\section*{Abstract}

The planar $\mathcal{N} = 2^*$ Super-Yang-Mills (SYM) theory is solved at large 't Hooft coupling using localization on $S^4$. The solution permits detailed investigation of the resonance phenomena responsible for quantum phase transitions in infinite volume, and leads to quantitative predictions for the semiclassical string dual of the $\mathcal{N} = 2^*$ theory.

\* Also at ITEP, Moscow, Russia
1 Introduction

Supersymmetric localization is a way to compute path integrals in interacting field theories directly, without making any approximations [1]. Our work exploits localization on $S^4$, in which case the field-theory path integral reduces to a finite-dimensional matrix model [2]. An interesting regime that can then be explored in detail is the planar, large-$N$ limit. The strong-coupling behavior of a planar theory is generally believed to have a simple (weakly-coupled) string description. Localization can be used to test gauge/string duality in a very precise way, while also giving us insight into possible dynamical effects in strongly coupled gauge theories.

Localization, as a method, has obvious limitations as it requires a sufficient amount of supersymmetry. Gauge/string duality is believed to have a much broader scope, but it too is formulated precisely only in a limited number of cases. The model studied in this paper, $\mathcal{N} = 2^*$ super-Yang-Mills theory, is special in this respect. Its partition function on $S^4$ and some select observables are calculable by localization [2], and at the same time it has a well-defined holographic dual — the type IIB string theory on the Pilch-Warner background [3].

The localized partition function of the $\mathcal{N} = 2^*$ SYM on $S^4$ is a matrix model which at large-$N$ can be studied by standard methods [4] of random matrix theory [5, 6, 7, 8, 9]. A number of observables computed with the help of localization can be successfully compared to string-theory predictions at strong coupling. These include Wilson loops for asymptotically large contours [6], and the free energy on $S^4$ [10].

Away from the strict strong-coupling limit, localization leads to somewhat unexpected results. It turns out that the planar $\mathcal{N} = 2^*$ SYM has a very complicated phase structure, undergoing an infinite number of quantum phase transitions as the ’t Hooft coupling changes from zero to infinity [7, 8]. While these are a common phenomenon in large-$N$ theories [11], the behavior in $\mathcal{N} = 2^*$ SYM is unique and differs in many respects from phase transitions in ordinary matrix models [8]. Similar phase transitions have also been found in QCD-like vector models [8, 12] as well as in three [12, 13] and five [14] dimensional theories.

The existence of an infinite number of phase transitions raises the question of how the non-trivial phase structure of $\mathcal{N} = 2^*$ SYM is reflected in its holographic dual. Ideally, one would like to tune the ’t Hooft coupling to its critical value, but going to finite coupling is notoriously difficult in holog-
Figure 1: The phase diagram of the planar $\mathcal{N} = 2^*$ theory [8]. If the strong coupling limit is approached at strictly infinite radius (on $\mathbb{R}^4$), as shown in black horizontal arrows, the theory undergoes an infinite number of phase transitions. In this paper we approach the same corner of the phase diagram along a different direction, by varying the compactification radius while keeping the coupling strictly infinite, as shown in green vertical arrows. Although the results may depend on the direction along which the critical point is approached, we do find the structures that cause phase transitions in the strong-coupling finite-volume solution.

...
corrections are equivalent to quantum corrections on the string side, therefore, they are potentially calculable by semiclassical string quantization in the dual supergravity background. As we shall see, irregular structures responsible for phase transitions in the decompactification limit are already present in the first-order approximation, which opens an avenue to study the critical behavior of $\mathcal{N} = 2^*$ SYM within semiclassical string theory.

2 Localization

The $\mathcal{N} = 2^*$ theory is the unique massive deformation of $\mathcal{N} = 4$ SYM that preserves half of the rigid supersymmetry. The field content consists of the gauge fields $A_\mu$, their scalar superpartners $\Phi$ and $\Phi'$, complex scalars $Z_{1,2}$ from the hypermultiplet, and the fermions. All the fields are in the adjoint of the gauge group which we take to be $SU(N)$. The Lagrangian of the $\mathcal{N} = 2^*$ theory is obtained from that of $\mathcal{N} = 4$ SYM by giving common mass $M$ to the hypermultiplet fields and adding certain Yukawa couplings necessary for supersymmetry.

The scalars in the vector multiplet can condense along the flat directions of the potential $V \sim \text{tr}[\Phi, \Phi']^2$:

$$\langle \Phi \rangle = \text{diag}(a_1, \ldots, a_N),$$

thus breaking the gauge symmetry to $U(1)^{N-1}$. The $ij$ components of the vector-multiplet fields then acquire masses $m^{v}_{ij} = |a_i - a_j|$ while hypermultiplet masses are $m^{h}_{ij} = |a_i - a_j | \pm M$. The vector multiplets get light as soon as the eigenvalues $a_i$ and $a_j$ approach one another – the diagonal states with $i = j$ are photons of the unbroken $U(1)^{N-1}$. In the matrix model these massless states do not lead to any dramatic effects, because the contribution of the light fields is counteracted by the Vandermonde repulsion of the matrix eigenvalues. The effects due to light hypermultiplets are much more dramatic. Massless hypermultiplets appear once the distance between two eigenvalues gets close to $M$. This resonance phenomenon plays an important rôle in shaping the phase structure of the $\mathcal{N} = 2^*$ SYM in the large-$N$ limit.

An important characteristic of the theory at large $N$ is the master field, characterized by the eigenvalue density:

$$\rho(x) = \left( \frac{1}{N} \sum_{i=1}^{N} \delta(x - a_i) \right).$$
Our goal is to study the exact master field of the $\mathcal{N} = 2^*$ theory on a four-sphere of radius $R$ at large ’t Hooft coupling $\lambda \equiv g^2_{YM}N \gg 1$. Up to some point, we will keep the full dependence on the dimensionless parameter $MR$, and will then separately study the decompactification limit $R \to \infty$, where most of the interesting phenomena occur. The resulting theory can be viewed as $\mathcal{N} = 2^*$ SYM in flat space in a particular vacuum state selected by compactification. The same vacuum is singled out by conformal perturbation theory in $\mathcal{N} = 4$ SYM, and also by AdS/CFT duality for the Pilch-Warner background (see [9] for a more detailed discussion of the vacuum selection in this context).

Supersymmetric localization reduces the path integral of $\mathcal{N} = 2^*$ SYM on $S^4$ to an $(N-1)$-dimensional eigenvalue integral [2]:

$$Z = \int d^{N-1}a \prod_{i<j} \frac{(a_i - a_j)^2 H_2(a_i - a_j)}{H(a_i - a_j - M) H(a_i - a_j + M)} e^{-\frac{\pi^2 N}{4} \sum_i a_i^2 |Z_{\text{inst}}|^2},$$

(2.3)

where

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

(2.4)

We will set the instanton contribution to zero, $Z_{\text{inst}} = 1$, because at large $N$ instantons are exponentially suppressed.

Localization allows one to compute some special correlation functions, for example the expectation value of the Wilson loop for the big circle of $S^4$. The localizable loop operator couples to the scalar $\Phi$ of the vector multiplet, in addition to the usual path-ordered vector coupling:

$$W(C) = \left\{ \frac{1}{N} \text{tr} \exp \oint_C ds \left( iA_\mu \dot{x}^\mu + \Phi |\dot{x}| \right) \right\}.\quad (2.5)$$

If $C$ is the equatorial circle of $S^4$, the Wilson loop can be computed by just substituting the constant classical value (2.1) for $\Phi$ and averaging over the eigenvalues with the weight given by the partition function of the matrix model (2.3):

$$W(C) = \left\{ \frac{1}{N} \sum_i e^{2\pi a_i} \right\} = \int_{-\mu}^\mu dx \rho(x) e^{2\pi x}.\quad (2.6)$$

The eigenvalue integral (2.3) is of the saddle-point type at large $N$, and the partition function is dominated by a single equilibrium configuration of
$a_i$’s when $N \to \infty$. The equilibrium condition can be written as a singular integral equation for the master field (2.2):

$$\int_{-\mu}^{\mu} dy \rho(y) S(x - y) = \frac{8\pi^2}{\lambda}x,$$  

(2.7)

where the kernel is given by

$$S(x) = \frac{1}{x} - K(x) + \frac{1}{2} K(x + M) + \frac{1}{2} K(x - M).$$

(2.8)

The function $K(x)$ is the logarithmic derivative of $H(x)$:

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

(2.9)

The equations above are written in units in which the radius of the sphere is set to one. The dependence on $R$ can be recovered by rescaling $M \to MR$, $x \to xR$ and so on. We will keep the dimensionless units in this paper, but it is important to remember that the decompactification limit $R \to \infty$ is now traded for the infinite mass limit $M \to \infty$.

