

Structure of Supersymmetric Gauge Theories

Exact instanton expansion of the ABJM partition functionYasuyuki Hatsuda^{1,*}, Sanefumi Moriyama^{2,*}, and Kazumi Okuyama^{3,*}¹*DESY Theory Group, DESY Hamburg, Notkestrasse 85, D-22603 Hamburg, Germany*²*Department of Physics, Osaka City University, Osaka 558-8585, Japan*³*Department of Physics, Shinshu University, Matsumoto 390-8621, Japan*

*E-mail: yasuyuki.hatsuda@desy.de, moriyama@sci.osaka-cu.ac.jp, okuyama@azusa.shinshu-u.ac.jp

Received July 21, 2015; Accepted July 27, 2015; Published 28 October 2015

.....
 We review recent progress in determining the partition function of the ABJM theory in the large- N expansion, including all of the perturbative and non-perturbative corrections. In particular, we focus on how these exact expansions are obtained from various beautiful relations to the Fermi gas system, topological string theory, the integrable model, and the supergroup.

Subject Index B04, B16, B25, B35, B83

	Contents	PAGE
1	Introduction	2
1.1	ABJ(M) matrix model	4
1.2	Supergroup structure	4
1.3	Complete large- N expansion	5
1.4	History	7
1.5	Pole cancellation mechanism	10
1.6	Contents	11
2	Fermi gas formalism	11
2.1	ABJM matrix model	11
2.2	ABJ matrix model	13
	2.2.1 Closed-string formalism	14
	2.2.2 Open-string formalism	14
3	Non-perturbative corrections at large N	18
3.1	WKB expansion	18
3.2	Exact values of partition function at finite N	19
3.3	Grand potential	20
3.4	Comparison with topological strings	23
3.5	Topological string theory	24
4	Related topics	27
4.1	Exact results with supersymmetry enhancement	27
4.2	't Hooft expansion and Borel resummation	29

4.3	Other theories	30
5	Summary and further directions	32
	Acknowledgements	33
	Funding	33
	References	33

1. Introduction

Chern–Simons theories play a central role in modern string theory. Two major subjects in the early days were open-string field theory [1] and topological string theory [2]. It was known not only that Chern–Simons theory satisfies the axiom of topological field theory but also that the topological theory can be lifted to topological string theory. Also, the beautiful gauge-invariant structure was used to construct the covariant open-string field theory. Both of these developments are deeply related to understanding the non-perturbative aspects of string theory.

The relation between Chern–Simons theory and non-perturbative string theory becomes even more important in supersymmetrization. It is known that, up to $\mathcal{N} = 3$, the supersymmetric Chern–Simons theory can be constructed for any gauge group and any representation [3–5]. In the context of the conformal symmetry, the condition that the levels are summed up to zero was proposed [6,7]. After some special arguments of the $\mathcal{N} = 4$ enhancements [8,9], it was finally found that, when the gauge group is $U(N)_k \times U(N)_{-k}$, with the subscripts denoting levels k and $-k$, and the matters are in the bifundamental representation, the supersymmetry is enhanced to $\mathcal{N} = 6$ [10], which is called ABJM theory. Later, it was also found [11,12] that, even when the ranks of two factors in the gauge group are different, $U(N_1)_k \times U(N_2)_{-k}$, Chern–Simons theory still enjoys the enhanced supersymmetry $\mathcal{N} = 6$, which is called ABJ theory. So we shall often refer to them inclusively as ABJ(M) theory.

The discovery of this maximally supersymmetric conformal Chern–Simons theory enables us to open up another window to the non-perturbative aspects of string theory. In the study of the non-perturbative effects in string theory, we are naturally led to a conjecture that there exists an 11D theory, which is dubbed M-theory [13]. It is expected that this theory is the *mother* theory with fundamental *membrane* excitation from which our perturbative string theory stems after compactifications, though only little is known about this *mysterious* theory. From the supergravity analysis, it is known that, besides the fundamental electric M2-branes, there exist solitonic magnetic M5-branes. Also, from the anti-de Sitter/conformal field theory (AdS/CFT) correspondence, we know that the degrees of freedom of N M2-branes and N M5-branes are $N^{3/2}$ and N^3 , respectively [14].

The ABJM theory describes a stack of N multiple M2-branes put on the target space $\mathbb{C}^4/\mathbb{Z}_k$, while the ABJ theory describes a composite system of both N_1 multiple M2-branes and $N_2 - N_1$ multiple fractional M2-branes for $N_2 > N_1$. For other superconformal Chern–Simons theories with fewer supersymmetries [6,7,15], we expect that they describe multiple M2-branes on various backgrounds with fewer supersymmetries. It is then natural to ask whether we can reproduce the degrees of freedom by studying the free energy of the ABJM theory in the large- N limit. Also, it is interesting to see how the large- N corrections behave and whether there is any interesting physical implication to this.

