Massive $\mathcal{N} = 2$ Gauge Theories at Large $N$

J.G. Russo$^{1,2}$ and K. Zarembo$^{3,4}$

$^1$ Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluís Companys, 23, 08010 Barcelona, Spain
$^2$ Department ECM, Institut de Ciències del Cosmos, Universitat de Barcelona, Martí Franquès, 1, 08028 Barcelona, Spain
$^3$ Nordita, KTH Royal Institute of Technology and Stockholm University, Roslagstullsbacken 23, SE-106 91 Stockholm, Sweden
$^4$ Department of Physics and Astronomy, Uppsala University SE-751 08 Uppsala, Sweden
jorge.russo@icrea.cat, zarembo@nordita.org

Abstract

Using exact results obtained from localization on $S^4$, we explore the large $N$ limit of $\mathcal{N} = 2$ super Yang-Mills theories with massive matter multiplets. We focus on three cases: $\mathcal{N} = 2^*$ theory, describing a massive hypermultiplet in the adjoint representation, $SU(N)$ super-Yang-Mills with $2N$ massive hypermultiplets in the fundamental, and super QCD with massive quarks. When the radius of the four-sphere is sent to infinity the theories at hand are described by solvable matrix models, which exhibit a number of interesting phenomena including quantum phase transitions at finite ’t Hooft coupling.

1 Introduction

In this paper, we investigate massive $\mathcal{N} = 2$ supersymmetric gauge theories in the multicolor limit by exploiting the results of supersymmetric localization. The path integral of any $\mathcal{N} = 2$ theory on $S^4$ can be localized to a finite-dimensional matrix integral [1], and our goal will be to study the resulting matrix models in the large-$N$ limit.

The multicolor limit of non-Abelian $SU(N)$ gauge theories is known to simplify their dynamics without distorting essential features of the non-perturbative behavior. Moreover, the

*Also at ITEP, Moscow, Russia.
relationship to string theory, which is now believed to be inherent to any non-Abelian theory, becomes most transparent within the large-$N$ expansion [2]. The cleanest manifestation of the large-$N$ gauge/string duality is the AdS/CFT correspondence [3] where the large-$N$ limit corresponds to free, non-interacting strings in a curved space. The dual field theory, $\mathcal{N} = 4$ superconformal Yang-Mills, has been studied quite thoroughly in the planar approximation. Much less is known about less supersymmetric and non-conformal theories.

As far as $\mathcal{N} = 2$ theories are concerned, the standard approach is based on the Seiberg-Witten theory [4]. The large-$N$ limit of the latter was investigated in [5, 6] for pure gauge $\mathcal{N} = 2$ super-Yang-Mills (SYM). The localization matrix models for various $\mathcal{N} = 2$ theories were analyzed in the large-$N$ limit in [7, 8, 9, 10, 11, 12, 13]. One of the outcomes of this analysis is a direct verification of the gauge/string duality in the non-conformal setting of the mass-deformed $\mathcal{N} = 2^*$ super-Yang-Mills (SYM) theory [12]. The large-$N$ vacuum structure of pure gauge $\mathcal{N} = 2$ SYM [5, 6] can be reproduced from localization as well [11]. Here we concentrate on various other massive theories, starting with $\mathcal{N} = 2^*$ SYM.

An important simplification of the multicolor limit is expected to arise in the non-perturbative sector: instanton contributions should become negligible at large-$N$ due to exponential suppression of the instanton weight. The instanton moduli integration may, in principle, overcome the exponential suppression thus leading to a large-$N$ phase transition [14]. We have searched for the instanton-induced phase transitions in a number of $\mathcal{N} = 2$ theories (secs. 3.7, 4.3), so far with negative results.

Instead, we have found a novel type of phase transitions, which take place in the infinite volume limit and are associated with appearance of new nearly massless particles in the spectrum. In the $\mathcal{N} = 2^*$ theory, the phase transition of this kind separates the weak-coupling phase from the strong-coupling phases [13]. As the ’t Hooft coupling is increased, the theory undergoes an infinite sequence of phase transitions with critical points accumulating at infinite coupling. The strongly coupled vacuum acquires a rather irregular, fractal structure at small scales in the field space. The agreement with the holographic description [12] is obtained after coarse-grained average over this small-scale structure. What implications such a fractal structure can have for the holographic duality is unclear to us. It is conceivable that the strong-coupling limit (beyond the leading order, discussed in [12]) is not unique and depends on how the infinite-volume limit is taken. As a first step towards a deeper understanding of these issues, we will perform a detailed study of the critical behavior near the first of these phase transitions.

It turns out that quantum, weak/strong coupling phase transitions are generic features of $\mathcal{N} = 2$ theories with two mass scales or with a dimensionless coupling. For instance, we will find that super-QCD (SQCD) in the Veneziano limit [15] undergoes a third-order phase transition as the quark mass varies. In fact, the localization matrix model of $\mathcal{N} = 2$ SQCD is much simpler, making a detailed analysis of the phase transition possible.

The weak-coupling expansion of massive $\mathcal{N} = 2$ theories that we will study is also of some interest, as it illustrates some generic features of asymptotically free QFTs with two well separated scales: the dynamically generated scale $\Lambda$ and the “kinematic” scale $M$. If $M \gg \Lambda$ perturbation theory should be a reasonable approximation, but only up to power-like correc-
tions:

\[ A = \text{perturbative} + \sum_{n=1}^{\infty} C_n \left( \frac{\Lambda}{M} \right)^{2n}. \] (1.1)

The coefficients \( C_n \) are in general not calculable, unless the theory can be solved exactly, and, at best, can be parameterized by vacuum condensates, like in the ITEP sum rules. Using localization techniques it will be possible to compute the coefficients of the OPE expansion exactly in certain multicolor \( \mathcal{N} = 2 \) theories.

## 2 Generalities

The theories we are going to study will contain a single vector multiplet \((A_\mu, \psi_\alpha^1, \psi_\alpha^2, \Phi + i\Phi')\) of the \( SU(N) \) gauge symmetry, and a number of matter hypermultiplets \((\phi, \chi_\alpha, \tilde{\chi}_\alpha, \tilde{\phi})\) either in the adjoint or in the fundamental representation of \( SU(N) \). Each hypermultiplet of mass \( M \) should be accompanied by the \( CPT \) conjugate of mass \(-M\). We also briefly discuss quiver-type theories with product gauge groups and bi-fundamental matter.

The \( SU(N) \) gauge symmetry of \( \mathcal{N} = 2 \) SYM theories is usually broken to \( U(1)^{N-1} \) by the vev of the adjoint scalar in the vector multiplet:

\[ \langle \Phi \rangle = \text{diag}(a_1, \ldots, a_N). \] (2.1)

The supersymmetric localization reduces the path integral of the theory compactified on \( S^4 \) to a finite dimensional integral over the Coulomb moduli, the eigenvalues of the scalar vev [1]

\[ Z = \int Da Z_{1\text{-loop}}(a) |Z_{\text{inst}}(a)|^2 e^{-S_{\text{cl}}(a)}, \] (2.2)

where \( Da \) is the usual Vandermonde measure on Hermitean matrices:

\[ Da = \prod_i da_i \delta \left( \sum_i a_i \right) \prod_{i<j} (a_i - a_j)^2. \] (2.3)

The classical action arises from the \( \mathcal{R} \text{tr} \Phi^2/4 \) coupling of the scalar to the curvature of \( S^4 \), which is necessary for maintaining supersymmetry (e.g. [16]), and is equal to

\[ S_{\text{cl}} = \frac{8\pi^2 N}{\lambda} \sum_i a_i^2, \] (2.4)

where \( \lambda = g_{\text{YM}}^2 N \) is the ’t Hooft coupling.

The one-loop factor \( Z_{1\text{-loop}} \) was computed in [1] and is expressed in terms of a single function

\[ H(x) \equiv \prod_{n=1}^{\infty} \left( 1 + \frac{x^2}{n^2} \right)^n e^{-\frac{x^2}{n^2}}. \] (2.5)

\[ ^1 \text{More precisely, the integral goes along a real section of the complex moduli space, see [1] for a discussion.} \]
Various properties of this function and of its logarithmic derivative

$$K(x) \equiv -\frac{H'(x)}{H(x)} = 2x \sum_{n=1}^{\infty} \left( \frac{1}{n} - \frac{n}{n^2 + x^2} \right)$$

are listed in appendix [A]. The various multiplets contribute as follows:

Vector multiplet: \[ \prod_{i<j} H^2(a_i - a_j) \]

Adjoint hypermultiplet: \[ \prod_{i<j} \frac{1}{H(a_i - a_j + M)H(a_i - a_j - M)} \]

Fundamental hypermultiplet: \[ \prod_i \frac{1}{H(a_i + M)} \]  

The one-loop factor \( Z_{1\text{-loop}} \) is the product of these factors over the matter content of the theory.

The instanton factor \( Z_{\text{inst}} \) is the Nekrasov partition function \([17]\) with the equivariant parameters set equal: \( \epsilon_1 = \epsilon_2 = 1 \). In the large-\( N \) limit the instantons are suppressed and for the most part we will just drop the instanton factor, except for sections \([3.7, 4.3]\) where we explicitly check that the one-instanton contribution is exponentially small at \( N \to \infty \).

Our conventions are such that the radius \( R \) of the four-sphere is set to one. It is easy to recover the dependence on \( R \), which we will occasionally do, by rescaling all the dimensionful quantities by \( R \): \( a_i \to a_i R \) and \( M \to MR \). This in particular means that the decompactification limit \( (R \to \infty) \) in the radius-one units corresponds to the infinite-mass limit \( M \to \infty \). The equivariant parameters of the instanton partition function, equal to one when \( R = 1 \), in arbitrary units are equal to \( 1/R \).

Apart from to the free energy,

$$F = -\frac{1}{N^2} \ln Z,$$

localization also allows one to compute the expectation value of the circular Wilson loop which, in addition to the gauge field, couples to the scalar from the vector multiplet:

$$W(C) \equiv \left( \frac{1}{N} \text{tr} \text{P} \exp \int_C d\tau \left( i\dot{x}^\mu A_\mu + |\dot{x}|\Phi \right) \right).$$

If the contour \( C \) runs along the big circle of the four-sphere, the path integral with the Wilson loop inserted still localizes to the matrix model. The localization amounts to replacing the fields by their classical values, \( A_\mu = 0 \) and \( \Phi \) given by \([2.1]\), and subsequently integrating over the Coulomb moduli, so the Wilson loop expectation value maps to the exponential operator in the matrix model:

$$W(C) = \left\{ \frac{1}{N} \sum_i e^{2\pi a_i} \right\}.$$
There are two possible ways to view this theory. One way is to start with $\mathcal{N} = 4$ SYM, and add specific dimension two and dimension three operators to the Lagrangian. The field content of $\mathcal{N} = 4$ SYM, in the $\mathcal{N} = 2$ terms, consists of a vector multiplet and two massless adjoint hypermultiplets. The unique massive deformation that preserves half of the supersymmetry is obtained by adding equal masses to the two hypermultiplets. The resulting theory constitutes the simplest relevant perturbation of the superconformal $\mathcal{N} = 4$ theory away from the conformal point. The difference between $\mathcal{N} = 2^*\text{SYM}$ and $\mathcal{N} = 4$ SYM disappears in the UV, for instance on the four-sphere of a very small radius $R \ll 1/M$.

In the opposite limit, say for $MR \gg 1$, the hypermultiplets can be integrated out, leaving behind pure $\mathcal{N} = 2$ SYM, to the leading approximation. As a relevant perturbation of a finite theory and due to $\mathcal{N} = 2$ supersymmetry, $\mathcal{N} = 2^*$ theory is also UV finite. Thus, one can alternatively view $\mathcal{N} = 2^*$ theory as a convenient UV regularization of pure $\mathcal{N} = 2$ SYM, with the hypermultiplet mass playing the rôle of the UV cutoff and the finite ’t Hooft coupling playing the rôle of the bare coupling. From that perspective, $\mathcal{N} = 2^*$ SYM describes a flow from $\mathcal{N} = 4$ SYM in the UV to $\mathcal{N} = 2$ SYM in the IR (fig. [1]). The hypermultiplet mass and the coupling constant combine into the dynamically generated scale of the $\mathcal{N} = 2$ theory:

$$\Lambda = Me^{-\frac{4\pi^2}{\lambda}}.$$  \hfill (3.1)

From this formula and (1.1) we infer the general form of the weak-coupling expansion in the $\mathcal{N} = 2^*$ SYM:

$$\mathcal{A} = \text{perturbative} + \sum_{n=1}^{\infty} C_n e^{-\frac{4\pi^2}{\lambda} n}.$$  \hfill (3.2)

This expansion can be interpreted as OPE, with the higher-order terms originating from irrelevant operators in the low-energy effective field theory obtained by integrating out the hypermultiplets [13].

It is important to realize that the flow picture only makes sense at weak coupling, when the scales $M$ and $\Lambda$ are well separated. When $\lambda$ is not very small, $M$ and $\Lambda$ are of the same order of magnitude and the effective field theory approximation breaks down. Moreover, at strong coupling the suitably defined dynamical scale is parametrically larger than $M$. This follows
from the supergravity analysis of the known holographic dual \[18\] of $\mathcal{N} = 2^*$ SYM \[19\], or can be derived directly from field theory using localization \[12\]. The interactions at $\lambda \to \infty$ are so strong that the mass perturbation distorts the dynamics at energy scales much bigger than $M$.

### 3.1 Localization and large-$N$

In order to use the localization results of \[1\], we compactify $\mathcal{N} = 2^*$ SYM on $S^4$ of radius $R$. We will be interested in the large-$N$ planar limit, in which the theory depends on two parameters, $\lambda$ and $MR$. Perhaps the most interesting regime is the decompactification limit $MR \to \infty$, but it is useful to keep $MR$ as an extra parameter, for instance in order to compare with $\mathcal{N} = 4$ SYM at $MR \to 0$.

The starting point of our analysis is the exact partition function \[1\]:

$$Z^{N=2^*} = \int d^{N-1}a \prod_{i<j} \frac{(a_i - a_j)^2 H^2(a_i - a_j)}{H(a_i - a_j - M)H(a_i - a_j + M)} e^{-\frac{8\pi^2 N}{\lambda} \sum_i a_i^2 |Z_{\text{inst}}|^2}. \quad (3.3)$$

From now on we will set $Z_{\text{inst}} = 1$, returning to instantons later to check if they are indeed exponentially suppressed at large $N$. The resulting matrix model, simple as it is, has unexpectedly rich dynamics. It captures the OPE expansion (3.2) at weak coupling \[13\], agrees with the string-theory predictions at strong coupling \[12\], while at intermediate $\lambda$ it features an infinite sequence of quantum phase transitions, which only happen in the infinite-volume limit. The phase diagram of the model is shown in fig. 2.

In the planar limit, the matrix integral (3.3) is governed by a saddle point. In terms of the eigenvalue density,

$$\rho(x) = \frac{1}{N} \sum_i \delta(x - a_i), \quad (3.4)$$

the saddle-point equations are equivalent to a singular integral equation:

$$\int_{-\mu}^{\mu} dy \rho(y) \left( \frac{1}{x-y} - \mathcal{K}(x-y) + \frac{1}{2} \mathcal{K}(x-y+M) + \frac{1}{2} \mathcal{K}(x-y-M) \right) = \frac{8\pi^2}{\lambda} x. \quad (3.5)$$

The density $\rho(x)$ is defined on an interval $(-\mu, \mu)$ and is unit normalized. The integral equation only holds for $x$ between $-\mu$ and $\mu$. Potentially possible two-cut solutions of the matrix model turn out to be dynamically disfavored \[11\]. For clarity, we switched to the units where $R = 1$.

Once the saddle-point equation is solved, the circular Wilson loop is computed as

$$W(C) = \int_{-\mu}^{\mu} dx \rho(x) e^{2\pi x} \equiv \langle e^{2\pi x} \rangle. \quad (3.6)$$

As far as the free energy is concerned, it is more convenient to calculate its first derivatives:

$$\frac{\partial F}{\partial \lambda} = -\frac{8\pi^2}{\lambda^2} \langle x^2 \rangle$$

$$\frac{\partial F}{\partial M} = \frac{1}{2} \langle (\mathcal{K}(x-y-M) - \mathcal{K}(x-y+M)) \rangle, \quad (3.7)$$
Figure 2: The phase diagram of $\mathcal{N} = 2^*$ theory on $S^4$. On the sphere of a very small radius the theory is equivalent to $\mathcal{N} = 4$ SYM. The opposite, decompactification limit corresponds to the theory defined on flat $\mathbb{R}^4$. Interpolation between $MR = 0$ and $MR = \infty$ at weak coupling describes the RG flow from $\mathcal{N} = 4$ SYM in the UV to $\mathcal{N} = 2$ SYM in the IR. The dynamical scale $\Lambda$ from (3.1) takes on a particular fixed value on each line of constant physics. In the decompactification limit, the theory undergoes an infinite sequence of phase transitions that accumulate at strong coupling. The strong-coupling regime is supposed to have a weakly-curved holographic description at any $R$, but the dual geometry is known only in the two extreme cases: at $MR \to 0$ the theory becomes equivalent to $\mathcal{N} = 4$ SYM, and the background should degenerate to $AdS_5 \times S^5$ in the foliation where the boundary of $AdS_5$ is the round four-dimensional sphere; at $MR = \infty$, the dual background is the Pilch-Warner solution of the ten-dimensional type IIB supergravity [18] that has a flat $\mathbb{R}^4$ boundary.

where the double brackets denote average over both $x$ and $y$.

We will analyze the saddle-point equation in detail at the extreme values of parameters, corresponding to the sides of the rectangle in fig. 2. We will also describe salient features of the solution for generic $M$ and $\lambda$, which can be obtained numerically.

