 alone. $S_{\text{eff}}(U)$ also has a Z_N symmetry under

$$U \rightarrow e^{\frac{2\pi i}{N}} U. \quad (8)$$

This is due to the global gauge transformations which are periodic in the euclidean time direction up to Z_N factors. Z_N invariance puts a further restriction on the form of the effective action, and a generic term in $S(U, U^\dagger)$ will have the form

$$\begin{aligned} &\text{Tr } U^{n_1} \text{Tr } U^{n_2} \dots \text{Tr } U^{n_k}, \\ &n_1 + \dots + n_k = 0 \pmod{N}, \quad k > 1. \end{aligned} \quad (9)$$

In the large N limit we can work with $U(N)$ rather than $SU(N)$, and in that case Z_N is replaced by $U(1)$.

We can expand S_{eff} in terms of a complete set of such operators. The first few terms are

$$\begin{aligned} S(U, U^\dagger) &= a \text{Tr } U \text{Tr } U^{-1} + \frac{b}{N} \text{Tr } U^2 \text{Tr } U^{-1} \text{Tr } U^{-1} \\ &\quad + \frac{c}{N^2} \text{Tr } U^3 \text{Tr } U^{-1} \text{Tr } U^{-1} \text{Tr } U^{-1} + \dots \end{aligned} \quad (10)$$

More generally we will write the effective action (10) in a form which will be convenient for later discussion,

$$S(U, U^\dagger) = \sum_{i=1}^p a_i \text{Tr } U^i \text{Tr } U^{\dagger i} + \sum_{\mathbf{k}, \mathbf{k}'} \alpha_{\mathbf{k}, \mathbf{k}'} \Upsilon_{\mathbf{k}}(U) \Upsilon_{\mathbf{k}'}(U^\dagger), \quad (11)$$

where \mathbf{k}, \mathbf{k}' are arbitrary vectors of non-negative entries, and

$$\Upsilon_{\mathbf{k}}(U) = \prod_j (\text{Tr } U^j)^{k_j}. \quad (12)$$

It is useful to define

$$\ell(\mathbf{k}) = \sum_j j k_j, \quad |\mathbf{k}| = \sum_j k_j. \quad (13)$$

The above parametrization of the general action is slightly redundant, since the second summand in (11) is already the most general gauge invariant action for U, U^\dagger , but writing it this way will be very useful. Reality of the action (11) requires $\alpha_{\mathbf{k}\mathbf{k}'} = \alpha_{\mathbf{k}'\mathbf{k}}^*$. In fact, using the explicit perturbative rules to compute $S(U, U^\dagger)$ in (11), one can show that the $\alpha_{\mathbf{k}\mathbf{k}'}$ are real; therefore

$$\alpha_{\mathbf{k}\mathbf{k}'} = \alpha_{\mathbf{k}'\mathbf{k}}. \quad (14)$$

On the other hand, invariance of $S(U, U^\dagger)$ under $U \rightarrow e^{i\theta}U$ requires that

$$\ell(\mathbf{k}) = \ell(\mathbf{k}'). \quad (15)$$

We now present evidence at both weak and strong λ that the above effective action is correct.

4.1 Perturbative analysis

In perturbation theory one can integrate out all fields, except the zero-mode A_0 of the time component of a gauge field, to get an effective action of U [46]. All fields other than this mode are massive in a free YM theory on S^3 . The scalar fields get their mass due to the curvature coupling. We can expand all other fields on S^3 , and due to the finite radius of S^3 all the harmonics are massive. Hence at small coupling (small λ) one may integrate out all the fields and derive an effective action for U . In [19] the perturbative (up to three loop order) effective action was calculated and it has the form (10).

4.2 Strong coupling analysis

The above discussion is perturbative, and there is no guarantee that the scalar fields remain massive in the expansion of the theory around $\lambda = \infty$. We will now show, using the AdS/CFT correspondence, that even at strong coupling (large λ), all the excitations of $N = 4$ SYM theory on S^3 are massive [29]. For illustration we consider the wave equation of a scalar field $\phi(r, t)$ in a general black hole background which is asymptotically $AdS_5 \times S^5$.

The $AdS_5 \times S^5$ metric is given by

$$ds^2 = \left(1 + \frac{r^2}{R^2}\right) d\tau^2 + \frac{dr^2}{1 + \frac{r^2}{R^2}} + r^2 d\Omega_3^2 + R^2 d\Omega_5^2. \quad (16)$$

Let us consider the situation when the asymptotic solution depends on the co-ordinates of S^5 and S^3 . Since S^5 and S^3 are compact spaces, their laplacians have a discrete spectrum. We focus on the radial part and consider a finite energy solution of energy E , $\phi(r, \theta_3, \theta_5, \tau) = f(r, \theta_3, \theta_5) \exp(E\tau)$. The wave equation in the asymptotic metric (16) is given by

$$\begin{aligned} & (3 + 5r^2)f'(r, \theta_3, \theta_5) + r(1 + r^2)f''(r, \theta_3, \theta_5) \\ & + \left(\frac{r}{1 + r^2}E^2 + \frac{1}{r}\Delta_{\Omega_3}^2 + r\Delta_{\Omega_5}^2\right) f(r, \theta_3, \theta_5) = 0, \\ & (3 + 5r^2)f'(r) + r(1 + r^2)f''(r) \\ & + \left(\frac{r}{1 + r^2}E^2 - \frac{1}{r}L_{\Omega_3}^2 - rM_{\Omega_5}^2\right) f(r) = 0, \end{aligned}$$

where $'$ is the partial derivative with respect to r and L_{Ω_3} is the contribution from S^3 harmonics, and $M_{\Omega_5}^2$ is the contribution from S^5 harmonic.

For $f(r) \sim r^\alpha$, as $r \rightarrow \infty$, (17) reduces to

$$5r^{\alpha+2} ((\alpha(\alpha - 1) + 5\alpha) - M_{\Omega_5}^2) = 0. \quad (17)$$

In the last equation we have neglected the term $E^2 r^\alpha$ and the S^3 harmonics part, as it is suppressed by a factor of order $\frac{1}{r}$. Hence $\alpha_1 = -2 + \sqrt{4 + M^2}$ or $\alpha_2 = -2 - \sqrt{4 + M^2}$ are two solutions of (17). Consequently, $f(r) \sim r^{\alpha_2}$ is the only solution which is normalizable.

Let us now analyze the situation near the black hole horizon which, in the euclidean continuation, acts like the origin of polar co-ordinates. Hence, we have the boundary condition,

$$\frac{df}{dr} = 0. \quad (18)$$

Near the origin, the scalar field laplacian in the black hole background will have two solutions for a given E . One of them diverges at the horizon and the other maintains the condition (18). For a generic E , a well-behaved solution in general approaches a non-renormalizable solution as $r \rightarrow \infty$. As in quantum mechanical problems, a normalizable solution exists only for those values of E for which the solution that behaves correctly at the lower endpoint also vanishes for $r \rightarrow \infty$. This eigenvalue condition determines a discrete value of E . Hence the mass gap in SYM theory on S^3 persists at the strong coupling. The basic physical reason for the discrete spectrum is that the asymptotic $AdS_5 \times S^5$ geometry gives rise to an infinitely rising potential for large r .

In order to make an estimate of the mass gap we note that the black hole metric depends on the combination GM , where $G \sim \frac{1}{N^2}$ is Newton's coupling and $M \sim N^2$ is the mass of the black hole. Further using standard formulas of black hole thermodynamics it is possible to express GM entirely in terms of the temperature of the black hole, which sets the scale of the mass gap.

We also expect the single negative eigenvalue in the spectrum of the euclidean Schwarzschild solution in asymptotically flat space-time to persist in the present case. Next we discuss the zero modes.

SO(6) non-invariance of the 10-dimensional black hole. As discussed in the introduction, our main interest is the study of the 10-dimensional small black hole to string transition in $AdS_5 \times S^5$. The metric of the small 10-dimensional black hole in $AdS_5 \times S^5$ is not symmetric under the $SO(6)$ transformations of S^5 . Hence the corresponding saddle point in the gauge theory would transform under the $SO(6)$ R -symmetry group and a natural question is whether the $SO(6)$ symmetry is spontaneously broken in the dual gauge theory with associated massless Nambu-Goldstone modes. If this were true, then we would have to include additional degrees of freedom in the effective action (10).

Fortunately, even though the small 10-dimensional black hole sits at a point in S^5 , the massless modes associated with motions about this point correspond to normalizable solutions of the small fluctuations equation. Let us discuss this point in more detail.

We have already discussed in the Sect. 3 that the small 10-dimensional black hole is invariant under an "unbroken" $SO(5)$ subgroup of $SO(6)$. The remaining broken generators of $SO(6)$ rotate the black hole in S^5 . The black

hole is labeled by its mass (equivalently temperature) and its position in S^5 , which we denote by the co-ordinates θ_5 . $SO(6)$ rotations can rotate the black hole to any point in S^5 . The action of the initial and final black hole is the same, because we get the final solution just by a co-ordinate rotation of the initial solution. As there is an orbit of black hole solutions with the same action, it is expected that there is a zero mode in the spectrum of the small oscillations operator around the black hole.

Let us clarify this point in more detail. Consider a black hole metric ($g_{\mu\nu}^0(\theta_5)$) as a function of θ_5 . As we mentioned before, an infinitesimal rotation in S^5 creates a new black hole solution which is given by $g_{\mu\nu}^1 = g_{\mu\nu}^0 + \delta g_{\mu\nu}$. As both the matrices $g_{\mu\nu}^0$ and $g_{\mu\nu}^1$ solve the equations of motion, their difference $\delta g_{\mu\nu}$ will be a zero mode. The existence of such a zero mode does not necessarily signal the onset of spontaneous symmetry breaking. The important point is whether the zero mode is normalizable or not. We will show that $\delta g_{\mu\nu}$ is a *normalizable* zero mode.

We make the assumption that the asymptotic geometry of an uncharged black hole solution is determined by its mass. Hence the asymptotic geometry of the black hole is given by that of a small AdS_5 black hole [23] with corrections $f_{\mu\nu}$,

$$ds^2 = \left(1 + \frac{r^2}{R^2} - \frac{m}{r^2}\right) dt^2 + \frac{dr^2}{\left(1 + \frac{r^2}{R^2} - \frac{m}{r^2}\right)} + r^2 d\theta_3^2 + R^2 d\theta_5^2 + f_{\mu\nu} dx^\mu dx^\nu, \quad (19)$$

where $f_{\mu\nu} \sim \frac{1}{r^3}$ as $r \rightarrow \infty$. Hence the difference of $g^0(\mu, \nu)$ and $g^1(\mu, \nu)$ can be written as

$$\delta g(\mu, \nu) = f_{\mu, \nu}^1 - f_{\mu, \nu}^0, \quad (20)$$

where f^0 and f^1 denotes the f 's corresponding to g^0 and g^1 . Now, $f_{\mu\nu} \sim \frac{1}{r^3}$ implies $\delta g_{\mu\nu} \sim \frac{1}{r^3}$. Hence $\delta g_{\mu\nu}$ is square integrable³,

$$\int d^4x \delta g_{\mu\nu}^2 \propto \int dr r^3 \frac{1}{r^6} \propto \int dr \frac{1}{r^3}. \quad (21)$$

Since the symmetry is not spontaneously broken, we should consider the full orbit of the classical field under $SO(6)$ (or its coset) using the method of collective coordinates [38]. Hence we have the situation in which the degrees of freedom correspond to two sets of zero modes: those corresponding to A_0 and those corresponding to $SO(6)$ symmetry. In the method of collective coordinates we make the following change of variables in the gauge theory path integral.