In contrast to Chern–Simons theory, let us briefly comment on Yang–Mills theory. Along with Chern–Simons theory, Yang–Mills theory also plays a crucial role in modern string theory. In particular, it is known that multiple D-branes are described by the maximally supersymmetric Yang–Mills theory with the gauge group $U(N)$ [16]. This was found by identifying the dynamical degrees of freedom of D-branes as open-string excitations appearing in each pair of multiple D-branes.

The free energy N^2 in the large- N limit is interpreted as the number of matrix elements and matches the supergravity computation [14].

Although the maximally supersymmetric Chern–Simons theory selects a special gauge group and a special representation, the maximally supersymmetric Yang–Mills theory is defined for general gauge groups and representations. Then, it is natural to expect that the gauge group and the matter contents of the ABJ(M) theory are special. In fact, it was observed in Ref. [17] that, when we look at the matrix model that appears after applying the localization techniques, the gauge group and the matter content have a good interpretation in the supergroup $U(N_1|N_2)$. Interestingly, the relation to the supergroup $U(N_1|N_2)$ persists even in the one-point function of the half-BPS Wilson loop. The underlying relation to the supergroup may go beyond an accidental relation after the localization techniques and plays some roles in understanding the ABJ(M) theory itself.

In the supersymmetric theory, it was found that, when we compute the partition function or the correlation function preserving supersymmetries, the infinite-dimensional path integral reduces to a finite-dimensional matrix integration [18]. For a more careful explanation of localization techniques, see Ref. [19] in this volume.

After the localization techniques, the partition function of the ABJ(M) theory is reduced to a matrix model [20]. Although many works have concentrated on the matrix model expression itself, solving the matrix model is an interesting and important subject. In physics, solving one theory classically means finding the exact solutions to the equation of motion, while solving one theory on the quantum-mechanical level means performing the path integral exactly. Since the localization techniques already reduce the infinite-dimensional path integral into a matrix integration, solving the ABJ(M) theory amounts to performing the matrix integration exactly. This is our main result in Refs. [21–24] and we shall explain it in more detail in the remaining part of this introduction.

Before explaining the details, let us try to discuss our motivation and expectation for the results. First of all, the full expression of the ABJ(M) partition function may provide some hidden structures of the M2-branes. After reproducing the predicted $N^{3/2}$ behavior in the large- N limit from the matrix model [25], it is natural to ask what the corrections look like. In particular, compared with the matrix interpretation of the N^2 behavior for the D-branes, it is perplexing to find $N^{3/2}$ behavior for the M2-branes. We expect that finding all the corrections will give us some hints to this mysterious $N^{3/2}$ behavior of the M2-branes. Prior to the discovery of the ABJ(M) theory, attempts were made to find the worldvolume theory of the M2-branes from a generalization of the Lie algebra [26,27]. Though it turned out that the algebra is very restricted and is not explored extensively any more, these works may suggest a hidden novel structure in M2-branes yet to be discovered.

Secondly, the ABJ(M) theory is one of the maximally supersymmetric theories. If we consider that the uniqueness of string theory stems from the maximum in the supersymmetrization, we expect that maximally supersymmetric theories will play a crucial role in understanding string theory. At this point, it should be important to study carefully all the fundamental quantities in the maximally supersymmetric theories.

Thirdly, let us stress the importance of the exact solution. The main method in studying physics is the perturbation theory. However, in a standard situation, the physics beyond the perturbation theory has a very rich structure. We expect that the exact instanton expansion will provide us with a window to understand non-perturbative effects.

In the remaining part of this introduction, let us briefly recapitulate the main results of a series of our works [21–24]. We will see that the results provide some insights to the above questions.

1.1. *ABJ(M) matrix model*

After applying the localization techniques, the partition function of the ABJ(M) theory reduces to a matrix model

$$Z_k(N_1, N_2) = \frac{(-1)^{\frac{1}{2}N_1(N_1-1) + \frac{1}{2}N_2(N_2-1)}}{N_1!N_2!} \times \int_{\mathbb{R}^{N_1+N_2}} \prod_{i=1}^{N_1} D\mu_i \prod_{j=1}^{N_2} Dv_j \frac{\prod_{i<j}^{N_1} \left(2 \sinh \frac{\mu_i - \mu_j}{2}\right)^2 \prod_{i<j}^{N_2} \left(2 \sinh \frac{v_i - v_j}{2}\right)^2}{\prod_{i,j}^{N_1, N_2} \left(2 \cosh \frac{\mu_i - v_j}{2}\right)^2}, \quad (1.1)$$