The solution of the saddle-point equations at infinitely strong coupling was found in [6], and is very simple. Let us repeat the derivation here. The restoring force on the right hand side of (2.7) is very small at large $\lambda$ and the eigenvalues spread over wider and wider intervals as $\lambda$ grows. It is consequently true that both $\mu \gg M$ and $\mu \gg 1$ when $\lambda \gg 1$. Since $x - y \sim \mu$, the difference operator in (2.8) can be replaced by the second derivative, and the kernel function $K(x)$ can be replaced by its asymptotics at large values of the argument,

$$K(x) \approx x \ln x^2 \quad (x \to \infty).$$

(2.10)

We thus have

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

(2.11)

The saddle-point equation reduces to that of the Gaussian matrix model:

$$\int_{-\mu}^{\mu} dy \rho(y) \frac{1 + M^2}{x - y} = \frac{8\pi^2}{\lambda}x,$$

(2.12)
whose solution is Wigner’s semicircle law:

\[ \rho_{\text{\infty}}(x) = \frac{2}{\pi \mu^2} \sqrt{\mu^2 - x^2} \]  

(2.13)

with

\[ \mu = \frac{\sqrt{\lambda (1 + M^2)}}{2\pi} \]  

(2.14)

Taking the decompactification limit \( M \to \infty \) leads to the Wigner distribution of width \( \sqrt{\lambda}M/2\pi \), in precise agreement with the D-brane probe analysis of the Pilch-Warner solution \([15]\). The same result can be also derived within the Seiberg-Witten theory \([16]\).

From this solution we can calculate the expectation value of the Wilson loop (2.6):

\[ W(C) \approx \text{const} \lambda^{-\frac{3}{4}} \exp{\sqrt{\lambda}(1+M^2)} \]  

(2.15)

The leading exponential behavior corresponds to the minimal area law in the dual gravitational description, and at \( M \to \infty \) one finds a precise match with the area law in the Pilch-Warner background \([6]\). The prefactor corresponds to the contribution of string fluctuations, potentially calculable by semiclassi-
csical string quantization. One of our goals is to compute this prefactor from the matrix model.

Using the leading-order solution (2.13), we would get \( \sqrt{2/\pi}(1 + M^2)^{-3/4} \) for the constant of proportionality in (2.15). We do not display this result in the equation, because it is actually incorrect. Indeed, the weight in the integral (2.6) is exponentially peaked at the largest eigenvalue, and the most important contribution comes from \( x \) very close to \( \mu \), namely from \( \mu - x \sim 1 \). The eigenvalue density there is very small, so that corrections to the leading-
order Wigner distribution (2.13) are \( O(1) \) and not \( O(1/\sqrt{\lambda}) \) as one might naively expect. Thus, to compute the Wilson loop, we need to know the endpoint behavior of the eigenvalue distribution exactly.

The importance of the endpoint region can be understood from a different perspective. Suppose that we want to calculate the next correction in \( 1/\sqrt{\lambda} \) to the eigenvalue density. How do we proceed? The first idea that comes to mind is to expand (2.11) to higher orders in \( 1/x \); however, careful inspection of the resulting equations shows that this idea does not work. In fact, the equation stays the same at the next order in \( 1/\sqrt{\lambda} \). What changes is the
boundary behavior of the density. The general solution to (2.12) reads

\[ \rho(x) = \frac{8\pi}{\lambda(1 + M^2)} \sqrt{\mu^2 - x^2} + \frac{\beta}{\mu\sqrt{\mu^2 - x^2}}. \quad (2.16) \]

The second term, which we normalized to obtain a $1/\sqrt{\lambda}$ correction to the Wigner distribution, satisfies the homogeneous form of the integral equation (2.12). The coefficient $\beta$ is thus not fixed by the equations; nor is it fixed by the normalization condition of the density, which is supposed to determine the endpoint position $\mu$.

This freedom to choose any coefficient for the second term may look worrisome. Moreover, the density vanishes at the endpoints for any finite $\lambda$, while the solution above blows up at the edges of the interval. A resolution of these apparent contradictions lies in the fact that the second term in (2.16) cannot be treated as a small correction near the endpoints. The two terms in (2.16) become comparable at $\mu - x \sim 1$, where the naive strong-coupling expansion breaks down. We shall see that the most interesting phenomena, responsible for the phase transitions in infinite volume, happen precisely in this regime. We will eventually fix the constant $\beta$ and then $\mu$ by matching the exact solution in the near-endpoint region to the asymptotic solution (2.16) in the bulk of the eigenvalue distribution.

3 Solution at strong coupling

3.1 Wiener-Hopf problem

As we concluded above, the strong-coupling expansion breaks down near the endpoints of the eigenvalue distribution. The integral equation, therefore, has to be analyzed separately in this regime. We recall that the support of the eigenvalue density becomes very large, $\mu \gg 1$, at strong coupling. We are interested in the behavior at $x$ close to $\mu$, with $\mu - x \sim 1$. To gain some intuition we can start with the leading-order solution (2.13). Introducing the variable

\[ \xi = \mu - x, \quad (3.1) \]

we find at $\xi \sim 1$:

\[ \rho_{\infty}(x) \simeq \frac{2^3}{\pi \mu^2} \sqrt{\xi}. \quad (3.2) \]
We thus expect that the exact density near the upper endpoint has the form
\[ \rho(x) = \frac{2^{\frac{3}{2}}}{\pi \mu^2} f(\xi), \] (3.3)

where \( f(\xi) \) is a scaling function that does not depend on \( \lambda \).

Since \( \rho_\infty \) is a good approximation in the bulk of the eigenvalue distribution, (3.3) should approach (3.2) away from the endpoint. This fixes the boundary conditions on \( f(\xi) \) at large \( \xi \):
\[ f(\xi) \approx \sqrt{\xi} \quad (\xi \to \infty) . \] (3.4)

At \( \xi \sim 1 \), \( f(\xi) \) deviates from \( \sqrt{\xi} \) by an \( O(1) \) amount.

An integral equation for \( f(\xi) \) can be obtained by the following trick: the exact saddle-point equation (2.7) can be re-written as
\[ \int_{-\mu}^{\mu} dy \left[ \rho(y) S(x - y) - \rho_\infty(y) \frac{1 + M^2}{x - y} \right] = 0, \] (3.5)

where we have made use of (2.12). This step entails no approximations. We can now notice that for \( x \) close to \( \mu \) the weight in the \( y \) integral is peaked near the endpoint. We can thus introduce the scaling variable \( \eta \), replace \( \rho(y) \) and \( \rho_\infty(y) \) by their scaling forms (3.3), (3.2), and extend the limit of integration over the scaling variable to infinity:
\[ \int_{0}^{\infty} d\eta \left[ f(\eta) S(\eta - \xi) - \frac{(1 + M^2)\sqrt{\eta}}{\eta - \xi} \right] = 0. \] (3.6)

The integral over \( \eta \) converges at the upper limit due to the asymptotic form of \( K(x) \) given in (2.10), (2.11) and the boundary condition (3.4).

The integral equation (3.6) is of the Wiener-Hopf type. It can be brought to the standard Wiener-Hopf form by introducing the regularized scaling function with a better behavior at infinity:
\[ g(\xi) = f(\xi) - \sqrt{\xi}. \] (3.7)

This function satisfies the equation
\[ \int_{0}^{\infty} d\eta g(\eta) S(\eta - \xi) = F(\xi), \] (3.8)
where

\[ F(\xi) = \int_0^\infty d\eta \sqrt{\eta} \left[ \frac{1 + M^2}{\eta - \xi} - S(\eta - \xi) \right]. \] (3.9)

Both integrals here converge – in the first case because \( g(\eta) \sim 1/\sqrt{\eta} \) at large \( \eta \), and in the second case because the kernel behaves as \( 1/(\eta - \xi)^3 \) at large \( \eta \), as a consequence of (2.11).

### 3.2 Exact solution

In order to solve the integral equation (3.8), let us first assume that the function \( g(\xi) \) is defined on the whole real axis, but is equal to zero for \( \xi < 0 \). Then, the integral on the left-hand side of (3.8) takes the convolution form:

\[ S \ast g(\xi) = F(\xi) + \theta(-\xi) X(\xi), \] (3.10)

where \( X(\xi) \) is an arbitrary function whose appearance reflects the fact that the original equation only holds for \( \xi > 0 \).

After the equation is written in the convolution form, it can be brought to an algebraic form by Fourier transform:

\[ g(\xi) = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} e^{-i\omega\xi} \hat{g}(\omega). \] (3.11)

After the Fourier transform we get:

\[ \hat{S}(\omega) \hat{g}(\omega) = \hat{F}(\omega) + X_-(\omega), \] (3.12)

where \( X_-(\omega) \) is a function that has no singularities in the lower half-plane of complex \( \omega \), because its Fourier image vanishes on the positive real semi-axis. The subindex + will be used similarly, but to indicate analyticity in the upper half-plane. In fact, \( \hat{g}(\omega) = \hat{g}_+(\omega) \). The explicit expressions for the remaining terms are:

\[ \hat{S}(\omega) = i\pi \text{sign} \omega \left( 1 + \frac{\sin^2 \frac{M\omega}{2}}{\sinh^2 \frac{\omega}{2}} \right), \] (3.13)

\[ \hat{F}(\omega) = i\pi \text{sign} \omega \left( M^2 - \frac{\sin^2 \frac{M\omega}{2}}{\sinh^2 \frac{\omega}{2}} \right) \frac{i^{\frac{3}{2}} \sqrt{\pi}}{2\omega \sqrt{\omega + i\epsilon}}, \] (3.14)
which can be obtained from the definitions (2.8), (2.9) and (3.9), as well as the following representation of $K''(x)$:

$$K''(x) = 4 \int_0^\infty d\omega \frac{\omega^2 \sin 2x \omega}{\sinh^2 \omega}.$$  (3.15)

Here, the sign function should be understood in terms of analytic regularization:

$$\text{sign} \omega = \lim_{\epsilon \to 0} \frac{\sqrt{\omega + i \epsilon}}{\sqrt{\omega - i \epsilon}},$$  (3.16)

where the square root $\sqrt{\omega \pm i \epsilon}$ has a branch cut extending from $\mp i \epsilon$ up to infinity along the negative/positive imaginary semi-axis.