3.2 Weak coupling

The weak-coupling limit of the $\mathcal{N} = 2^*$ matrix model was studied in [10]. Here we add a few more remarks on the free energy and on the validity of the weak-coupling approximation. At small $\lambda$, the classical force term on right-hand-side of the saddle-point equation squeezes the eigenvalue distribution towards zero making $\mu$ very small. We can then expand the kernel in powers of $x - y \sim \mu$. To the leading order we are left with just the Hilbert kernel. The solution

\[ \text{where the double brackets denote average over both } x \text{ and } y. \]

\[ \text{We will analyze the saddle-point equation in detail at the extreme values of parameters, corresponding to the sides of the rectangle in fig. 2. We will also describe salient features of the solution for generic } M \text{ and } \lambda, \text{ which can be obtained numerically.} \]

\[ \text{3.2 Weak coupling} \]

The weak-coupling limit of the $\mathcal{N} = 2^*$ matrix model was studied in [10]. Here we add a few more remarks on the free energy and on the validity of the weak-coupling approximation. At small $\lambda$, the classical force term on right-hand-side of the saddle-point equation squeezes the eigenvalue distribution towards zero making $\mu$ very small. We can then expand the kernel in powers of $x - y \sim \mu$. To the leading order we are left with just the Hilbert kernel. The solution
to the integral equation is then described by the Wigner’s semicircle:
\[ \rho(x) = \frac{2}{\pi \mu^2} \sqrt{\mu^2 - x^2}, \]
(3.8)
with the width of the eigenvalue distribution given by \( \mu = \sqrt{\lambda}/2\pi \).

The next order in the expansion can be taken into account without doing any new calculation. Indeed, the next term is linear in \( x - y \), the \( y \) term integrates to zero due to the \( SU(N) \) constraint, and the \( x \) term renormalizes the coupling constant on the right-hand-side:
\[ \frac{8\pi^2}{\lambda} \to \frac{8\pi^2}{\lambda_R} = \frac{8\pi^2}{\lambda} - K'(M). \]
(3.9)
At large \( M \), we can use (A.5) to approximate \( K(M) \). The effective coupling \( \lambda_R \) then coincides with the running \( \mathcal{N} = 2 \) Yang-Mills coupling renormalized at the scale set by the radius of the four-sphere:
\[ \frac{4\pi^2}{\lambda_R} \simeq \frac{4\pi^2}{\lambda} - \ln M - 1 - \gamma = -\ln \Lambda - 1 - \gamma \quad (M \gg 1). \]
(3.10)
The lines of constant \( \lambda_R \) are the lines of constant physics in the \( \mathcal{N} = 2 \) theory.

The solution of the saddle-point equations is again the semicircle with the width determined by the renormalized coupling:
\[ \mu = \frac{\sqrt{\lambda_R}}{2\pi}. \]
(3.11)
Having the eigenvalue density, we can compute the expectation value of the circular Wilson loop and the free energy. We find:
\[ W(C) = 1 + \frac{\lambda_R}{8} + \frac{\lambda_R^2}{192} + O\left(\lambda_R^3\right) = 1 + \frac{1}{8} \lambda + \left(\frac{1}{192} + \frac{K'(M)}{64\pi^2}\right) \lambda^2 + O(\lambda^3). \]
(3.12)
For the derivatives of the free energy, given by (3.7), we have:
\[ \frac{\partial F}{\partial \lambda} = -\frac{\lambda_R}{2\lambda} \]
(3.13)
\[ \frac{\partial F}{\partial M} = -K(M) - \frac{\lambda_R}{16\pi^2} K''(M), \]
(3.14)
where in the first equation we used
\[ \langle x^2 \rangle = \frac{\mu^2}{4} = \frac{\lambda_R}{16\pi^2}, \]
(3.15)
and in the second equation expanded the kernel to the second order in \( x - y \). Integrating we obtain:
\[ F = -\frac{1}{2} \ln \lambda_R + \ln H(M) + O\left(\lambda_R^2\right) = -\frac{1}{2} \ln \lambda + \ln H(M) - \frac{K'(M)}{16\pi^2} \lambda + O\left(\lambda^2\right). \]
(3.16)
In particular, using the asymptotic behavior of $H(x)$ (see appendix [A]), we note that at large $MR$ (here we recover the dependence on $R$):

$$\ln Z \rightarrow N^2 M^2 R^2 \ln MR \quad (MR \gg 1).$$

This reproduces the expected UV divergence of the partition function originating from zero modes of the one-loop determinant$^2$.

One can notice from the calculations above that the perturbative expansion is reorganized in power series in $\lambda R$. At $M \sim 1$ this is not a big change, as $\lambda_R \sim \lambda$, but when $M$ becomes big, $\lambda_R$ contains a large logarithm. When written in terms of $\lambda_R$, expressions above resum large logs of the form $\lambda(\ln M)^n$ and, for the Wilson loop, $\lambda^2(\ln M)^n$. The limit of $M \rightarrow \infty$ at fixed $\lambda_R$ corresponds to $\mathcal{N} = 2$ SYM with fixed dynamical scale $\Lambda$. The results above are valid when the radius of the sphere is small in the $\mathcal{N} = 2$ units, and consequently $\lambda_R \ll 1$. At finite radius (finite $\lambda_R$), the width of the eigenvalue distribution $\mu$ is no longer small compared to one, but is still much smaller than $M$. The integral equation valid in this regime is obtained from (3.5) by expanding only the last two terms in the kernel:

$$\int_{-\mu}^{\mu} dy \rho(y) \left( \frac{1}{x-y} - \mathcal{K}(x-y) \right) = \frac{8\pi^2}{\lambda_R} x. \quad (3.17)$$

This equation describes pure $\mathcal{N} = 2$ SYM on $S^4$ and was analyzed in detail in [11]$^3$.

### 3.3 Strong coupling

The limit of $\lambda \gg 1$ is supposed to have a holographic description in terms of a weakly-curved supergravity dual. For the theory defined on flat $\mathbb{R}^4$ the dual geometry is known [18]. The large-$N$ solution of $\mathcal{N} = 2^*$ SYM at strong coupling [12], continued to infinite volume of $S^4$, completely agrees with available predictions of holography. We briefly review these results, for completeness.

As $\lambda$ is increased, the attractive linear force $8\pi^2 x/\lambda$ in the saddle-point equation (3.5) becomes weaker, and the eigenvalue distribution expands to larger and larger values of $x$. Eventually, at sufficiently big $\lambda$, the width of the eigenvalue distribution becomes much larger than the bare mass scale of the $\mathcal{N} = 2^*$ theory: $\mu \gg M$. In this case we can assume that $|x-y| \gg M$ and rescale $x \rightarrow x\mu$. This justifies the approximation$^4$

$$\frac{1}{2} \mathcal{K}(x-y+M) + \frac{1}{2} \mathcal{K}(x-y-M) - \mathcal{K}(x-y) \approx \frac{1}{2} \mathcal{K}''(x-y) M^2 \approx \frac{M^2}{x-y}. \quad (3.18)$$

At the end, the only effect of the complicated $\mathcal{K}$-terms in the equations is a multiplicative renormalization of the Hilbert kernel:

$$\frac{1}{x-y} \rightarrow \frac{1 + M^2}{x-y}. \quad (3.19)$$

---

$^2$We thank Arkady Tseytlin for comments on this point.

$^3$In comparing to [11] it is necessary to take into account that our definition of $\Lambda$ differs from that in [11] by a factor of $e^{-1-\gamma}$.

$^4$It is also true that $|x-y| \gg 1$, which is important for the last step in the approximation.
The equation is again solved by the Wigner distribution (3.8), but now with
\[ \mu = \frac{\sqrt{\lambda (1 + M^2)}}{2\pi}. \] (3.20)

The Wilson loop vev then behaves as
\[ \ln W(C) \simeq \sqrt{\lambda (1 + M^2)}, \] (3.21)
and the free energy is given by\(^5\)
\[ F = -\frac{1 + M^2}{2} \ln \left( \frac{\lambda (1 + M^2)}{16\pi^2} \right) e^{2\gamma + \frac{1}{2}} + \frac{1}{4} - \ln 4\pi. \] (3.22)

The strongly-coupled \( \mathcal{N} = 2^* \) theory, on flat space, is dual to type IIB supergravity on the Pilch-Warner background [18]. In the equations above, we can reach the flat-space limit by rescaling \( M \to MR, \mu \to \mu R \) and taking \( R \to \infty \), which amounts to just dropping the 1 in \( 1+M^2 \). The results are then in perfect agreement with supergravity. In particular, the probe analysis of the Pilch-Warner solution shows that the eigenvalues of the Higgs vev are distributed on an interval in the two-dimensional moduli space with the semicircular Wigner density [19]. The width of the distribution, \( \mu = \sqrt{\lambda M/2\pi} \), is the same as (3.20) with \( M \gg 1 \). The semicircular shape of the eigenvalue distribution is actually a generic prediction of the probe analysis, valid for many other supergravity backgrounds [20]. It is a challenge for localization to reproduce this universality on the field-theory side.

The vev of the circular Wilson loop (3.21) obeys the perimeter law at strong coupling. Extrapolating this behavior to generic Wilson loops, we may expect that the Wilson loop vev for a contour of length \( L \) behaves as \( \ln W(C) \simeq \sqrt{\lambda LM}/2\pi \), assuming \( LM \gg 1 \). Taking the standard value \( \sqrt{\lambda}/2\pi \) for the dimensionless string tension, and using the same regularization prescription as in \( \text{AdS}_5 \times S^5 \), one reproduces this result from the minimal area law in the Pilch-Warner geometry [12]. To compare the supergravity predictions with localization for finite \( MR \), it is necessary to find the supergravity solution whose boundary is \( S^4 \) rather than flat \( \mathbb{R}^4 \). First steps towards constructing such solution were taken in [21].

### 3.4 Conformal perturbation theory

In the limit of mass going to zero, the theory flows to \( \mathcal{N} = 4 \) SYM. The matrix model then becomes Gaussian, and the eigenvalues form the Wigner distribution (3.8) of width \( \mu = \sqrt{\lambda}/2\pi \). If \( \lambda \) is large, the circular Wilson loop [22, 23] and the free energy [11] computed in the Gaussian matrix model agree with the minimal area law and the on-shell action of type IIB supergravity on \( \text{AdS}_5 \times S^5 \). It is of interest to calculate the first correction in \( M^2 \), at any \( \lambda \). The expansion in \( M \) can be interpreted as conformal perturbation theory.

---

\(^5\)We shifted the result in [12] by a constant in order to uniformly normalize the free energy across the whole phase diagram, and in particular to make it agree with the free energy of \( \mathcal{N} = 4 \) SYM at \( M = 0 \), in the normalization of [11]. Of course adding a constant to the free energy does not change any physics.
Expanding (3.5) to the leading order in $M$, we get:

$$\int_{-\mu}^{\mu} dy \frac{\rho(y)}{x-y} = \frac{8\pi^2}{\lambda} x - \frac{M^2}{2} \int_{-\mu}^{\mu} dy \rho(y) K''(x-y).$$

(3.23)

Treating the second term on the right hand side as perturbation we can plug in the leading order solution for the $\rho(y)$ and, using the Fourier representation (A.6), we obtain

$$\int_{-\mu}^{\mu} dy \frac{\rho(y)}{x-y} = \frac{8\pi^2}{\lambda} x - \frac{4\pi M^2}{\sqrt{\lambda}} \int_{0}^{\infty} d\omega \frac{\omega J_1\left(\frac{\sqrt{\lambda}\omega}{\pi}\right) \sin 2\omega x}{\sinh^2 \omega}.$$

(3.24)

The Hilbert kernel is inverted by applying

$$\int_{-\mu}^{\mu} dx \frac{1}{\sqrt{\mu^2 - x^2}} \frac{1}{z-x}$$

(3.25)

to both sides of the equation. We thus find:

$$\rho(x) = \frac{8\pi}{\lambda} \sqrt{\mu^2 - x^2} \left[ 1 - \frac{\sqrt{\lambda} M^2}{\pi} \int d\omega \frac{\omega^2 J_1\left(\frac{\sqrt{\lambda}\omega}{\pi}\right) J_0\left(\frac{\sqrt{\lambda}(1-s)\omega}{\pi}\right) \cos 2s\omega x}{\sinh^2 \omega} \right] + O(M^4).$$

(3.26)

To determine $\mu$, we need to impose the normalization condition on the density. Integrating both sides of (3.26) from $-\mu$ to $\mu$, and replacing $\mu$ by $\sqrt{\lambda}/2\pi$ in the second term we get:

$$1 = \frac{4\pi^2 \mu^2}{\lambda} - 2M^2 \int d\omega \int ds \frac{\omega J_1\left(\frac{\sqrt{\lambda}\omega}{\pi}\right) J_0\left(\frac{\sqrt{\lambda}(1-s)\omega}{\pi}\right) J_1\left(\frac{\sqrt{\lambda}s\omega}{\pi}\right)}{s \sinh^2 \omega} + O(M^4).$$

(3.27)

The $s$-integral here can be computed explicitly, and we finally obtain:

$$\mu = \frac{\sqrt{\lambda}}{2\pi} \left( 1 + M^2 \int_{0}^{\infty} d\omega \frac{\omega J_1^2\left(\frac{\sqrt{\lambda}\omega}{\pi}\right)}{\sinh^2 \omega} + O(M^4) \right)$$

(3.28)

This expression is first order in $M^2$, but is non-perturbative in $\lambda$. To make contact with the results of the two previous sections, we consider the limiting cases of $\lambda \ll 1$ and $\lambda \gg 1$. If $\lambda$ is large, the main contribution to the integral comes from $\omega \sim 1/\sqrt{\lambda}$. The $\sinh^2 \omega$ in the denominator can be then approximated by $\omega^2$, after which the whole expression integrates to $1/2$. We thus find

$$\mu \approx \frac{\sqrt{\lambda}}{2\pi} \left( 1 + \frac{M^2}{2} \right) \quad (\lambda \to \infty),$$

(3.29)

in agreement with the strong-coupling result (3.20) expanded to the first order in $M^2$. 

11
The weak-coupling limit is even simpler, we just need to expand the Bessel function under the integral. This gives

\[
\mu \approx \frac{\sqrt{\lambda}}{2\pi} \left( 1 + \frac{3\zeta(3)\lambda M^2}{8\pi^2} \right) \quad (\lambda \to 0),
\]

which matches with (3.11), (3.10), if we take into account that \( K'(M) = 6\zeta(3)M^2 + O(M^4) \) according to (A.3).

The weak-coupling expansion in fact can be carried out to all orders in \( \lambda \):

\[
\mu = \frac{\sqrt{\lambda}}{2\pi} \left[ 1 + M^2 \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(2k)! (2k+1)! \zeta(2k+1)}{(k-1)! (k+1)! k!^2} \left( \frac{\lambda}{16\pi^2} \right)^k + O(M^4) \right].
\]

It is necessary to stress that the general arguments on the structure of perturbation series given in the introduction are not valid in the conformal limit. These arguments should apply in the opposite, IR regime, when the sphere is big and effective field theory gives an accurate description of physics. The \( M^2 \) correction, calculated above, should be interpreted as the first order of conformal perturbation theory around the \( \mathcal{N} = 4 \) point. As such, it should inherit the perturbative structure of the \( \mathcal{N} = 4 \) SYM, well understood due to integrability [24].

In a finite theory, such as \( \mathcal{N} = 4 \) SYM, planar perturbation theory should have a finite radius of convergence [25] which is determined by the combinatorics of planar graphs and thus should not depend much on the particular observable. Since the spectrum of local gauge-invariant operators in \( \mathcal{N} = 4 \) SYM is quite well understood, we can draw some conclusions on the radius of convergence from the spectral problem. Computation of the anomalous dimensions of local operators can be conveniently mapped to a spin-chain problem [26]. Due to integrability, the only necessary input is the dispersion relation of the elementary magnon excitations of the spin chain and their two-body S-matrix. The exact dispersion relation of the \( \mathcal{N} = 4 \) magnon is [27]

\[
\epsilon_{\text{mag}}(p) = \sqrt{1 + \frac{\lambda}{\pi^2} \sin^2 \frac{p}{2}}.
\]

This expression has manifestly finite radius of convergence in \( \lambda \). The “staggered” magnon with momentum at the edge of the Brillouin zone: \( p = \pi \) has the smallest radius of convergence. Its energy has a square root branch point at

\[
\lambda_c = -\pi^2.
\]

This should be the radius of convergence of generic observable.

Quite remarkably, the radius of convergence of perturbation series in (3.31) is exactly the same. Indeed, perturbative coefficients in (3.31) behave as const \( \cdot \pi^{-2k} \) at large \( k \) indicating a pole at \( \lambda = -\pi^2 \). From the point of view of the integral representation (3.28), the singularity at \( \lambda = -\pi^2 \) occurs because of the exponential growth of the Bessel function of an imaginary argument, which saturates the convergence of the integral at large \( \omega \) when \( \lambda \) approaches \( \lambda_c \). Using the large-argument asymptotics of the Bessel function we get:

\[
\mu \approx \frac{\sqrt{\lambda M^2}}{\pi^2 + \lambda} \quad (\lambda \to -\pi^2).
\]
The quantities that may have a more direct \( N = 4 \) interpretation are the free energy and the circular Wilson loop, to the computation of which we now proceed. To calculate the free energy we integrate (3.26) with the \( x^2 \) weight, which gives:

\[
\langle x^2 \rangle = \frac{\pi^2 \mu^4}{\lambda} + \frac{M^2}{2} \int_0^\infty d\omega \int_0^1 ds \frac{J_1\left(\frac{\sqrt{\lambda} \omega}{\pi}\right) J_0\left(\frac{\sqrt{\lambda} (1-s) \omega}{\pi}\right)}{\omega \sinh^2 \omega} \frac{\partial^2}{\partial s^2} \frac{J_1\left(\frac{\sqrt{\lambda} \omega}{\pi}\right)}{s} + O(M^4). \tag{3.35}
\]

Taking \( \mu \) from (3.28), calculating the \( s \)-integral, and using (3.7), we get:

\[
\frac{\partial F}{\partial \lambda} = -\frac{1}{2\lambda} - \frac{8\pi^2 M^2}{\lambda} \int_0^\infty d\omega \frac{J_1\left(\frac{\sqrt{\lambda} \omega}{\pi}\right) \left( J_1\left(\frac{\sqrt{\lambda} \omega}{\pi}\right) - \frac{\sqrt{\lambda} \omega}{2\pi} J_0\left(\frac{\sqrt{\lambda} \omega}{\pi}\right) \right)}{\omega \sinh^2 \omega} + O(M^4). \tag{3.36}
\]

This equation can be integrated to

\[
F = -\frac{1}{2} \ln \lambda - \frac{4\pi^2 M^2}{\lambda} \int_0^\infty d\omega \frac{\lambda \omega^2}{4\pi^2} - J_1^2\left(\frac{\sqrt{\lambda} \omega}{\pi}\right) + O(M^4). \tag{3.37}
\]

The leading term here is the free energy of \( N = 4 \) SYM. It can be understood holographically as arising from the log-divergence of the on-shell supergravity action on \( AdS_5 \times S^5 \) upon taking into account a factor of \( \sqrt{\lambda} \) in the radius-energy relation in the AdS/CFT correspondence.