where $\prod_{i<j}^{N_1}$ and $\prod_{i<j}^{N_2}$ denote the product taking over all of the pairs (i, j) satisfying $1 \leq i < j \leq N_1$ and $1 \leq i < j \leq N_2$, respectively, while the product $\prod_{i,j}^{N_1, N_2}$ denotes $\prod_{i=1}^{N_1} \prod_{j=1}^{N_2}$. The integration is defined by

$$D\mu_i = \frac{d\mu_i}{2\pi} e^{\frac{ik}{4\pi}(\mu_i)^2}, \quad Dv_j = \frac{dv_j}{2\pi} e^{-\frac{ik}{4\pi}(v_j)^2}. \quad (1.2)$$

The derivation of this partition function was originally done in Ref. [20]. Here we shall not go into the details of the derivation and instead take this ABJ(M) matrix model (1.1) as our starting point.

1.2. *Supergroup structure*

This matrix model comes from the maximally superconformal Chern–Simons theory. Let us briefly comment on the beauty of this matrix model [17]. In (1.1) the most complicated part of this partition function is probably the measure

$$\frac{\prod_{i<j}^{N_1} \left(2 \sinh \frac{\mu_i - \mu_j}{2}\right)^2 \prod_{i<j}^{N_2} \left(2 \sinh \frac{v_i - v_j}{2}\right)^2}{\prod_{i,j}^{N_1, N_2} \left(2 \cosh \frac{\mu_i - v_j}{2}\right)^2}. \quad (1.3)$$

This measure can be greatly simplified by dropping one factor of the gauge group:

$$\prod_{i<j}^{N_1} \left(2 \sinh \frac{\mu_i - \mu_j}{2}\right)^2. \quad (1.4)$$

This measure is the famous one of the Chern–Simons matrix model. If the reader is not familiar with it, she can replace the hyperbolic function by a rational function

$$\prod_{i<j}^{N_1} (\mu_i - \mu_j)^2. \quad (1.5)$$

This is nothing but the $U(N_1)$ -invariant Vandermonde measure appearing in the diagonalization of the $N_1 \times N_1$ Hermitian matrix in the Gaussian matrix model. We can restore the hyperbolic function by exponentiating the variables μ into $e^{\pm\mu}$. If we want to restore the dependence of the variables v , all we have to do is the supersymmetrization. Namely, we replace the invariant measure of $U(N_1)$ by that of the supergroup $U(N_1|N_2)$:

$$\frac{\prod_{i<j}^{N_1} (\mu_i - \mu_j)^2 \prod_{i<j}^{N_2} (v_i - v_j)^2}{\prod_{i,j}^{N_1, N_2} (\mu_i + v_j)^2}. \quad (1.6)$$

After exponentiating μ_i, v_j into $e^{\pm\mu_i}, e^{\pm v_j}$, respectively, we come back to the original measure of the ABJ(M) matrix model. The explanation is summarized in Table 1.

Table 1. The measure of the ABJ(M) matrix model can be regarded as a simultaneous deformation of the supersymmetry deformation and the Chern–Simons deformation from the standard Vandermonde measure.

$$\begin{array}{ccc}
 \frac{\prod_{i<j}^{N_1} \left(2 \sinh \frac{\mu_i - \nu_j}{2}\right)^2 \prod_{i<j}^{N_2} \left(2 \sinh \frac{\nu_i - \nu_j}{2}\right)^2}{\prod_{i,j}^{N_1, N_2} \left(2 \cosh \frac{\mu_i - \nu_j}{2}\right)^2} & & \\
 \nearrow & & \nwarrow \\
 \prod_{i<j}^{N_1} \left(2 \sinh \frac{\mu_i - \nu_j}{2}\right)^2 & & \frac{\prod_{i<j}^{N_1} (\mu_i - \nu_j)^2 \prod_{i<j}^{N_2} (\nu_i - \nu_j)^2}{\prod_{i,j}^{N_1, N_2} (\mu_i + \nu_j)^2} \\
 \nwarrow & & \nearrow \\
 \prod_{i<j}^{N_1} (\mu_i - \nu_j)^2 & &
 \end{array}$$

In this sense, we can say that the ABJ(M) matrix model is the generalization of the most fundamental Gaussian matrix model with simultaneous deformation of the supersymmetrization and the Chern–Simons deformation. As a message, we would like to stress that, aside from the physical interpretation, even purely from a mathematical viewpoint, the ABJ(M) matrix model is one of the most fundamental matrix models and deserves careful study.