The solution of the Wiener-Hopf problem (3.12) is based on the factorization formula

$$\hat{S}\left(\omega\right) = \frac{1}{G_+(\omega)G_-(\omega)}.$$  (3.17)

In our case, the functions $G_{\pm}$ have the form:

$$G_{\pm}(\omega) = \left( \frac{M^2 + 1}{i\pi} \right)^{\frac{1}{2}} (\omega \pm i \epsilon)^{\frac{1}{2}} e^{\mp \frac{\omega \epsilon}{2\pi}} \frac{\Gamma\left( \frac{M \pm i}{2\pi} \omega \right) \Gamma\left( -\frac{M \pm i}{2\pi} \omega \right)}{\Gamma^2\left( \frac{i \omega}{2\pi} \right)},$$  (3.18)

where

$$\phi = 2M \arctan M - \ln\left( M^2 + 1 \right).$$  (3.19)

These formulae follow from factorization of trigonometric and hyperbolic functions in terms of the gamma function:

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin \pi x}.$$  (3.20)

The phase term $e^{\mp \frac{i \omega \epsilon}{2\pi}}$ is introduced to cancel the bad asymptotics of the combination of gamma functions at large imaginary $\omega$, which can be inferred from the Stirling formula. Without these factors the functions $G_{\pm}(\omega)$ would exponentially grow in their respective domains of analyticity.

The inverse kernel of the integral equation $\hat{S}^{-1}(\omega)$ has a cut along the imaginary axis, due to the sign function, which we break in two parts by regularization (3.16). It also has simple poles at $\omega = \omega_n$ and $\omega = \bar{\omega}_n$, $n \neq 0$, where

$$\omega_n = \frac{2\pi \left( Mn + \pm |n| \right)}{M^2 + 1}.$$  (3.21)
The singularities of the inverse kernel, $\hat{S}^{-1}(\omega)$, are two cuts along positive and negative imaginary semi-axes, and simple poles at $\omega = \omega_n$ and $\omega = \bar{\omega}_n$. The functions $G_-$ and $G_+$ inherit singularities in respectively the upper and lower half-planes.

Figure 2: The singularities of the inverse kernel, $\hat{S}^{-1}(\omega)$, are two cuts along positive and negative imaginary semi-axes, and simple poles at $\omega = \omega_n$ and $\omega = \bar{\omega}_n$. The functions $G_-$ and $G_+$ inherit singularities in respectively the upper and lower half-planes.

The factorization (3.17) assigns the cut in the upper half-plane and the poles at $\omega = \omega_n$ to $G_-$, while $G_+$ has a cut in the lower half-plane and poles at $\omega = \bar{\omega}_n$ (fig. 2). The poles of $G_+$ lie on the straight lines that make an angle $\alpha = \arccot M$ with the real axis. After the Fourier transform back to $\xi$ space, the poles will create resonances. For generic $M$ the resonances are damped because $\text{Im}\omega_n \sim \text{Re}\omega_n$, but when $M$ becomes large, the poles pinch the real axis and cause oscillations in the $\xi$ space, with periods that are integer multiples of $M$. This behavior is a manifestation of the nearly massless hypermultiplets that we have discussed in sec. 2. We shall study the large-$M$ behavior of the solution in much detail in the next section.

Returning to the Wiener-Hopf equation, we can use factorization of the kernel to rewrite (3.12) as

$$\hat{g}_+(\omega) \over G_+(\omega) = G_-(\omega) \hat{F}(\omega) + G_-(\omega)X_-(\omega). \quad (3.22)$$

Defining the projection on the positive/negative-frequency part of a function
via contour integration:

\[ \mathcal{F}_+(\omega) = \pm \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi i} \frac{\mathcal{F}(\omega')}{\omega' - \omega + i\epsilon}, \]

(3.23)

we can project out the \(-\) term in the last equation, and thus find the solution to the Wiener-Hopf problem:

\[ \hat{g}_+(\omega) = G_+(\omega) (G_- \hat{F})_+(\omega). \]

(3.24)

It is important here that \(G_+(\omega)^{-1}\) is also an analytic function in the upper half-plane of complex \(\omega\), and thus the left hand side of \(\hat{g}_+(\omega)\) is a \(+\) function.

In order to compute the positive projection of \((G_- \hat{F})(\omega)\), let us discuss the analytic structure of this function first. Since the branch cut of \(G_-\) cancels in the product \(G_- \hat{F}\), the latter is a meromorphic function with simple poles at \(\omega = \omega_n, n = \pm 1, \pm 2, \ldots\) and double poles at \(\omega = -2\pi i m, m = 1, 2, \ldots\). All the double poles lie in the lower half-plane, and the integral in (3.23) can be done by closing the contour of integration in the upper half-plane and picking up the poles at \(\omega\) and \(\omega_n\):

\[ \hat{g}(\omega) = \frac{\hat{F}(\omega)}{S(\omega)} - G_+(\omega) \sum_{n \neq 0} \frac{\hat{F}(\omega_n)}{\omega - \omega_n} \text{res } G_-(z). \]

(3.25)

The first term is the naive Fourier transform that would solve the integral equation on the whole real axis. Since we need a solution identically equal to zero for \(\xi < 0\), its Fourier transform must be analytic in the upper half plane. The rôle of the last term is to subtract the singularities of the first term in order to make the solution a \(+\) function.

Explicitly, we get:

\[ \hat{g}(\omega) = \frac{i \omega^{1/2} \sqrt{\pi}}{2 \omega \sqrt{i \omega + i\epsilon}} \left[ \frac{M^2 \sinh^2 \frac{\omega}{2} - \sin^2 \frac{M\omega}{2}}{\sinh^2 \frac{\omega}{2} + \sin^2 \frac{M\omega}{2}} \right] \Gamma (\frac{M+i}{2\pi i} \omega) \Gamma (\frac{-M+i}{2\pi i} \omega) \]

\[ + \left( M^2 + 1 \right)^2 \omega e^{-\frac{\omega}{2\pi i}} \left[ \frac{\Gamma (\frac{M+i}{2\pi i} \omega) \Gamma (\frac{-M+i}{2\pi i} \omega) \Gamma (\frac{M-i}{2\pi i} \omega) \Gamma (\frac{-M-i}{2\pi i} \omega) \right] \]

\[ \times \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \left( \frac{e^{\frac{i\omega n}{M+i}}}{\omega - \frac{2\pi n}{M+i}} \frac{\Gamma (\frac{M+i}{M+i} \frac{n}{2})}{\Gamma^2 (\frac{M+i}{2\pi i} \frac{n}{M+i})} + \frac{e^{-\frac{i\omega n}{M+i}}}{\omega + \frac{2\pi n}{M+i}} \frac{\Gamma (\frac{M-i}{M+i} \frac{n}{2})}{\Gamma^2 (\frac{M-i}{2\pi i} \frac{n}{M+i})} \right) \]

(3.26)

This is our final expression which in general cannot be further simplified.
The solution in the $\xi$ space is the inverse Fourier transform of (3.26). Since $\hat{g}(\omega)$ has a relatively simple structure of singularities in the lower half-plane, its Fourier transform can be computed using the residue theorem, which results in a double infinite sum representation for $g(\xi)$. The final expression (A.5) and the details of the derivation are given in the appendix. This expression is very convenient for numerical evaluation of the function $g(\xi)$, but many quantities of interest, such as the Wilson loop expectation value, can be calculated directly from the Fourier representation.

The Wilson loop is an eigenvalue average with the exponential weight. Writing the defining equation (2.6) in terms of the endpoint variable $\xi$, and using (3.3) and (3.7), we get

$$W(C) = \frac{2^{3/2}}{\pi \mu^{3/2}} e^{2\pi\mu} \int_0^\infty d\xi \left( g(\xi) + \sqrt{\xi} \right) e^{-2\pi\xi},$$

(3.27)

where as before, the integration domain is extended to infinity, allowed by the strong coupling limit. This is a Laplace integral, and the part with $g(\xi)$ is simply $\hat{g}(2\pi i)$. Therefore, the Wilson loop is:

$$W(C) = \frac{2^{3/2}}{\pi \mu^{3/2}} e^{2\pi\mu} \left( \hat{g}(2\pi i) + \frac{1}{2^{5/2}\pi} \right).$$

(3.28)

The last term in the brackets is the contribution of the uncorrected Wigner distribution, which gives the prefactor quoted after (2.15). The Wiener-Hopf term $\hat{g}(2\pi i)$ is the correction produced by the distortion of the eigenvalue distribution near the endpoint. We study it in more detail in the next subsection.