Again, we can check consistency with the weak-coupling results by expanding the Bessel function:

\[
F \sim -\frac{1}{2} \ln \lambda - \frac{3\zeta(3) \lambda M^2}{8\pi^2} \quad (\lambda \to 0). \tag{3.38}
\]

This is in agreement with (3.16), when (A.3) is used for \( \ln H(M) \) and \( K'(M) \). Checking the strong-coupling limit is a bit trickier, since the integral in (3.37) logarithmically diverges on the upper limit if \( \sinh^2 \omega \) is replaced by \( \omega^2 \). Cutting off the resulting integral at \( \omega \sim 1 \) we get, with the logarithmic accuracy:

\[
F \sim -\frac{1}{2} \left(1 + M^2\right) \ln \lambda, \tag{3.39}
\]

in precise agreement with (3.22). Systematic strong-coupling expansion of (3.37) requires matching contributions from \( \omega \sim 1/\sqrt{\lambda} \) and \( \omega \sim 1 \). To the first two non-vanishing orders,

\[
\left. \frac{\partial F}{\partial M^2} \right|_{M=0} = -\frac{1}{2} \ln \frac{\lambda}{16\pi^2} - \frac{3}{4} - \gamma - \frac{2\pi^2}{3\lambda} + O\left(\frac{1}{\lambda^{3/2}}\right). \tag{3.40}
\]

In general the expansion goes in powers of \( 1/\sqrt{\lambda} \), as expected from string theory in \( AdS_5 \times S^5 \). It is curious though that the term of order \( M^2/\sqrt{\lambda} \) is absent.

The radius of convergence of perturbation series for the free energy is also \( \pi^2 \) as in (3.33). The singularity at \( \lambda = -\pi^2 \) is a logarithmic branch cut:

\[
F = \text{analytic} - \frac{8M^2}{\pi^3} \left( \pi^2 + \lambda \right) \ln \left( \pi^2 + \lambda \right) \quad (\lambda \to -\pi^2). \tag{3.41}
\]
Finally, one can also compute the vev of the circular Wilson loop, $W = \langle e^{2\pi x} \rangle$. It is useful to consider a more general expectation value:

$$\langle e^{2\pi ax} \rangle = \frac{4\pi \mu I_1 (2\pi a\mu)}{a\lambda} - M^2 \int_0^\infty d\omega \frac{\omega^2 J_1 \left( \frac{\sqrt{\lambda}}{\pi} \right)}{\sinh^2 \omega} \int_0^1 ds J_0 \left( \frac{\sqrt{\lambda} (1-s)}{\pi} \right) \left[ \frac{J_1 \left( \frac{\sqrt{\lambda}}{\pi} + i a \sqrt{\lambda} \right)}{s\omega + i\pi a} \right] + (a \to -a),$$

which can be brought to a simpler form:

$$\langle e^{2\pi ax} \rangle = \frac{2I_1 \left( a\sqrt{\lambda} \right)}{a\sqrt{\lambda}} + 2\pi aM^2 \int_0^\infty d\omega \frac{\omega J_1 \left( \frac{\sqrt{\lambda}}{\pi} \right) \left( \pi a I_0 \left( a\sqrt{\lambda} \right) J_1 \left( \frac{\sqrt{\lambda}}{\pi} \right) - \omega I_1 \left( a\sqrt{\lambda} \right) J_0 \left( \frac{\sqrt{\lambda}}{\pi} \right) \right)}{(\omega^2 + \pi^2 a^2) \sinh^2 \omega}.$$  \hspace{1cm} (3.42)

Eq. (3.36) can be obtained by expanding this expression to second order in $a$. Setting $a = 1$, we get:

$$W(C) = \frac{2I_1 \left( \sqrt{\lambda} \right)}{\sqrt{\lambda}} + 2\pi M^2 \int_0^\infty d\omega \frac{\omega J_1 \left( \frac{\sqrt{\lambda}}{\pi} \right) \left( \pi I_0 \left( \sqrt{\lambda} \right) J_1 \left( \frac{\sqrt{\lambda}}{\pi} \right) - \omega I_1 \left( \sqrt{\lambda} \right) J_0 \left( \frac{\sqrt{\lambda}}{\pi} \right) \right)}{(\omega^2 + \pi^2) \sinh^2 \omega}.$$  \hspace{1cm} (3.43)

It is straightforward to check that the weak-coupling result (3.12) is reproduced when $W(C)$ is expanded in $\lambda$. The leading asymptotics at strong coupling is

$$W(C) \simeq \sqrt{\frac{2}{\pi}} \lambda^{-\frac{3}{4}} e^{\sqrt{\lambda}} \left( 1 + \frac{\sqrt{\lambda} M^2}{2} \right) \quad (\lambda \to \infty),$$

again in agreement with the previous result, eq. (3.21). Near the critical point $\lambda = -\pi^2$, the Wilson loop, like the free energy, has a logarithmic branch point:

$$W(C) \simeq -4M^2 J_1 (\pi) \ln (\pi^2 + \lambda) \quad (\lambda \to -\pi^2).$$  \hspace{1cm} (3.46)

The circular Wilson loop in $\mathcal{N} = 4$ theory is given by the first term in (3.44) \cite{22, 23} and has an infinite radius of convergence, which happens because of massive diagram cancellations. It is known that only rainbow graphs without internal vertices contribute \cite{22}. In this sense the circular Wilson loop at $M = 0$ is not a generic observable. At the first order of conformal perturbation theory, the circular loop starts to receive contributions from generic diagrams. This shifts the radius of convergence to the expected $\lambda_c = -\pi^2$ point. Making a more direct contact with the AdS/CFT integrability, and in particular reformulating conformal perturbation theory for the free energy (3.37) and for the Wilson loop (3.44) in the $\mathcal{N} = 4$ language is an interesting open problem.
### 3.5 Decompactification limit

The compactification on the sphere can be considered just as a convenient way to impose an IR cutoff, necessary for computing the path integral by localization. If we take this point of view, the radius of the sphere $R$ should be sent to infinity at the end of the calculation. All the dimensionful parameters, including $\mu$ and $M$, scale linearly with $R$ and after $R$ is sent to infinity and eliminated from the equations, the dimensionful quantities regain their canonical dimensions. In the limit $R \to \infty$, the Hilbert kernel drops out from the saddle-point equation (3.5), and $K(x)$ can be approximated by its asymptotics at infinity (A.5). It is convenient to differentiate the resulting equation twice, which gives:

$$
\int_{-\mu}^{\mu} dy \rho(y) \left( \frac{2}{x-y} - \frac{1}{x-y+M} - \frac{1}{x-y-M} \right) = 0.
$$

The Hilbert kernel of the original equation produces an $1/(x-y)^3$ term, which scales as $1/R^2$ and can therefore be neglected in the decompactification limit.

As shown in [13], the boundary conditions at the ends of the interval change from the square root at finite $R$ to the inverse square root in the strict $R \to \infty$ limit. The integral equation (3.47) with the inverse-square-root boundary conditions has normalizable solutions for any $\mu$ (the norm can be adjusted by simply multiplying $\rho(x)$ with a constant). An extra condition, which fixes $\mu$, follows from the integrated form of (3.47), equivalent to the original saddle-point equation differentiated once:

$$
\int_{-\mu}^{\mu} dx \rho(x) \ln \left( \frac{M^2}{x^2} - 1 \right) = \frac{16\pi^2}{\lambda}.
$$

The set of equations (3.47), (3.48) can be easily solved at weak coupling, when $\mu \ll M$. The last two terms in the kernel then approximately cancel leading to a very simple equation

$$
\int_{-\mu}^{\mu} dy \rho(y) \frac{1}{x-y} = 0,
$$

whose properly normalized solution is

$$
\rho(x) = \frac{1}{\pi \sqrt{\mu^2 - x^2}} \quad (\lambda \to 0).
$$

An extra condition (3.48) then determines $\mu$:

$$
\mu = 2M e^{-\frac{4\pi^2}{\lambda}} = 2\Lambda \quad (\lambda \to 0),
$$

where $\Lambda$ is the dynamically generated scale in the IR limit of the pure $\mathcal{N} = 2$ theory. The solution (3.50) was derived from localization in [11] and reproduces earlier results obtained by taking the large-$N$ limit within Seiberg-Witten theory [5, 6]. Interestingly, the same distribution of
eigenvalues arises in one of the supergravity solutions proposed as a holographic dual of pure $\mathcal{N} = 2$ SYM \cite{29}.

The system of equations (3.47), (3.48) can be solved analytically without making any approximations \cite{13} with the help of the method proposed in \cite{30, 31}. The solution is actually valid as long as $\lambda$ is not too big. When $\lambda$ reaches a critical value $\lambda_c \approx 35$, the theory undergoes a transition to a new phase. Similar phase transitions happen at larger couplings: $\lambda_c^{(2)} \approx 83$, $\lambda_c^{(3)} \approx 150$, and so on, with an infinite sequence of critical points accumulating at strong coupling.

The phase transitions are caused by new light states that appear in the spectrum. Each pole in the kernel of the integral equation (3.47) corresponds to a massless, or nearly massless particle. The pole at $x = y$ arises due to the photons of the unbroken $U(1)^{N-1}$, while the poles at $x = y \pm M$ correspond to massless hypermultiplets.\footnote{More precisely, to very light hypermultiplets, whose masses scale as $1/N$ and $1/N^2$ in the large-$N$ limit (cf. \cite{3}).} Of course at weak coupling, when $M \gg \mu$, all hypermultiplets are quite heavy. This is reflected in the integral equation (3.47) by the absence of hypermultiplet poles in the region of integration, where $|x - y| < M$ for any $x$ and $y$ lying within the interval $(-\mu, \mu)$. The largest possible value of $|x - y|$ is equal to $2\mu$. When $\lambda$ becomes bigger, $\mu$ also grows and eventually exceeds $M/2$. When $\mu$ reaches $M/2$ the first resonance appears in the spectrum, causing transition to a new phase. At $\mu = M$ the secondary resonance appears, leading to another phase transition, and so on. The $n$-th critical point is determined by the condition

$$\mu\left(\lambda_c^{(n)}\right) = \frac{nM}{2}.$$ (3.52)

Using the strong-coupling solution (3.20) we can estimate asymptotically

$$\lambda_c^{(n)} \approx \frac{\pi^2 n^2}{\mu} \quad (n \to \infty).$$ (3.53)

After reviewing the exact solution found in \cite{13} in the weak-coupling phase, we will study the critical behavior near the first phase transition, and then will analyze the structure of the strong-coupling phase in more detail.

### 3.5.1 Exact solution in weak-coupling phase

The solution found in \cite{13} is written down in terms of the resolvent:

$$G(z) = \int_{-\mu}^{\mu} \frac{dy \rho(y)}{(z - y)^2 - \frac{M^2}{4}}.$$ (3.54)

The resolvent is an analytic function on the complex plane with two distinct cuts ($\pm M/2 - \mu, \pm M/2 + \mu$), centered at $\pm M/2$. When $\lambda$ is small, the cuts are very short and are well separated. With $\lambda$ growing, the endpoints of the cuts move closer to the origin and eventually collide at $z = 0$, after which the two cuts coalesce. That happens when $\mu = M/2$. This is another way to see how the phase transition arises in the solution of the saddle-point equations.
The eigenvalue density can be found from the resolvent by taking discontinuity across one of the cuts:

$$\rho(x) = \frac{M}{2\pi i} \left( G\left(x + \frac{M}{2} - i0\right) - G\left(x + \frac{M}{2} + i0\right)\right).$$  \hfill (3.55)

Qualitatively, it has the same shape as \([3.50]\), with the inverse square root singularities at the endpoints and a minimum at \(z = 0\), although the precise functional form of the density is more complicated than a simple square root.

As long as the cuts do not overlap, that is, before the phase transition the resolvent can be found exactly by integrating the equation \([13]\):

$$dz = -\frac{dG}{2\sqrt{G^3(1 + \xi G)(1 + \eta G)(1 + \bar{\eta} G)}}.$$  \hfill (3.56)

The inverse function \(z(G)\) can thus be expressed in terms of elliptic integrals. The parameters that characterize the solution\(^7\) are given by

$$\begin{align*}
\xi &= \frac{M^2}{12} \left( E_2 + \theta_3^4 + \theta_4^4 \right) \\
\eta &= \frac{M^2}{12} \left( E_2 - 2\theta_3^4 + \theta_4^4 \right) \\
\bar{\eta} &= \frac{M^2}{12} \left( E_2 + \theta_3^4 - 2\theta_4^4 \right),
\end{align*}$$  \hfill (3.57)

where \(E_2 \equiv E_2(-r^2)\) is the Eisenstein series of index two, and \(\theta_a \equiv \theta_a(0|ir)\) are the theta-constants, which both depend on the 't Hooft coupling through the modular parameter

$$r = e^{-\frac{4\pi^2}{\lambda}}.$$  \hfill (3.58)

Notations and conventions for the theta-functions and Eisenstein series are listed in the appendix\([B]\) where we also collect some of their useful properties. An equivalent representation in terms of elliptic integrals is given in appendix\([C]\). It is clear that the resolvent only depends on symmetric combinations of the three parameters \(\xi, \eta, \bar{\eta}\). As a consequence, correlation functions \(<x^{2n}>\) are symmetric polynomials in \(\xi, \eta, \bar{\eta}\) of degree \(n\) \([13]\). It is shown in appendix\([B]\) that any such symmetric polynomial can be expressed through the Eisenstein series only, with all theta-constants canceling out.

The width of the eigenvalue distribution is given by

$$\mu = -i M \frac{\theta_4'(v)}{2 \theta_4(v)},$$  \hfill (3.59)

where the parameter \(v\) is a solution of the transcendental equation

$$\frac{\theta_3^2(v)}{\theta_4^2(v)} = \frac{2\theta_3^4 - \theta_4^4 - E_2}{3\theta_3^2\theta_4^2}.$$  \hfill (3.60)

\(^7\)Those are related to the parameters \(\xi_i\) in \([13]\) as follows: \(\xi_1 = -\xi, \xi_2 = -\bar{\eta}, \xi_3 = \eta\).
The theta-functions of a real argument and pure imaginary modulus satisfy

\[ \theta_{1,2}^4 = -\theta_{1,2}^4, \quad \theta_{3,4}^4 = \theta_{4,3}^4. \]

Consequently, \( \eta \) and \( \bar{\eta} \) are complex conjugate to one another and \( \xi \) is real positive. It is more difficult to see that \( \mu \) is real and positive. One can show using Landen transformations that the solution of (3.60) is of the form \( v = \pi/4 + is \) with real negative \( s \) and that for such \( v \), \( \mu \) given by (3.59) is indeed real [13].

Expanding (3.59), (3.60) at small \( \lambda \) we get:

\[ \frac{\mu}{M} = 2e^{-4s^2/\lambda} - 4e^{-12s^2/\lambda} - 16e^{-28s^2/\lambda} - 58e^{-44s^2/\lambda} - 324e^{-50s^2/\lambda} + \ldots \] (3.61)

The first term reproduces (3.51). The rest of the expansion has precisely the form (3.2) anticipated on general grounds from the OPE in the effective field theory. The OPE coefficients can in principle be computed to any desired order and are just numbers (potentially they could have been power series in \( \lambda \)).

For the free energy, the OPE can actually be resummed into a rather compact expression. The free energy can be calculated from (3.7), using

\[ \langle x^2 \rangle = \frac{\xi + \eta + \bar{\eta}}{3} - \frac{M^2}{12} = \frac{M^2}{12} (E_2 - 1). \] (3.62)

The first equality follows from comparing the Laurent expansion of the resolvent (3.54) at \( z = \infty \) with that of the exact solution (3.56). The second equality is a consequence of (3.57). It is straightforward to generalize this computation to higher correlators \( \langle x^{2n} \rangle \) with \( n > 1 \). Explicit expressions for the first few are listed in appendix B. Integrating (3.7) in \( \lambda \) we find:

\[ F = -M^2 \ln \left( M e^{\gamma - \frac{1}{2}} \prod_{n=1}^{\infty} \left[ 1 - (-1)^n e^{-8s^2_n} \right]^2 \right). \] (3.63)

Up to an unessential, \( \lambda \)-independent constant the free energy can be written as

\[ F = \text{const} + 2M^2 \sum_{k=1}^{\infty} (-1)^k \sigma_{-1}(k) e^{-s_k^2/\lambda}. \] (3.64)

The divisor function \( \sigma_{-1}(k) \) has a nice combinatorial interpretation, which suggests that the quantum field theory calculation reproducing the OPE coefficients for the free energy might reduce to simple combinatorics.

The above formulas are non-perturbative in \( \lambda \), and have remarkably simple modular properties. It is actually tempting to extend them to strong coupling with the help of modular transformations (see appendix B). Unfortunately, this makes little sense, as the strong-coupling regime is described by a totally different solution. The phase transition that happens in between invalidates analytic continuation to \( \lambda \) larger than \( \lambda_c \). We now turn to the detailed discussion of the phase transition.
3.5.2 Phase transition

The phase transition, according to (3.52), happens when \( \mu(\lambda_c) = M/2 \). As shown in [13], the parameter \( \xi \) vanishes at the critical point \( \xi(\lambda_c) = 0 \). The critical coupling thus corresponds to the solution of the transcendental equation \( E_2 + \theta_3^4 + \theta_4^4 = 0 \). Solving this equation numerically we find:

\[
rc = 0.328107 \ldots \quad \lambda_c = 35.4252 \ldots (3.65)
\]

A convenient small parameter in the vicinity of the critical point is

\[
\Delta = \mu - \frac{M}{2}. (3.66)
\]

In appendix [C] we show that in the weak-coupling phase

\[
\Delta = -\frac{2\xi^2}{3|\eta|} + o\left(\xi^2\right), (3.67)
\]

or, in terms of the coupling constant:

\[
\Delta \approx -C_\Delta M(\lambda_c - \lambda)^{\frac{2}{3}}, (3.68)
\]

with

\[
C_\Delta = \frac{2\sqrt{2\pi^3}\theta_3^4\theta_4^4}{3\lambda_c^2} = 0.0031347 \ldots (3.69)
\]

The theta-constants here are evaluated at the critical coupling given by (3.65).