For simplicity of presentation we denote the fields of the gauge theory that transform under $SO(6)$ by $\phi(x)$ and consider

$$\phi(x) = \phi_0(x)^{[\Omega^5]} + \eta(x) \quad (22)$$

and the gauge condition,

$$\left(\eta, \phi_0^{[\Omega^5]}\right) = 0, \quad (23)$$

where $\phi_0(x)^{[\Omega^5]}$ is the orbit under $SO(6)$ of the classical configuration $\phi_0(x)$. The path integral measure now becomes

$$D\phi(x) = d\Omega^5 D\eta(x) \delta\left(\eta, \phi_0^{[\Omega^5]}\right) \Delta, \quad (24)$$

where Δ is the Faddeev–Popov determinant. Then by standard means we can see that the zero mode is eliminated by the delta function and the collective coordinate (compact group measure) factors out of the path integral and the remaining action is a functional of the classical field $\phi_0(x)$. Integrating out the fluctuations η , we will obtain an effective action entirely in terms of the unitary matrix U . The coefficients of the effective action will now depend on the VEVs of the scalar fields.

4.3 Comments on the effective theory

It should be mentioned that the effective action (10) is constructed only from the zero mode of A_0 on a compact manifold. Hence this effective action will not be able to describe physical situations which depend on the co-ordinates of the compact manifold S^3 . However, on the supergravity side all the zero angular momentum black hole solutions are invariant under the $SO(4)$ symmetry of S^3 . The black hole may be localized in S^5 , but it does not depend on the co-ordinates of S^3 . This fortunate circumstance enables us to use (10) as a reliable effective action to describe some aspects of the string theory in $AdS_5 \times S^5$.

The saddle points of (10) corresponding to the $N = 4$ SYM theory are in one to one correspondence with the bulk supergravity (more precisely IIB string theory) saddle points. For example, the $AdS_5 \times S^5$ geometry corresponds to a saddle point such that $\langle \text{Tr } U^n \rangle = 0 \forall n \neq 0$. Hence the eigenvalue density function is a uniform function on the circle. Now, depending on the co-efficients in (10) the saddle point $\langle \text{Tr } U^n \rangle$ can have a non-uniform gapped or ungapped eigenvalue density profile. Changing the values of the coefficients, by varying the temperature, may open or close the gap and lead to non-analytic behavior in the temperature dependence of the free energy at $N = \infty$. We will interpret this phenomenon as the string–black hole transition. As we shall see, this non-analytic behavior can be smoothed out by a double scaling technique in the vicinity of the phase transition.

5 Exact integral transform for the partition function

We start with the most general effective action given in (11). The partition function is given by

$$Z = \int [dU] e^{S(U, U^\dagger)}. \quad (25)$$

³ This argument seems to be independent of α' corrections, as the asymptotic geometry is always weakly curved for any black hole situated in a asymptotic AdS space with $l_s \ll R$.

We will assume in the following that $a_i > 0$ in (11). This amounts to the assumption that $\rho_i = \langle \frac{1}{N} \text{Tr } U^i \rangle = 0$ is always a saddle point of the effective action. It corresponds to the $AdS_5 \times S^5$ saddle point of IIB string theory. In [30] it was shown that, at sufficiently low temperatures, $a_1 > 0$.

We now use the standard gaussian trick to write

$$\begin{aligned} \exp \left\{ \sum_{i=1}^p a_i \text{Tr } U^i \text{Tr } U^{\dagger i} \right\} &= \left(\frac{N^2}{2\pi} \right)^p \int \prod_{i=1}^p \frac{dg_i d\bar{g}_i}{a_i} \\ &\times \exp \left\{ -N^2 \sum_{i=1}^p \frac{g_i \bar{g}_i}{a_i} + N \sum_{i=1}^p (g_i \text{Tr } U^i + \bar{g}_i \text{Tr } U^{\dagger i}) \right\}. \end{aligned} \quad (26)$$

Using this trick a second time, we have

$$\begin{aligned} \exp \left(-N^2 \sum_{j=1}^p \frac{g_j \bar{g}_j}{a_j} \right) &= \left(\frac{N^2}{\pi} \right)^p \int \prod_{j=1}^p a_j d\mu_j d\bar{\mu}_j \\ &\times \exp \left\{ -N^2 \sum_{j=1}^p a_j \mu_j \bar{\mu}_j + iN^2 \sum_j (\mu_j \bar{g}_j + \bar{\mu}_j g_j) \right\}. \end{aligned} \quad (27)$$

In order to deal with an arbitrary polynomial P of $\text{Tr } U^i, \text{Tr } U^{\dagger i}$, we use the following identity in (26):

$$\begin{aligned} &\exp \left\{ P(\text{Tr } U^i, \text{Tr } U^{\dagger i}) + \sum_{i=1}^p a_i \text{Tr } U^i \text{Tr } U^{\dagger i} \right\} \\ &= \left(\frac{N^2}{2\pi} \right)^p \int \prod_{i=1}^p \frac{dg_i d\bar{g}_i}{a_i} \exp \left\{ -N^2 \sum_{i=1}^p \frac{g_i \bar{g}_i}{a_i} \right\} \\ &\times \exp \left\{ P \left(\frac{\partial}{N \partial g_i}, \frac{\partial}{N \partial \bar{g}_i} \right) \right\} \\ &\times \exp \left\{ N \sum_{i=1}^p (g_i \text{Tr } U^i + \bar{g}_i \text{Tr } U^{\dagger i}) \right\} \quad (28) \\ &= \left(\frac{N^2}{2\pi} \right)^p \int \prod_{i=1}^p \frac{dg_i d\bar{g}_i}{a_i} \exp \left\{ N \sum_{i=1}^p (g_i \text{Tr } U^i + \bar{g}_i \text{Tr } U^{\dagger i}) \right\} \\ &\times \exp \left\{ P \left(-\frac{\partial}{N \partial g_i}, -\frac{\partial}{N \partial \bar{g}_i} \right) \right\} \exp \left\{ -N^2 \sum_{i=1}^p \frac{g_i \bar{g}_i}{a_i} \right\}. \end{aligned} \quad (29)$$

In the last line we have integrated by parts. Then we use (27) to write

$$\begin{aligned} &\exp \left\{ P \left(-\frac{\partial}{N \partial g_i}, -\frac{\partial}{N \partial \bar{g}_i} \right) \right\} \exp \left\{ -N^2 \sum_{i=1}^p \frac{g_i \bar{g}_i}{a_i} \right\} \\ &= \left(\frac{N^2}{\pi} \right)^p \exp \left\{ P \left(-\frac{\partial}{N \partial g_i}, -\frac{\partial}{N \partial \bar{g}_i} \right) \right\} \int \prod_{j=1}^p a_j d\mu_j d\bar{\mu}_j \\ &\times \exp \left\{ -N^2 \sum_{j=1}^p a_j \mu_j \bar{\mu}_j + iN^2 \sum_j (\mu_j \bar{g}_j + \bar{\mu}_j g_j) \right\} \end{aligned}$$

$$\begin{aligned} &= \left(\frac{N^2}{\pi} \right)^p \int \prod_{j=1}^p a_j d\mu_j d\bar{\mu}_j \exp \left\{ -N^2 \sum_{j=1}^p a_j \mu_j \bar{\mu}_j \right. \\ &\left. + iN^2 \sum_j (\mu_j \bar{g}_j + \bar{\mu}_j g_j + P(iN\mu_j, iN\bar{\mu}_j)) \right\}. \end{aligned} \quad (30)$$

Since the effective action (11) is a polynomial in $\text{Tr } U^i, \text{Tr } U^{\dagger i}$, we can use the procedure discussed above to write the partition function (25) as

$$Z = \left(\frac{N^4}{2\pi^2} \right)^p \int \prod_{i=1}^p dg_i d\bar{g}_i d\mu_i d\bar{\mu}_i \exp(N^2 S_{\text{eff}}), \quad (31)$$

where

$$\begin{aligned} S_{\text{eff}} &= - \sum_{j=1}^p a_j \mu_j \bar{\mu}_j + i \sum_j (\mu_j \bar{g}_j + \bar{\mu}_j g_j) \\ &+ \sum_{\mathbf{k}, \mathbf{k}'} \alpha_{\mathbf{k}, \mathbf{k}'} (-i)^{|\mathbf{k}| + |\mathbf{k}'|} \Upsilon_{\mathbf{k}}(\bar{\mu}) \Upsilon_{\mathbf{k}'}(\mu) + F(g_k, \bar{g}_k). \end{aligned} \quad (32)$$

In the above formula we have introduced the definition

$$\Upsilon_{\mathbf{k}}(\mu) = \prod_j \mu_j^{k_j}, \quad (33)$$

and the free energy $F(g_k, \bar{g}_k)$ is defined by

$$\begin{aligned} &\exp(N^2 F(g_k, \bar{g}_k)) \\ &= \int [dU] \exp \left\{ N \sum_{i \geq 1} (g_i \text{Tr } U^i + \bar{g}_i \text{Tr } U^{\dagger i}) \right\}. \end{aligned} \quad (34)$$

It is important to note that, given the effective action $S(U, U^\dagger)$ of the gauge theory, S_{eff} can be exactly calculated.

One notes that $F(g_i, \bar{g}_i)$ depends only on the $p-1$ phases, since one of the phases of the g_i can be absorbed by a rotation of U in the unitary integral in (34). The full integrand (31) can be shown to be independent of one phase of g_i by a redefinition of the auxiliary variables $\mu_j, \bar{\mu}_j$.

The significance of (31) is that the partition function (25) can be expressed as an exact integral transformation of the linear matrix model (34). The phase structure and the critical behavior of the linear matrix model is well understood, and hence we can study these to learn about the critical behavior and the phase structure of (25). In the next section we will discuss the phase structure of (34).

6 Critical behavior in matrix model

The eigenvalues of an unitary matrix U are the complex numbers $e^{i\theta_i}$. In the large N limit, we can consider an eigenvalue density $\rho(\theta)$, defined on the unit circle by

$$\rho(\theta) = \frac{1}{N} \sum_{i=1}^N \delta(\theta - \theta_i) = \sum_n \exp(in\theta) \frac{1}{N} \text{Tr } U^n. \quad (35)$$

⁴ The phase structure of a generic unitary matrix model has been discussed in [39].

The density function is non-negative and normalized,

$$\int \rho(\theta) d\theta = 1, \tag{36}$$

$$\rho(\theta) \geq 0. \tag{37}$$

It is well known that in the limit of $N \rightarrow \infty$, $\rho(\theta)$ can develop gaps, i.e. it can be non-zero only in bounded intervals. For example, in the case of a single gap when $\rho(\theta)$ is non-zero only in the interval $(-\frac{\theta_0}{2}, \frac{\theta_0}{2})$, it is given by the classical formula

$$\rho(\theta) = f(\theta) \sqrt{\sin^2 \frac{\theta_0}{2} - \sin^2 \frac{\theta}{2}}. \tag{38}$$

A well known example of a $\rho(\theta)$ which does not have a gap is

$$\rho(\theta) = \frac{1}{2\pi} (1 + a \cos(\theta)), \quad a < 1. \tag{39}$$

At $a = 1$, $\rho(\pi) = 0$, and a gap will begin to open. For $a > 1$ the functional form of $\rho(\theta)$ is as given by (38).