### 3.3 General structure of solution

For moderate and small values of the mass $-M$ of order 1 and below $-g(\xi)$ is a featureless function whose asymptotic behavior is prescribed by the boundary conditions of the Wiener-Hopf problem:

$$g(\xi) \xrightarrow{\xi \to 0} B\sqrt{\xi}, \quad g(\xi) \xrightarrow{\xi \to \infty} \frac{C}{\sqrt{\xi}},$$

(3.29)

where the coefficients $B$ and $C$ depend on the mass. The plot in fig. 3 shows $g(\xi)$ at $M = 1/2$ for illustration.
In view of (3.3), (3.7), the constant $B$ determines the endpoint behavior of the eigenvalue density:

$$\rho(x) = \frac{2^{\frac{3}{2}} (1 + B)}{\pi \mu^{\frac{3}{2}}} \sqrt{\mu - x} \quad (x \rightarrow \mu). \quad (3.30)$$

The constant $C$ will be later used to compute $1/\sqrt{\lambda}$ corrections in the bulk of the eigenvalue distribution. We also introduce

$$A = 2^{5/2} \hat{g}(2\pi i) + 1, \quad (3.31)$$

which determines the normalization factor in the expectation value of the Wilson loop:

$$W(C) = \frac{A}{2\pi^2 \mu^{3/2}} e^{2\pi \mu}. \quad (3.32)$$

For pure Wigner distribution $A = 1$, but $A$ is also a non-trivial function of the mass.

The constants $B$ and $C$ can be read off from the asymptotic behavior of $\hat{g}(\omega)$:

$$\hat{g}(\omega) \xrightarrow{\omega \rightarrow \infty} \frac{i^{\frac{3}{2}} \sqrt{\pi} B}{2\omega^{\frac{3}{2}}}, \quad \hat{g}(\omega) \xrightarrow{\omega \rightarrow 0} \frac{\sqrt{\pi} C}{\sqrt{\omega}}. \quad (3.33)$$

The constant $A$ is also expressed explicitly in terms of $\hat{g}(\omega)$. Therefore to compute these constants we do not need to perform the inverse Fourier
transform back to the \( \xi \) representation. Explicitly:

\[
A = \frac{2\pi M (M^2 + 1)^2 \cos \phi}{\sinh \pi M} \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \left( e^{\frac{i\phi n}{M+i}} \frac{\Gamma\left(\frac{M+i}{M-i} n\right)}{\Gamma^2\left(\frac{i}{M-i} n\right)}\right) \tag{3.34}
\]

\[
B = M^2 + 2 (M^2 + 1)^{\frac{3}{2}} \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \left( e^{\frac{i\phi n}{M+i}} \frac{\Gamma\left(\frac{M+i}{M-i} n\right)}{\Gamma^2\left(\frac{i}{M-i} n\right)}\right) - \frac{M^2 + 1}{\pi} \arctan M \tag{3.35}
\]

\[
C = \frac{M^2 + 1}{2\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \left( e^{\frac{i\phi n}{M+i}} \frac{\Gamma\left(\frac{M+i}{M-i} n\right)}{\Gamma^2\left(\frac{i}{M-i} n\right)}\right) \tag{3.36}
\]

The origin of the last term in \( B \) is explained in appendix B. The first of these equations solves the problem of computing the prefactor in the Wilson loop expectation value.

For small \( M \), we can approximate

\[
e^{\frac{i\phi n}{M+i}} \frac{\Gamma\left(\frac{M+i}{M-i} n\right)}{\Gamma^2\left(\frac{i}{M-i} n\right)} \approx \frac{(-1)^n n!}{2\pi} \frac{\pi M n}{\tan \frac{\pi M n}{M-i}}. \tag{3.37}
\]

The constant \( A \) is then saturated by the \( n = 1 \) term, which develops a pole at \( M \to 0 \), while the main contributions to \( B \) and \( C \) come from terms in the sum with very large \( n \sim 1/M \). Replacing the sums by the integrals we get:

\[
A \approx 1, \quad B \approx \frac{M^2}{2}, \quad C \approx \frac{M^2}{4\pi} \quad (M \to 0). \tag{3.38}
\]

The constants \( A \) and \( B \) measure, in different ways, deviations from the Wigner distribution near the endpoint. We see that these deviations vanish in the \( M \to 0 \) limit.

Conversely, deviations from the naive strong coupling result grow with \( M \) and, as we shall see, become parametrically large at \( M \to \infty \). This limit is equivalent to the flat space limit, in which the sphere inflates to an infinite radius, corresponding to the top right hand corner of the phase diagram (fig. 1) approached from below. It is here that we hope to detect signs of non-trivial phase structure, making this regime particularly interesting. It turns out that the solution indeed develops such structure at large \( M \), which we investigate in detail in the next section.
Figure 4: $g(\xi)$ evaluated numerically for $M = 10$. See additional plots on page 26.

4 Decom pactification limit

One of our main motivations for solving the localization matrix model at strong coupling is to study the flat space/infinite mass limit, $MR \rightarrow \infty$. The corner of the phase diagram in fig. 1 is an accumulation point of an infinite number of phase transitions, and even though we approach the critical point from a different direction, we may expect to see signatures of the non-trivial phase structure in the eigenvalue density. There are three well separated scales in the problem in the decompactification limit: the IR cutoff scale $1/R$ (equal to 1 in the units we use), the mass scale $M$ and the symmetry breaking scale $\mu \sim \sqrt{\lambda} M$. The physics behind the phase transitions is governed by the second of these, and we expect the non-trivial structures to occur distance $\sim M$ away from the endpoints of the eigenvalue distribution. But $M \ll \mu$ and, in spite of the fact that $M \gg 1$, the distances of order $M$ are still within the range of the scaling limit applicable near the endpoints. We thus expect that most of the interesting phenomena associated with the phase transitions are described by the solution of the Wiener-Hopf problem. We thus need to study the large-$M$ limit of the solution.

As $M$ grows, the shape of the scaling function $g(\xi)$ dramatically changes. At first, a smooth profile similar to that in fig. 3 starts to be modulated with period approximately equal to $M$. At yet larger masses, the amplitude of modulation grows, and the scaling function develops a structure of regularly spaced peaks of diminishing amplitude (fig. 4). The peaks become sharper and sharper with growing $M$ and in the strict $M \rightarrow \infty$ limit morph into cusps.
of infinite height. This is precisely the phenomenon observed in [7, 8] where
the flat space theory was studied at arbitrary coupling and the appearance
of cusps was identified as a cause of the phase transitions.

We now have an analytic expression for the scaling function which is valid
at any $M$, and thus can study the large-$M$ limit rather explicitly. Moreover,
the complicated expression (3.26) simplifies somewhat in this limit. As should
be clear from the preceding discussion, there are two distinct regimes, of $\xi \sim 1$
and of $\xi \sim M$, which should be analyzed separately. It turns out that an
additional scale arises in the UV at $\xi \sim M^2$.

4.1 Small $\xi$

In the limit $M \to \infty$ and $\omega \sim O(1)$, the sums in (3.26) can be replaced by
integrals, which leads to massive cancellations. This is not unexpected, since
the solution to the Wiener-Hopf problem is designed so as to subtract the
singularities of the scaling function in the upper half of the complex plane.
These singularities are a series of poles (fig. 2), which at $M \to \infty$ collapse
onto the real axis and collide with the poles in lower half plane. In effect
all the singularities happen to be subtracted, leaving behind a subleading
contribution without poles, whose only singularity is the square-root cut in
the lower half-plane:

$$
\hat{g}(\omega) \approx \frac{M}{2} \sqrt{\frac{i \pi}{\omega}} e^{\frac{\pi}{2 \pi i} (\ln \frac{2 \pi i}{\omega} - 1)} \Gamma^2 \left(1 + \frac{\omega}{2 \pi i}\right) \quad (M \to \infty, \ \omega \sim 1). \quad (4.1)
$$

A derivation of this result is given in appendix C.

This limit describes the first peak in fig. 4. The Fourier transform of
(4.1) has more or less the same shape, shown in fig. 3, as the whole scaling
function at small $M$. The function grows as $M \sqrt{\xi}$, reaches a maximum
and then decays as $M/(2\sqrt{\xi})$. The important difference with the small-$M$
regime is that now the scaling function is parametrically big, proportional to
$M$, in contradistinction to small $M$ when the scaling function is just a small
correction to the leading Gaussian result. The fact that $g(\xi)$ is the leading
term now means that at large-$M$ the full density also has a peak; at smaller
$M$ the peak is diluted by the growth of the square root from the Gaussian
approximation, with density growing monotonically away from the endpoint.

Notice that we cannot extract the value of the constant $C$ defined in
(3.29), (3.33) from this calculation since we tacitly assume that $\omega \gg 1/M$. 

17
and consequently cannot take the limit $\omega \to 0$. To compute $C$, we need to consider the regime $\xi \sim M \iff \omega \sim 1/M$, which we do in the next section.