The critical density \( \rho_c(x) \) (the eigenvalue density right at the critical point) can be calculated by setting \( \xi = 0 \) in (3.55). Even then the resolvent is still an elliptic integral. However, certain simplifications occur near the endpoints of the eigenvalue distribution, when the resolvent becomes very large and the 1’s in \((1 + \eta G)\) and \((1 + \bar{\eta} G)\) can be dropped. Then (3.55) can be easily integrated:

\[
G(z) \approx (3|\eta|)^{-\frac{2}{3}}(z - M)^{-\frac{2}{3}}. (3.70)
\]

This approximation is valid as long as \(|z - M| \ll M\). The constant of integration was chosen such that the branch point lies exactly at \( z = \mu_c + M/2 = M \). Taking discontinuity of the resolvent across the cut, we find for the density, according to (3.55):

\[
\rho_c(x) \approx C_\rho M^{1/3} \left(\frac{M}{2} - x\right)^{-\frac{2}{3}}, (3.71)
\]

with

\[
C_\rho = \frac{M^{\frac{2}{3}}}{2 \cdot \frac{3^4}{2} \frac{1}{\pi |\eta|^{\frac{2}{3}}}} = \frac{2^{\frac{1}{3}}}{3^\frac{1}{2} \pi \theta_3^\frac{2}{3} \theta_4^\frac{2}{3}} = 0.198421 \ldots (3.72)
\]

Interestingly, and in contradistinction to more conventional matrix models, the endpoint singularity hardens at the critical point. The scaling exponent changes from \(-1/2\) away from criticality to \(-2/3\) at the transition point.

\[\text{This happens because } G(0) = -1/\xi. \text{ At the critical point the two cuts of the resolvent collide at } z = 0, \text{ and since the resolvent is singular at the endpoints of the cuts, } G(0) \text{ blows up.}\]
3.5.3 Critical behavior

It was found numerically \cite{13} that after the phase transition the density develops two cusps at 
\( x = \pm (M - \mu) \). The dynamical reason can be understood from the saddle-point equation (3.47):
when \( \mu > M/2 \) the force due to the \( 1/(x - y \pm M) \) terms in the equation becomes repulsive on
the interval between the poles at \( x = \pm (M - \mu) \) and the edge of the eigenvalue distribution.
This force pushes eigenvalues toward \( x = \pm (M - \mu) \) forming the cusps.

We did not succeed in finding an analytic solution for the eigenvalue density in the strong-coupling phase, but as a first step we can study the critical behavior in the vicinity of the transition point, where the equations substantially simplify.

Away from the critical point, but still in the vicinity of the phase transition, the density should not be too different from \( \rho_c \). In the bulk of the eigenvalue distribution, the difference is linear in \( \lambda_c - \lambda \sim |\Delta|^{2/3} \). The deviation must be parametrically larger near the endpoints, to accommodate the change in the endpoint exponents. The characteristic scale in the vicinity of the endpoints is \( |\Delta| \), making
\[
    u = \frac{M}{2} - x \left| \frac{x - y}{\Delta} \right|^{3/2}
\]
the appropriate scaling variable. We will study the density in the limit of \( \Delta \to 0 \) with \( u \) finite. It is reasonable to expect that the density assumes a universal shape in this scaling regime.

We introduce two scaling functions, \( f_+(u) \) and \( f_-(u) \), which describe the endpoint behavior
of the density respectively above and below the phase transition, such that near \( x = \mu \) the density assumes the following form:
\[
    \rho(x) = C\rho M^{-\frac{2}{3}}|\Delta|^{-\frac{2}{3}} f_\pm \left( \frac{M - x}{\frac{2}{|\Delta|}} \right), \quad \pm = \text{sign } \Delta.
\]

The constant \( C\rho \) is defined in (3.72) and is introduced here to match with the critical solution (3.71). The scaling functions \( f_\pm(u) \) are defined on the semi-infinite intervals \( (\pm 1, \infty) \), and should not depend on any parameters. Away from the endpoints the density asymptotes to the critical solution. Comparing (3.74) with (3.71), we find that at large \( u \):
\[
    f_\pm(u) \sim u^{-\frac{2}{3}} \quad (u \to \infty).
\]

Both functions have an inverse square-root singularity at the endpoints of the intervals on which they are defined: \( f_\pm \sim 1/\sqrt{u} \pm 1 \). The function \( f_+ \), in addition, is expected to have a cusp at \( u = 1 \).

We first compute the scaling function \( f_-(u) \), which can be extracted from the exact solution (3.56) in the weak-coupling phase. To this end, we define the scaling resolvent:
\[
    g(w) = \frac{\sqrt{3}}{2\pi} \int_{-1}^{\infty} dv \frac{f_-(v)}{v - w}.
\]

The overall numerical factor is introduced for future convenience. The resolvent has a cut from 1 to \( \infty \) along the real axis, with discontinuity equal to the scaling function:
\[
    g(u \pm i0) = -\frac{r(u)}{2} \pm 2^{-\frac{2}{3}} \sqrt{3} i f_-(u) \quad (u \in \mathbb{R}, u > 1).
\]
On the other hand, we can substitute the scaling form of the density \((3.74)\) into the definition of the full resolvent \((3.54)\) to find the relationship between the latter and its scaling form:

\[
G(M - |\Delta|w) = \left(\frac{2}{3|\eta|\Delta}\right)^{\frac{2}{3}} g(w) = \frac{1}{\xi} g(w), \quad (3.78)
\]

where we assume that \(w \sim 1\) at \(\Delta \to 0\), and used \((3.72)\) and \((3.67)\) to express the coefficient of proportionality in terms of \(\xi\), one of the constants that defines the exact solution \((3.56)\).

In the scaling limit the constant \(\xi\) goes to zero, while two other constants, \(\eta\) and \(\bar{\eta}\), stay finite. We see that the product \(\xi G\) also remains finite, which means that \(\eta G, \bar{\eta} G \gg 1\). In this approximation, \((3.56)\) can be explicitly integrated:

\[
3|\eta||\Delta|w = (2\xi G - 1) \sqrt{\frac{\xi G + 1}{G^3}}. \quad (3.79)
\]

Using \((3.78)\) and once again \((3.67)\), we arrive at the cubic equation for \(g(w)\):

\[
4\left(w^2 - 1\right) g^3 + 3g = 1. \quad (3.80)
\]

This equation is universal, it does not depend on any parameters. Setting \(w = u \pm i0\), and substituting \((3.77)\) in \((3.80)\) we find

\[
f_-(u) = 2^{\frac{2}{3}} \sqrt{\frac{1}{u^2 - 1} + r^2}
\]

\[
4\left(u^2 - 1\right) r^3 + 3r = 1. \quad (3.81)
\]

It is easy to see that the scaling function has the correct end-point behavior:

\[
f_-(u) \simeq \frac{2^{\frac{1}{6}}}{\sqrt{u} - 1} \quad (u \to 1^+). \quad (3.82)
\]

As expected, \(f_-(u)\) asymptotes to the critical solution \(u^{-2/3}\) at infinity:

\[
f_-(u) \simeq u^{-\frac{2}{3}} \quad (u \to \infty). \quad (3.83)
\]

Interestingly, a closed equation for \(f_+(u)\) can be obtained by taking the scaling limit of \((3.47)\). This is possible because the density has an inverse square root singularity at the endpoints and, as soon as \(x\) is close enough to \(\mu\), the contribution from the bulk of the eigenvalue distribution scales away in the \(\Delta \to 0\) limit. Substituting the scaling form of the density \((3.74)\) into the integral equation, we obtain:

\[
\int_{\pm1}^{\infty} dv f_+(v) \left(\frac{2}{v - u} + \frac{1}{v + u}\right) = 0, \quad (3.84)
\]

which has to be solved with the boundary condition \((3.75)\).
The resulting equation is reminiscent of the saddle-point equations in the exactly solvable \( O(n) \) matrix model \([32, 33, 34, 35]\) with \( n = 1 \). The \( O(1) \) model reduces to solving a cubic equation \([33]\), quite similar to the formula (3.81) for \( f_- \) that we obtained by indirect arguments. Neither these arguments, nor, seemingly, a more direct approach of \([33, 34]\) can be used to solve for \( f_+ \), because of the cusp singularity in the middle of the eigenvalue distribution. It is nevertheless possible to understand the analytic structure of the cusp without finding an exact solution.

The integral equation for \( f_+(u) \) can be simplified by a substitution\(^9\)

\[
 f_+(u) = \begin{cases} 
 4 \frac{h(u)}{3} & 1 > u > -1 \\
 h(u) & u > 1.
 \end{cases}
\] (3.85)

If the function \( h(u) \) satisfies

\[
 2 \int_{-1}^{\infty} dv \frac{h(v)}{u-v} = \int_{1}^{\infty} dv \frac{h(v)}{u+v},
\] (3.86)

the function \( f_+(u) \), constructed via (3.85), can be checked to satisfy the integral equation (3.84).

The right-hand side of the integral equation (3.86) is a continuous function on the whole interval \((-1, \infty)\). Its inverse Hilbert transform is consequently also continuous. Therefore, \( h(u) \) has no singularities anywhere between \(-1\) and \( \infty \). This becomes clear once the Hilbert kernel in (3.86) is inverted:

\[
 h(u) = \frac{1}{2\pi \sqrt{u+1}} \int_{1}^{\infty} dv \frac{\sqrt{v-1} h(v)}{u+v}.
\] (3.87)

The correct boundary conditions at \( u = -1 \) and \( u = \infty \) also become manifest. Indeed, \( h(u) \sim 1/\sqrt{u+1} \) at \( u \to -1 \). Assuming that \( h(u) \sim \text{const} \cdot u^{-\alpha} \) at \( u \to \infty \), we find that \( 2 \sin \pi (\alpha - 1/2) = 1 \) and thus \( \alpha = 2/3 \). The unique solution with \( \text{const} = 1 \) can be obtained by iterations, which actually converge rather fast. Solving the equation numerically, we find that \( h(u) \) is a monotonously decreasing function and \( h(u) \sim 3C_+(u+1)^{-1/2} \) at \( u \to 1 \) with

\[
 C_+ = 0.650367 \ldots
\] (3.88)

Although \( h(u) \) is continuous on the interval \((-1, \infty)\), the change of variables (3.85) induces a cusp singularity in \( f_+(u) \) at \( u = 1 \). The structure of the cusp is easy to infer from (3.85): to the left of the cusp, \( f_+ \) has an inverse square root singularity, while from the right it approaches a finite limiting value given by

\[
 C_+ = h(1) = 0.531021 \ldots
\] (3.89)

The scaling function thus has the following singularity structure:

\[
 f_+(u) \sim \begin{cases} 
 2C_+ \sqrt{u+1} & (u \to -1^+) \\
 C_+ \sqrt{1-u} & (u \to 1^-) \\
 \tilde{C}_+ & (u \to 1^+).
 \end{cases}
\] (3.90)

The scaling functions \( f_+ \) and \( f_- \) are depicted in fig. \ref{fig:scaling_functions}.\(^9\)

---

\(^9\)Suggested to us by D. Volin.
Finite distance away from the phase transition we can solve the saddle-point equations numerically. The structure of the density remains qualitatively the same all the way up to the next critical point. The density has two cusps at \( \pm (M - \mu) \) of the following structure: the density is finite on the inside of the cusp, and diverges as an inverse square root on the outside (fig. 4). The cusp is in resonance with the opposite edge of the eigenvalue density (the distance corresponds exactly to the massless hypermultiplet). The repulsive force, that only acts on the eigenvalues between the cusp and the other edge of the eigenvalue density, causes the pileup on the outside of the cusp. Since the cusp is an image of the edge, the pileup is in general weaker than the edge singularity as can be seen from eq. (3.90), or from fig. 4.

3.5.4 Remarks on thermodynamics.

An interesting question to address is the order of the phase transition. The thermodynamic singularity at the transition point can be conveniently characterized by the second moment of the eigenvalue density \( \langle x^2 \rangle \), which according to (3.7) plays the rôle of a heat capacity. In the weak-coupling phase \( \langle x^2 \rangle \) is given by eq. (3.62). Above the transition we computed it numerically. We plot these results in fig. 5. It is clear from the plot that the transition is rather smooth.

The log-log plot of the difference between the true \( \langle x^2 \rangle \) and \( \langle x^2 \rangle_{\text{weak}} \), the analytic continuation of the weak-coupling result past the transition point, is shown in fig. 6. As expected the difference grows as a power \(^{10}\) of the distance to the critical point: \( \langle x^2 \rangle - \langle x^2 \rangle_{\text{weak}} \sim (\lambda - \lambda_c)^\alpha \).

\(^{10}\)Our accuracy is insufficient to exclude possible logarithmic corrections to the power-law scaling.
Figure 4: The eigenvalue density in the strong-coupling phase. The plot shows the numerical solution of the saddle-point equations at $\mu = 0.7M$, $\lambda \approx 54$.

Within the numerical precision of our calculations, the scaling exponent is consistent with $\alpha = 3$, which indicates that the first two derivatives of $f'(\lambda) \sim \langle x^2 \rangle$ are continuous, while the third derivative experiences a finite jump:

$$\langle x^2 \rangle - \langle x^2 \rangle_{\text{weak}} \approx C_2 M^2 (\lambda - \lambda_c)^3 \quad (\lambda \to \lambda_c^+) .$$

For the constant $C_2$ we get an estimate $C_2 \approx 1.5 \cdot 10^{-5}$. We thus conclude that the transition is of the fourth order with zero critical exponents. The fourth derivative of the free energy experiences a finite jump at the transition point.

In this regard it is interesting to look at higher derivatives of the free energy in the weak-coupling phase. They can all be computed analytically. For convenience we consider $\beta \equiv 8\pi^2 / \lambda$ as an independent variable. The first derivatives of the free energy in $\beta$ follows from (3.7), (3.62):

$$\frac{\partial F}{\partial \beta} = \frac{M^2}{12} (E_2 - 1) ,$$

where the argument of $E_2$ is $-e^{-\beta}$. Using the standard Ramanujan identities for differentiation of Eisenstein series, we further find:

$$\frac{\partial^2 F}{\partial \beta^2} = \frac{M^2}{144} \left( E_4 - E_2^2 \right)$$

$$\frac{\partial^3 F}{\partial \beta^3} = \frac{M^2}{864} \left( E_2^3 - 3E_2E_4 + 2E_6 \right)$$

(3.93)
Figure 5: The second moment of the eigenvalue density as a function of the 't Hooft coupling for $M = 10$. Below the phase transition the free energy is calculated analytically from (3.62). The curve above the transition is obtained numerically. The analytic continuation of the weak-coupling result past $\lambda_c$ is shown as the thin purple line.

It then follows from (B.5) that

$$M^4 \frac{d^3 F}{d\beta^3} = 2\xi \eta \bar{\eta}$$  \hspace{1cm} (3.94)

The parameter $\xi$ vanishes at the critical point and, with it, the third derivative of the free energy. The critical point can thus be identified with the inflection point of $E_2(-e^{-\beta})$.

Another interesting quantity is the order parameter of the phase transition. Although no real order parameter can exist, as no symmetry is broken at the transition point, the transition is nevertheless caused by IR physics, and we can try to identify a parameter that controls the correlation length that diverges. The first guess is $1/\xi$, since $\xi$ vanishes at the critical point. Let us recall the definition of $\xi$:

$$\frac{1}{\xi} = G(0) = \left(\frac{1}{x^2 - \frac{M^2}{4}}\right).$$  \hspace{1cm} (3.95)

The hypermultiplet masses are proportional to $|x - y \pm M|$, and in this sense, $\xi$ is an average mass squared of the hypermultiplet, with averaging defined in a specific way. At the transition point, the first massless hypermultiplet appears in the spectrum and $1/\xi$ diverges. Consequently, $1/\sqrt{\xi}$ represents the correlation length and $\sqrt{\xi}$ can be indeed identified with the mass gap that closes at the point of phase transition. Fig. 7 shows $\sqrt{\xi}$ as a function of $\lambda$. As can be seen from (3.67), (3.68), $\sqrt{\xi}$ vanishes at the critical point with the critical exponent 1/2.

As the coupling $\lambda$ is increased, the system undergoes an infinite sequence of phase transitions occurring whenever $\mu$ crosses thresholds at $nM/2$, with integer $n$. All transitions are of a similar nature, with a pair of new cusps created at the boundary of the eigenvalue distribution.
Figure 6: The log-log plot of the deviation of the exact $\langle x^2 \rangle$ from the analytic continuation of the weak-coupling result. The dots are obtained by numerically solving the saddle-point equation. The solid line is a fit to $a(\lambda - \lambda_c)^3 + b(\lambda - \lambda_c)^4$. The best-fit values of the coefficients are $a \approx 1.5 \cdot 10^{-3}$, $b \approx 5 \cdot 10^{-5}$. $M$ is set to 10.

different phases are described in more detail in [13]. As the coupling grows the cusps proliferate, but become less and less pronounced. The enveloping curve of the limiting density at very strong coupling approaches the Wigner semicircular shape, reproducing the result (3.8), (3.20) of the strong-coupling analysis, albeit in a somewhat irregular fashion, since on top of the enveloping curve the density has a complicated non-analytic fine structure.

### 3.6 Arbitrary $\lambda$ and $M$

We have so far analyzed the saddle-point equations (3.5) near the edges of the phase diagram in fig. 2. Although solving for the density in the analytic form is a difficult problem, the solution can be constructed numerically with good accuracy for any given $M$ and $\lambda$. Let us describe qualitatively the behavior of the eigenvalue density across the whole phase diagram.

The eigenvalues tend to spread more and more with $M$ or $\lambda$ growing: $\partial \mu / \partial M > 0$ and $\partial \mu / \partial \lambda > 0$. The $\mu = M/2$ and $\mu = M$ contours in the $\lambda - M$ plane are displayed in fig. 8. These contours can be regarded as crossover lines, since at $M \to \infty$ they approach the critical values of the coupling $\lambda = \lambda_c^{(1)}$, $\lambda = \lambda_c^{(2)}$, at which the infinite-volume system undergoes phase transitions. The phase transitions disappear at finite $M$, since the IR effects responsible for the critical behavior are regulated by the finite volume of the four-sphere (recall that $M$ should be understood as $MR$). At very large but finite $M$, $\mu = M n/2$ are the crossovers lines, on which the fourth derivative of the free energy rapidly changes. Likewise, the cusps in the eigenvalue density become sharp peaks of finite height. These features quickly disappear away from the infinite-volume limit and are not really visible at moderate values of $M$.