1.3. Complete large- N expansion

Now let us summarize the result found in a series of works [21–25,28–39] for both the ABJM case and the ABJ case. Let us reparametrize the arguments of the partition function $Z_k(N_1, N_2)$ in (1.1) as $Z_{k,M}(N)$ with $(N_1, N_2) = (N, N + M)$ and define the grand canonical partition function $\Xi_{k,M}(\mu)$ with

$$\Xi_{k,M}(\mu) = \sum_{N=0}^{\infty} e^{\mu N} |Z_{k,M}(N)|, \tag{1.7}$$

and the (modified) grand potential $J_{k,M}(\mu)$ with [22]

$$\Xi_{k,M}(\mu) = \sum_{n=-\infty}^{\infty} e^{J_{k,M}(\mu + 2\pi i n)}, \tag{1.8}$$

by regarding the rank N as the number of particles and introducing the chemical potential μ dual to N . Note that the phase dependence of the partition function $Z_{k,M}(N)$ was studied carefully in Ref. [20] and, to make contact with the topological string theory, we have to take the absolute value of the partition function in defining the grand potential [38]. It is important to notice that the large- N physics in the canonical ensemble is governed by the large- μ limit in the grand canonical ensemble; thus, below we focus on this expansion.

Note that the grand canonical partition function defined originally in (1.7) is symmetric under the $2\pi i$ shift of the chemical potential

$$\mu \rightarrow \mu + 2\pi i. \tag{1.9}$$

Due to the periodicity, the grand partition function contains an oscillatory behavior in an apparent sense [22]. The (modified) grand potential in (1.8) is defined by getting rid of this oscillation.

In the inverse transformation, the partition function is given in terms of the grand canonical partition function $\Xi_{k,M}(\mu)$ by

$$|Z_{k,M}(N)| = \int_{-\pi i}^{\pi i} \frac{d\mu}{2\pi i} \Xi_{k,M}(\mu) e^{-\mu N}, \tag{1.10}$$

and in the grand potential $J_{k,M}(\mu)$ (1.8) by

$$|Z_{k,M}(N)| = \int_{-\infty i}^{\infty i} \frac{d\mu}{2\pi i} e^{J_{k,M}(\mu) - \mu N}. \tag{1.11}$$

This difference plays crucial roles, e.g., in discussing the generalizations of the ABJM theory in Ref. [40].

Let us first display our main result, and explain how it was found and what the physical implication is later in the subsequent subsections. The large- μ results of the grand potential can be summarized as follows. The grand potential is split into the ‘‘perturbative’’ part and the ‘‘non-perturbative’’ part:

$$J_{k,M}(\mu) = J_{k,M}^{\text{pert}}(\mu^{\text{eff}}) + J_{k,M}^{\text{np}}(\mu^{\text{eff}}), \tag{1.12}$$

where the ‘‘effective’’ chemical potential μ^{eff} will be explained later. The perturbative part is given by

$$J_{k,M}^{\text{pert}}(\mu) = \frac{C_k}{3} \mu^3 + B_{k,M} \mu + A_k, \tag{1.13}$$

where the coefficients C_k and $B_{k,M}$ are [25,38]

$$C_k = \frac{2}{\pi^2 k}, \quad B_{k,M} = \frac{1}{3k} - \frac{k}{12} + \frac{k}{2} \left(\frac{M}{k} - \frac{1}{2} \right)^2, \tag{1.14}$$

while A_k is [33,41]

$$A_k = \frac{2\zeta(3)}{\pi^2 k} \left(1 - \frac{k^3}{16} \right) + \frac{k^2}{\pi^2} \int_0^\infty dx \frac{x}{e^{kx} - 1} \log(1 - e^{-2x}). \tag{1.15}$$

For integral k , the integration is performed explicitly.

The non-perturbative part is given by

$$J_{k,M}^{\text{np}}(\mu) = F^{\text{top}}(\mathbf{T}; g_s) + \frac{1}{2\pi i} \frac{\partial}{\partial g_s} \left[g_s F^{\text{NS}} \left(\frac{\mathbf{T}}{g_s}; \frac{1}{g_s} \right) \right]. \tag{1.16}$$