For the constants $A$ and $B$ defined in sec. 3.3 we get

$$A \approx \frac{2\pi M}{e^2}, \quad B \approx M \quad (M \to \infty). \quad (4.2)$$

These results can also be obtained by applying the formulae from appendix C directly to (3.34), (3.35). We see again that the deviations from the simple Gaussian model are parametrically large in the decompactification limit.

The regime described here completely determines the Wilson loop expectation value at strong coupling, including the normalization factor. The Wilson loop is dual to a fundamental string in the dual supergravity background. Consequently, from the matrix model point of view, the fundamental string probes the extreme vicinity of the endpoint of the eigenvalue distribution. The features in the eigenvalue density responsible for phase transitions are simply not visible to the fundamental string probes.

### 4.2 Oscillatory behavior

When $\omega \sim \mathcal{O}(1/M)$, corresponding to $\xi \sim \mathcal{O}(M)$, the sum is dominated by terms with small $n$. In this case we simply take the naïve large-$M$ limit of (3.26) (replacing eg. $M \pm i$ by $M$ wherever it appears). For instance, the prefactor in front of the sum in (3.26) becomes

$$\frac{M^3 \omega^2}{2 \sin \frac{M \omega}{2}} e^{-\frac{i \omega M}{2}}, \quad (4.3)$$

where we have used the identity (3.20) for the gamma functions.

Computation of the sum involves one subtlety. It turns out that we need to keep the leading correction to the phase $\phi$:

$$\epsilon \equiv \pi - \frac{\phi}{M} \approx \frac{2 \ln M}{M},$$

vanishing in the $M \to \infty$ limit. Yet this quantity leaves a finite imprint in the final answer. For the sum we get

$$\text{sum} \approx -\frac{1}{M^2} \sum_{n=1}^{\infty} \left( \frac{e^{-i \epsilon n}}{\omega - \frac{2\pi n}{M}} + \frac{e^{i \epsilon n}}{\omega + \frac{2\pi n}{M}} \right) = \frac{1}{\pi M} \sum_{n=1}^{\infty} \frac{\frac{M \omega}{2\pi} \cos \epsilon n - i \sin \epsilon n}{n^2 - \left(\frac{M \omega}{2\pi}\right)^2} \quad (4.4)$$
In the \textit{cos} term we can set $\epsilon = 0$ right away, because the sum converges. The \textit{sin} term does not contribute at first sight, but the sum converges slowly, and the limit $\epsilon \to 0$ does not commute with summation. Indeed,

$$
\lim_{\epsilon \to 0} \sum_{n=1}^{\infty} \frac{\sin \epsilon n}{n} = \frac{\pi}{2},
$$

and not zero. Taking this into account, we find that

$$
\text{sum} = \frac{1}{M^2 \omega} - \frac{e^{iM\omega/2}}{2M \sin M \omega/2}.
$$

Combining all terms together, we get for the scaling function in the large-mass limit:

$$
\hat{g}(\omega) \approx \frac{i^{3/2} \sqrt{\pi}}{2\omega \sqrt{\omega + i\epsilon}} \left[ \frac{M \omega e^{-iM\omega/2}}{2 \sin M \omega/2} - 1 \right].
$$

Some remarks are in order here. In the limit we are considering, the poles of the Green’s functions $G_{\pm}$, shown in fig. 2, collapse onto the real line and merge pairwise. The unprojected part of the solution (3.26) as a consequence has double poles in this limit. But the poles of the exact scaling function in the upper half-plane are eliminated by the $+$ projection, and we expect the limiting scale function to have only single poles ascending from the lower half-plane. This is exactly what happens – the double poles get cancelled and the limiting solution has a sequence of single poles along the real line. Their origin in the lower half-plane defines the epsilon-prescription for integrating the scaling function over frequencies.

The contour of integration in the inverse Fourier transform (3.11) thus passes all the poles from above. For $\xi > 0$, the contour can be closed in the lower half-plane, picking up the poles along the real axis and wrapping the branch cut along the negative imaginary axis. The scaling function becomes a sum of two terms:

$$
g(\xi) = h_0(\xi) + h_1(\xi)
$$

Let us consider the branch cut contribution first:

$$
h_1(\xi) = \frac{\sqrt{M}}{2\sqrt{\pi}} \int_{0}^{\infty} du \frac{e^{-u \xi}}{u^{3/2}} \left( 1 - \frac{u}{e^u - 1} \right) = -\sqrt{\xi} - \sqrt{\frac{M}{2}} \zeta \left( \frac{1}{2}, 1 + \frac{\xi}{M} \right),
$$

where $\zeta(s, a)$ is the Hurwitz zeta function. Despite its appearance, the function $h_1(\xi)$ is always positive, starts off as a constant at $\xi = 0$ and decays monotonically with $\xi$, asymptoting to the $1/\sqrt{\xi}$ tail at infinity.
Figure 5: (a) The scaling function \( g(\xi) \) (upper, blue curve) and the periodic functions \( h_0(\xi) \) (lower, purple curve). At larger \( \xi \), \( h_0(\xi) \) becomes better and better approximation to \( g(\xi) \). (b) The eigenvalue density near the endpoint forms a comb-like structure with an infinite series of resonances on top of the leading-order square-root distribution, shown as a purple curve.

The poles give

\[
h_0(\xi) = -i \sum_{n \neq 0} \text{Res} \left[ g(\omega) e^{-i\omega \xi}, \ \omega = \frac{2\pi n}{M} \right] = \sqrt{\frac{iM}{8}} \sum_{n \neq 0} \frac{e^{-2\pi i n\xi/M}}{\sqrt{n}}.
\]  

(4.10)

This function is obviously periodic with period \( M \). In order to expose the periodicity, it is convenient to decompose \( \xi/M \) on the integer and fractional parts:

\[
\left\{ \frac{\xi}{M} \right\} = \frac{\xi}{M} \mod 1, \quad \left[ \frac{\xi}{M} \right] = \frac{\xi}{M} - \left\{ \frac{\xi}{M} \right\},
\]  

(4.11)

keeping in mind that \( h_0 \) only depends on the fractional part. The sum can again be expressed in terms of the Hurwitz zeta function:

\[
h_0(\xi) = \frac{\sqrt{M}}{2} \zeta \left( \frac{1}{2}, \left\{ \frac{\xi}{M} \right\} \right).
\]  

(4.12)

Since \( h_0(\xi) \) is periodic and \( h_1(\xi) \) decreases at infinity, their sum asymptotes to \( h_0(\xi) \) at \( \xi \gg M \). This is illustrated in fig. 5(a). The sum of \( h_0 \) and \( h_1 \) can be actually simplified with the help of the zeta-function identities:

\[
g(\xi) = \frac{\sqrt{M}}{2} \sum_{k=0}^{\left[ \frac{\xi}{M} \right]} \frac{1}{\sqrt{k + \left\{ \frac{\xi}{M} \right\}}} - \sqrt{\xi},
\]  

(4.13)
Figure 6: Cusp-like structure of \( g(\xi) \) in the regime \( \xi \sim \mathcal{O}(M) \) for large \( M \). Here we compare our analytic result (4.13) with numerics (red) for \( M = 100 \).

and, as we can see in fig. 6, it agrees well with the numeric evaluation of the sum (3.26) for \( M = 100 \). As for the scaling form of the density, we get a particularly simple expression:

\[
f(\xi) = \sqrt{\frac{M^2}{2}} \sum_{k=0}^{\left\lfloor \frac{\xi}{M} \right\rfloor} \frac{1}{\sqrt{\left\{ \frac{\xi}{M} \right\} + k}}.
\]  

(4.14)

The function \( h_0(\xi) \), and with it the scaling function \( g(\xi) \), blows up as \( M/(2\sqrt{\xi}) \) at \( \xi \to 0 \). This behavior matches with the \( 1/\sqrt{\xi} \) tail at the upper end of the small-\( \xi \) regime, which describes the first peak of the density. The change of the endpoint exponent from \( +1/2 \) to \( -1/2 \) is characteristic of the infinite-volume limit, and is universally observed in all massive theories that can be solved on \( S^4 \) by localization \([17, 7, 8, 9]\). But \( h_0(\xi) \) is also periodic, and consequently has inverse square-root singularities in all integer points \( \xi = nM \). These are the resonances that arise due to the presence of nearly massless hypermultiplets in the spectrum. Our analysis applies to strictly infinite coupling. Varying the coupling will cause the resonances to move, resulting in phase transitions each time a full interval is traversed and a new cusp (dis)appears in the density function.

The density behaves as

\[
f(\xi) \approx \frac{M}{2\sqrt{\xi - nM}} \quad (\xi \to nM^+) ,
\]  

(4.15)
to the right of each resonance, and approaches a finite limiting value from 
the left. This structure is qualitatively similar to the one previously observed 
at finite coupling in the vicinity of the first phase transition [8]. But now we 
have an analytic solution that describes the whole resonance structure.