The endpoint singularity of the eigenvalue density is of the usual square-root type across the whole phase diagram. At small $M$ the density has a maximum at zero and monotonically...
decreases towards the endpoints of the distribution. But with $M$ growing, two additional maxima develop near the endpoints. In some range of parameters (for sufficiently large $\lambda$ and not so big $M$), the density has three clearly visible peaks. The peak at zero diminishes in size and at some point disappears, while the peaks near the endpoints become more and more pronounced, and asymptotically form the inverse square-root spikes of the infinite-volume density, cf. (3.50). Qualitative explanation of this behavior, which is illustrated in fig. 9, is given in [11]. In the large-$M$, large-$\lambda$ corner of the phase diagram, the density has a more complicated shape with many minima and maxima, due to proximity of the phase transition points in infinite volume.

3.7 Instantons

Instanton contributions are usually believed to be negligible at large $N$ because of the exponential suppression of the instanton weight\footnote{Assuming the standard 't Hooft scaling of the gauge coupling. If the gauge coupling is kept fixed at large $N$, instantons are not suppressed \cite{36}.}.

\[
Z^\text{inst}_k \sim e^{-\frac{sN^2 k N}{\lambda}}.
\] (3.96)

This estimate does not take into account the instanton moduli integration, which can considerably modify the instanton weight and may even overcome the exponential suppression, leading to an instanton-induced large-$N$ phase transition \cite{14}. It is reasonable to assume that all instanton contributions are either suppressed or simultaneously blow up, independently of the instanton number. It is thus sufficient to investigate the moduli space integration for a single instanton with topological charge $k = 1$.\footnote{Assuming the standard 't Hooft scaling of the gauge coupling. If the gauge coupling is kept fixed at large $N$, instantons are not suppressed \cite{36}.}
Figure 8: The lines $\mu = M/2$ and $\mu = M$. The former approaches $M = \infty$ at $\lambda \approx 35.4$, the latter approaches $M = \infty$ at $\lambda \approx 83$, in consistency with the phase transitions found at these values of $\lambda$ for the decompactified ($MR = \infty$) theory.

The one-instanton contribution to the partition function of the $\mathcal{N} = 2^*$ theory is given by

$$Z_{\text{inst}}^1 = -e^{-8\pi^2 N \lambda + i\theta} M^2 \sum_{l=1}^{N} \prod_{j \neq l} \frac{(a_l - a_j + i)^2 - M^2}{(a_l - a_j)(a_l - a_j + 2i)},$$

(3.97)

It also admits an integral representation:

$$Z_{\text{inst}}^1 = e^{-8\pi^2 \lambda + i\theta} \frac{2M^2}{M^2 + 1} \int \frac{dz}{2\pi} \prod_{j=1}^{N} \frac{(z - a_j)^2 - M^2}{(z - a_j)^2 + 1},$$

(3.98)

where the contour of integration encircles the poles at $a_j + i$ counterclockwise.

At large $N$ the $z$ integral is of the saddle point type and, with exponential accuracy,

$$Z_{\text{inst}}^1 \sim e^{-8\pi^2 N \lambda} \int dz \ e^{NS_{\text{ms}}(z)} \sim e^{-NS_{\text{inst}}},$$

(3.99)

where the moduli space action is

$$S_{\text{ms}}(z) = \int_{-\mu}^{\mu} dx \rho(x) \ln \frac{(z - x)^2 - M^2}{(z - x)^2 + 1},$$

(3.100)

and the effective instanton action $S_{\text{inst}}$ is determined by the value of $S_{\text{ms}}(z)$ at the dominant saddle point:

$$S_{\text{inst}} = \frac{8\pi^2}{\lambda} - S_{\text{ms}}(z_\ast).$$

(3.101)

We thus need to classify all possible saddle points, which need not lie on the contour of integration. The simplest saddle point is $z_\ast = \infty$ where $S_{\text{ms}}(\infty) = 0$. If this saddle-point
saturates the integral, the instanton action is not modified, and the naive estimate (3.96) of the instanton weight holds true. This was found to happen in the conformal $\mathcal{N} = 2$ SYM with $2N$ fundamental hypermultiplets [5]. In the $\mathcal{N} = 2^*$ case, $z_\ast = \infty$ is the only saddle point as long as $\mu > M$. One can look, for instance, at the behavior of the integrand on the real axis. For $|z| > M + \mu$, the integrand is exponentially small, going asymptotically to one at $z \to \infty$. Between $-M - \mu$ and $M + \mu$, the integrand rapidly oscillates, since $S_{\text{ms}}(z + i0)$ has a non-zero imaginary part. The oscillatory behavior can be understood from (3.98) as well, where the numerator changes sign each time $z$ passes through $a_j \pm M$. We denote the critical value of the mass, at which $\mu = M$, by $M_{c_1}$.

If $\mu < M$, the integrand oscillates on two separate intervals $[-M - \mu, -M + \mu]$ and $[M - \mu, M + \mu]$. The gap between these intervals contains a new saddle-point at $z_\ast = 0$. The moduli space action at this saddle-point is

$$S_{\text{ms}}(0) = \int_{-\mu}^{\mu} dx \rho(x) \ln \frac{M^2 - x^2}{1 + x^2} + i\pi.$$  \hspace{1cm} (3.102)

Let us calculate it in the weak-coupling regime discussed in sec. 3.2. The eigenvalue density then is highly peaked at zero, and the integrand can be Taylor expanded in $x$:

$$S_{\text{ms}}(0) = \ln M^2 + i\pi - \frac{M^2 + 1}{M^2} \langle x^2 \rangle + \ldots = \ln M^2 + i\pi - \frac{M^2 + 1}{M^2} \frac{\lambda}{16\pi^2} + O(\lambda^2).$$  \hspace{1cm} (3.103)

The real part of the action can be positive or negative, depending on $M$ and $\lambda$. If the action is negative, the saddle point at infinity still gives the dominant contribution. When $\text{Re} S_{\text{ms}}(0)$
Figure 10: The phase structure of the instanton weight. The solid line is $M_{c_1}$ ($\mu = M$). The dashed line represents $M = M_{c_2}$ and meets with $M_{c_1}$ at $\lambda \approx 40.5$. In the shaded region the instanton action is just $8\pi^2/\lambda$.

changes sign the integral switches from the saddle point at $z = \infty$ to the saddle point at $z = 0$. We denote the critical value of the mass by $M_{c_2}$.

There are thus two critical lines on the $(\lambda, M)$ plane, $M_{c_1}(\lambda)$ and $M_{c_2}(\lambda)$, at which the instanton weight discontinuously changes its behavior. At weak coupling, $\mu = \sqrt{\lambda}/2\pi$. Hence,

$$M_{c_1} \simeq \frac{\sqrt{\lambda}}{2\pi} \quad (\lambda \to 0).$$

(3.104)

In the other limiting case, $M \to \infty$, condition $\mu = M$ determines the second critical point $\lambda_{c_2}^{(2)} \approx 83$ that we have found in the decompactification limit (see fig. 8). Consequently,

$$M_{c_1} \to \infty \quad \text{at} \quad \lambda \to \lambda_{c_2}^{(2)}.$$  

(3.105)

From (3.103) we find:

$$M_{c_2} \simeq 1 + \frac{\lambda}{4\pi^2} \quad (\lambda \to 0).$$

(3.106)

The two critical lines are shown in fig. 10. In the shaded region on the plot the instanton action is not renormalized and the naive estimate of the instanton weight is quantitatively correct.

We thus find:

$$S_{\text{inst}} = \frac{8\pi^2}{\lambda} \quad \text{if} \quad \lambda > \lambda_{c_2}^{(2)} \quad \text{or} \quad M < \max(M_{c_1}, M_{c_2}).$$

(3.107)

Otherwise:

$$S_{\text{inst}} = \frac{8\pi^2}{\lambda} - \int_{-\mu}^{\mu} dx \rho(x) \ln \frac{M^2 - x^2}{1 + x^2}. \quad (3.108)$$

The imaginary part of $S_{\text{ms}}(0)$ in the latter case leads to renormalization of the theta-angle: $\theta \to \theta + i\pi N$. 

30
Let us now examine $S_{\text{inst}}$ in the region where the moduli space corrections are non-trivial. We first consider the limits where we can compute the instanton action analytically. At weak coupling,

$$S_{\text{inst}} \approx \frac{8\pi^2}{\lambda} - 2 \ln M \quad (\lambda \to 0).$$

(3.109)

Not surprisingly, at large $M$ and small $\lambda$ the moduli space corrections renormalize the coupling, combining into $8\pi^2/\lambda_R$, where $\lambda_R$ is the running coupling of the pure $\mathcal{N} = 2$ SYM. Potentially, $\lambda_R$ can be rather big or even negative for sufficiently large $M$. However, the approximation (3.103) used in deriving this result is valid only when the renormalized coupling is small. Otherwise the eigenvalue density is no longer peaked at zero and Taylor expansion in $x$, used in deriving this result, is no longer accurate. The analysis for arbitrary $\lambda_R$ \cite{11} indicates that the instanton action always remains positive, even when $\ln M$ is larger than $4\pi^2/\lambda$.

We can also calculate the instanton action in the decompactification limit. Using (3.48) we find:

$$S_{\text{inst}} = \int_{-\mu}^{\mu} dx \rho(x) \ln \left(1 + \frac{1}{x^2}\right) \quad (M \to \infty).$$

(3.110)

This expression is manifestly positive. Since $\mu \propto M \to \infty$ in the decompactification limit, and the integrand is highly peaked near $x = 0$, we can compute the integral as

$$S_{\text{inst}} = \rho(0) \int_{-\infty}^{\infty} dx \ln \left(1 + \frac{1}{x^2}\right) = 2\pi \rho(0) \quad (M \to \infty).$$

(3.111)

The density at zero can be calculated from (3.55). Taking into account that \cite{13}

$$G\left(\frac{M}{2} - i0\right) = \frac{1}{\eta}, \quad G\left(\frac{M}{2} + i0\right) = \frac{1}{\bar{\eta}},$$

(3.112)

we get from (3.55), (3.57):

$$S_{\text{inst}} = \frac{iM (\eta - \bar{\eta})}{\eta \bar{\eta}} = \frac{36i (\theta_4^1 - \theta_3^4)}{M (E_2 - 2\theta_4^1 + \theta_3^3) (E_2 - 2\theta_4^1 + \theta_3^4)} \quad (M \to \infty),$$

(3.113)

where the argument of the Eisenstein series is $-r^2$ and the modulus parameter of the theta-constants is $ir$ with $r = \exp(-4\pi^2/\lambda)$.

We verified numerically that the instanton action is positive definite throughout the whole phase diagram. We thus conclude that instantons are always exponentially suppressed in the large-$N$ $\mathcal{N} = 2^*$ SYM on the four-sphere.

4 Massive deformations of superconformal QCD

Another way to make an $\mathcal{N} = 2$ theory UV finite is to couple a vector multiplet to $2N$ hypermultiplets in the fundamental representation. When the hypermultiplets are massless this theory
is superconformal and will be referred to as SCFT. The large-$N$ limit of its partition function on $S^4$ was analyzed in [8]. Here we study super-QCD-type theories obtained by relevant perturbations of $\mathcal{N} = 2$ SCFT by various assignments of hypermultiplet masses.

The localization partition function for $\mathcal{N} = 2$ super-QCD with an arbitrary mass assignment is given by

$$Z_{\text{SQCD}} = \int d^{N-1} a \prod_{i<j} (a_i - a_j)^2 H^2(a_i - a_j) e^{-\frac{8\pi^2}{x} \sum a_i^2} \left| Z_{\text{inst}} \right|^2.$$  \hspace{1cm} (4.1)

The flavor index $f$ runs from 1 to $N$, making the theory finite in the UV.

At large $N$, the integral is dominated by a saddle-point, determined by the equation

$$\int d\mu \rho(y) \left(\frac{1}{x-y} - \mathcal{K}(x-y)\right) = \frac{8\pi^2}{\lambda} x - \frac{1}{2N} \sum_{f=1}^{N} \left( \mathcal{K}(x + M_f) + \mathcal{K}(x - M_f) \right).$$  \hspace{1cm} (4.2)

The qualitative structure of the solution depends on the assumptions made about the hypermultiplet masses. We mostly concentrate on two representative cases: (i) All masses equal, $M_f = M$. For short, we refer to this theory as $\mathcal{N} = 2$ SCFT*. (ii) Another example that we consider in detail is partially massless theory with $M_f = M$ for $f = 1, \ldots, N - N_0$ and $M_f = 0$ for the remaining $N_0$ flavors. Later we will also consider $\mathcal{N} = 2$ SQCD with $N_f < N$ flavors, which can be obtained from the UV finite theory by making $N - N_f$ masses infinitely heavy: $M_f = \Lambda_{\text{cutoff}} \rightarrow \infty$ for $f > N_f$. For simplicity we will assume that all the remaining quark masses are equal: $M_f = M$ for $f = 1, \ldots, N_f$.

Like $\mathcal{N} = 2^*$ theory, $\mathcal{N} = 2$ SQCD can be viewed as a UV completion of pure SYM theory. The latter is obtained by making all hypermultiplets infinitely heavy, $M_f \rightarrow \infty$, and simultaneously sending $\lambda$ to zero at fixed $\Lambda = e^{-\frac{4\pi^2}{x} \prod_{f=1}^{N} M_f^2}$.  \hspace{1cm} (4.3)

By expanding $\mathcal{K}(x \pm M_f)$ in (4.2) to the linear order in $x \ll M_f$, we obtain:

$$\int d\mu \rho(y) \left(\frac{1}{x-y} - \mathcal{K}(x-y)\right) = \frac{8\pi^2}{\lambda_r} x,$$  \hspace{1cm} (4.4)

with the renormalized coupling defined as

$$\frac{4\pi^2}{\lambda_r} \equiv \frac{4\pi^2}{\lambda} - \ln \Lambda - 1 - \gamma.$$  \hspace{1cm} (4.5)

This equation describes the large $N$ limit of pure $N = 2$ Yang-Mills theory, and was studied in detail in [11].

The weak-coupling limit at fixed $M_f$ can be analyzed along the same lines as in sec. 3.2 with quite similar results. We will not repeat these calculations here. The massless limit of superconformal theory was studied in [8]. Here we concentrate on two other possible limits, the strong-coupling regime and the decompactification limit. The qualitative behavior of $\mathcal{N} = 2$ SQCD in these regimes is quite different from the $\mathcal{N} = 2^*$ case.

32
Figure 11: Eigenvalue density obtained analytically and numerically for $\mathcal{N} = 2$ SYM with $2N$ fundamental hypermultiplets of equal mass at large $\lambda$. For clarity, the numerical eigenvalue density, which terminates at a finite value of $x$, is shown only at $x > 0$.

### 4.1 Strong coupling

At strong coupling, $\lambda \gg 1$, the eigenvalue density extends to a large interval $(-\mu, \mu)$ which grows logarithmically with $\lambda$. In contradistinction to the $\mathcal{N} = 2^*$ case, the eigenvalue density has a well-defined limiting shape $\rho_\infty(x)$, and one can just set $\lambda = \infty$ and $\mu = \infty$ in the saddle-point equation, to a first approximation, much like in the massless case $[8]$. The saddle-point equation is then solved by Fourier transform:

$$i\pi \text{sign}\omega \left(1 + \frac{1}{2\sinh^2\frac{\omega}{2}}\right)\tilde{\rho}_\infty(\omega) = i\pi \text{sign}\omega \frac{1}{N} \sum_{f=1}^{N} \cos M_f\omega.$$

Hence

$$\tilde{\rho}_\infty(\omega) = \frac{1}{\cosh \omega} \frac{1}{N} \sum_{f=1}^{N} \cos M_f\omega.$$

and

$$\rho_\infty(x) = \frac{1}{4N} \sum_{f=1}^{N} \left[\frac{1}{\cosh \left(\frac{\pi}{2}(x + M_f)\right)} + \frac{1}{\cosh \left(\frac{\pi}{2}(x - M_f)\right)}\right].$$

The eigenvalue density in the $\mathcal{N} = 2$ SCFT$^*$ (all masses equal) has two hills peaked at $x \sim \pm M$ and decays exponentially at $\pm$ infinity. Figure 11 compares this analytic result with the numerical solution at $\lambda = 2000$, $M = 3$. For the partially massless theory, the density has three peaks, at $x \sim 0$ and $x \sim \pm M$.

The asymptotic density $[4.8]$ has two exponential tails that extend all the way to infinity. This is not so if $\lambda$ is large but finite. The density then has to terminate at some $\mu$. This is clearly visible in the numerical solution. By matching the endpoint behavior of the density to the asymptotic solution at infinite $\lambda$, it is possible to estimate the endpoint position and the
Wilson loop vev \([8]\). The leading order is \(M\)-independent:

\[
\mu \approx \frac{2}{\pi} \ln \lambda \quad (\lambda \to \infty).
\] (4.9)

The Wilson loop vev receives the biggest contribution from the vicinity of the endpoint, and is estimated as

\[
W(C) \approx \text{const} \frac{\lambda^3}{(\ln \lambda)^{3/2}} \quad (\lambda \to \infty).
\] (4.10)

Again the masses do not affect the leading-order behavior, but the constant of proportionality will depend on \(M_f\). It is in principle calculable by the Wiener-Hopf method \([8]\), but we will not attempt this calculation here.