The matrix model under investigation has a complicated effective action. The saddle point distribution of the eigenvalues of the matrix U may or may not have a gap, depending on the values of parameters g_k in (34). In the large N expansion, the functional dependence of $F(g_k, \bar{g}_k)$ on g_k, \bar{g}_k depends on the phase, and we quote from the known results [31, 32, 34, 43],

$$N^2 F(g_k, \bar{g}_k) = N^2 \sum_k \frac{k g_k \bar{g}_k}{4} + e^{-2Nf(g_k, \bar{g}_k)} \sum_{n=1}^{\infty} \frac{1}{N^n} F_n^{(1)}, \quad \text{ungapped}$$

$$N^2 F(g_k, \bar{g}_k) = N^2 \sum_k \frac{k g_k \bar{g}_k}{4} + \sum_{n=0}^{\infty} N^{-\frac{2}{3}n} F_n^{(2)},$$

$$g - g_c \sim o\left(N^{-\frac{2}{3}}\right)$$

$$N^2 F(g_k, \bar{g}_k) = N^2 G(g_k, \bar{g}_k) + \sum_{n=1}^{\infty} \frac{G^{(n)}}{N^2}, \quad \text{gapped}. \tag{40}$$

In the above, we have assumed for simplicity that the eigenvalue distribution has only one gap. (In principle we cannot exclude the possibility of a multi-gap solution. But in this paper, since we are interested in the critical phenomena that results when the gap opens (or closes), we will concentrate on the single gap solution.) Near the boundary of phases, the functions $F_n(g)$ and $G_n(g)$ diverge. It is well known that in the leading order N , $F(g_k, \bar{g}_k)$ has a third order discontinuity at the phase boundary. This non-analytic behavior is responsible for the large N GWW type transition. In the $o(N^{-\frac{2}{3}})$ scaling region near the phase boundary (the middle expansion in (40)) this non-analytic behavior can be smoothed by the method of double scaling. This smoothing is important for our calculation of the double scaled partition function near the critical surface.

In (40) $f(g_k, \bar{g}_k), F_n^{(1)}, F_n^{(2)}$ and $G^n(g_k, \bar{g}_k)$ are calculable functions using standard techniques of orthogonal polynomials. As an example, $G(g_k, \bar{g}_k)$ can be expressed as

$$G(g_k, \bar{g}_k) = \frac{1}{N} \log h_0 + \int_0^1 d\xi (1 - \xi) \log f_0(\xi), \tag{41}$$

where $f_0(\xi)$ and h_0 are determined in terms of g_k, \bar{g}_k by a recursion relation of orthogonal polynomials for the unitary matrix model. It should be noted that in the ungapped phase, all perturbative ($\frac{1}{N^2}$) corrections to the leading free energy vanish. This follows from the fact that in the character expansion (strong coupling expansion) the ungapped free energy becomes an exact result. We also note that at $g_k = 0 = \bar{g}_k, f = 0$ and the non-perturbative term is absent.

6.1 Gap opening critical operator at $m = 1$ critical point

We now derive the form of the critical operator that opens the gap and corresponds to the scaling region of width $o(N^{-\frac{2}{3}})$.

From (40) we can easily find the density of eigenvalues in the ungapped phase,

$$\rho(\theta) = \frac{1}{2\pi} \left(1 + \sum_k (k g_k \exp(ik\theta) + k \bar{g}_k \exp(ik\theta)) \right) \text{ and } \rho_k = k g_k. \tag{42}$$

For a set of real g_k , the lagrangian (34) is invariant under $U \rightarrow U^\dagger$. We will assume that the gap opens at $\theta = \pi$ according to $\rho(\pi - \theta) \sim (\pi - \theta)^2$, which characterizes the first critical point⁵. At the boundary of the gapped-ungapped phase (critical surface) we have $\rho(\pi) = 0$. In terms of the critical Fourier components ρ_k^c , it is the equation of a plane with normal vector $\vec{D}_k = (-1)^k$,

$$\sum_{k=-\infty}^{\infty} (-1)^k (\rho_k^c + \bar{\rho}_k^c) = -1. \tag{43}$$

Now since $\rho_k^c = k g_k^c$ (up to non-perturbative corrections), we get the equation of a plane,

$$\sum_{k=-\infty}^{\infty} (-1)^k k (g_k^c + \bar{g}_k^c) = -1, \tag{44}$$

where g_k^c are the values of g_k at the critical plane. Since the metric induced in the space of g_k from the space of ρ_k is $G_{k,k'} = k^2 \delta_{k,k'}$, the vector that defines this plane is

$$C_k = \frac{(-1)^k}{k}. \tag{45}$$

We mention that the exact values of g_k^c where the thermal history of the small black hole intersects the critical surface

⁵ In general, the m th critical point is characterized by $\rho(\pi - \theta) \sim (\pi - \theta)^{2m}$.

are not known to us as we do not know the coefficients of the effective lagrangian. However, this information, which depends on the details of the dynamics, does not influence the critical behavior. The information where the small black hole crosses the critical surface is given by the saddle point (50), which is in turn determined by the $o(N^2)$ part of the partition function.

Below we will show that the critical behavior is determined by the departure from the critical surface and not by where the thermal history intersects it, and we conclude that the $o(1)$ part of the double scaled partition function is always determined in terms of the solution of the Painlevé II equation.

If we go slightly away from the critical surface by setting $g_k = g_k^c + \delta g_k$ and $\bar{g}_k = \bar{g}_k^c + \delta \bar{g}_k$, then the gap opens, provided $\rho(\pi) < 0^6$. This condition is easily ensured by the choice $\delta g_k + \delta \bar{g}_k = t N^{-\frac{2}{3}} C_k$, $t < 0$, which is normal to the critical plane (44).

The operator that corresponds to $\rho(\pi) = 0$ at the first critical point is

$$\hat{O} = \sum_{k=1}^{\infty} (g_k^c \text{Tr } U^k + \bar{g}_k \text{Tr } U^{\dagger k}) . \quad (46)$$

The gap at $\theta = \pi$ opens if we add a perturbation that leads to a small negative value for the ungapped solution of $\rho(\pi)$. Such a perturbation is necessarily in the direction of the vector C_k , because a perturbation that lies in the critical plane does not contribute to the opening of the gap. Hence we will set $(g_k - g_k^c) = N^{-\frac{2}{3}} \tilde{t}_k$. As we shall explain in Appendix A, $\tilde{t}_k = t C_k$, where $t = \tilde{C} \cdot \tilde{t}$ is an arbitrary parameter and \tilde{C} is the unit vector corresponding to C . Therefore the *relevant* gap opening perturbation to be added to the action is

$$\hat{O}_t = N^{-\frac{2}{3}} t \sum_{k=1}^{\infty} C_k (\text{Tr } U^k + \text{Tr } U^{\dagger k}) . \quad (47)$$

The factor $N^{-\frac{2}{3}}$ is indicative that the perturbation is relevant and has exponent $-\frac{2}{3}$. N acts like an infrared cutoff.

In the double scaling limit, near the critical surface, $F_0^{(2)}$ in (40) is a function of the parameter t (see Appendix A). It is known that $F_0^{(2)}(t)$ (from now on we will call it $F(t)$) satisfies the following differential equation:

$$\frac{\partial^2 F}{\partial t^2} = -f^2(t) , \quad (48)$$

where $f(t)$ satisfies the Painlevé II equation,

$$\frac{1}{2} \frac{\partial^2 f}{\partial t^2} = t f + f^3 . \quad (49)$$

The exact analytic form of $F(t)$ is not known, but $F(t)$ is a smooth function in the domain $(-\infty, \infty)$. Smoothness of $F(t)$ guarantees the smoothening of a large N transition in the double scaling limit.

In the gapped phase of the matrix model, $F(g_k, \bar{g}_k)$ has a standard expansion in integer powers of $\frac{1}{N^2}$, which becomes divergent as one approaches the critical surface. In the double scaling region (40) $(g - g_c) \sim \mathcal{O}(N^{-\frac{2}{3}})$, and the perturbation series (40) is organized in an expansion in powers of $N^{-\frac{2}{3}}$. The reason for the origin of such an expansion is not clear from the viewpoint of the bulk string theory. However, it is indeed possible to organize the perturbation series, in the scaling region, in terms of integral powers of a renormalized coupling constant. We will come back to this point later. In the ungapped phase the occurrence of $o(e^{-N})$ terms is also interesting. Here too we lack a clear bulk understanding of the non-perturbative terms which naturally remind us of the D-branes.

7 Saddle point equations at large N

In this section we will use the results of the previous section to write down the large N saddle point equations for the multi-trace matrix model (31). We treat μ_j and $\bar{\mu}_j$ as independent complex variables. This is natural, as the saddle point of the theory may occur at complex values of the variable, though at the end we will find that for real $\alpha_{\mathbf{k}, \mathbf{k}'}$ in (11) we have saddle points in imaginary μ_i and real g_i . From (11) we deduce the saddle point at large N by including the leading $o(N^2)$ contribution of $F(g_k, \bar{g}_k)$ to the free energy. The equations for saddle points are given by

$$\begin{aligned} \frac{\partial S_{\text{eff}}}{\partial g_j} &= i\bar{\mu}_j + \frac{1}{2j} \bar{g}_j = 0 , \\ \frac{\partial S_{\text{eff}}}{\partial \bar{g}_j} &= i\mu_j + \frac{1}{2j} g_j = 0 , \\ \frac{\partial S_{\text{eff}}}{\partial \mu_j} &= -a_j \bar{\mu}_j + i\bar{g}_j + \sum_{\mathbf{k}, \mathbf{k}'} \alpha_{\mathbf{k}, \mathbf{k}'} (-i)^{|\mathbf{k}|+|\mathbf{k}'|} \frac{k'_j}{\mu_j} \Upsilon_{\mathbf{k}}(\bar{\mu}) \Upsilon_{\mathbf{k}'}(\mu) \\ &= 0 , \\ \frac{\partial S_{\text{eff}}}{\partial \bar{\mu}_j} &= -a_j \mu_j + i g_j + \sum_{\mathbf{k}, \mathbf{k}'} \alpha_{\mathbf{k}, \mathbf{k}'} (-i)^{|\mathbf{k}|+|\mathbf{k}'|} \frac{k_j}{\bar{\mu}_j} \Upsilon_{\mathbf{k}}(\bar{\mu}) \Upsilon_{\mathbf{k}'}(\mu) \\ &= 0 . \end{aligned} \quad (50)$$

These equations correspond to the ungapped phase. Equations similar to (50) can also be written using $F(g_k, \bar{g}_k)$ in the gapped phase.

By the AdS/CFT correspondence the solutions to (50) are dual to supergravity/string theory solutions, like $AdS_5 \times S^5$ and various $AdS_5 \times S^5$ black holes. The number and types of saddle points and their thermal histories depends on the dynamics of the gauge theory (i.e. on the numerical values of the parameter a_j and $\alpha_{\mathbf{k}, \mathbf{k}'}$, which in turn are complicated functions of λ and β). These issues have been discussed in the framework of simpler models in [16], where the first order confinement/deconfinement transition and its relation with the Hawking–Page type transition in the bulk has also been discussed. Here we will not address these issues but focus on the phenomenon of an *unstable saddle point* of (50) crossing the critical surface (44).

⁶ To calculate $\rho(\theta)$ we have used the ungapped solution in (40).

By solving (50) we can write g_j in terms of μ_j and the coefficients $a_j(\beta), \alpha_{\mathbf{k},\mathbf{k}'}(\beta)$. Using the critical values of g_j (44), we get the relation between $a_j(\beta), \alpha_{\mathbf{k},\mathbf{k}'}(\beta)$ at the critical surface,

$$g_j^c(ja_j - 1) + \frac{\hat{g}_j^c}{j} + \sum_{\mathbf{k},\mathbf{k}'} 2^{2-|\mathbf{k}|-|\mathbf{k}'|} (-1)^{|\mathbf{k}+|\mathbf{k}'|} \alpha_{\mathbf{k},\mathbf{k}'} \frac{k_j}{g_j^c} \Upsilon_{\mathbf{k}+\mathbf{k}'}(g_j^c) = 0, \quad j = 1, \dots, p. \tag{51}$$

Whether the above relation is achieved for some values of the co-efficients $a_j(\beta), \alpha_{\mathbf{k},\mathbf{k}'}(\beta)$ is a difficult question which again needs a detailed understanding of the gauge theory dynamics. The coefficients $a_j(\beta), \alpha_{\mathbf{k},\mathbf{k}'}(\beta)$ have been perturbatively calculated in [19], and it can be shown that at some specific $\beta < \beta_{\text{HG}}$ ⁷ the condition (51) is satisfied.