To move beyond the regime $\xi \sim \mathcal{O}(M)$, we should recall that the poles in 
$\hat{g}(\omega)$ are really located slightly off the real axis, at $\omega = \pm \frac{2\pi n}{M^2}$. This displacement means the phase in the definition (4.10) of $h_0(\xi)$ acquires an imaginary 
component $2\pi i/\xi$, which is unimportant for $\omega \sim \mathcal{O}(1/M)$, but for large 
$\xi \sim \mathcal{O}(M^2)$ causes the peaks to decay exponentially. At these large scales the 
oscillations in the density die out and the leading contribution is the $1/\sqrt{\xi}$ 
tail of the function $h_1(\xi)$:

$$h_1(\xi) \approx \frac{M}{4\sqrt{\xi}} \quad (\xi \to \infty)$$ \hspace{1cm} (4.16)

This behavior determines the constant $C$ defined in (3.29). Comparing (4.7) 
at $\omega \to 0$ to (3.33) we find:

$$C \approx \frac{M}{4} \quad (M \to \infty).$$ \hspace{1cm} (4.17)

This result can also be obtained directly from (3.36) by methods outlined in 
appendix C.

To summarize, the overall picture of the eigenvalue density that emerges 
in the decompactification limit is as follows.

- Square root behavior at the extreme endpoint given by

$$\rho(\xi) = \frac{2^3}{\pi \mu^2} M\sqrt{\xi} \quad (\xi \sim 1).$$ \hspace{1cm} (4.18)

The density reaches a peak while $\xi \ll M$ (corresponding to the overall 
peak in the small mass solution), and decays thereafter as $M/(2\sqrt{\xi})$.

- When the density is scaled to larger $\xi \sim M$, the peak is not resolved any 
more and becomes a cusp, thus changing the endpoint behavior of the 
density from $\sqrt{\xi}$ to $1/\sqrt{\xi}$. Moreover, the density develops secondary 
cusps at the resonance points $\xi = nM$, and thus acquires a comb-like 
shape with cusps separated by $M$ which are superimposed on the leading 
square root function. We were able to find the precise analytic form
of the density in this regime:

\[ \rho(\xi) = \frac{\sqrt{2M}}{\pi \mu^{3/2}} \sum_{k=0}^{\left\lfloor \frac{\xi}{M} \right\rfloor} \frac{1}{\sqrt{\left\{ \frac{\xi}{M} \right\} + k}}. \]  

(4.19)

- At yet larger \( \xi \), of order \( M^2 \), the amplitude of these resonances decays exponentially, leaving a small \( 1/\sqrt{\xi} \) correction to the leading-order Wigner distribution.

- In the bulk of the eigenvalue distribution the density is given by (2.16). The constant \( \beta \) that controls the overall size of the correction is determined in the next section.

5 Strong-coupling expansion

We now return to the question of \( 1/\sqrt{\lambda} \) corrections. As we have shown before, the functional form of the correction to the density is fixed by the saddle-point equations, but its overall normalization is not. The normalization constant can be determined by matching the bulk density to the exact scaling solution near the endpoints of the distribution.
Using the parametrization $\xi = \mu - x$, the bulk solution \((2.16)\) takes the following form near the endpoint:

$$
\rho_{\text{bulk}} = \frac{2^\frac{3}{2}}{\pi \mu^\frac{3}{2}} \left( \sqrt{\xi} + \frac{\pi \beta}{4\sqrt{\xi}} \right),
$$

(5.1)

where we used \((2.14)\) for the endpoint position. This has to match the asymptotic behavior of the endpoint solution at large $\xi$, for which we get combining \((3.3), (3.7)\) with \((3.29)\):

$$
\rho_{\text{end}} = \frac{2^\frac{3}{2}}{\pi \mu^\frac{3}{2}} \left( \sqrt{\xi} + g(\xi) \right) \to \frac{2^\frac{3}{2}}{\pi \mu^\frac{3}{2}} \left( \sqrt{\xi} + \frac{C}{\sqrt{\xi}} \right).
$$

(5.2)

Comparing the two expressions we conclude that

$$
\beta = \frac{4C}{\pi}.
$$

(5.3)

For the bulk eigenvalue density we thus get

$$
\rho_{\text{bulk}}(x) = \frac{8\pi}{\lambda (1 + M^2)} \sqrt{\mu^2 - x^2} + \frac{4C}{\pi \mu \sqrt{\mu^2 - x^2}},
$$

(5.4)

The constant $C$ is given by an infinite sum \((3.36)\), which at small and large $M$ asymptotes to

$$
C \simeq \frac{M^2}{4\pi} \quad (M \to 0), \quad C \simeq \frac{M}{4} \quad (M \to \infty).
$$

(5.5)

We can now determine the correction to the endpoint position, by imposing the normalization condition on the density:

$$
\int_{-\mu}^{\mu} dx \rho_{\text{bulk}}(x) = 1,
$$

(5.6)

which becomes

$$
\frac{4\pi^2 \mu^2}{\lambda (M^2 + 1)} + \frac{4C}{\mu} = 1,
$$

(5.7)

and we find:

$$
\mu = \frac{\sqrt{\lambda (1 + M^2)}}{2\pi} - 2C + O(\lambda^{-1/2}).
$$

(5.8)
The first strong-coupling correction to $\mu$ is thus simply related to the slope of the scaling function at large $\xi$.

This latter relation can be checked by numerically solving the exact saddle-point equation (2.7). The method we used becomes unstable at strong coupling unless $M$ is sufficienly small, so we restricted numerical analysis to $M < 0.5$, where we can use the small-$M$ asymptotics (3.38) for $C$. We fitted the data points to:

$$\mu(M) = a\sqrt{M^2 + 1} + bM^2.$$  \hspace{1cm} (5.9)

The results are shown in fig. 8 and are in perfect agreement with (5.8):

| Analytic Values | Numerical Fit   |
|-----------------|----------------|
| a               | 44.7214        |
|                 | $44.7221 \pm 0.0004$ |
| b               | -0.1591        |
|                 | $-0.1593 \pm 0.0045$ |

In fig. 9 we compare the numerical results for the density with the Wiener-Hopf solution in the endpoint region.

The correction to the endpoint position affects the normalization of the Wilson loop expectation value, through its exponential dependence on $\mu$ in (3.32). Taking this correction into account, we get for the Wilson loop expectation value:

$$W(C) \approx \sqrt{\frac{2}{\pi}} \frac{A e^{-4\pi C}}{\lambda^\frac{3}{4} (1 + M^2)^{\frac{3}{4}}} e^{\sqrt{\lambda (1 + M^2)}},$$  \hspace{1cm} (5.10)
where the constants \( A \) and \( C \) are given in (3.34), (3.36). In the decompactification limit, we get:

\[
W(C) \simeq \sqrt{\frac{8\pi}{MR}} \lambda^{-\frac{3}{4}} e^{\left(\sqrt{\lambda-\pi}\right)MR^{-2}},
\]  

(5.11)

where we have reinstated the dependence on the radius of the four-sphere through the rescaling \( M \to MR \).

The leading exponential term should be universal, implying that any sufficiently big Wilson loop in the \( \mathcal{N} = 2^* \) theory on \( \mathbb{R}^4 \) should obey the perimeter law:

\[
\ln W(C) = P(\lambda)ML \quad (ML \gg 1),
\]  

(5.12)

where \( L \) is the length of the contour \( C \). The coefficient \( P(\lambda) \) governs the self-energy of an infinitely heavy quark immersed in the \( \mathcal{N} = 2^* \) vacuum. From (5.11) we find that at strong coupling the self-energy coefficient behaves as

\[
P(\lambda) = \frac{\sqrt{\lambda}}{2\pi} - \frac{1}{2} + O\left(\frac{1}{\sqrt{\lambda}}\right).
\]  

(5.13)

The leading order term was computed in [6] and successfully compared to the area law in the Pilch-Warner geometry. The second term constitutes a prediction for the first quantum correction to the minimal area. It should be possible to compute this correction by semiclassical quantization of the string dual to the straight Wilson line.
The prefactor in (5.11) is in principle calculable by quantizing the string suspended on the big circle in the spherical geometry. The supergravity solution in this case is also known [10], but unfortunately only in the five-dimensional form. To compute the Wilson loop one needs to know the string action, determined by the ten-dimensional uplift of the solution.

The prefactor in the Wilson loop expectation value (5.11) is a contour-dependent quantity. The known ten-dimensional dual of the $\mathcal{N} = 2^*$ theory on $\mathbb{R}^4$ is of little help for computing this number. However, the square-root scaling of the prefactor with the size of the contour may be universal and apply to any sufficiently big Wilson loop. The leading finite-size correction to (5.12) is then logarithmic with precisely known coefficient: $\delta_{\text{fin.size}} \ln W(C) = - (1/2) \ln L$.

6 Conclusions

We studied the strong-coupling planar limit of $\mathcal{N} = 2^*$ theory compactified on $S^4$. The supergravity approximation should be accurate in this regime, and since the supergravity dual of $\mathcal{N} = 2^*$ SYM on $S^4$ is explicitly known [1], our calculations can potentially be compared to semiclassical string theory on the dual supergravity background. The exponent in the Wilson loop expectation value (5.10) should then correspond to the area of the surface bounded by the big circle of $S^4$. The prefactor can be identified with the one-loop determinant due to string fluctuations around the minimal surface.