The free energy has a finite strong-coupling limit, and can be calculated from the asymptotic solution (4.8). We have

\[
\frac{\partial F}{\partial M_f} = \frac{1}{N} \left( \mathcal{K}(x - M_f) - \mathcal{K}(x + M_f) \right)
\] (4.11)

\[
\frac{\partial F}{\partial \lambda} = -\frac{8\pi^2}{\lambda^2} \left( x^2 \right).
\] (4.12)

Using the eigenvalue density (4.8) we find

\[
\langle x^2 \rangle_\infty = 1 + \frac{1}{N} \sum_f M_f^2
\] (4.13)

and

\[
\langle \mathcal{K}(x + m) \rangle = -\frac{1}{2N} \sum_f \frac{\partial}{\partial m} \left( \ln \mathcal{F}(m + M_f) + \ln \mathcal{F}(m - M_f) \right),
\] (4.14)

where the function \(\mathcal{F}(x)\) is defined as

\[
\mathcal{F}(x) = H(x) \frac{\Gamma \left( \frac{1+ix}{4} \right) \Gamma \left( \frac{1-ix}{4} \right)}{\Gamma \left( \frac{3+ix}{4} \right) \Gamma \left( \frac{3-ix}{4} \right)}
\] (4.15)

Hence

\[
N^2 F^{\text{SQCD}} = \ln \left( \prod_{f} \mathcal{F}(M_f + M_f') \mathcal{F}(M_f - M_f') \right) + \frac{8\pi^2 N}{\lambda} \sum_f \left( 1 + M_f^2 \right) + O \left( \frac{1}{\lambda^2} \right).
\] (4.16)

For \(\mathcal{N} = 2\) SCFT\(^*\) this gives, up to a constant:

\[
F^{\text{SCFT}\,*} = \ln \mathcal{F}(2M) + \frac{8\pi^2(1 + M^2)}{\lambda} + O \left( \frac{1}{\lambda^2} \right).
\] (4.17)
4.2 Decompactification

The decompactification limit can be analyzed much in the same way as in sec. 3.5. First we recover the dependence on $R$ by rescaling all dimensionful quantities and then send $R$ to infinity. Once the resulting equation is differentiated twice, the kernel becomes algebraic, because $\mathcal{K}(x)$ can be replaced by its large-argument asymptotics (A.5). We will analyze separately two special cases, the $\mathcal{N} = 2$ SCFT* with equal hypermultiplet masses and partially massless theory.

4.2.1 SCFT*

In the $\mathcal{N} = 2$ SCFT* case, the steps described above lead to the following simple equation:

$$2 \int_{-\mu}^{\mu} dy \frac{\rho(y)}{x - y} = \frac{1}{M + x} - \frac{1}{M - x}.$$  \hspace{1cm} (4.18)

This looks like the saddle-point equation for a one-matrix model with a logarithmic potential. Slightly more general matrix model with an additional quartic potential was considered in [37], as a model for open strings in zero dimensions. The model has a rich phase structure and exhibits quite non-trivial critical behavior, governed by an interplay between the logarithmic and polynomial terms in the potential. More general mass assignment in SQCD can probably mimic additional terms in the effective matrix-model potential and thus can lead to an interesting critical behavior.

In our case, the effective potential is actually upside-down, which should not worry us too much, as the boundary conditions here are different compared to usual matrix models. In the matrix model language, we need to find the solution squeezed between two infinite walls at $x = \pm \mu$. The unique normalizable solution with such boundary conditions exists for any $\mu$. In contradistinction to the usual matrix models, normalization does not fix the endpoint positions $\pm \mu$, which are rather determined by the integrated form of (4.18), equivalent to the original saddle-point equation after the first differentiation:

$$\int_{-\mu}^{\mu} dy \rho(y) \ln \frac{y^2}{M^2} = -\frac{8\pi^2}{\lambda}.$$  \hspace{1cm} (4.19)

The unique normalizable solution of (4.18), at fixed $\mu$, is given by

$$\rho(x) = \frac{M\sqrt{M^2 - \mu^2}}{\pi} \frac{1}{\sqrt{\mu^2 - x^2}} \frac{1}{M^2 - x^2}.$$  \hspace{1cm} (4.20)

In order to find $\mu$ as a function of $\lambda$ and $M$, we substitute $\rho$ into (4.19). Using

$$\int \ln \frac{x^2}{M^2} = 2\ln \frac{\mu}{M + \sqrt{M^2 - \mu^2}},$$  \hspace{1cm} (4.21)

12 We would like to thank V. Kazakov for comments on this point.
we obtain
\[ \mu = \frac{M}{\cosh \frac{4\pi^2}{\lambda}}. \] (4.22)

Note that \( \mu \) never exceeds \( M \). As a result there are no phase transitions in this model. The reason can be easily understood from the saddle-point equation (4.18). The effective potential in the analog matrix model is unbounded from below and as soon as \( \mu \) approaches \( M \) the eigenvalues start to fall down the infinite potential well. The attractive force acting towards \( x = \pm M \) becomes stronger and stronger when \( \mu \) approaches \( M \) and can overcome mutual repulsion between the eigenvalues, compressing larger and larger number of them towards the endpoint of the distribution. A natural question is how to reconcile this behavior with the strong-coupling behavior studied in the previous section. In the latter case, \( \mu \) grows like \( \ln \lambda \) and can certainly exceed \( M \). However, the eigenvalues sitting at \( x > M \) represent the exponential tail of the distribution which vanishes as \( M \to \infty \). This tail is automatically cutoff when the limit \( M = \infty \) is taken before considering \( \lambda \to \infty \), and in this case the eigenvalue distribution always ends at \( \mu < M \). In fact, in the limit where both \( M \to \infty \) and \( \lambda \to \infty \), taken in any order, the eigenvalue density approaches two delta functions peaked at \( x = \pm M \).

The free energy can be found from (4.12). For the second moment of the eigenvalue density we get:
\[ \langle x^2 \rangle = M^2 - M\sqrt{M^2 - \mu^2} = M^2 \left(1 - \tanh \frac{4\pi^2}{\lambda}\right), \] (4.23)
which gives:
\[ F = -2M^2 \ln \left(1 + e^{-\frac{\pi^2}{\lambda}}\right) + \text{const}. \] (4.24)

Strikingly, the free energy of the \( \mathcal{N} = 2 \) SCFT* is given by the first \( n = 1 \) term of the free energy (3.63) of \( \mathcal{N} = 2^* \) SYM.

The weak-coupling expansion of (4.22), (4.24) has the expected OPE form (3.2). For instance,
\[ F = 2M^2 \sum_{k=1}^{\infty} \frac{(-1)^k}{k} e^{-\frac{\pi^2}{\lambda}k}. \] (4.25)

The simplicity of the expansion coefficients again suggests that there may be a more direct way to calculate them, without the use of localization.

Computing the circular Wilson loop, we find:
\[ W(C) = \langle e^{2\pi R x} \rangle \approx \frac{1}{2\pi\sqrt{MR}} \frac{\cosh \frac{3\pi^2}{2\lambda}}{\sinh \frac{4\pi^2}{\lambda}} \exp \left(\frac{2\pi MR}{\cosh \frac{4\pi^2}{\lambda}}\right). \] (4.26)

Since \( R \to \infty \), the main contribution comes from the region near \( x \approx \mu \). We thus conclude that large Wilson loops obey perimeter law with the coefficient given by \( \mu \) in (4.22). At strong coupling, the coefficient just asymptotes to \( M \). The prefactor, as a function of the coupling constant, grows linearly at large \( \lambda \). This is different from the result (4.10), found by taking the limit \( \lambda \to \infty \) first. The discrepancy is not surprising, since the Wilson loop is sensitive to the exponential tail of the eigenvalue density, which is cut off in the \( MR = \infty \) limit that we are considering now.
An interesting feature of the Wilson loop vev (4.26) is that the exponent does not have a \( \sqrt{\lambda} \) coefficient as one would expect from a string world-sheet interpretation. Recall that, for \( \mathcal{N} = 4 \) SYM, the coefficient \( \sqrt{\lambda} \) in the exponent arises from the squared radius of AdS space in units of \( \alpha' \). Likewise, perimeter law in the \( \mathcal{N} = 2^* \) SYM at strong coupling bears a factor of \( \sqrt{\lambda} \), with the coefficient in exact agreement with the area law in the geometry of the holographic dual [12]. Strings in the supergravity dual of massive \( \mathcal{N} = 2 \) SCFT*, which is not known, are likely to have quantum and highly interacting worldsheet even in the limit of large \(^t\)t Hooft coupling.

### 4.2.2 Partially massless theory

If \( N_0 \) flavors are left massless, double differentiation of the saddle-point equation leads to

\[
2 \int_{-\mu}^{\mu} \frac{dy}{x-y} \rho(y) = \frac{\nu}{M+x} - \frac{\nu}{M-x} + \frac{2(1-\nu)}{x},
\]  

(4.27)

where \( \nu \) denotes the fraction of massive flavors:

\[
\nu = \frac{N - N_0}{N}.
\]  

(4.28)

In the large-\( N \) limit, \( \nu \) is a real number between zero and one. The force term on the right-hand side now has a singularity on the interval \((-\mu, \mu)\). The inversion of the Hilbert kernel thus becomes ambiguous, and depends on how the singularity is regularized. The limiting procedure \((R \to \infty)\), that was used in deriving (4.27), actually dictates a very concrete regularization prescription. Indeed, the \( 1/x \) driving term on the right-hand side arises from approximating \( K(xR) \) by \( xR \log|xR| \) in the limit \( R \to \infty \), eq. (A.5). This approximation is clearly inapplicable when \( x \to 0 \). In fact, before the limit was taken, the original function \( K(xR) \) had been non-singular at zero. Consequently, \( 1/x \) is smoothened out on the scale \( \epsilon \sim 1/R \). Since \( K(x) \) is an odd function, the smoothened \( 1/x \) will automatically retain anti-symmetry under \( x \to -x \). Such regularization is equivalent to the principal-value prescription

\[
\frac{2}{x} \to \frac{1}{x + i\epsilon} + \frac{1}{x - i\epsilon}.
\]

If \( 1/x \) is understood in this way, the normalizable solution to (4.27) is given by

\[
\rho(y) = \frac{\nu M \sqrt{M^2 - \mu^2}}{\mu} \frac{1}{\sqrt{\mu^2 - x^2}} \frac{1}{M^2 - x^2} + (1-\nu)\delta(x).
\]  

(4.29)

The analog of (4.19) now reads\(^{13}\)

\[
\int_{-\mu}^{\mu} dy \rho(y) \ln (x-y)^2 - \nu \ln (M^2 - x^2) - (1-\nu) \ln x^2 = -\frac{8\pi^2}{\lambda}.
\]  

(4.30)

\(^{13}\)Eq. (4.27) guarantees that the right-hand-side is constant independent of \( x \). In (4.19) we have thus set \( x = 0 \) without losing any information. Here we cannot set \( x = 0 \) directly, because of the log singularity, and thus prefer to keep \( x \) as a parameter.
Evaluating the left-hand side on the solution \((4.29)\), we find:

\[
\mu = \frac{M}{\cosh \frac{8\pi^2}{3\nu}}.
\]  

(4.31)

The only effect of the massless multiplets, compared to \(\mathcal{N} = 2\) SCFT*, is a rescaling of \(\lambda\). In particular, \(\mu\) still obeys the bound \(\mu < M\). The Wilson loop is given by the same formula as in the previous section with \(\lambda\) rescaled. Likewise, for the free energy we get:

\[
F = -2M^2 \nu^2 \ln \left(1 + e^{-\frac{8\pi^2}{3\nu}}\right) + \text{const.}
\]  

(4.32)

The OPE expansion now goes in powers of \(\exp(-8\pi^2/\lambda\nu)\). This is easy to understand. The low-energy sector of the model, left upon integrating out massive fields, is \(\mathcal{N} = 2\) SQCD with \(2N_0\) massless hypermultiplets. The beta functions of this theory is

\[
\beta_{\text{SQCD}} = -\frac{g_{\text{YM}}^4 N(N - N_0)}{4\pi^2} = -\frac{\lambda^2 \nu}{4\pi^2},
\]  

(4.33)

and the dynamically generated scale is given by

\[
\Lambda_{\text{SQCD}} = M e^{-\frac{8\pi^2}{3\nu}}.
\]  

(4.34)

The OPE goes in powers of \(\Lambda_{\text{SQCD}}^2/M^2\), which now translates into \(\exp(-8\pi^2/\lambda\nu)\).

### 4.3 Instantons

The one-instanton contribution to the partition function of the \(\mathcal{N} = 2\) \(SU(N)\) super Yang-Mills theory with \(2N\) hypermultiplets can be obtained from the general formulas given in \[17\]:

\[
Z^\text{inst}_1 = e^{-\frac{2\pi^2 N}{\lambda} \sum_{i=1}^{N} \sum_{j \neq l} \frac{\Pi_j (a_l + i)^2 - M_j^2}{(a_l - a_j)(a_l - a_j + 2i)}},
\]  

(4.35)

and has an integral representation:

\[
Z^\text{inst}_1 = 2 e^{-\frac{2\pi^2 N}{\lambda} \sum_{i=1}^{N} \sum_{j \neq l} \frac{\Pi_j (z^2 - M_j^2)}{(z - a_j)^2 + 1}},
\]  

(4.36)

where the contour of integration encircles the poles at \(a_j + i\) counterclockwise.

The large-\(N\) limit of the instanton contribution was examined in \[8\] for \(M_j = 0\). It was found that the exponential part of the instanton weight is not modified by the moduli integration. We extend this analysis to the massive SCFT* with equal hypermultiplet masses.
The moduli-space action,

\[ S_{ms}(z) = \ln(z^2 - M^2) - \int_{-\mu}^{+\mu} dx \rho(x) \ln((z-x)^2 + 1), \]  

(4.37)

has two saddle-points at \( z = 0 \) and \( z = \infty \). Since \( S_{ms}(\infty) = 0 \), the instanton action is not renormalized if the dominant saddle-point is the one at infinity. At zero,

\[ S_{ms}(0) = \ln \frac{M^2}{x^2 + 1} - i\pi. \]  

(4.38)

For the asymptotic solution at strong coupling, eq. (4.8), the real part of the saddle-point action is always negative:

\[ \left\langle \ln \frac{M^2}{x^2 + 1} \right\rangle_{\infty} = 2 \ln \tanh \frac{\pi M}{4} < 0. \]  

(4.39)

We thus conclude that at strong coupling the trivial saddle-point at infinity is dominant. In the decompactification limit,

\[ \left\langle \ln \frac{M^2}{x^2 + 1} \right\rangle = \left\langle \ln \frac{x^2}{x^2 + 1} \right\rangle + \frac{8\pi^2}{\lambda} \quad (M \to \infty), \]  

(4.40)

in virtue of (4.19). This expression is positive, so the saddle-point at \( z = 0 \) is dominant. The same steps that led to (3.111), now give for the instanton action

\[ S_{\text{inst}} = 2\pi \rho(0) = \frac{2}{M} \sinh \frac{4\pi^2}{\lambda}. \]  

(4.41)

There is a line \( M_c(\lambda) \) in the \((\lambda, M)\) plane below which the suppression is determined solely by the instanton action \( 8\pi^2/\lambda \). In the region above \( M_c \) the instanton action is renormalized by the moduli-space integration. As in sec. 3.7.1, \( M_c(0) = 1 \). At \( \lambda \to \infty \), \( M_c \) goes to infinity. We have computed \( S_{\text{inst}} \) numerically in the region above \( M_c \) and found that it is always positive definite. In conclusion, the one-instanton contribution is exponentially suppressed in the large-\( N \) limit.

## 5 Super-QCD with \( 2N_f \) massive hypermultiplets

Another interesting theory is \( \mathcal{N} = 2 \) SQCD, \( \mathcal{N} = 2 \) supersymmetric gauge theory with \( 2N_f \) massive hypermultiplets of equal mass \( M \). We would like to study this theory in the Veneziano limit, \( N \to \infty, N_f \to \infty \) with \( N_f/N \) fixed \cite{15}. We shall assume that \( N_f < N \), in which case the theory is asymptotically free, and interpolates between pure \( \mathcal{N} = 2 \) SYM at \( N_f = 0 \) and superconformal YM at \( N_f = N \).

A neat way to define the partition function of \( \mathcal{N} = 2 \) SQCD is to start with a UV finite partition function \((4.1)\) and make \( N - N_f \) quarks infinitely heavy, while keeping the mass of the

\[ M_c \] of the SCFT\(^*\) is analogous to \( M_{c_2} \) of \( \mathcal{N} = 2^* \) SYM.
remaining $N_f$ quarks fixed. The mass of the heavy flavors, which we denote by $\Lambda_0$, serves as a UV cutoff. Using (A.4),

$$\ln H(a + \Lambda_0) + \ln H(a - \Lambda_0) \simeq 2 \ln H(\Lambda_0) - 2 (\ln \Lambda_0 + \gamma + 1) a^2$$

the contribution of the heavy fields can be absorbed in the renormalization of the 't Hooft coupling:

$$\frac{4\pi^2 N}{\lambda_R} = \frac{4\pi^2 N}{\lambda} - (N - N_f) (\ln \Lambda_0 + \gamma + 1).$$

(5.1)

Introducing the Veneziano parameter

$$\zeta = \frac{N_f}{N},$$

(5.2)

and the dynamically generated scale

$$\Lambda = \Lambda_0 e^{-\frac{4\pi^2}{\lambda_R (1 - \zeta)}},$$

(5.3)

the partition function can be written as

$$Z_{N_f}^{\text{SQCD}} = \int d^{N-1} a \frac{\prod_{i<j} (a_i - a_j)^2 H^2(a_i - a_j)}{\prod_i H^{N_f}(a_i + M) H^{N_f}(a_i - M)} e^{2(N-N_f)(\ln \Lambda + \gamma + 1) \sum_i a_i^2} |Z_{\text{inst}}|^2.$$  

(5.4)

The saddle-point equation takes the following form:

$$2 \int_{-\mu}^{\mu} dy \rho(y) \left( \frac{1}{x - y} - \mathcal{K}(x - y) \right) = -4 (1 - \zeta) (\ln \Lambda + \gamma + 1) x - \zeta \mathcal{K}(x + M) - \zeta \mathcal{K}(x - M).$$

(5.5)

The model depends on two parameters $(\Lambda, M)$, and we may consider several limiting cases in which the saddle-point equation simplifies and can be explicitly solved. For instance, if $M \gg \Lambda$, the hypermultiplets can be integrated out leaving behind pure $\mathcal{N} = 2$ SYM with

$$\Lambda_{\text{eff}} = \Lambda^{1-\zeta} M^\zeta.$$  

(5.6)

In the perturbative regime of a very small sphere $(\Lambda \ll 1, M \ll 1)$, the linear force is strong and attractive and pushes all eigenvalues towards the origin. The solution is Wigner’s semicircle (3.8) with $\mu = \sqrt{\lambda_R / 2\pi}$.