We would like to mention that there is no fine tuning associated with the relation (44) or (51) being satisfied. This is because we have one tunable parameter, the temperature, and one relation (44) to satisfy. Hence one may hope that in the most general situation the relation (44) will be satisfied. In the next section we will discuss the double scaled partition function near the critical point.

In a later section we will use the AdS/CFT correspondence to argue that in the strongly coupled gauge theory, a 10-dimensional “small black hole” saddle point reaches the critical surface (51). The interpretation of this phenomenon in the bulk string theory as a black hole to excited string transition will also be discussed.

8 Double scaled partition function at crossover

We will assume that the matrix model (34) has a saddle point which makes a gapped to ungapped transition as we change the parameters of the theory ($\alpha_{\mathbf{k},\mathbf{k}'}^c, a_j$) by tuning the temperature β^{-1} . We will also assume that this saddle point has one unstable direction, which corresponds to opening the gap as we lower the temperature. These assumptions are motivated by the fact that the small (euclidean) Schwarzschild black hole crosses the critical surface and merges with $AdS_5 \times S^5$, and that it is an unstable saddle point of the bulk theory. To calculate the double scaled partition function near this transition point, we basically follow the method used in [16]. We expand the effective action (34) around the 1st critical point, and we simultaneously expand the original couplings a_j, g_j, \bar{g}_j and $\alpha_{\mathbf{k},\mathbf{k}'}$ around their critical values $a_j^c, \beta_j^c, g_j^c = 0$, and $\alpha_{\mathbf{k},\mathbf{k}'}^c$. For clarity we define

$$P(\mu, \bar{\mu}, \alpha) = \sum_{\mathbf{k},\mathbf{k}'} \alpha_{\mathbf{k},\mathbf{k}'} (-i)^{|\mathbf{k}+|\mathbf{k}'|} \Upsilon_{\mathbf{k}}(\bar{\mu}) \Upsilon_{\mathbf{k}'}(\mu). \tag{52}$$

We also introduce the column vectors,

$$\mu = \begin{pmatrix} \mu_j \\ \bar{\mu}_j \end{pmatrix}, \quad A = \begin{pmatrix} a_j \\ \alpha_{\mathbf{k},\mathbf{k}'} \end{pmatrix}, \quad g = \begin{pmatrix} g_j \\ \bar{g}_j \end{pmatrix} \tag{53}$$

and expand the above mentioned vector variables

$$g - g^c = N^{-\frac{2}{3}} \tilde{t}, \quad \mu - \mu^c = N^{-\frac{4}{3}} n, \quad A - A^c = \tilde{g} N^{-\frac{2}{3}} \alpha, \tag{54}$$

where $\tilde{g} = N^{\frac{2}{3}}(\beta - \beta_c)$ and $\alpha = \frac{\partial A}{\partial \beta} |_{\beta=\beta_c}$. The expansion of the co-efficients a_j and $\alpha_{\mathbf{k},\mathbf{k}'}$ are proportional to the deviation of the tuning parameter β from its critical value, i.e. $\tilde{g} = N^{\frac{2}{3}}(\beta_c - \beta)$.

The expanded action takes the following form:

$$N^2 S_{\text{eff}} = -\frac{1}{2} N^{-\frac{2}{3}} n^t \mathcal{L} n + n^t (\mathcal{J} t - \tilde{g} \mathcal{H} \alpha) + F(C \cdot \tilde{t}) + O(N^{-\frac{4}{3}}). \tag{55}$$

In the above we have, following the discussion in Appendix A, used the fact that the $o(1)$ function F depends on the scaled variable through the combination $t = C \cdot \tilde{t}$. Recall that C is the constant vector normal to the critical plane and the matrices $\mathcal{L}, \mathcal{J}, \mathcal{H}$ are given by

$$\mathcal{L} = \begin{pmatrix} -\frac{\partial^2 P}{\partial \mu_j \partial \mu_k} & a_j^{(c)} \delta_{jk} - \frac{\partial^2 P}{\partial \mu_j \partial \bar{\mu}_k} \\ a_j^{(c)} \delta_{jk} - \frac{\partial^2 P}{\partial \mu_j \partial \bar{\mu}_k} & -\frac{\partial^2 P}{\partial \bar{\mu}_j \partial \bar{\mu}_k} \end{pmatrix}, \quad \mathcal{H} = \begin{pmatrix} -\bar{\mu}_j \delta_{jk} & \frac{\partial^2 P}{\partial \mu_j \partial \alpha_{\mathbf{k},\mathbf{k}'}} \\ -\mu_j \delta_{jk} & \frac{\partial^2 P}{\partial \bar{\mu}_j \partial \alpha_{\mathbf{k},\mathbf{k}'}} \end{pmatrix}, \quad \mathcal{J} = \frac{1}{2} \begin{pmatrix} i\mathcal{F} & \mathcal{F} \\ i\mathcal{F} & -\mathcal{F} \end{pmatrix}. \tag{56}$$

In the above we have introduced the diagonal matrix

$$\mathcal{F}_{jk} = \frac{1}{j} \delta_{jk}, \quad j, k = 1, \dots, p. \tag{57}$$

All quantities appearing in the matrices are calculated at the first critical point. Here $o(N^2)$ part of the action does not depend on n, \tilde{t} and hence we do not show this part of the action explicitly.

We now do the gaussian integration over n_k in the functional integral

$$Z \sim \int d\tilde{t} \left(\det \left(N^{-\frac{2}{3}} \mathcal{L} \right) \right)^{-\frac{1}{2}} \times \exp \left\{ \frac{1}{2} N^{\frac{2}{3}} (\tilde{t} - \tilde{g} \mathcal{C} \alpha)^t \mathcal{M} (\tilde{t} - \tilde{g} \mathcal{C} \alpha) + F(C \cdot \tilde{t}) + O(N^{-\frac{2}{3}}) \right\}, \tag{58}$$

The matrices appearing here can be easily obtained,

$$\mathcal{D} = \frac{1}{2} \begin{pmatrix} \mathcal{F} & 0 \\ 0 & \mathcal{F} \end{pmatrix}, \quad \mathcal{M} = \mathcal{J}^t \mathcal{L}^{-1} \mathcal{J} + \mathcal{D}, \quad \mathcal{C} = -\mathcal{M}^{-1} \mathcal{J}^t \mathcal{L}^{-1} \mathcal{H}. \tag{59}$$

⁷ β_{HG}^{-1} is the temperature of the Hagedorn transition.

Notice that the hessian associated with S_{eff} is given by

$$H = \begin{pmatrix} -\mathcal{L} & \mathcal{J} \\ \mathcal{J} & \mathcal{D} \end{pmatrix}. \quad (60)$$

In order to discuss the further evaluation of the integral (58), we must take into account the fact that we are evaluating the integral near an unstable saddle point. That the saddle point has precisely one unstable direction is motivated by the fact that in the bulk theory the euclidean 10-dimensional black hole has one negative eigenvalue. This statement strictly speaking should apply to the saddle point in the gapped phase. However, since the GWW phase transition is third order, an unstable saddle point in the gapped phase should continue to be unstable at the crossover.

In order to render the gaussian integral (58) along the unstable direction well defined, we should make an analytic continuation. Once this is done we can easily see that as $N \rightarrow \infty$ the integral in (58) is localized at

$$\tilde{t} = \tilde{g}\mathcal{C}\alpha. \quad (61)$$

This follows from a matrix generalization of the gaussian representation of the delta function.

Putting the above expression in (58) we get the final result

$$Z \sim i(\det(H))^{-\frac{1}{2}} \exp F(\tilde{g}\mathcal{C} \cdot \mathcal{C}\alpha), \quad (62)$$

where $\mathcal{C} \cdot \mathcal{C}\alpha$ is a constant independent of \tilde{g} . We have assumed that the hessian H does not have a zero mode, but the one negative eigenvalue accounts for the i in front of (62).

The $o(1)$ part of the partition function, (62), is universal in the sense that the appearance of the function $F(\tilde{g} \times \text{constant})$ does not depend on the exact values of the parameters of the theory. In the double scaling limit the partition function becomes a function of a single scaling variable \tilde{g} . Exact values of the couplings and the $o(N^2)$ part of the partition function determine where the thermal history crosses the critical surface (44). However the form of the function F and the double scaling limit of (54) are independent of the exact values of g_k^c . They only depend on the fact that one is moving away perpendicular to the critical surface. This is the reason why in [16] we obtained exactly the same equation when $g_1^c \neq 0$, but all other $g_k^c = 0$.

8.1 Condensation of winding modes at the crossover

We will now discuss the condensation of the winding Polyakov lines in the crossover region. Specifically we will discuss the expectation value of the critical operator (46). In the leading order in large N we have already seen in (42) that $\rho_k^c = kg_k^c$. In order to calculate subleading corrections, it can be easily seen that all the ρ_k 's condense in the scaling region,

$$\left\langle N^{\frac{2}{3}} (\rho_k - \rho_k^{\text{ug}}) \right\rangle = C_k \frac{dF}{dt}, \quad (63)$$

where $\rho_k^{\text{ug}} = kg_k$. This smoothness of the expectation value of the ρ_k 's follows from the smooth nature of $F(t)$. The exact form of $F(t)$ is not known but it is known that it is a smooth function with the following asymptotic expansion:

$$\begin{aligned} F(t) &= \frac{t^3}{6} - \frac{1}{8} \log(-t) - \frac{3}{128t^3} + \frac{63}{1024t^6} + \dots, \quad -t \gg 1, \\ F(t) &= \frac{1}{2\pi} e^{-\frac{4\sqrt{2}}{3}t^{\frac{3}{2}}} \\ &\quad \times \left(-\frac{1}{8\sqrt{2}t^{\frac{3}{2}}} + \frac{35}{384t^3} - \frac{3745}{18432\sqrt{2}t^{\frac{9}{2}}} + \dots \right), \quad t \gg 1. \end{aligned} \quad (64)$$

The derivative of $F(t)$ diverges as $t \rightarrow -\infty$ and goes to zero as $t \rightarrow \infty$. This behavior tallies with the condensation of the winding mode in one phase (the gapped phase) and the non-condensation of winding modes in the ungapped phase. The condensation of the winding modes also indicates that the $U(1)$ symmetry (which is the Z_N symmetry of the $SU(N)$ gauge theory in the large N limit) is broken at the crossover but restored in the limit $t \rightarrow \infty$.

9 Higher critical points and the introduction of chemical potentials

Besides the first critical point, single trace unitary matrix models can have higher critical points. The m th critical point is characterized by

$$\rho_m(\theta) \sim (\theta - \pi)^{2m}, \quad \theta \rightarrow \pi, \quad (65)$$

and hence it is specified by the following relations:

$$\rho^{(2n)}(\pi) = 0, \quad 0 \leq n < m. \quad (66)$$

Writing the above in terms of the g_k we get

$$\sum_{k=-\infty}^{\infty} (-1)^k k^{2n-1} (g_k^c + \bar{g}_k^c) = 0, \quad 0 \leq n < m. \quad (67)$$

A particular choice for the density of eigenvalues with this behavior is

$$\rho_m(\theta) = c_m \left(2 \cos \frac{\theta}{2} \right)^{2m}, \quad (68)$$

where

$$c_m = \frac{2^{2m} (m!)^2}{2\pi (2m)!}. \quad (69)$$

By expanding in Fourier modes, one finds

$$\rho_m(\theta) = \frac{1}{2\pi} \left(1 + 2 \sum_{k=1}^m \frac{(m!)^2}{(m-k)!(m+k)!} \cos k\theta \right). \quad (70)$$

Using the relation between the density of eigenvalues in the ungapped phase and the matrix model potential one recovers the critical potential of Periwal and Shevitz.