Here the functions F^{top} and F^{NS} are defined by the limits

$$\begin{aligned} F^{\text{top}}(\mathbf{T}; \tau) &= \lim_{\tau_1 \rightarrow \tau, \tau_2 \rightarrow -\tau} F^{\text{ref}}(\mathbf{T}; \tau_1, \tau_2), \\ F^{\text{NS}}(\mathbf{T}; \tau) &= \lim_{\tau_1 \rightarrow \tau, \tau_2 \rightarrow 0} 2\pi i \tau_2 F^{\text{ref}}(\mathbf{T}; \tau_1, \tau_2) \end{aligned} \tag{1.17}$$

of a refined function

$$F^{\text{ref}}(\mathbf{T}; \tau_1, \tau_2) = \sum_{j_L, j_R} \sum_{n=1}^\infty \sum_d N_{j_L, j_R}^d \frac{\chi_{j_L}(q_L) \chi_{j_R}(q_R) e^{-nd \cdot \mathbf{T}}}{n \left(q_1^{n/2} - q_1^{-n/2} \right) \left(q_2^{n/2} - q_2^{-n/2} \right)}, \tag{1.18}$$

with the $\text{su}(2)$ character

$$\chi_j(q) = \frac{q^{2j+1} - q^{-(2j+1)}}{q - q^{-1}}, \tag{1.19}$$

and various variables

$$q_1 = e^{2\pi i \tau_1}, \quad q_2 = e^{2\pi i \tau_2}, \quad q_L = e^{\pi i (\tau_1 - \tau_2)}, \quad q_R = e^{\pi i (\tau_1 + \tau_2)}. \quad (1.20)$$

The functions in (1.17) have their origin in topological string theory. Namely, the function $F^{\text{ref}}(\mathbf{T}; \tau_1, \tau_2)$ is the free energy of the refined topological string theory, while $F^{\text{top}}(\mathbf{T}; \tau)$ and $F^{\text{NS}}(\mathbf{T}; \tau)$ are the free energy of the original topological string theory and the free energy in the so-called Nekrasov–Shatashvili (NS) limit. The integers N_{j_L, j_R}^d are the BPS indices, which are topological invariants counting modes on the manifolds. The relation between two Kähler parameters T and the chemical potential μ is given by

$$T_1 = \frac{4\mu}{k} + 2\pi i \left(\frac{1}{2} - \frac{M}{k} \right), \quad T_2 = \frac{4\mu}{k} - 2\pi i \left(\frac{1}{2} - \frac{M}{k} \right), \quad (1.21)$$

while the string coupling g_s is identified to the level k as

$$g_s = 2/k. \quad (1.22)$$

Note that, in (1.12), we need to substitute for the effective chemical potential

$$\mu^{\text{eff}} = \begin{cases} \mu - 2(-1)^{\frac{k}{2}-M} e^{-2\mu} {}_4F_3 \left(1, 1, \frac{3}{2}, \frac{3}{2}; 2, 2, 2; 16(-1)^{\frac{k}{2}-M} e^{-2\mu} \right), & \text{for even } k, \\ \mu + e^{-4\mu} {}_4F_3 \left(1, 1, \frac{3}{2}, \frac{3}{2}; 2, 2, 2; -16e^{-4\mu} \right), & \text{for odd } k. \end{cases} \quad (1.23)$$

It is interesting to find that the ABJ(M) matrix model has a deep relation to the topological string theory [42].

This result, with the effective chemical potential plugged in, shows that the grand potential is given by

$$J_{k,M}^{\text{pert}}(\mu^{\text{eff}}) + J_{k,M}^{\text{np}}(\mu^{\text{eff}}) = \sum_{m,\ell=0}^{\infty} f_{m,\ell}(\mu) e^{-\left(\frac{4m}{k} + 2\ell\right)\mu}, \quad (1.24)$$

with $f_{m,\ell}(\mu)$ being polynomials of μ . Here the effect of $e^{-\frac{4\mu}{k}}$ is identified with the worldsheet instanton, which is a fundamental string wrapping the holomorphic cycle $\mathbb{C}\mathbb{P}^1$ in $\mathbb{C}^4/\mathbb{Z}_k$ on the gravity side [25,43], and that of $e^{-2\mu}$ with the membrane instantons, which is a D2-brane wrapping the Lagrangian submanifold $\mathbb{R}\mathbb{P}^3$ [29]. In particular, the coefficients $f_{m,\ell}(\mu)$ with $\ell = 0$ or $m = 0$ take the following form [32]:

$$f_{m,0}(\mu) = d_m(k), \quad f_{0,\ell}(\mu) = a_\ell(k)\mu^2 + b_\ell(k)\mu + c_\ell(k). \quad (1.25)$$

Besides the worldsheet instantons and the membrane instantons, the full instanton series (1.24) contains the bound states of these two instantons.

1.4. History

Here we shall explain how the result (1.12) with (1.13) and (1.16) was found in chronological order. From the AdS/CFT correspondence, the $N^{3/2}$ behavior of the degrees of freedom of the N multiple M2-branes [14] in the large- N limit was known from the gravity side about two decades ago. The study from the gauge theory side to reproduce the $N^{3/2}$ behavior and investigate the corrections, on the other hand, was started quite recently [25] after the discovery of the worldvolume theory of the multiple M2-branes [10]. For the explanation, we shall mainly concentrate on the ABJM case.