Interestingly, all the non-trivial features that appear at strictly infinite volume, such as consecutive phase transitions, are visible already at the first order of the strong-coupling expansion, and it would be really interesting to explore their counterparts on the string theory side.

As we have seen, the Wilson loop expectation value is sensitive to the immediate vicinity of the endpoint in the eigenvalue distribution, namely to distances of order one in $x$ space. In contrast, the spikes that arise in the decompactification limit are located at distances of order $O(M)$. Thus Wilson loops, or semiclassical strings, are not sensitive to the phase transitions, simply because they probe a different corner in the parameter space.

Perhaps better probes are D-branes. It is known that the eigenvalue distribution as a whole can be derived from the D-brane probe analysis [15]:

---

1 An explicit analytic solution is known at infinite radius of the sphere [3], otherwise the problem reduces to a set of ODEs that can be integrated numerically [10].
the question is to resolve the region distance $M$ away from the endpoints, where the density has non-trivial features in the decompactification limit. We emphasize that while the decompactification limit implies that $M \gg 1$, it always remains true that $M \ll \mu$ because of the strong coupling. The D-brane probe analysis identifies the eigenvalue distribution with a particular locus in the dual geometry. We can now pinpoint exactly which parts of the eigenvalue distribution are responsible for the phase transitions, and we can even compute the fine structure of the eigenvalue density in this region. We can thus say that the critical behavior is associated with a particular location of the 10d space-time. It would be very interesting to understand what triggers the phase transitions in string theory.

**Acknowledgments**

We would like to thank L. Anderson and J. Russo for discussions. This work was supported by the Marie Curie network GATIS of the European Union’s FP7 Programme under REA Grant Agreement No 317089. The work of K.Z. was supported by the ERC advanced grant No 341222 and by the Swedish Research Council (VR) grant 2013-4329.

**A Exact expression for $g(\xi)$**

Here, we are going to derive a sum representation for $g(\xi)$, which is useful for studying the function numerically.

The Fourier integral (3.11) can be solved by applying the residue theorem, where we sum over the remaining poles of $\hat{g}(\omega)$ in the lower half plane. Notice that we also have a branch cut in the lower complex half-plane, due to the $1/\sqrt{\omega + i\epsilon}$ term in our solution (3.26). Our strategy is to inverse Fourier transform the branch cut and the part with poles separately. In the coordinate space, the final expression is then the convolution of these terms.

Let us factorize $\hat{g}(\omega) \equiv \hat{a}(\omega)\hat{b}(\omega)$, with $\hat{a}(\omega) = \frac{\sqrt{i\pi}}{\sqrt{\omega + i\epsilon}}$. In the coordinate space, the latter becomes:

$$a(\xi) = \frac{1}{\sqrt{\xi}}\theta(\xi),$$  \hfill (A.1)

where $\theta(\xi)$ is the Heaviside step function.
Regarding \( \hat{b}(\omega) \), the only poles in the lower half plane are due to

\[
\Gamma\left( -\frac{(M + i)\omega}{2\pi} \right) \Gamma\left( \frac{(M - i)\omega}{2\pi} \right),
\]

which comes from the expression for \( G_+ (\omega) \), (3.18). These poles are the complex conjugate of (3.21), i.e.

\[
\bar{\omega}_n = \frac{2\pi (Mn - i|n|)}{M^2 + 1}, \quad n = \pm 1, \pm 2, \ldots \quad (A.2)
\]

Applying the residue theorem to \( \hat{b}(\omega) \), we obtain:

\[
b(\xi) = -i \sum_{n=-\infty}^{\infty} \text{Res} \left( \hat{b}(\omega) e^{-i\bar{\omega}_n \xi}, \bar{\omega}_n \right) \theta(\xi).
\]

Hence,

\[
g(\xi) = \int_{-\infty}^{\infty} d\eta \, a(\eta) b(\xi - \eta)
\]

\[
= \int_{0}^{\xi} d\eta \frac{1}{\sqrt{\xi}} \sum_{n=-\infty}^{\infty} \text{Res} \left( \hat{b}(\omega), \bar{\omega}_n \right) e^{-i\bar{\omega}_n (\xi - \eta)}
\]

As the sum is convergent, the integral and the sum commute. The integral gives:

\[
\int_{0}^{\xi} d\eta \frac{1}{\sqrt{\eta}} e^{-i\bar{\omega}_n (\xi - \eta)} = e^{-i\bar{\omega}_n \xi} \sqrt{\frac{i\pi}{\bar{\omega}_n}} \text{erf} \left( \sqrt{-i\bar{\omega}_n \xi} \right) \quad (A.4)
\]

The residue term for \( n = 1, 2, \ldots \) is explicitly:

\[
\text{Res} \left( \hat{b}(\omega), \bar{\omega}_n \right) = -\frac{(M^2 + 1) (-1)^n}{2\pi n} \frac{\sinh^2 \left( \frac{\pi n}{M+i} \right)}{\sinh \left( \frac{\pi(1+M)n}{M+i} \right)}
\]

\[
- \frac{(M^2 + 1)^2 (-1)^n i\pi}{M + i} \frac{\Gamma \left( \frac{(M+i)n}{M-i} \right)}{n! \Gamma \left( \frac{i\pi}{M-i} \right)^2} e^{\frac{in\phi}{M-i}} A \left( \frac{2\pi n}{M+i} \right)
\]

where \( A(\omega) \) is the sum in (3.26). For negative values of \( n \), the residue is the negative complex conjugate of the expression above. Hence we need only the imaginary part of the positive \( n \) sum. The final expression for \( g(\xi) \) is thus:

\[
g(\xi) = 2 \sum_{n=1}^{\infty} J \left[ \text{Res} \left( \hat{b}(\omega), \bar{\omega}_n \right) e^{-i\bar{\omega}_n \xi} \sqrt{\frac{i\pi}{\bar{\omega}_n}} \text{erf} \left( \sqrt{-i\bar{\omega}_n \xi} \right) \right] \quad (A.5)
\]


**B Anomalous contribution to B**

In computing $B$ as defined in (3.33), we need to take the $\omega \to \infty$ limit of the scaling function (3.26). The naive limit gives the first two terms in (3.35), but (3.26) contains an infinite sum and one has to be careful and do the summation first, before taking $\omega \to \infty$. It turns out that the summation and taking the limit do not commute, and $B$ receives an anomalous contribution.

To isolate this contribution we can divide the sum into two parts, from 1 to $N_0$ and from $N_0$ to infinity for some $N_0 \geq 1$. The anomalous contribution can only come from the second part:

$$\delta B_{\text{anom}} = \lim_{\omega \to \infty} \left( \frac{1}{\omega} \sum_{n=1}^{\infty} \frac{1}{\omega - \frac{2\pi n}{M-i}} - \frac{1}{\omega + \frac{2\pi n}{M+i}} \right)$$

Here we used that $n \geq N_0 \gg 1$ to simplify the summand. This expression can also be written as

$$\delta B_{\text{anom}} = i \left( M^2 + 1 \right) \lim_{\omega \to \infty} \sum_{n=N_0}^{\infty} \left[ \frac{1}{(M-i)\omega - 2\pi n} + \frac{1}{(M+i)\omega + 2\pi n} \right].$$

The naive $\omega \to \infty$ limit would give zero, but we need to first sum and then take the limit, and this gives a finite result:

$$\delta B_{\text{anom}} = i \left( M^2 + 1 \right) \lim_{\omega \to \infty} \frac{2\pi N_0}{2\pi N_0 + (M+i)\omega} = \frac{-M^2 + 1}{\pi} \arctan M.$$  

**C Large $M$ limit of scaling function**

Here we give the details on the derivation of the limiting expression (4.1) for the scaling function from the exact one (3.26), in the case when $M \to \infty$ and $\omega$ stays finite. We assume that $\omega$ is real throughout the derivation.