As the plane (67) is determined by more than two equations, a generic curve in the space of the couplings g_k will not necessarily intersect the plane. Hence by tuning one parameter, the history of a saddle point may not reach the higher critical points. But one may consider a situation where, along with temperature, some additional chemical potentials are also turned on [35]. Using these chemical potentials (like, say, the R -charge) we may be able to reach higher multi-critical points.

In Appendix B, we have considered a more general effective action which includes general source terms in addition to (11),

$$\tilde{S}(U, U^\dagger) = S(U, U^\dagger) + N \sum_{k \geq 1} (b_k \text{Tr } U^k + \bar{b}_k \text{Tr } U^{\dagger k}) . \quad (71)$$

Using the above action, we have calculated the double scale partition function near higher critical points. Similar to our result in (62), the $o(1)$ part of the double scaled partition function becomes a universal function determined by the mKdV hierarchy. It should be mentioned that the calculation is performed near the m th multi-critical point characterized by

$$g_n = 0, \quad n > m . \quad (72)$$

According to the comments at the end of Sect. 8 the final form of the double scaled partition function (B.31) and the double scaling limit (B.18) is universal and independent of the particular choice of (72).

10 Applications to the small 10-dimensional black hole

We now apply what we have learned about the matrix model (gauge theory) GWW transition and its smoothening in the critical region to the black hole–string transition in the bulk theory. The first step is to identify the matrix model phase in which the black hole or, for that matter, the supergravity saddle points occur. We will argue that they belong to the gapped phase of the matrix model. This inference is related to the way perturbation theory in $\frac{1}{N}$ is organized in the gapped and ungapped phase as discussed in (40). Note that it is only in the gapped phase that the $\frac{1}{N}$ expansion is organized in powers of $\frac{1}{N^2}$, exactly in the way perturbation theory is organized around classical supergravity solutions in closed string theory. Hence at the strong gauge theory coupling ($\lambda \gg 1$), it is natural to identify the small 10-dimensional black hole with a saddle point of the equations of motion like (50) but obtained by using $F(g_k, \bar{g}_k)$ corresponding to the gapped phase⁸. One can associate a temperature with this saddle point which would satisfy $l_s^{-1} \gg T \gg R^{-1}$.

As the temperature increases towards l_s^{-1} , one traces out a curve (thermal history) in the space of the parameters $a_i, \alpha_{k,k'}$ of the effective theory. One can also say that a thermal history is traced in the space of $\rho_i = \langle \frac{1}{N} \text{Tr } U^i \rangle$, which depends on the parameters of the effective theory. We will now make the reasonable assumption that the thermal history, at a temperature $T_c \sim l_s^{-1}$, intersects the critical surface (43) (equivalently the plane (44)) and then as the temperature increases further it reaches the point $\rho_i = \langle \frac{1}{N} \text{Tr } U^i \rangle = 0$, which corresponds to $AdS_5 \times S^5$. Once the thermal history crosses the critical surface, the gauge theory saddle points are controlled by the free energy of the ungapped phase in (40). The saddle points of (50) which were obtained using this free energy do not correspond to supergravity backgrounds, because the temperature, on crossing the critical surface, is very high: $T \gtrsim l_s^{-1}$. Besides this, the free energy in the gapped phase has unconventional exponential factors (except at $g_k = 0$, which corresponds to $AdS_5 \times S^5$). It is likely that these saddle points define in the correspondence exact conformal field theories/non-critical string theories in the bulk. Neglecting the exponential corrections $\exp(-N)$, it seems reasonable, by inspecting the saddle point equations, that in this phase the spectrum would be qualitatively similar to that around $\rho_i = 0$. Since this corresponds to $AdS_5 \times S^5$, we expect the fluctuations to resemble a string spectrum.

As we saw in the previous section, our techniques are good enough only to compute the $o(1)$ part of the partition function in the vicinity of the critical surface which depends on the renormalized coupling. The exact solution of the free energy (in the single trace model) in the transition region in (40) enabled us to define a double scaling limit in which the non-analyticity of the partition function could be smoothened out, by a redefinition of the string coupling constant according to $\tilde{g} = N^{\frac{2}{3}}(\beta_c - \beta)$. This smooth crossover corresponds to the black hole crossing over to a state of strings corresponding to the ungapped phase.

We have also computed the VEV of the scaling operator, and hence at the crossover the winding modes $\rho_i = \langle \frac{1}{N} \text{Tr } U^i \rangle$ condense, (63). They also have a smooth parametric dependence across the transition. This phenomenon in the bulk theory may have the interpretation of a smooth topology change of a black hole spacetime to a spacetime without any black hole and only with a gas of excited string states. However in the crossover region a geometric spacetime interpretation is unlikely. We may be dealing with the exact description of a non-critical string in five dimensions in which only the zero mode along the S^3 directions is taken into account. This interpretation is inspired by the fact that the free energy $F(t)$ also describes the non-critical type 0B theory, as was already discussed in [16, 33].

the saddle point can transit through the critical surface. Using the results of [16], it is easy to see that this is precisely what happens for the perturbative gauge theory discussed in [19]. We note that in the corresponding bulk picture, since $l_s \gg R_{AdS}$, the supergravity approximation is not valid. It would be interesting to understand the bulk interpretation in this case.

⁸ A saddle point of the weakly coupled gauge theory may also exist in the gapped phase. With a change in the temperature

11 Implications for the lorentzian black hole, the information puzzle, and related matters

All our discussion has been in the context of the euclidean time, both in the bulk and the boundary theory. Since the boundary theory is governed by a well defined positive hamiltonian the analytic continuation from euclidean to lorentzian signature is well understood and simple. Hence the partition function gives a way of computing the density of states at a particular energy using the formula

$$Z(\beta) = \int_0^\infty dE \rho(E) e^{-\beta E}, \quad (73)$$

where $\rho(E) = \text{Tr} \delta(H - E)$ is the density of states at energy E . Since the partition function, in an appropriate scaling limit, is a smooth function of the renormalized coupling constant \tilde{g} , at the crossover between the gapped and un-gapped phase, (73) implies that $\rho(E)$ inherits the same property. Since $\rho(E)$ is as well a quantity that has meaning when the signature of time is lorentzian, it would imply that the black hole–string crossover in the lorentzian signature is also smooth. This is an interesting conclusion especially because we do not know the AdS/CFT correspondence for the small lorentzian black hole. The lorentzian section of the black hole has a horizon and singularity. Since the gauge theory should also describe this configuration, a smooth density of states in the crossover would imply that the black hole singularity was resolved in the gauge theory.

We believe in this conclusion, but an understanding of this can only be possible if we have an explicit model in the gauge theory of the small lorentzian black hole. Work in this direction is in progress, drawing lessons from [29, 36, 37, 40, 41, 47].

This program was originally motivated by an attempt to understand and resolve the information puzzle in black hole physics. In the AdS/CFT correspondence we know that the $SU(N)$ gauge theory is defined by a hermitian hamiltonian defined on $S^3 \times R$. The $N \rightarrow \infty$ limit and the $\lambda \rightarrow \infty$ limits make contact with the semi-classical gravity limit of the type IIB string theory in the bulk. In this limit, one can represent the quantum gravity theory path integral as an integral which splits into a sum over distinct topologies. In particular in the euclidean framework the path integral splits as a sum of contributions from histories with and without a black hole. However this representation arises by a naive consideration of the large N limit. We know that as long as N is finite the notion of summing over distinct topologies does not exist. A careful understanding of the double scaling limit has indeed made it possible to treat finite N effects in a saddle point expansion around large N and smoothened the GWW transition. Since we have identified this gauge theory phenomenon with a smooth black hole–string crossover, we conclude that topology change is indeed possible in the bulk string theory.

In light of our results we are not convinced about Hawking’s proposed solution to the information puzzle [42],

which uses the notion of representing the quantum gravity path integral as a sum over all topologies. At large but finite N (or equivalently at small but finite string coupling) this notion is not necessarily valid.

Appendix A: Discrete recursion relations, $m = 1$ critical point and Painlevé II

In this appendix we discuss the appearance of the $m = 1$ critical point in the discrete recursion relations in the presence of general couplings g_k , where k is a positive integer. The main point can be explicitly illustrated in the case of two couplings g_1 and g_2 , and the generalization to more general potentials is straightforward. We briefly review how we find scaling regions in matrix models and how double scaling limits are implemented. We follow closely the work of Periwal–Shevitz [32]. The action we consider is

$$g_1 \left(\text{Tr} U + \text{Tr} U^\dagger \right) + g_2 \left(\text{Tr} U^2 + \text{Tr} U^{\dagger 2} \right) = \mu_1 V_1 + \mu_2 V_2, \quad (A.1)$$

where $V_{1,2}$ are the first critical potentials found in [32]:

$$\begin{aligned} V_1 &= \frac{1}{2} \left(\text{Tr} U + \text{Tr} U^\dagger \right), \\ V_2 &= \frac{4}{3} \left(\text{Tr} U + \text{Tr} U^\dagger \right) + \frac{1}{12} \left(\text{Tr} U^2 + \text{Tr} U^{\dagger 2} \right), \end{aligned} \quad (A.2)$$

and

$$\mu_1 = 2g_1 - 16g_2, \quad \mu_2 = 12g_2. \quad (A.3)$$

For those interested in the details, we have modified the critical potentials by making the transformation $g_k \rightarrow (-1)^k g_k$, $U \rightarrow -U$. This is a symmetry of the action that guarantees that the gap opens at $\theta = \pi$. In the original paper [32] the gap opens at $\theta = 0$. Obviously the gap can open anywhere on the circle, but we simply have to be consistent once a convention is chosen. The Periwal–Shevitz [32] equation with two couplings g_1 and g_2 in our convention takes the following form:

$$\begin{aligned} -R_n \frac{n+1}{N} &= (1 - R_n^2) \left[-(R_{n+1} + R_{n-1})g_1 \right. \\ &\quad \left. - 2g_2 (R_{n-1}R_{n-1}^2 + R_{n-1}^2 R_n + 2R_{n-1}R_n R_{n+1} \right. \\ &\quad \left. + R_n R_{n+1}^2 - R_{n+2} - R_{n-2} + R_{n+1}^2 R_{n+2}) \right]. \end{aligned} \quad (A.4)$$

We will show that this equation besides the $m = 2$ fixed point also has the $m = 1$ fixed point. The latter is well known to be described by a Painlevé II equation with just one coupling. (The derivation of Painlevé II from the one coupling case has been discussed in the original paper [32]).