A more interesting case is the opposite, decompactification regime when $\Lambda \gg 1, M \gg 1$. In the decompactification limit the linear force is strong but repulsive. Eigenvalues are pushed away from the origin and extend over a large interval $(-\mu, \mu)$. As a result, $x, y$ are on average large. One can approximate the $\mathcal{K}$ function by its asymptotic form (A.5). Differentiating the saddle-point equation twice we get:

$$2 \int_{-\mu}^{\mu} dy \rho(y) \left( \frac{1}{x - y} - \mathcal{K}(x - y) \right) = \frac{\zeta}{x + M} + \frac{\zeta}{x - M}.$$  

(5.7)
This is a very simple equation, but it exhibits the same phenomenon as we encountered in the decompactification limit of the $\mathcal{N} = 2^*$ theory. Namely, the driving term has poles at $x = \pm M$ which may or may not lie within the eigenvalue distribution. The poles are associated with massless hypermultiplets which appear in the spectrum as soon as the poles cross the boundary of the eigenvalue distribution. The model thus has two phases: the weak-coupling phase with $\mu < M$, in which all hypermultiplets are heavy, and the strong-coupling phase at $\mu > M$, in parametrically light hypermultiplets appear in the spectrum.

We begin with the strong-coupling phase ($\mu > M$). The normalized eigenvalue density is then given by

$$\rho(x) = \frac{1 - \zeta}{\pi \sqrt{\mu^2 - x^2}} + \frac{\zeta}{2} \frac{\delta(x + M)}{\mu} + \frac{\zeta}{2} \frac{\delta(x - M)}{\mu}. \quad (5.8)$$

To find $\mu$, we use the integrated form of (5.7):

$$\int_{-\mu}^{\mu} dy \rho(y) \ln \frac{y^2}{M^2} = (1 - \zeta) \ln \frac{\Lambda^2}{M^2}. \quad (5.9)$$

Substituting the solution (5.8) we obtain:

$$\mu = 2\Lambda. \quad (5.10)$$

This generalizes the solution of the pure $\mathcal{N} = 2$ SYM in the 't Hooft limit [5, 6, 11] to the solution of $\mathcal{N} = 2$ SQCD in the Veneziano limit. Interestingly, the width of the eigenvalue distribution does not depend on the hypermultiplet mass.

The solution was obtained under the assumption that the width of the eigenvalue distribution exceeds the hypermultiplet mass: $\mu > M$. This condition breaks down when $M$ reaches $\mu = 2\Lambda$. As a result, the system undergoes a transition to the weak-coupling regime. The phase transition thus happens at

$$M_c = 2\Lambda. \quad (5.11)$$

We now proceed with solving the model in the phase with $\mu < M$. The saddle-point equation (5.7) is then very similar to (4.18). The solution defined on the interval $(-\mu, \mu)$ and having unit normalization is given by

$$\rho(x) = \frac{1}{\pi \sqrt{\mu^2 - x^2}} \left(1 - \zeta + \frac{\zeta M \sqrt{M^2 - \mu^2}}{M^2 - x^2}\right). \quad (5.12)$$

To find $\mu$ we substitute the density into (5.9) and use (4.21):

$$\left\{ \ln \frac{x^2}{M^2} \right\} = 2 (1 - \zeta) \ln \frac{\mu}{2M} + 2\zeta \ln \frac{\mu}{M + \sqrt{M^2 - \mu^2}}. \quad (5.13)$$

\[15\]Again, the equation is the same as in the matrix model of [37], but with different boundary conditions.
This results in a transcendental equation for \( \mu \), whose solution can be written in a parametric form:

\[
\mu = M \sqrt{1-u^2},
\]

\[
\left( \frac{2\Lambda}{M} \right)^{2-2\zeta} = (1+u)^{1-2\zeta} (1-u).
\]

With increasing \( M \), \( \mu \) grows from \( \mu = 2\Lambda \) at the critical point to infinity, asymptotically approaching

\[
\mu \approx 2\Lambda^{1-\zeta} M^\zeta = 2\Lambda_{\text{eff}},
\]
in accord with effective-field theory expectations. Indeed, at very large \( M \) the model should be equivalent to pure \( \mathcal{N} = 2 \) SYM with the effective scale \((5.6)\).

As we have shown, the model undergoes a phase transition at \( M = 2\Lambda \). To determine the order of the transition we can calculate the free energy, or better its first derivative:

\[
\frac{\partial F}{\partial \ln \Lambda} = -2 (1 - \zeta) (x^2).
\]

For the second moment of the density we get:

\[
\langle x^2 \rangle = 2 (1 - \zeta) \Lambda^2 + \zeta M^2 \quad (M < 2\Lambda),
\]

in the strong-coupling phase, and

\[
\langle x^2 \rangle = \frac{M^2}{2} (1 - u) [1 + \zeta + (1 - \zeta) u] \quad (M > 2\Lambda),
\]

in the weak-coupling phase. To compare the two expressions it is convenient to rewrite the strong-coupling answer \((5.17)\) in terms of the variable \( u \) defined in \((5.15)\):

\[
\langle x^2 \rangle = \frac{M^2}{2} \left[ 2\zeta + (1 - \zeta) (1 + u)^{\frac{1-2\zeta}{1-\zeta}} (1 - u)^{\frac{1}{1-\zeta}} \right] \quad (M < 2\Lambda).
\]

The phase transition happens at \( u = 0 \). Taylor expanding around the critical point, we find:

\[
\langle x^2 \rangle = \frac{M^2}{2} \times \begin{cases} 
1 + \zeta - 2\zeta u - (1 - \zeta) u^2 & u > 0 \\
1 + \zeta - 2\zeta u - \frac{1-2\zeta-\zeta^2}{1-\zeta} u^2 + \ldots & u < 0
\end{cases}
\]

The first two Taylor coefficients coincide. Consequently, the free energy is continuous up to the third derivative, which has a finite jump at the transition point. The phase transition is thus of the third order.

The Wilson loop satisfies perimeter law with the exponent dictated by \( \mu \):

\[
\ln W(C) \approx \mu L
\]

for a contour\(^{16}\) of length \( L \). The coefficient of proportionality is given by \((5.10)\) at \( M < 2\Lambda \) and by \((5.14), (5.15)\) at \( M > 2\Lambda \). The dependence of \( \mu \) on \( M \) is plotted in fig. 12. It is clear

\(^{16}\)Using localization we can compute the circular Wilson loop on the sphere, for which the exponent is equal to \( 2\pi\mu R \) and length is equal to \( 2\pi R \). We extrapolate this perimeter-law behavior to any sufficiently large contour.
from the plot, and also from eqs. (5.10), (5.14), (5.15), that $\mu$ is continuous across the phase transition but its first derivative experiences a jump.

The plot in fig. 12 is for $\zeta = 1/2$. In this particular case one can express $\mu$ in terms of $M$ and $\Lambda$ explicitly:

$$\mu^2 = 4\Lambda(M - \Lambda) \quad (\zeta = 1/2, \ M > 2\Lambda) \quad (5.22)$$

The second moment of the eigenvalue density takes the form:

$$\langle x^2 \rangle = \Lambda^2 + \frac{M^2}{2} \quad (\zeta = 1/2, \ M < 2\Lambda) \quad (5.23)$$

and

$$\langle x^2 \rangle = \Lambda(2M - \Lambda) \quad (\zeta = 1/2, \ M > 2\Lambda). \quad (5.24)$$

Its second derivative is discontinuous at the transition point.

The phase transition in $\mathcal{N} = 2$ SQCD shares a lot of similarities with the phase transition found in the $\mathcal{N} = 2^*$ theory, although the SQCD case is technically much simpler. In both cases the phase transition is caused by the emergence of massless modes, which produce resonance peaks in the eigenvalue density. In the present case, these peaks are delta functions $\delta(x \pm M)$. As a result, the transition is less continuous than in the $\mathcal{N} = 2^*$ case.

6 Conclusions

Massive $\mathcal{N} = 2$ gauge theories exhibit a wealth of non-perturbative phenomena in the large-$N$ limit. In specific limits the physics is described by solvable matrix models. We have found that theories with dimensionless couplings, such as $\mathcal{N} = 2^*$ SYM, or with two mass scales, such as SQCD, undergo large-$N$ phase transition as the couplings or mass ratios change. On the other hand, we found that $SU(N)$ super-Yang-Mills with $2N$ massive hypermultiplets in the
fundamental exhibits a continuous interpolation between the weak and strong coupling regimes, without phase transitions. The models studied in this paper have the expected OPE expansion, albeit with rather simple coefficients. This fact is probably due to supersymmetry. We have also given explicit formulas for the 1/2 BPS circular Wilson loop and free energy in the strong coupling limit, for the different models. This may allow for a direct comparison with formulas obtained from holographic dual candidates.

When phase transitions occur, the weak and strong coupling regimes correspond to different, disconnected branches of the large-$N$ master field. This could potentially pose a problem for the holographic description, for instance in the context of the $\mathcal{N} = 2^*$ theory, where an infinite number of phase transitions accumulate at strong coupling. However, the mere existence of phase transitions does not preclude the string description from being exactly equivalent to field theory. The phase transitions should be then visible on the string-theory side as well. We have no clear idea what mechanism can trigger phase transitions in the string sigma-model, but presumably they are related to the singularities that the supergravity background \cite{18, 19} in which the string propagates has in the far IR.

The localization result of \cite{1} is a plug-in formula valid in principle for any $\mathcal{N} = 2$ theory on $S^4$. Its generalization to the squashed four-sphere is also known \cite{38}. It would be interesting to investigate the large-$N$ limit of $\mathcal{N} = 2$ theories in both cases in more generality, or at least to go through a larger set of examples. As a first step in this direction we briefly comment on the large-$N$ limit of certain quiver models in appendix D.

Acknowledgments

We would like to thank N. Bobev, A. Buchel, N. Drukker, G. Festuccia, N. Gromov, V. Kazakov, I. Kostov, Y. Makeenko, K. Skenderis and D. Volin for discussions. The work of K.Z. was supported in part by People Programme (Marie Curie Actions) of the European Union’s FP7 Programme under REA Grant Agreement No 317089. J.R. acknowledges support by MCYT Research Grant No. FPA 2010-20807.

A Functions $H(x)$ and $K(x)$

Here we collect some formulas for the functions $H(x)$ and $K(x)$, defined in (2.5) and (2.6), respectively. The former is related to the Barnes $G$-function:

$$H(x) = e^{-(1+\gamma)x^2} G(1 + ix) G(1 - ix).$$  \hfill (A.1)

The latter can be expressed through the $\psi$-function (the logarithmic derivative of the $\Gamma$-function):

$$K(x) = x \left( \psi(1 + ix) + \psi(1 - ix) + 2\gamma \right),$$  \hfill (A.2)

where $\gamma = -\psi(1)$ is the Euler constant.

The function $K(x)$ is meromorphic on the whole complex plane, and has an infinite series of poles along the imaginary axis. The residue at the $n$th pole grows linearly with $n$. It can be
viewed as a generating function of odd zeta-values:

\[ K(x) = 2 \sum_{k=1}^{\infty} (-1)^{k+1} \zeta(2k+1) x^{2k+1}. \]  

(A.3)

The expansion is convergent for \( |x| < 1 \), until the first pole of the \( \psi \)-function at \( x = \pm i \).

At large real values of the argument, the functions \( H(x) \) and \( K(x) \) behave as

\[
\begin{align*}
\ln H(x) &= -\frac{1}{2} x^2 \ln x^2 + \left( \frac{1}{2} - \gamma \right) x^2 + O(\ln x^2) \\
K(x) &= x \ln x^2 + 2\gamma x + O(x^{-1}).
\end{align*}
\]

(A.4)

(A.5)

These are the first terms of an asymptotic expansion with zero radius of convergence.

Another useful formula is the Fourier-transform representation of \( K''(x) \):

\[ K''(x) = 4 \int_0^\infty d\omega \frac{\omega^2 \sin 2\omega x}{\sinh^2 \omega}. \]  

(A.6)

B  Theta functions and Eisenstein series

Here we list some functions that appear in the exact solution of the \( \mathcal{N} = 2^* \) matrix model. Those are the four theta-functions:

\[
\begin{align*}
\theta_1(z|q) &= 2q^{\frac{1}{4}} \sum_{n=0}^{\infty} (-1)^n q^{n(n+1)} \sin (2n+1) z \\
\theta_2(z|q) &= 2q^{\frac{1}{4}} \sum_{n=0}^{\infty} q^{n(n+1)} \cos (2n+1) z \\
\theta_3(z|q) &= 1 + 2 \sum_{n=0}^{\infty} q^n \cos 2nz \\
\theta_4(z|q) &= 1 + 2 \sum_{n=0}^{\infty} (-1)^n q^n \cos 2nz,
\end{align*}
\]

(B.1)

and the Eisenstein series:

\[
\begin{align*}
E_2(q) &= 1 - 24 \sum_{n=1}^{\infty} \frac{nq^n}{1 - q^n} \\
E_4(q) &= 1 + 240 \sum_{n=1}^{\infty} \frac{n^3q^n}{1 - q^n} \\
E_6(q) &= 1 - 504 \sum_{n=1}^{\infty} \frac{n^5q^n}{1 - q^n}.
\end{align*}
\]

(B.2)

The higher Eisenstein series, including \( E_4 \) and \( E_6 \), can be expressed in terms of the theta-constants:

\[
\begin{align*}
2E_4(q^2) &= \theta_2^8(0|q) + \theta_3^8(0|q) + \theta_4^8(0|q) \\
2E_6(q^2) &= (\theta_2^4(0|q) - \theta_2^4(0|q)) \left(2\theta_3^8(0|q) + \theta_2^6(0|q)\theta_4^4(0|q)\right).
\end{align*}
\]

(B.3)
Together with the identity
\[ \theta_4^4(0|q) = \theta_2^4(0|q) + \theta_4^4(0|q), \]  
this allows one to express \( \xi, \eta \) and \( \bar{\eta} \) from (3.57) through the Eisenstein series only:
\[ \begin{align*}
\xi + \eta + \bar{\eta} &= \frac{M^2}{4} E_2 \\
\xi \eta + \xi \bar{\eta} + \eta \bar{\eta} &= \frac{M^4}{48} (E_4 - E_2^2) \\
\xi \eta \bar{\eta} &= \frac{M^6}{1728} (2E_6 - 3E_2E_4 + E_4^3). \tag{B.5}
\end{align*} \]

This allows one to express \( \xi, \eta \) and \( \bar{\eta} \) from (3.57) through the Eisenstein series only:
\[ \begin{align*}
\xi + \eta + \bar{\eta} &= \frac{M^2}{4} E_2 \\
\xi \eta + \xi \bar{\eta} + \eta \bar{\eta} &= \frac{M^4}{48} (E_4 - E_2^2) \\
\xi \eta \bar{\eta} &= \frac{M^6}{1728} (2E_6 - 3E_2E_4 + E_4^3). \tag{B.5}
\end{align*} \]

Here, as in the main text, the argument of the Eisenstein series is \( -r^2 \), with \( r \) given by (3.58).

The Eisenstein series \( E_2 \) is also not independent; it is given in terms of \( E_4 \) and \( E_6 \) by the logarithmic derivative of the modular discriminant \( \Delta' = 1/1728 (E_4^3 - E_6^2) \) (the derivation being \( \frac{1}{2 \pi i} \frac{d}{d \tau} \)). It follows that it is not a modular form: the modular transformation has an anomalous piece,
\[ E_2(\alpha \cdot \tau) = (c \tau + d)^2 E_2(\tau) + \frac{6c}{\pi i} (c \tau + d), \quad \alpha \cdot \tau = \frac{a \tau + b}{c \tau + d}, \quad \alpha \in SL(2, \mathbb{Z}) \]  
Complete elliptic integrals \( K \equiv K(m), E \equiv E(m) \) are expressed through theta-constants as
\[ \begin{align*}
m &= \frac{\theta_4^4}{\theta_3^4} \tag{B.7} \\
K &= \frac{\pi}{2} \theta_3^2 \tag{B.8} \\
E &= \frac{\pi}{6} \frac{E_2 + \theta_4^4 + \theta_3^4}{\theta_3^2}, \tag{B.9}
\end{align*} \]

with
\[ r = -i e^{-\frac{\pi K'}{K}}, \tag{B.10} \]

where \( K' = K(1 - m) \). The incomplete elliptic integrals \( E(\varphi|m) \) and \( F(\varphi|m) \) are given by
\[ \begin{align*}
\sin \varphi &= \frac{\theta_3(0) \theta_1(v)}{\theta_2(0) \theta_4(v)} \tag{B.11} \\
KE(\varphi) - EF(\varphi) &= \frac{\pi \theta_4'(v)}{2 \theta_4(v)}. \tag{B.12}
\end{align*} \]

## C Scaling behavior of endpoint position

In this appendix we derive (3.68), (3.69) from the solution of the \( N = 2^* \) matrix model in the weak-coupling phase.

\( q = e^{i \pi \tau}, \) as usual.
First, it is convenient to express the parameters of the solution in terms of elliptic integrals, using eqs. (B.7)–(B.12). We have:

\[
\begin{align*}
\xi &= \frac{M^2}{\pi^2} KE , \\
\eta &= \frac{M^2}{\pi^2} K(E - K) , \\
\bar{\eta} &= \frac{M^2}{\pi^2} K [E - (1 - m) K] ,
\end{align*}
\]  
(C.1)

and the modular parameter \( m \) is determined by the equation

\[
\frac{K'}{K} = \frac{4\pi}{\lambda} - \frac{i}{2}.
\]  
(C.2)

The width of the eigenvalue distribution is given by

\[
\mu = -\frac{iM}{\pi} \left( KE(\varphi) - EF(\varphi) \right), \quad \sin^2 \varphi = \frac{K - E}{mK}.
\]  
(C.3)

The phase transition happens when \( \xi(\lambda_c) = 0 \). From (C.1) we see that at the critical point elliptic \( E \) turns to zero: \( E(m_c) = 0 \). The second equation in (C.3) then implies that

\[
\sin^2 \varphi_c = \frac{1}{m_c}.
\]  
(C.4)

For such \( \varphi_c \), incomplete elliptic integrals can be expressed through the complete ones:

\[
F(\varphi_c) = K - iK', \quad E(\varphi_c) = i(E' - K').
\]  
(C.5)

Substituting these equalities into the first equation in (C.3), and using Legendre’s identity,

\[
KE' + EK' - KK' = \frac{\pi}{2},
\]  
(C.6)

along with \( E = 0 \), we obtain that \( \mu(\lambda_c) = M/2 \), which demonstrates the equivalence of the two conditions for the critical coupling, \( \xi = 0 \) and \( \mu = M/2 \).