We start by examining the infinite sums appearing in (3.26):

$$A_{\pm} = \sum_{n=1}^{\infty} \frac{a_{\pm} \left( \frac{n}{M \pm i} \right)}{\omega \pm \frac{2\pi n}{M \pm i}},$$

where

$$a_{\pm}(x) = \frac{e^{\mp i(M \pm i)\pi x}}{(M \pm i)^2 x^2} \frac{\Gamma((M \mp i)x)}{\Gamma((M \pm i)x) \Gamma^2(\mp i x)}.$$
In terms of these sums,

\[
\hat{g}(\omega) = i \frac{3}{2} \sqrt{\pi \omega} \left[ \frac{M^2 \sinh^2 \frac{\omega}{2} - \sin^2 \frac{M\omega}{2}}{\sinh^2 \frac{\omega}{2} + \sin^2 \frac{M\omega}{2}} \right] + \left( M^2 + 1 \right)^2 \omega e^{\frac{i\omega}{2\pi}} \Gamma \left( \frac{M+i\omega}{2\pi} \right) \Gamma \left( -\frac{M+i\omega}{2\pi} \right) \Gamma \left( \frac{M+i\omega}{2\pi} \right) \Gamma \left( -\frac{M+i\omega}{2\pi} \right) \left( A_- + A_+ \right). \tag{C.3}
\]

Since \( a_\pm(x) \) has a finite limiting value at zero:

\[
a_\pm(0) = -\frac{1}{M^2 + 1}, \tag{C.4}
\]

the sums \( \text{(C.1)} \) appear linearly divergent if \( M \) is sent to infinity independently in each term. The main contribution consequently comes from very large \( n \sim M \), because then \( a_\pm(x) \) become slowly varying functions of their argument, namely

\[
a_\pm(x) \xrightarrow{M \to \infty} e^{\pm 2i\pi \left( \ln(x) + 1 \right)} \frac{M^2 x^2 \Gamma}{(\mp i\pi)}, \tag{C.5}
\]

assuming \( x > 0 \). We are not going to use these approximate expressions, because of the necessity to keep the next-to-leading order accuracy. It will suffice to know that the exact \( a_\pm(z) \) is an analytic function in the upper half plane, decreases as \( 1/z \) in its domain of analyticity, and satisfies the following functional identity:

\[
(M + i)a_+(x) = (M - i)a_-(x) \frac{\sin \pi (M + i)x}{\sin \pi (M - i)x} e^{2\pi x}. \tag{C.6}
\]

Normally, the sum of \( f(n/M) \), where \( M \) is a big parameter, is well approximated by the integral with the help of the Euler-Maclaurin formula, but here we need to be more careful. The summands in \( \mathcal{A}_\pm \) have poles at \( 2\pi n/(M \pm i) = \omega \) that collapse onto the contour of integration in the \( M \to \infty \) limit. In the vicinity of the poles the summand is not a slowly varying function of \( n/M \), and the summation has to be performed exactly. As a result a more general formula applies:

\[
\mathcal{A}_\pm \simeq (M \pm i) \int_0^{\infty} \frac{dx \ a_\pm(x)}{\omega \pm 2\pi x} - \frac{a_\pm(0)}{2\omega} + \frac{M \pm i}{2} \theta(\mp \omega) a_\pm \left( \frac{\mp \omega}{2\pi} \right) \cot \left( \frac{(M \pm i)\omega}{2} \right). \tag{C.7}
\]
The second term is the Euler-Maclaurin correction, which we need to make this formula correct throughout the next-to-leading order. The last term is the result of summation around $n \sim M \omega/2\pi$. We excluded this region from the integral by the principal-value prescription, to avoid double-counting. This formula can be viewed as a contour-deformation prescription that takes into account the discreteness of the sum in the residue term. The formula can be brought to the form

$$A_{\pm} \approx (M \pm i) \int_{0}^{\infty} \frac{dx}{\omega + 2\pi x + i\epsilon} \frac{a_{\pm}(x)}{\omega} + \frac{M \pm i}{2} \theta(\pi \omega) a_{\pm} \left( \frac{\pi \omega}{2\pi} \right) e^{i \frac{M \pi \omega}{2\omega}} \sin \frac{M x}{\omega},$$  \hspace{1cm} (C.8)

that facilitates rotation of the contour of integration into the domain of analyticity of the integrand.

Substitution of (C.8) into (C.3) leads to massive cancellations. Let us first concentrate on the integral terms. Using the identity (C.6) we can transform their sum as

$$\begin{align*}
(M + i) \int_{0}^{\infty} \frac{dx}{\omega + 2\pi x + i\epsilon} a_{+}(x) + (M - i) \int_{0}^{\infty} \frac{dx}{\omega - 2\pi x + i\epsilon} a_{-}(x) \\
= (M + i) \int_{-\infty}^{\infty} \frac{dx}{\omega + 2\pi x + i\epsilon} a_{+}(x) \\
+ (M - i) \int_{0}^{\infty} \frac{dx}{\omega - 2\pi x + i\epsilon} \left( 1 - \frac{\sin(\pi M x + ix)}{\sin(\pi M x - ix)} e^{ix} \right).
\end{align*}$$ \hspace{1cm} (C.9)

We are going to argue that both terms are negligible in the large-$M$ limit. The first integral actually vanishes identically, which follows from the contour argument since the intergand has no singularities in the upper half plane. The second integral contains a rapidly oscillating function that depends on the slow variable $x$ and the fast variable $\pi M x$. The dependence on the fast variable is periodic, and integration over $x$ goes through many periods of oscillations before the slow dependence on $x$ can substantially alter the integrand. In any integral of this type, the integrand $\mathcal{F}(\pi M x, x)$ can be replaced by its average:

$$\int dx \mathcal{F}(\pi M x, x) \approx \int dx \langle \mathcal{F}(\Omega, x) \rangle, \quad \langle \mathcal{F}(\Omega, x) \rangle \equiv \int_{0}^{2\pi} \frac{d\Omega}{2\pi} \mathcal{F}(\Omega, x).$$ \hspace{1cm} (C.10)
It is easy to show that
\[
\left( 1 - \frac{\sin(\Omega + ix)}{\sin(\Omega - ix)} e^{2\pi x} \right) = 0
\]
for \( x > 0 \). We can therefore drop the last integral in (C.9) too.

A relatively long but straightforward calculation shows that the residue term in (C.8) cancels with the first term in the square brackets in (C.3) with the requisite, \( O(M^0) \) accuracy. Thus only the Euler-Maclaurin term contributes and we are left with
\[
\hat{g}(\omega) = \frac{i^3 \sqrt{\pi} M^2}{2\omega^2} e^{-\frac{i\omega}{2\pi}} \frac{\Gamma \left( \frac{M-i\omega}{2\pi} \right) \Gamma \left( -\frac{M+i\omega}{2\pi} \right)}{\Gamma^2 \left( \frac{1+i\omega}{2\pi} \right)}.
\]
(C.11)

Using the basic gamma-function identity (3.20), and the Stirling formula we get at large \( M \):
\[
\hat{g}(\omega) \sim \frac{i^3 \sqrt{\pi} M}{4\sqrt{\omega}} \frac{e^{\frac{\omega}{\pi} \left( \ln \frac{\omega}{2\pi} - 1 \right)}}{\Gamma^2 \left( 1 + \frac{\omega}{2\pi} \right)} \frac{e^{-\frac{iM\omega}{2} + \frac{i|\omega|}{2}}}{\sin \left( \frac{M\omega}{2} + \frac{i|\omega|}{2} \right)}.
\]
(C.12)

The last factor still depends on \( M \) through the periodic dependence on the fast variable \( M\omega/2 \). The Fourier transform back to the \( \xi \)-representation will effectively average over rapid oscillations, so the last factor can be replaced with
\[
\left\{ \frac{e^{-\frac{i\Omega + |\omega|}{2}}}{\sin \left( \Omega + \frac{i|\omega|}{2} \right)} \right\} = \frac{2}{i},
\]
(C.13)
which gives the final result (4.1) quoted in the main text.

References

[1] E. Witten, “Topological Quantum Field Theory”, Commun.Math.Phys. 117, 353 (1988).

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

[3] K. Pilch and N. P. Warner, “N=2 supersymmetric RG flows and the IIB dilaton”, Nucl.Phys. B594, 209 (2001), hep-th/0004063.
[4] E. Brezin, C. Itzykson, G. Parisi and J. B. Zuber, “Planar Diagrams”, Commun. Math. Phys. 59, 35 (1978).

[5] J. G. Russo, “A Note on perturbation series in supersymmetric gauge theories”, JHEP 1206, 038 (2012), 1203.5061

[6] 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.

[7] J. G. Russo and K. Zarembo, “Evidence for Large-N Phase Transitions in N=2* Theory”, JHEP 1304, 065 (2013), 1302.6968.

[8] J. Russo and K. Zarembo, “Massive N=2 Gauge Theories at Large N”, JHEP 1311, 130 (2013), 1309.1004.

[9] J. Russo and K. Zarembo, “Localization at Large N”, 1312.1214

[10] N. Bobev, H. Elvang, D. Z. Freedman and S. S. Pufu, “Holography for N = 2* on S^4”, JHEP 1407, 001 (2014), 1311.1508.

[11] D. Gross and E. Witten, “Possible Third Order Phase Transition in the Large N Lattice Gauge Theory”, Phys.Rev. D21, 446 (1980). • S. R. Wadia, “A Study of U(N) Lattice Gauge Theory in 2-dimensions”, 1212.2906.

[12] A. Barranco and J. G. Russo, “Large N phase transitions in supersymmetric Chern-Simons theory with massive matter”, JHEP 1403, 012 (2014), 1401.3672

[13] L. Anderson and K. Zarembo, “Quantum Phase Transitions in Mass-Deformed ABJM Matrix Model”, 1406.3366 • J. G. Russo, G. A. Silva and M. Tierz, “Supersymmetric U(N) Chern-Simons-matter theory and phase transitions”, 1407.4794

[14] J. A. Minahan and A. Nedelin, “Phases of planar 5-dimensional supersymmetric Chern-Simons theory”, 1408.2767

[15] 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.

[16] M. Billo, M. Frau, F. Fucito, A. Lerda, J. Morales et al., “Modular anomaly equations in N=2* theories and their large-N limit”, 1406.7255

[17] J. Russo and K. Zarembo, “Large N Limit of N=2 SU(N) Gauge Theories from Localization”, JHEP 1210, 082 (2012), 1207.3806