As usual, to find scaling regions we first solve the planar theory. However, we have to solve it for any n ; in other words, in the planar case R_n becomes a function $R(\xi)$, where $\xi = n/N$, which completely determines the planar

limit of the theory. The equation that determines $R(\xi)$ is obtained by ignoring in (A.4) the above shifts in the $R's$. This yields the planar string equation:

$$R\xi = (1 - R^2) (2(g_1 - 2g_2)R + 12g_2R^3) . \quad (A.5)$$

If we take the scaling region to be close to the endpoint of the ξ interval, i.e. 1, we introduce the scaling variable

$$\xi = 1 - a^2t , \quad (A.6)$$

as is standard in matrix models, and a is a small “lattice” parameter that is necessary to study the scaling region. Since in these theories the critical value of $R = 0$, we have to write the function R in terms of some scaling function with appropriate exponents:

$$R = a^\gamma f(t) . \quad (A.7)$$

Since we want to consider only the first critical point $m = 1$, this implies that $\gamma = 1$, and the scaling behavior of R is

$$R = af(t) . \quad (A.8)$$

Substituting in the planar string equation we obtain

$$af(1 - a^2t) = (1 - a^2f^2) (2(g_1 - 2g_2)af + 12g_2a^3f^3) . \quad (A.9)$$

The terms of order a determine the criticality condition, which as expected is the gap opening condition:

$$g_1 - 2g_2 = \frac{1}{2} . \quad (A.10)$$

The terms of order a^3 now provide the planar string equation that determines the functional form of f as a function of t to leading order in $1/N$:

$$-a^3tf(t) = -a^3f^3(2(g_1 - 2g_2) - 12g_2) ; \quad (A.11)$$

all other terms are irrelevant to this order, and what this equation does is to determine $f(t)$, and also it provides the first term in the expansion of the P-II equation in powers of fractional powers of t . The condition (A.10) determines the first critical point of the theory, $m = 1$, which implies that near $\xi = 1$ (A.5) has a second order zero in R . If we require that the zero is of order 4 (after dividing by a common R on both sides) we obtain the conditions for the $m = 2$ critical point governed by the scaling action V_2 above. Since in our problem we have a single control parameter, i.e. the temperature, we focus on the $m = 1$ condition (A.10) and study next the double scaling limit. To make contact with the arguments of Sect. 6 we will study this limit for generic coupling g_1, g_2 ; this way we include also the perturbations of a given model on the “critical surface” (A.10) by the gap opening operator (43).

So far the parameter a is just a small number, and for the time being it has no dependence on N . To get the N dependence we do the double scaling limit, by expanding the full string equation, and see what is the relation between N and a that leads to a differential equation containing the

string coupling constant, i.e. containing higher genus terms in the expansion and thus generating a string perturbation theory. Let us do it in general, but of course we have to keep track of the fact that we have already determined the scaling behavior of both ξ and $R(\xi)$, and we have to include it in (A.4):

$$\begin{aligned} af(\xi)(1 - a^2t) &= (1 - a^2f(\xi)^2) \\ &\times \left(2ag_1f(\xi) - 4ag_2f(\xi) + 12a^3g_2f(\xi)^3 \right) \\ &+ (1 - a^2f(\xi)^2) \left(20a^3g_2f(\xi)f'(\xi)^2 \right) \\ &+ a(g_1 - 8g_2)f''(\xi) + 20a^3g_2f(\xi)^2f''(\xi) \\ &\times \frac{1}{N^2} + \dots \end{aligned} \quad (A.12)$$

Now we are ready to get the relation between N and a . In going from derivatives with respect to ξ to derivatives with respect to t , we obtain, including the factor of $1/N$, a term of the form

$$\frac{1}{Na^2} \frac{d}{dt} \quad (A.13)$$

for each derivative. Since the first non-trivial terms with derivatives contains two of them, this means that

$$\frac{1}{(Na^2)^2} \frac{d^2}{dt^2} . \quad (A.14)$$

The final result up to two derivatives (it is easy to show that higher ones are irrelevant) is

$$\begin{aligned} -a^3tf(t) &= -(1 - 12g_2)a^3f(t)^3 \\ &+ (g_1 - 8g_2)a \frac{1}{(Na^2)^2} \frac{d^2f}{dt^2} \\ &+ 20g_2a^3 \frac{1}{(Na^2)^2} (ff'' + f^2\ddot{f}) , \end{aligned} \quad (A.15)$$

where the dots are derivatives with respect to t . To get the double scaling limit, notice that we want that up to a numerical constant

$$a \left(\frac{1}{Na^2} \right)^2 = g_{st}^2 a^3 . \quad (A.16)$$

Hence, up to g_{st} we obtain

$$a \sim N^{-1/3} . \quad (A.17)$$

Note that the terms in the third line of (A.15) will vanish like a^2 after we divide out by a^3 unless we force a strange scaling of g_2 , but this is something we cannot do in the above procedure. The equation that survives is of course Painlevé II after some simple numerical rescalings. The computation has been carried out only for the two coupling case, but it is easy to generalize to a more general action. We have also included the case where we have a shift of the couplings of the model with respect to the critical surface. Of course the answer is the same, and the reason is that any

of the terms $\text{Tr}(U^k + U^{-k})$ that appear in the gap opening operator have a component along the first scaling operator. For the two coupling theory this is the origin of the term $-12g_2$ in the f^3 piece and the term $-8g_2a$ in the term \dot{f} . We get Painlevé II unless we do some unnatural fine tuning in the coupling g_2 , a freedom we do not have at our disposal given that we have just one control parameter. Obviously, even if we consider a more general potential, the same will happen with the gap opening operator. The operator identified with gap opening in the text should more precisely be called the “bare” gap opening operator. After renormalization around any critical point, and in particular near $m = 1$ it will be dominated by the first scaling operator. We know also from [32] that the integrable hierarchy behind the unitary matrix model is modified mKdV (mKdV), and their flows can be identified with the expectation values of the scaling operators of the theory (including of course the irrelevant ones at the $m = 1$ critical surface).

One may wonder what happens with the expectation values of ρ_n at the crossover region. This is however no problem, since we can renormalize these operators with more freedom than we have above, in fact, the way to argue that generically, at the initial conditions of the mKdV hierarchy that starts with Painlevé II, and the continuum limit of the ρ_n gets an expectation value, is to use the renormalized Wilson loop operator of the matrix model, as it is done in [43]. The expansion of the Wilson loop $\langle w(t) \rangle$ has as coefficients for each power of t^{n+1} precisely the expectation value of the corresponding σ_n , which are the continuum limits of the ρ_n , and what follows from the double scaling limit of the loop equations is that to leading order those expectation values are not zero and are given by a power of f to leading planar order with corrections. This power of course is not zero, and hence it says that the corresponding derivative of the free energy with respect to the scaling parameter t_n that produces the expectation value of ρ_n is not zero, even when we set $t_n = 0$ after taking the derivative.

Appendix B: Partition function near multi-critical points

Here we will calculate the double scaled partition function near higher multi-critical points. We start with (71) and denote

$$Z = \int [dU] e^{\tilde{S}(U, U^\dagger)}, \quad (\text{B.1})$$

where $\tilde{S}(U, U^\dagger)$ has the form (71). We will assume in the following that $a_i > 0$. We closely follow the discussion of Sect. 8 and use the standard gaussian trick discussed in Sect. 5 to write

$$Z = \left(\frac{N^4}{2\pi^2} \right)^p \int \prod_{i=1}^p dg_i d\bar{g}_i d\mu_i d\bar{\mu}_i \exp N^2 S_{\text{eff}}, \quad (\text{B.2})$$

where

$$S_{\text{eff}} = - \sum_{j=1}^p a_j \mu_j \bar{\mu}_j + i \sum_j (\mu_j \bar{g}_j + \bar{\mu}_j g_j) + \sum_{\mathbf{k}, \mathbf{k}'} \alpha_{\mathbf{k}, \mathbf{k}'} (-i)^{|\mathbf{k}| + |\mathbf{k}'|} \Upsilon_{\mathbf{k}}(\bar{\mu}) \Upsilon_{\mathbf{k}'}(\mu) + F(g_k + b_k, \bar{g}_k + \bar{b}_k). \quad (\text{B.3})$$

We now write g_k as

$$g_l = \frac{1}{2l} (\beta_l - i\gamma_l), \quad (\text{B.4})$$

and we also write

$$b_k = \frac{1}{2k} (\tilde{g}_k - i\hat{\gamma}_k). \quad (\text{B.5})$$

Performing a change of the variables in the integral,

$$g_k \rightarrow g_k + b_k, \quad \bar{g}_k \rightarrow \bar{g}_k + \bar{b}_k, \quad (\text{B.6})$$

we get

$$S_{\text{eff}} = \sum_{j=1}^p \left(-a_j \mu_j \bar{\mu}_j + \frac{i}{2j} \left((\beta_j - \tilde{g}_j) (\mu_j + \bar{\mu}_j) + i(\gamma_j - \hat{\gamma}_j) (\mu_j - \bar{\mu}_j) \right) + \sum_{\mathbf{k}, \mathbf{k}'} \alpha_{\mathbf{k}, \mathbf{k}'} (-i)^{|\mathbf{k}| + |\mathbf{k}'|} \Upsilon_{\mathbf{k}}(\bar{\mu}) \Upsilon_{\mathbf{k}'}(\mu) + F(\beta_k, \gamma_j) \right). \quad (\text{B.7})$$

We will assume that we are analyzing the theory in the ungapped phase, in the proximity of the even multi-critical point $m = 2k$. In this case we have

$$N^2 F(\beta, \gamma) = N^2 F_{\text{ug}}(\beta, \gamma) + N^2 F_{\text{scaling}}(\beta, \gamma), \quad (\text{B.8})$$

where $F_{\text{ug}}(\beta, \gamma)$ is the planar free energy in the ungapped phase (40), and $F_{\text{scaling}}(\beta, \gamma)$ satisfies

$$\lim_{N \rightarrow \infty} N^2 F_{\text{scaling}}(\beta, \gamma) = F^{(m)}(t_l), \quad (\text{B.9})$$

where $F^{(m)}(t_l)$ is the double scaled free energy at the m th multi-critical point determined by the solution to the mKdV hierarchy [34].

To find the saddle point at large N we only have to consider the contribution of the free energy $F(\beta, \gamma)$ in the ungapped phase. The equations for the saddle point are

given by

$$\begin{aligned}
\frac{\partial S_{\text{eff}}}{\partial \beta_j} &= \frac{i}{2j} (\mu_j + \bar{\mu}_j) + \frac{1}{2j} \beta_j = 0, \\
\frac{\partial S_{\text{eff}}}{\partial \gamma_j} &= -\frac{1}{2j} (\mu_j - \bar{\mu}_j) + \frac{1}{2j} \gamma_j = 0, \\
\frac{\partial S_{\text{eff}}}{\partial \mu_j} &= -a_j \bar{\mu}_j + \frac{i}{2j} (\beta_j - \tilde{g}_j + i\gamma_j - i\hat{\gamma}_j) \\
&\quad + \sum_{\mathbf{k}, \mathbf{k}'} \alpha_{\mathbf{k}, \mathbf{k}'} (-i)^{|\mathbf{k}|+|\mathbf{k}'|} \frac{k'_j}{\mu_j} \Upsilon_{\mathbf{k}}(\bar{\mu}) \Upsilon_{\mathbf{k}'}(\mu) = 0, \\
\frac{\partial S_{\text{eff}}}{\partial \bar{\mu}_j} &= -a_j \mu_j + \frac{i}{2j} (\beta_j - \tilde{g}_j - i\gamma_j + i\hat{\gamma}_j) \\
&\quad + \sum_{\mathbf{k}, \mathbf{k}'} \alpha_{\mathbf{k}, \mathbf{k}'} (-i)^{|\mathbf{k}|+|\mathbf{k}'|} \frac{k_j}{\bar{\mu}_j} \Upsilon_{\mathbf{k}}(\bar{\mu}) \Upsilon_{\mathbf{k}'}(\mu) = 0.
\end{aligned} \tag{B.10}$$

In the first two equations we have used that, in the un-gapped phase,

$$\frac{\partial F_{\text{ug}}}{\partial \beta_j} = \frac{1}{2j} \beta_j, \quad \frac{\partial F_{\text{ug}}}{\partial \gamma_j} = \frac{1}{2j} \gamma_j. \tag{B.11}$$

We will assume that there is a solution to these equations corresponding to the m th multi-critical even point of the model (B.3), which is characterized by

$$\gamma_j = 0, \quad \beta_j = \beta_j^{(m)}, \tag{B.12}$$

where the critical values of the couplings $\beta_j^{(m)}$ can be read from the particular solution (70). We find that this solution leads to the conditions

$$\mu_j^{(m)} = \bar{\mu}_j^{(m)} = \frac{i}{2} \beta_j^{(m)}. \tag{B.13}$$

One finds the equations for the critical submanifolds in the original couplings, a_j , \tilde{g}_k , and $\alpha_{\mathbf{k}, \mathbf{k}'}$,

$$\begin{aligned}
&\beta_j^{(m)} (j a_j - 1) + \frac{\tilde{g}_j^c}{j} \\
&\quad + \sum_{\mathbf{k}, \mathbf{k}'} 2^{2-|\mathbf{k}|-|\mathbf{k}'|} (-1)^{|\mathbf{k}|+|\mathbf{k}'|} \alpha_{\mathbf{k}, \mathbf{k}'} \frac{k_j}{\beta_j^{(m)}} \Upsilon_{\mathbf{k}+\mathbf{k}'}(\beta_j^{(m)}) = 0, \\
&\quad j = 1, \dots, p,
\end{aligned} \tag{B.14}$$

where \tilde{g}_j^c is the critical value of \tilde{g}_j , and we have set $\hat{\gamma}_j^c = 0$ for simplicity.