Since the partition function was reduced to a matrix integration from the localizations of the supersymmetric theories [20,44,45], we can apply 't Hooft expansion, the standard matrix model technique, to it. The 't Hooft expansion is to consider the 't Hooft limit

$$N \rightarrow \infty, \quad \lambda = N/k : \text{fixed}, \tag{1.26}$$

and keep track of the $1/N$ corrections. Then, the free energy $F(N, \lambda) = \log Z_{k,0}(N)$ is given by

$$F(N, \lambda) = N^2 F_0(\lambda) + F_1(\lambda) + N^{-2} F_2(\lambda) + \dots, \tag{1.27}$$

which can be regarded as the genus expansion. From this viewpoint, it is not difficult to imagine that the 't Hooft expansion has a good interpretation in string theory.

The first important work of the ABJM matrix model is the reproduction of the $N^{3/2}$ behavior [25] in the 't Hooft limit. In Refs. [25,29] it was further noted that the expansion has infinitely many $1/N$ corrections,

$$\begin{aligned} F_0(\lambda) &= \lambda^{-\frac{1}{2}} + \mathcal{O}\left(e^{-\sqrt{\lambda}}\right), \\ F_1(\lambda) &= \lambda^{\frac{3}{2}} + \log \lambda + \mathcal{O}\left(e^{-\sqrt{\lambda}}\right), \\ F_2(\lambda) &= \lambda^{\frac{7}{2}} + \lambda^2 + \lambda^{\frac{1}{2}} + \mathcal{O}\left(e^{-\sqrt{\lambda}}\right), \\ &\dots, \end{aligned} \tag{1.28}$$

where we only present the result schematically. In particular, the coefficients are not displayed and the 't Hooft coupling constant is slightly shifted from the original one, but we will only come back to these points later. The study of the $1/N$ corrections is most easily done by observing the relation between the ABJ(M) matrix model and the topological string theory on local $\mathbb{P}^1 \times \mathbb{P}^1$ [42] and studying it with the holomorphic anomaly equation [46] from the topological string theory. Here it is interesting to notice a triangular structure in the expansion (1.28). Namely, by going to one higher genus from $F_g(\lambda)$ to $F_{g+1}(\lambda)$, we always gain only one more perturbative term and the perturbative terms in $F_{g+1}(\lambda)/F_g(\lambda)$ range between λ^2 and $\lambda^{\frac{1}{2}}$ in the large- λ limit. This means that the large- N expansion can be alternatively obtained by taking the large- N limit with $\lambda^2 N^{-2a}$ ($0 \leq a \leq 1$) fixed.¹ If we look at the coefficients carefully, it is not difficult to observe that, after introducing [29]

$$\lambda^{\text{ren}} = \lambda - \frac{1}{24} - \frac{\lambda^2}{3N^2}, \tag{1.29}$$

the infinite expansion can be partially taken care of. So the remaining question is whether we can perform the other summation by dropping the non-perturbative terms and solving the perturbative holomorphic anomaly equation. In Ref. [30] it was found that if we drop the non-perturbative corrections all of the perturbative expansions can be summed up to the Airy function:

$$Z_k^{\text{pert}}(N) = e^{A_k} C_k^{-\frac{1}{3}} \text{Ai} \left[C_k^{-\frac{1}{3}} (N - B_{k,0}) \right]. \tag{1.30}$$

In Ref. [30] the constant A_k was not taken care of but this constant was later pointed out in Ref. [33] and studied carefully in Ref. [41]. There are many arguments on the meaning of the Airy function

¹ See Ref. [47] for more discussions on the limits.

appearing in the summation of the perturbative corrections.² Technically, it suggests the viewpoint of the grand canonical ensemble. Namely, the integral representation of the Airy function is

$$\text{Ai}(N) = \int_C \frac{d\mu}{2\pi i} e^{\frac{1}{3}\mu^3 - \mu N}, \tag{1.31}$$

where the integral contour C comes from the infinity in the region with angle $-\pi/2 < \theta < -\pi/3$ seen from the origin and goes to the infinity in the region with angle $\pi/3 < \theta < \pi/2$. Comparing this expression with the inverse transformation (1.11), it is not difficult to find that the complicated Airy function is mapped to a simple cubic function (1.13). This suggests that the full expression is clearer in the grand canonical ensemble.