To study the critical behavior we need the first correction to (C.5). We can regard either \( \varphi - \varphi_c \) or elliptic \( E \) as a small parameter, since both vanish at the critical point. The two are related by the second equation in (C.3) and (C.4). Since \( \sin^2 \varphi = 1/m \) is a square-root branch point of \( F(\varphi) \) and \( E(\varphi) \), their expansion in \( \varphi - \varphi_c \) contains non-analytic terms:

\[
\begin{align*}
F(\varphi) &\approx K - iK' - \frac{\sin^2 \varphi_c - \sin^2 \varphi}{\cos \varphi_c} \approx K - iK' - \sqrt{\frac{E}{(m - 1) K}} \\
E(\varphi) &\approx i(E' - K') + E - m \left( \frac{\sin^2 \varphi_c - \sin^2 \varphi}{3 \cos \varphi_c} \right)^{\frac{3}{2}} \approx i(E' - K') + E - \sqrt{\frac{E^3}{9 (m - 1) K^3}}.
\end{align*}
\]  
(C.7)
Substituting this into (C.3), we get for (3.66):

\[ \Delta z = -\frac{2ME^2}{3\pi\sqrt{(1 - m)K}}. \]  

(C.8)

Trading elliptic integrals for the parameters \( \xi \) and \( \eta \) with the help of (C.1), we arrive at eq. (3.67) in the main text.

Now using

\[ \left( \frac{\partial E}{\partial m} \right)_{E=0} = -\frac{K}{2m} \]

\[ \left( \frac{\partial}{\partial m} \frac{K'}{K} \right)_{E=0} = \frac{\pi}{4m(1 - m)K^2} \]  

(C.9)

we get from (C.2):

\[ E \approx \frac{8(1 - m)K^3}{\lambda_c^2} (\lambda_c - \lambda). \]  

(C.10)

Substituting this into (C.8) and expressing \( m \) and \( K \) through the theta-constants as in (B.7), (B.8) we get eq. (3.68).

**D Quiver models**

In this appendix we briefly comment on the large \( N \) behavior of certain quiver models which interpolate between different theories considered in the main text, as well as \( \mathcal{N} = 4 \) SYM, pure \( \mathcal{N} = 2 \) SYM and the \( \mathcal{N} = 2 \) superconformal SYM with \( 2N \) massless hypermultiplets in the fundamental representation.

We consider the following quivers:

A) \( SU(N)_1 \times SU(N)_2 \) with one bi-fundamental hyper and \( N \) fundamental hypers for each of the \( SU(N) \)'s. The localization partition function of this superconformal theory is

\[
Z_A = \int d^{N-1}a \, d^{N-1}b \frac{\prod_{i<j} (a_i - a_j)^2 (b_i - b_j)^2 H^2(a_i - a_j)H^2(b_i - b_j)}{\prod_{i,j} H(a_i - b_j) \prod_i H^N(a_i)H^N(b_i)} e^{\frac{8\pi^2N\frac{a_1^2 + b_1^2}{\lambda_1} + \frac{a_2^2 + b_2^2}{\lambda_2}}{}} \times |Z_{\text{inst}}^A(a, b; g_1^2, g_2^2)|^2. \]  

(D.1)

If one of the two couplings is set to zero (for instance, \( \lambda_2 \)), the \( b_i \) variables are frozen at \( b_i = 0 \), and the theory becomes equivalent to \( \mathcal{N} = 2 \) \( SU(N) \) SCYM with \( 2N \) flavors.

B) \( SU(N)_1 \times SU(N)_2 \) with 2 bi-fundamental hypers. The localization formula gives

\[
Z_B = \int d^{N-1}a \, d^{N-1}b \frac{\prod_{i<j} (a_i - a_j)^2 (b_i - b_j)^2 H^2(a_i - a_j)H^2(b_i - b_j)}{\prod_{i,j} H^2(a_i - b_j)} e^{\frac{8\pi^2N\frac{a_1^2 + b_1^2}{\lambda_1} + \frac{a_2^2 + b_2^2}{\lambda_2}}{}} \times |Z_{\text{inst}}^B(a, b; g_1^2, g_2^2)|^2. \]  

(D.2)
This theory is also superconformal. Decoupling one of the gauge groups again gives $\mathcal{N} = 2$ SCYM with $2N$ flavors.

B*) The massive deformation of the B model:

$$Z_{B^*} = \int d^{N-1}a \, d^{N-1}b \, \prod_{i<j} (a_i - a_j)^2 (b_i - b_j)^2 H^2(a_i - a_j) H^2(b_i - b_j) \, e^{-8\pi^2 N \sum_{i,j} x_i^2 + x_j^2} \prod_{i,j} H(a_i - b_j + M) H(a_i - b_j - M) \times |Z_{\text{mat}}^{B^*}(a, b; g_1^2, g_2^2)|^2.$$  \hspace{1cm} (D.3)

The theory with $\lambda_1 = \lambda_2$ should have similar dynamics as $\mathcal{N} = 2^* SU(N)$ SYM. The hypermultiplets decouple in the IR and the theory flows to two copies of pure $\mathcal{N} = 2$ SYM.

### D.1 Model A

After introducing two densities for the $a_i$ and $b_i$ eigenvalues, the saddle-point equations become

$$\mu_1 \int_{-\mu_1} dy \rho_1(y) \left( \frac{1}{x - y} - \mathcal{K}(x - y) \right) + \frac{1}{2} \int_{-\mu_2} d\hat{x} \rho_2(\hat{x}) \mathcal{K}(x - \hat{x}) = \frac{8\pi^2}{\lambda_1} x - \frac{\mathcal{K}(x)}{2}, \hspace{1cm} (D.4)$$

$$\mu_2 \int_{-\mu_2} dy \rho_2(y) \left( \frac{1}{\hat{x} - y} - \mathcal{K}(\hat{x} - y) \right) + \frac{1}{2} \int_{-\mu_1} dx \rho_1(x) \mathcal{K}(\hat{x} - x) = \frac{8\pi^2}{\lambda_2} \hat{x} - \frac{\mathcal{K}(\hat{x})}{2}. \hspace{1cm} (D.5)$$

When $\lambda_2 \ll 1$, the $\rho_2$ density is peaked at zero, while $\rho_1$ satisfies the equations for $\mathcal{N} = 2$ SCYM, as expected.

Another simplifying case is $\lambda_1 = \lambda_2$, when the model possesses a $\mathbb{Z}_2$ symmetry. The solution with unbroken $\mathbb{Z}_2$ has $\rho_1 = \rho_2$, and we are left with the equation

$$\mu_1 \int_{-\mu_1} dy \rho_1(y) \left( \frac{2}{x - y} - \mathcal{K}(x - y) \right) = \frac{16\pi^2}{\lambda_1} x - \mathcal{K}(x). \hspace{1cm} (D.6)$$

This equation is similar but not equivalent to the one for $\mathcal{N} = 2$ SCFT with $2N$ massless hypermultiplets.

### D.2 Model B

The saddle-point equations are

$$\mu_1 \int_{-\mu_1} dy \rho_1(y) \left( \frac{1}{x - y} - \mathcal{K}(x - y) \right) + \mu_2 \int_{-\mu_2} d\hat{x} \rho_2(\hat{x}) \mathcal{K}(x - \hat{x}) = \frac{8\pi^2}{\lambda_1} x, \hspace{1cm} (D.7)$$

$$\mu_2 \int_{-\mu_2} dy \rho_2(y) \left( \frac{1}{\hat{x} - y} - \mathcal{K}(\hat{x} - y) \right) + \mu_1 \int_{-\mu_1} dx \rho_1(x) \mathcal{K}(\hat{x} - x) = \frac{8\pi^2}{\lambda_2} \hat{x}. \hspace{1cm} (D.7)$$

49
Interestingly, at the symmetric point $\lambda_1 = \lambda_2 \equiv \lambda$ the complicated interaction terms cancel, and the saddle-point equations reduce to those for the Gaussian matrix model. The solution is
\[ \rho_1(x) = \rho_2(x) = \frac{2}{\pi \mu^2} \sqrt{\mu^2 - x^2}, \quad \mu = \frac{\sqrt{\lambda}}{2\pi}. \]

### D.3 Massive B* model

For the massive case, the B* model, the saddle-point equations read
\[
\int_{-\mu_1}^{\mu_1} dy \rho_1(y) \left( \frac{1}{x - y} - \mathcal{K}(x - y) \right) + \frac{1}{2} \int_{-\mu_2}^{\mu_2} d\hat{x} \rho_2(\hat{x}) (\mathcal{K}(x - \hat{x} + M) + \mathcal{K}(x - \hat{x} - M)) = \frac{8\pi^2}{\lambda_1} x,
\]
\[
\int_{-\mu_2}^{\mu_2} d\hat{y} \rho_2(\hat{y}) \left( \frac{1}{\hat{x} - \hat{y}} - \mathcal{K}(\hat{x} - \hat{y}) \right) + \frac{1}{2} \int_{-\mu_1}^{\mu_1} dx \rho_1(x) (\mathcal{K}(\hat{x} - x + M) + \mathcal{K}(\hat{x} - x - M)) = \frac{8\pi^2}{\lambda_2} \hat{x}.
\]

If $\lambda_2 \to 0$, the second gauge group decouples leaving behind $\mathcal{N} = 2$ SYM with $2N$ massive hypermultiplets. On the other hand, if $\lambda_1 = \lambda_2$, the symmetric solution satisfies the saddle-point equation for $\mathcal{N} = 2^* \text{SYM} (3.5)$. Thus the $B^*$ model interpolates between the two massive models discussed in the main text.

### References

[1] V. Pestun, “Localization of gauge theory on a four-sphere and supersymmetric Wilson loops”, Commun.Math.Phys. 313, 71 (2012), [0712.2824].

[2] G. ’t Hooft, “A Planar Diagram Theory for Strong Interactions”, Nucl.Phys. B72, 461 (1974).

[3] J. M. Maldacena, “The large N limit of superconformal field theories and supergravity”, Adv. Theor. Math. Phys. 2, 231 (1998), [hep-th/9711200] • S. S. Gubser, I. R. Klebanov and A. M. Polyakov, “Gauge theory correlators from non-critical string theory”, Phys. Lett. B428, 105 (1998), [hep-th/9802109] • E. Witten, “Anti-de Sitter space and holography”, Adv. Theor. Math. Phys. 2, 253 (1998), [hep-th/9802150].

[4] N. Seiberg and E. Witten, “Monopole Condensation, And Confinement In $\mathcal{N} = 2$ Supersymmetric Yang-Mills Theory”, Nucl. Phys. B426, 19 (1994), [hep-th/9407087] • N. Seiberg and E. Witten, “Monopoles, duality and chiral symmetry breaking in $\mathcal{N} = 2$ supersymmetric QCD”, Nucl. Phys. B431, 484 (1994), [hep-th/9408099].

[5] M. R. Douglas and S. H. Shenker, “Dynamics of SU(N) supersymmetric gauge theory”, Nucl.Phys. B447, 271 (1995), [hep-th/9503163].

[6] F. Ferrari, “The Large N limit of N=2 superYang-Mills, fractional instantons and infrared divergences”, Nucl.Phys. B612, 151 (2001), [hep-th/0106192].

[7] S.-J. Rey and T. Suyama, “Exact Results and Holography of Wilson Loops in N=2 Superconformal (Quiver) Gauge Theories”, JHEP 1101, 136 (2011), [1001.0016].
[8] F. Passerini and K. Zarembo, “Wilson Loops in N=2 Super-Yang-Mills from Matrix Model”, JHEP 1109, 102 (2011), 1106.5763
[9] J.-E. Bourgine, “A Note on the integral equation for the Wilson loop in N = 2 D=4 superconformal Yang-Mills theory”, J.Phys.A A45, 125403 (2012), 1111.0384 • B. Fraser and S. P. Kumar, “Large rank Wilson loops in N=2 superconformal QCD at strong coupling”, JHEP 1203, 077 (2012), 1112.5182
[10] J. G. Russo, “A Note on perturbation series in supersymmetric gauge theories”, JHEP 1206, 038 (2012), 1203.5061
[11] J. Russo and K. Zarembo, “Large N Limit of N=2 SU(N) Gauge Theories from Localization”, JHEP 1210, 082 (2012), 1207.3806
[12] A. Buchel, J. G. Russo and K. Zarembo, “Rigorous Test of Non-conformal Holography: Wilson Loops in N=2* Theory”, JHEP 1303, 062 (2013), 1301.1597
[13] J. G. Russo and K. Zarembo, “Evidence for Large-N Phase Transitions in N=2* Theory”, JHEP 1304, 065 (2013), 1302.6968
[14] D. J. Gross and A. Matytsin, “Instanton induced large N phase transitions in two-dimensional and four-dimensional QCD”, Nucl. Phys. B429, 50 (1994), hep-th/9404004.
[15] G. Veneziano, “Some Aspects of a Unified Approach to Gauge, Dual and Gribov Theories”, Nucl.Phys. B117, 519 (1976).
[16] G. Festuccia and N. Seiberg, “Rigid Supersymmetric Theories in Curved Superspace”, JHEP 1106, 114 (2011), 1105.0689
[17] N. A. Nekrasov, “Seiberg-Witten prepotential from instanton counting”, Adv. Theor. Math. Phys. 7, 831 (2004), hep-th/0206161 • N. Nekrasov and A. Okounkov, “Seiberg-Witten theory and random partitions”, hep-th/0306238
[18] K. Pilch and N. P. Warner, “N=2 supersymmetric RG flows and the IIB dilaton”, Nucl.Phys. B594, 209 (2001), hep-th/0004063.
[19] A. Buchel, A. W. Peet and J. Polchinski, “Gauge dual and noncommutative extension of an N=2 supergravity solution”, Phys.Rev. D63, 044009 (2001), hep-th/0008076
[20] J. E. Carlisle and C. V. Johnson, “Holographic RG flows and universal structures on the Coulomb branch of N=2 supersymmetric large N gauge theory”, JHEP 0307, 039 (2003), hep-th/0306168
[21] A. Buchel, “Compactifications of the N = 2* flow”, Phys.Lett. B570, 89 (2003), hep-th/0302107 • A. Buchel, “Localization and holography in N=2 gauge theories”, 1304.5652
[22] J. K. Erickson, G. W. Semenoff and K. Zarembo, “Wilson loops in N = 4 supersymmetric Yang-Mills theory”, Nucl. Phys. B582, 155 (2000), hep-th/0003055
[23] N. Drukker and D. J. Gross, “An exact prediction of N = 4 SUSYM theory for string theory”, J. Math. Phys. 42, 2896 (2001), hep-th/0010274.
[24] N. Beisert, C. Ahn, L. F. Alday, Z. Bajnok, J. M. Drummond et al., “Review of AdS/CFT Integrability: An Overview”, Lett.Math.Phys. 99, 3 (2012), 1012.3982
E. Brezin, C. Itzykson, G. Parisi and J. B. Zuber, “Planar Diagrams”, Commun. Math. Phys. 59, 35 (1978).

J. A. Minahan and K. Zarembo, “The Bethe-ansatz for $N = 4$ super Yang-Mills”, JHEP 0303, 013 (2003), hep-th/0212208 • N. Beisert, C. Kristjansen and M. Staudacher, “The dilatation operator of $N = 4$ conformal super Yang-Mills theory”, Nucl. Phys. B664, 131 (2003), hep-th/0303060 • N. Beisert and M. Staudacher, “The $N = 4$ SYM Integrable Super Spin Chain”, Nucl. Phys. B670, 439 (2003), hep-th/0307042.

N. Beisert, V. Dippel and M. Staudacher, “A novel long range spin chain and planar $N = 4$ super Yang-Mills”, JHEP 0407, 075 (2004), hep-th/0405001 • N. Beisert, “The su(2|2) dynamic S-matrix”, Adv. Theor. Math. Phys. 12, 945 (2008), hep-th/0511082.

A. W. Peet and J. Polchinski, “UV / IR relations in AdS dynamics”, Phys.Rev. D59, 065011 (1999), hep-th/9809022 • M. Bianchi, D. Z. Freedman and K. Skenderis, “How to go with an RG flow”, JHEP 0108, 041 (2001), hep-th/0105276.

F. Bigazzi, A. Cotrone and A. Zaffaroni, “$N=2$ gauge theories from wrapped five-branes”, Phys.Lett. B519, 269 (2001), hep-th/0106160.

J. Hoppe, as cited in [31].

V. A. Kazakov, I. K. Kostov and N. A. Nekrasov, “$D$-particles, matrix integrals and KP hierarchy”, Nucl. Phys. B557, 413 (1999), hep-th/9810035.

I. Kostov, “$O(n)$ Vector Model On A Planar Random Lattice: Spectrum Of Anomalous Dimensions”, Mod.Phys.Lett. A4, 217 (1989). • M. Gaudin and I. Kostov, “$O(n)$ Model On A Fluctuating Planar Lattice: Some Exact Results”, Phys.Lett. B220, 200 (1989). • I. K. Kostov and M. Staudacher, “Multicritical phases of the $O(n)$ model on a random lattice”, Nucl.Phys. B384, 459 (1992), hep-th/9203030.

B. Eynard and J. Zinn-Justin, “The $O(n)$ model on a random surface: Critical points and large order behavior”, Nucl.Phys. B386, 558 (1992), hep-th/9204082.

B. Eynard and C. Kristjansen, “Exact solution of the $O(n)$ model on a random lattice”, Nucl.Phys. B455, 577 (1995), hep-th/9506193.

B. Eynard and C. Kristjansen, “More on the exact solution of the $O(n)$ model on a random lattice and an investigation of the case $|n| > 2$”, Nucl.Phys. B466, 463 (1996), hep-th/9512052 • L. Chekhov and C. Kristjansen, “Hermitian matrix model with plaquette interaction”, Nucl.Phys. B479, 683 (1996), hep-th/9605013.

T. Azeyanagi, M. Hanada, M. Honda, Y. Matsuo and S. Shiba, “A new look at instantons and large-$N$ limit”, 1307.0809.

V. Kazakov, “A Simple Solvable Model of Quantum Field Theory of Open Strings”, Phys.Lett. B237, 212 (1990).

N. Hama and K. Hosomichi, “Seiberg-Witten Theories on Ellipsoids”, JHEP 1209, 033 (2012), 1206.6359 • F. Fucito, J. F. Morales, R. Poghossian and D. R. Pacifi, “Exact results in $N=2$ gauge theories”, 1307.6612.