We now expand the effective action around the critical point, and we expand simultaneously the original couplings a_j , \tilde{g}_j , $\hat{\gamma}_j$ and $\alpha_{\mathbf{k}, \mathbf{k}'}$ around a point a_j^c , \tilde{g}_j^c , $\hat{\gamma}_j^c = 0$, and $\alpha_{\mathbf{k}, \mathbf{k}'}^c$ on the critical submanifold determined by (B.14). We denote

$$P(\mu, \bar{\mu}, \alpha) = \sum_{\mathbf{k}, \mathbf{k}'} \alpha_{\mathbf{k}, \mathbf{k}'} (-i)^{|\mathbf{k}|+|\mathbf{k}'|} \Upsilon_{\mathbf{k}}(\bar{\mu}) \Upsilon_{\mathbf{k}'}(\mu). \tag{B.15}$$

We introduce the column vectors of variables,

$$\begin{aligned}
\xi(N)n &= \begin{pmatrix} \mu_j - \mu_j^{(m)} \\ \bar{\mu}_j - \bar{\mu}_j^{(m)} \end{pmatrix}, \quad \alpha = \begin{pmatrix} a_j - a_j^c \\ \alpha_{\mathbf{k}, \mathbf{k}'} - \alpha_{\mathbf{k}, \mathbf{k}'}^c \end{pmatrix}, \\
g &= \begin{pmatrix} \beta_j - \beta_j^{(m)} \\ \gamma_j \end{pmatrix}, \quad b = \begin{pmatrix} \tilde{g}_j - \tilde{g}_j^c \\ \hat{\gamma}_j \end{pmatrix},
\end{aligned} \tag{B.16}$$

where $\xi(N)$ is an appropriate scaling factor. When we expand the action in (B.3) around the m th multi-critical point, we obtain

$$\sum_l (g_l \text{Tr } U^l + \bar{g}_l \text{Tr } U^{+l}) = V^{(m)} + \sum_n N^{\frac{n-2m}{2m+1}} t_n \tilde{V}_n, \tag{B.17}$$

where $V^{(m)}$ is the critical potential associated to the m th multi-critical point, and the \tilde{V}_n are scaling operators which can be explicitly written by using the results of [44]. In this way we find the relation between the variables g introduced in (B.16) and the scaling operators of the multi-critical model,

$$g_a = \sum_{n \geq 0} \mathcal{G}_{an} N^{\frac{n-2m}{2m+1}} t_n, \tag{B.18}$$

where \mathcal{G} is a matrix that can be explicitly determined from the expressions for the perturbations of the density of eigenvalues. Equation (B.18) determines the scaling properties of the g_a . Notice that we can use the freedom to rotate U to get rid of one of the $2p$ parameters g_i, \bar{g}_i , so we will only have $2p-1$ times.

We now do a gaussian integration over n . The relevant part of the action reads

$$\begin{aligned}
N^2 S_{\text{eff}} &= -\frac{1}{2} N^2 \xi(N)^2 n^t \mathcal{L} n \\
&\quad + N^2 \xi(N) n^t (\mathcal{J} g - \mathcal{J} b + \mathcal{H} \alpha) + \dots,
\end{aligned} \tag{B.19}$$

where the matrices $\mathcal{L}, \mathcal{J}, \mathcal{H}$ are given by

$$\begin{aligned}
\mathcal{L} &= \begin{pmatrix} -\frac{\partial^2 P}{\partial \mu_j \partial \bar{\mu}_k} & a_j^{(c)} \delta_{jk} - \frac{\partial^2 P}{\partial \mu_j \partial \bar{\mu}_k} \\ a_j^{(c)} \delta_{jk} - \frac{\partial^2 P}{\partial \mu_j \partial \bar{\mu}_k} & -\frac{\partial^2 P}{\partial \bar{\mu}_j \partial \mu_k} \end{pmatrix}, \\
\mathcal{H} &= \begin{pmatrix} -\bar{\mu}_j \delta_{jk} & \frac{\partial^2 P}{\partial \mu_j \partial \alpha_{\mathbf{k}, \mathbf{k}'}} \\ -\mu_j \delta_{jk} & \frac{\partial^2 P}{\partial \bar{\mu}_j \partial \alpha_{\mathbf{k}, \mathbf{k}'}} \end{pmatrix}, \\
\mathcal{J} &= \frac{1}{2} \begin{pmatrix} i\mathcal{F} & \mathcal{F} \\ i\mathcal{F} & -\mathcal{F} \end{pmatrix},
\end{aligned} \tag{B.20}$$

and we have introduced the diagonal matrix

$$\mathcal{F}_{jk} = \frac{1}{j} \delta_{jk}, \quad j, k = 1, \dots, p. \tag{B.21}$$

All quantities involved in these matrices are evaluated at the critical point. The gaussian integration leads to

$$\begin{aligned}
&N^{2p} (\det(\mathcal{L}))^{-\frac{1}{2}} \\
&\quad \times \exp \left\{ \frac{1}{2} N^2 (g - \mathcal{E} b - \mathcal{C} \alpha)^t \mathcal{M} (g - \mathcal{E} b - \mathcal{C} \alpha) \right. \\
&\quad \left. + F^{(m)}(t_\ell) + \dots \right\},
\end{aligned} \tag{B.22}$$

where we have assumed that \mathcal{L} does not have zero modes, and we have used and the fact that the gaussian integration gives an overall factor N^{-2p} which combines with the overall N^{4p} in (B.2). Notice that the scaling $\xi(N)$ does not appear in this equation. The choice of $\xi(N)$ must be done in such a way that the rest of the terms involving n in the expansion of $N^2 S_{\text{eff}}$ vanish in the limit $N \rightarrow \infty$. The matrices appearing here can be easily obtained from the above data. Then we have

$$\begin{aligned} \mathcal{D} &= \frac{1}{2} \begin{pmatrix} \mathcal{F} & 0 \\ 0 & \mathcal{F} \end{pmatrix}, \\ \mathcal{M} &= \mathcal{J}^t \mathcal{L}^{-1} \mathcal{J} + \mathcal{D}, \\ \mathcal{C} &= -\mathcal{M}^{-1} \mathcal{J}^t \mathcal{L}^{-1} \mathcal{H}, \\ \mathcal{E} &= \mathcal{M}^{-1} \mathcal{J}^t \mathcal{L}^{-1} \mathcal{J}. \end{aligned} \quad (\text{B.23})$$

Notice that the hessian associated to S_{eff} is given by

$$H = \begin{pmatrix} -\mathcal{L} & \mathcal{J} \\ \mathcal{J} & \mathcal{D} \end{pmatrix}. \quad (\text{B.24})$$

We now introduce scaling variables for the couplings g and α . The scaling of g is determined. In this way we obtain for (B.22)

$$\begin{aligned} &\exp \left\{ \frac{1}{2} \sum_{n,p} N^{\frac{2+n+p}{2m+1}} (t_n - t_n^0) \mathcal{A}_{np} (t_p - t_p^0) + F^{(m)}(t_\ell) \right. \\ &\quad \left. + \dots \right\}, \end{aligned} \quad (\text{B.25})$$

where

$$\begin{aligned} \mathcal{A} &= \mathcal{G}^t \mathcal{M} \mathcal{G}, \\ t_n^0 &= N^{\frac{2m-n}{2m+1}} \sum_{\ell} \left((\mathcal{G}^{-1} \mathcal{C})_{n\ell} \alpha_\ell + (\mathcal{G}^{-1} \mathcal{E})_{nj} b_j \right). \end{aligned} \quad (\text{B.26})$$

As we see, the scaling of the original coupling constants packaged in α and b is determined by the scaling of the couplings in the m th critical point.

In the limit $N \rightarrow \infty$, the integral localizes in

$$t_n = t_n^0. \quad (\text{B.27})$$

To see this in detail, we use the following fact. Let B_ϵ be an $n \times n$ matrix whose entries go to $+\infty$ as $\epsilon \rightarrow 0$. Then, one has the following:

$$\lim_{\epsilon \rightarrow 0} (\det(B_\epsilon))^{\frac{1}{2}} e^{-\frac{1}{2} x^t B_\epsilon x} = \pi^{\frac{n}{2}} \delta(x). \quad (\text{B.28})$$

In our case we find that

$$\begin{aligned} &\exp \left\{ \frac{1}{2} \sum_{n,p} N^{\frac{2+n+p}{2m+1}} (t_n - t_n^0) \mathcal{A}_{np} (t_p - t_p^0) \right\} \rightarrow \\ &N^{-\frac{\sum_{n>0} (n+1)}{2m+1}} \frac{\pi^{p-\frac{1}{2}}}{\det(\mathcal{G})(\det(-\mathcal{M}))^{\frac{1}{2}}} \delta(t - t_0), \end{aligned} \quad (\text{B.29})$$

as $N \rightarrow \infty$. Remember that there are only $2p-1$ times involved. After changing variables in the integral from g, \bar{g} to t , we inherit a jacobian

$$N^{\frac{\sum_{n>0} (n-2m)}{2m+1}} \det(\mathcal{G}). \quad (\text{B.30})$$

Putting all these ingredients together, we finally obtain

$$Z \sim N (\det(H))^{-\frac{1}{2}} \exp F^{(m)}(t_n^0), \quad (\text{B.31})$$

up to factors of π . We have assumed here that H has no zero modes. The factor of N comes from the fact that the quotient between the factors of N in (B.29) and (B.30) gives a power of N given simply by minus the number of times involved, which is $-2p+1$. This combines with the factor N^{2p} in (B.22) to give an overall factor of N . In the above derivation we have assumed that \mathcal{M} (and therefore H has no zero eigenvalues).