As pointed out in Ref. [28], there are dissatisfactions with the 't Hooft expansion. In the 't Hooft expansion, the large- N limit is taken by fixing the 't Hooft coupling $\lambda = N/k$. However, in approaching the M-theory regime in the large- N limit, the parameter we want to fix is the level k itself, which characterizes the M-theory background $\mathbb{C}^4/\mathbb{Z}_k$:

$$N \rightarrow \infty, \quad k : \text{fixed}. \tag{1.32}$$

These ideas were incarnated in Ref. [32], by rewriting the partition function of the ABJM theory with gauge group $U(N)_k \times U(N)_{-k}$ into that of a Fermi gas system with N non-interacting particles whose Hamiltonian is given by

$$e^{-H} = \frac{1}{\sqrt{2 \cosh \frac{q}{2}}} \frac{1}{2 \cosh \frac{p}{2}} \frac{1}{\sqrt{2 \cosh \frac{q}{2}}}, \tag{1.33}$$

with the canonical commutation relation $[q, p] = i\hbar$ where the Chern–Simons level plays the role of the Planck constant, $\hbar = 2\pi k$.

This beautiful rewriting directly suggests another analysis for the study of the ABJM matrix model. Namely, we can study the grand canonical ensemble of the ABJM matrix model from the WKB expansion around $\hbar = 0$ [32]. In this expansion, we first expand the whole grand potential in the small- k region,

$$J_k(\mu) = \frac{1}{k} J^{(0)}(\mu) + k J^{(2)}(\mu) + k^3 J^{(4)}(\mu) + \dots, \tag{1.34}$$

and further study each term separately in the large- μ expansion. After that, we collect all terms of the same order in μ and hope that we can sum up the k expansion. In this sense, we can separate the small- k expansion from the large- μ expansion and take the large- μ limit by keeping the level k fixed.

Both the 't Hooft expansion and the WKB expansion are powerful enough to study the matrix model to some extent. On one hand, the 't Hooft expansion detects not only the perturbative part of the Airy function but also the worldsheet instanton part, though is sensible to neither the membrane instanton part nor the bound state part. This is because, after plugging in the stationary condition for μ ,

$$N = \frac{\partial J_k}{\partial \mu} \simeq C_k \mu^2, \tag{1.35}$$

the worldsheet instanton factor and the membrane instanton factor in (1.24) become

$$e^{-\frac{4\mu}{k}} \simeq e^{-2\pi\sqrt{2\lambda}}, \quad e^{-2\mu} \simeq e^{-\pi N\sqrt{2/\lambda}}. \tag{1.36}$$

² See a recent work [48] for interesting generalizations and discussions.

On the other hand, the WKB expansion detects only the perturbative part and the membrane instanton part, but not the worldsheet instanton part or the bound state part. So, unfortunately, for the bound state of the worldsheet instantons and the membrane instantons, neither the 't Hooft expansion nor the WKB expansion is applicable. The only method we have is the numerical analysis and the pole cancellation mechanism, which we will explain in the next subsection.

1.5. Pole cancellation mechanism

In determining the non-perturbative effects, the so-called pole cancellation mechanism plays an important role. Let us recapitulate it briefly. For this purpose we display the explicit forms of the first few worldsheet instantons in (1.25):

$$d_1(k) = \frac{1}{\sin^2 \frac{2\pi}{k}}, \quad d_2(k) = -\frac{1}{2 \sin^2 \frac{4\pi}{k}} - \frac{1}{\sin^2 \frac{2\pi}{k}}, \dots, \tag{1.37}$$

and membrane instantons [22,35]:

$$\begin{aligned} a_1(k) &= -\frac{4 \cos \frac{\pi k}{2}}{\pi^2 k}, & b_1(k) &= \frac{2 \cos^2 \frac{\pi k}{2}}{\pi \sin \frac{\pi k}{2}}, & c_1(k) &= \dots, \\ a_2(k) &= -\frac{8 + 10 \cos \pi k}{\pi^2 k}, & b_2(k) &= \frac{4(1 + \cos \pi k)}{\pi^2 k} + \frac{17 + 24 \cos \pi k}{2\pi \sin \pi k}, & c_2(k) &= \dots, \\ & & \dots & & & \end{aligned} \tag{1.38}$$

It is not difficult to find that the coefficients of the m th worldsheet instanton contain poles when $\frac{2m}{k} \in \mathbb{Z}$, while the coefficients of the ℓ th membrane instanton contain poles when $2\ell k \in \mathbb{Z}$. However, it is interesting to observe that, whenever one coefficient is divergent, the other coefficient is always divergent as well and, in total, the coefficient is completely finite. This property was first observed in Ref. [22] and its importance was stressed in Ref. [35].