We can also analyze the more general case in which \mathcal{M} (which is a $p \times p$ matrix) has ℓ nonzero eigenvalues d_n , $n = 1, \dots, \ell$, and $2p-\ell$ zero eigenvalues. Let R^{-1} be the orthogonal $2p \times 2p$ matrix that diagonalizes \mathcal{M} , i.e. $R^{-1t} \mathcal{M} R^{-1} = \text{diag}(d_n, 0)$. Define now the following eigenvectors of \mathcal{M} :

$$r = N^{\frac{2m}{2m+1}} R g, \quad (\text{B.32})$$

which in terms of the scaling operators means

$$r_n = \sum_q \mathcal{R}_{nq} t_q N^{\frac{q}{2m+1}}, \quad (\text{B.33})$$

where $\mathcal{R} = R \mathcal{G}$. Then, the exponent in the gaussian (B.22) becomes

$$\frac{1}{2} N^{\frac{2}{2m+1}} \sum_{n=1}^{\ell} d_n \left(r_n - N^{\frac{2m}{2m+1}} c_n \right)^2 + N^{\frac{2+2m}{2m+1}} \sum_{n=\ell+1}^{2p} r_n \zeta_n, \quad (\text{B.34})$$

where

$$\begin{aligned} \zeta_n &= \sum_q R_{nq}^{-1t} (\mathcal{J}^t \mathcal{L}^{-1} \mathcal{H} \alpha - \mathcal{J}^t \mathcal{L}^{-1} \mathcal{J} b)_q, \\ n &= \ell + 1, \dots, 2p, \\ c_n &= -d_n^{-1} \sum_q R_{nq}^{-1t} (\mathcal{J}^t \mathcal{L}^{-1} \mathcal{H} \alpha - \mathcal{J}^t \mathcal{L}^{-1} \mathcal{J} b)_q, \\ n &= 1, \dots, \ell. \end{aligned} \quad (\text{B.35})$$

As $N \rightarrow \infty$, the first term in (B.34) gives a delta function constraint of the form

$$\sum_{q \geq 0} \mathcal{R}_{nq} t_q N^{\frac{q}{2m+1}} = c_n, \quad n = 1, \dots, \ell; \quad (\text{B.36})$$

therefore, there are only $2p-1-\ell$ independent times involved. From the behavior of the above equation as $N \rightarrow \infty$ it follows that we have to solve for the times with the higher

scaling dimension in terms of the constants c_n . This in turn determines the scaling properties of c_n :

$$t_q = t_q^0 \equiv N^{\frac{2m-q}{2m+1}} \sum_{n=1}^{\ell} \mathcal{R}_{qn}^{-1} c_n, \quad q = 2p-1-\ell, \dots, 2p-2, \quad (\text{B.37})$$

where we have inverted the $\ell \times \ell$ submatrix \mathcal{R}_{qn} , $q, n = 2p-1-\ell, p-2$. This fixes the values of ℓ times in the free energy as functions of the scaled parameters c_n , $n = 1, \dots, \ell$. The other times lead to an integral transform. To see this, let us define

$$\bar{t}_q = N^{\frac{2m+2+q}{2m+1}} \sum_{n=\ell+1}^{2p} \mathcal{R}_{nq} \zeta_n. \quad (\text{B.38})$$

This equation determines the scaling of ζ_n . Notice that the scaling properties induced on c_n and ζ_n are very different. Up to overall factors, we end up with the integral

$$\begin{aligned} & \int \prod_{n=0}^{2p-2} dt_n \prod_{q=2p-1-\ell}^{2p-2} \delta(t_q - t_q^0) \exp \left\{ \sum_{q=0}^{2p-2} t_q \bar{t}_q + F^{(m)}(t_q) \right\} \\ &= e^{\sum_{q=2p-1-\ell}^{2p-2} t_q^0 \bar{t}_q} \int \prod_{n=0}^{2p-2-\ell} dt_n \exp \left\{ \sum_{q=0}^{2p-2-\ell} t_q \bar{t}_q \right. \\ & \quad \left. + F^{(m)}(t_0, \dots, t_{2p-2-\ell}, t_{2p-1-\ell}^0, \dots, t_{2p-2}^0) \right\}. \quad (\text{B.39}) \end{aligned}$$

For hermitian matrix models, a similar result was obtained in [45]. Notice that the integral transform will change the critical exponents of the model, as noted in [45].

To illustrate our formalism we can look to the example of free YM theories at finite temperature [17, 46]:

$$S(U, U^\dagger) = \sum_{j=1}^{\infty} a_j \text{Tr} U^j \text{Tr} U^{\dagger j}, \quad (\text{B.40})$$

where

$$a_j = \frac{1}{j} (z_B(x^j) + (-1)^{j+1} z_F(x^j)). \quad (\text{B.41})$$

The equation for the critical surface reduces to

$$\beta_j^{(m)} (j a_j - 1) + \frac{\tilde{g}_j^c}{j} = 0, \quad (\text{B.42})$$

and by tuning the value of \tilde{g}_j^c we can reach any critical point. Notice that, if we do not include the b_k terms in the original action, only the first critical point $m = 1$ can be realized in the model. In that case, one has

$$a_1(T) = 1, \quad (\text{B.43})$$

which defines the Hagedorn temperature $T = T_H$. Also, if we do not include the source terms involving b_k , we can turn on only a single scaling operator in the theory and we recover the $m = 1$ model. When one includes the b_k, \bar{b}_k couplings one can also recover all the evolution times of the double scaled matrix model.

Acknowledgements. We would like to thank Ofer Aharony, Jose Barbon, Avinash Dhar, Rajesh Gopakumar, David Gross, Veronica Hubeny, Elias Kiritsis, Gautam Mandal, Hong Liu, Mukund Rangamani and Ashoke Sen for useful discussions during the course of this work. We are especially thankful to Shiraz Minwalla for many valuable and critical discussions during the course of this work. PB and SRW would like to thank the CERN theory division for a visit where part of this work was done. Various versions of this work were presented at various string meetings: Third Regional Crete Meeting in String theory (Crete, June 2005), IIT Kanpur String theory workshop (IIT Kanpur, October 2005), Inauguration Meeting of the Institute for Quantum Spacetime (Seoul, October 2005), Batsheva de Rothschild Seminar on Innovative Aspects of String Theory, India-Israel string theory meeting (Ein-Boqeq, March 2006), 12th Regional Conference on Mathematical Physics (Islamabad, March 2006), Workshop on Black holes, Blackrings and Topological String Theory (Munich, April 2006), IPM String School and Workshop (Tehran, April 2006). We thank the organizers of these conferences for the invitation to present this work.

References

1. L. Susskind, Some speculations about black hole entropy in string theory, arXiv:hep-th/9309145
2. G.T. Horowitz, J. Polchinski, Phys. Rev. D **55**, 6189 (1997) [arXiv:hep-th/9612146]
3. A. Sen, Mod. Phys. Lett. A **10**, 2081 (1995) [arXiv:hep-th/9504147]
4. M.J. Bowick, L. Smolin, L.C.R. Wijewardhana, Gen. Relat. Grav. **19**, 113 (1987)
5. M.J. Bowick, L. Smolin, L.C.R. Wijewardhana, Phys. Rev. Lett. **56**, 424 (1986)
6. A. Dabholkar, Phys. Rev. Lett. **94**, 241301 (2005) [arXiv:hep-th/0409148]
7. A. Dabholkar, R. Kallosh, A. Maloney, JHEP **0412**, 059 (2004) [arXiv:hep-th/0410076]
8. A. Dabholkar, F. Deneff, G.W. Moore, B. Pioline, JHEP **0510**, 096 (2005) [arXiv:hep-th/0507014]
9. T. Mohaupt, arXiv:hep-th/0512048
10. A. Giveon, D. Kutasov, E. Rabinovici, A. Sever, Nucl. Phys. B **719**, 3 (2005) [arXiv:hep-th/0503121]
11. A. Giveon, D. Kutasov, JHEP **0601**, 120 (2006) [arXiv:hep-th/0510211]
12. A. Giveon, E. Rabinovici, A. Sever, JHEP **0307**, 055 (2003) [arXiv:hep-th/0305140]
13. D. Kutasov, Accelerating branes and the string/black hole transition, arXiv:hep-th/0509170
14. Y. Nakayama, K.L. Panigrahi, S.J. Rey, H. Takayanagi, JHEP **0501**, 052 (2005) [arXiv:hep-th/0412038]
15. Y. Nakayama, S.J. Rey, Y. Sugawara, JHEP **0509**, 020 (2005) [arXiv:hep-th/0507040]
16. L. Álvarez-Gaumé, C. Gomez, H. Liu, S. Wadia, Phys. Rev. D **71**, 124023 (2005) [arXiv:hep-th/0502227]
17. B. Sundborg, Nucl. Phys. B **573**, 349 (2000) [arXiv:hep-th/9908001]
18. A.M. Polyakov, Int. J. Mod. Phys. A **17S1**, 119 (2002) [arXiv:hep-th/0110196]
19. O. Aharony, J. Marsano, S. Minwalla, K. Papadodimas, M. Van Raamsdonk, Phys. Rev. D **71**, 125018 (2005) [arXiv:hep-th/0502149]

20. H. Liu, Fine structure of Hagedorn transitions, arXiv:hep-th/0408001
21. M. Spradlin, A. Volovich, Nucl. Phys. B **711**, 199 (2005) [arXiv:hep-th/0408178]
22. J. Hallin, D. Persson, Phys. Lett. B **429**, 232 (1998) [arXiv:hep-ph/9803234]
23. S.W. Hawking, D.N. Page, Commun. Math. Phys. **87**, 577 (1983)
24. D.J. Gross, E. Witten, Phys. Rev. D **21**, 446 (1980)
25. S. Wadia, A Study of $U(N)$ Lattice Gauge Theory In Two-Dimensions, EFI-79/44-CHICAGO
26. S.R. Wadia, Phys. Lett. B **93**, 403 (1980)
27. V.E. Hubeny, M. Rangamani, JHEP **0205**, 027 (2002) [arXiv:hep-th/0202189]
28. G.T. Horowitz, V.E. Hubeny, JHEP **0006**, 031 (2000) [arXiv:hep-th/0005288]
29. E. Witten, Adv. Theor. Math. Phys. **2**, 505 (1998) [arXiv:hep-th/9803131]
30. P. Basu, S.R. Wadia, Phys. Rev. D **73**, 045022 (2006) [arXiv:hep-th/0506203]
31. Y.Y. Goldschmidt, J. Math. Phys. **21**, 1842 (1980)
32. V. Periwal, D. Shevitz, Phys. Rev. Lett. **64**, 1326 (1990)
33. I.R. Klebanov, J. Maldacena, N. Seiberg, arXiv:hep-th/0309168
34. V. Periwal, D. Shevitz, Nucl. Phys. B **344**, 731 (1990)
35. D. Yamada, L.G. Yaffe, arXiv:hep-th/0602074
36. J.M. Maldacena, JHEP **0304**, 021 (2003) [arXiv:hep-th/0106112]
37. P. Kraus, H. Ooguri, S. Shenker, Phys. Rev. D **67**, 124022 (2003) [arXiv:hep-th/0212277]
38. L. Gervais, B. Sakita, Phys. Rev. D **11**, 2943 (1975)
39. G. Mandal, Mod. Phys. Lett. A **5**, 1147 (1990)
40. L. Fidkowski, V. Hubeny, M. Kleban, S. Shenker, JHEP **0402**, 014 (2004) [arXiv:hep-th/0306170]
41. G. Festuccia, H. Liu, Excursions beyond the horizon: Black hole singularities in Yang–Mills theories. I, arXiv:hep-th/0506202
42. S. Hawking, Black holes and the information paradox, Prepared for GR17: 17th International Conference on General Relativity and Gravitation, Dublin, Ireland, 18–24 July 2004
43. C. Crnkovic, M.R. Douglas, G.W. Moore, Nucl. Phys. B **360**, 507 (1991)
44. C. Crnkovic, M.R. Douglas, G.W. Moore, Int. J. Mod. Phys. A **7**, 7693 (1992) [arXiv:hep-th/9108014]
45. I.R. Klebanov, A. Hashimoto, Nucl. Phys. B **434**, 264 (1995) [arXiv:hep-th/9409064]
46. O. Aharony, J. Marsano, S. Minwalla, K. Papadodimas, M. Van Raamsdonk, Adv. Theor. Math. Phys. **8**, 603 (2004) [arXiv:hep-th/0310285]
47. O. Aharony, S. Minwalla, T. Wiseman, arXiv:hep-th/0507219