This property is probably due to the supersymmetries. Since there is no phase transition in theories with enough supersymmetries, the partition function should be finite. From the viewpoint of applications, this property can be used to fix the first few instanton coefficients. In fact, the first few coefficients of the membrane instantons listed above were found from this pole cancellation mechanism. The existence of the membrane instantons was first observed in Ref. [49]. It is surprising that the first quantitative study of the membrane instanton is performed in this way.

We can proceed to the bound states and find a similar pole cancellation mechanism there. Namely, instead of the simple pole cancellation mechanism between the worldsheet instanton and the membrane instanton, for higher instanton effects the bound states of the worldsheet instanton and the membrane instanton appear as well. Hence, we need to consider the cancellation among all of the bound states besides the pure worldsheet instanton and the pure membrane instanton.

In Ref. [23] we find that the effect of the bound states can be incorporated into that of the worldsheet instantons by shifting the chemical potential as in (1.23). As a bonus, after shifting the chemical potential μ into the effective one μ_{eff} , we find that the quadratic polynomial coefficients of the pure membrane instanton $f_{0,\ell}(\mu)$ (1.25) become linear polynomials, where the coefficients of the linear terms and constant terms obey a derivative relation. Finally, in Ref. [24], all of these structures are re-expressed in terms of the refined topological string theory as in (1.12).

1.6. Contents

The main purpose of this paper is to review the beautiful aspects of the ABJ(M) matrix model and its generalizations. After reviewing the history of the progress in Sects. 1.4 and 1.5, it is not difficult to imagine that the explanation would be simplest if we start with reviewing the Fermi gas formalism. We shall do this in the next section first for the ABJM matrix model and then generalize to the ABJ matrix model. After that, in Sect. 3, we shall proceed to various analyses, including the WKB expansion, the exact evaluation of the partition function, the numerical study of the grand potential, the pole cancellation mechanism, and the viewpoint from the topological string theory. Finally, in Sect. 4, we summarize some recent progress in the supersymmetry-enhanced case and more general cases.

2. Fermi gas formalism

In this section, let us explain how the Fermi gas formalism is applied to the ABJM and ABJ matrix model.

2.1. ABJM matrix model

In this subsection, we shall explain the Fermi gas formalism for the ABJM matrix model [32]. Namely, we shall study (1.1) by restricting ourselves to the case of $N_2 = N_1 = N$ and define

$$Z_k(N) = Z_k(N, N). \tag{2.1}$$

Using the Cauchy identity

$$\det\left(\frac{1}{u_i + v_j}\right) = \frac{\prod_{i < j} (u_i - u_j)(v_i - v_j)}{\prod_{i,j} (u_i + v_j)}, \tag{2.2}$$

one can show that (1.1) is rewritten as

$$Z_k(N) = \int \frac{D^N \mu}{N!} \frac{D^N \nu}{N!} \det\left(\frac{1}{2 \cosh \frac{\mu_i - \nu_j}{2}}\right)_{\substack{1 \leq i \leq N \\ 1 \leq j \leq N}} \det\left(\frac{1}{2 \cosh \frac{\nu_i - \mu_j}{2}}\right)_{\substack{1 \leq i \leq N \\ 1 \leq j \leq N}}. \tag{2.3}$$

Note that the two determinant factors are identical, though we write them separately, so that the structure is clearer. Using a determinantal formula from Ref. [38], we can combine the two determinants,

$$Z_k(N) = \int \frac{D^N \mu}{N!} \det\left(\frac{1}{2 \cosh \frac{\mu_i - \nu}{2}} \circ \frac{1}{2 \cosh \frac{\nu - \mu_j}{2}}\right)_{\substack{1 \leq i \leq N \\ 1 \leq j \leq N}}, \tag{2.4}$$

where \circ stands for the integration:

$$\phi \circ \psi = \int D\nu \phi(\nu) \psi(\nu). \tag{2.5}$$

In other words, we regard two ingredients $(2 \cosh \frac{\mu - \nu}{2})^{-1}$ and $(2 \cosh \frac{\nu - \mu}{2})^{-1}$ as matrices with continuous indices μ and ν . As in the discrete case, where we perform the matrix multiplication by contracting the indices, here the contractions of the continuous indices are done by the integrations given in (1.2).

After expanding the determinant of (2.4) and rescaling the integration variables $\mu = x/k$, $\nu = y/k$, we find that it can be expressed as a sum over permutations S_N :

$$Z_k(N) = \frac{1}{N!} \sum_{\sigma \in S_N} (-1)^\sigma \int \frac{d^N x}{(2\pi)^N} \prod_{i=1}^N \rho(x_i, x_{\sigma(i)}), \tag{2.6}$$

with $\rho(x, x')$ given by

$$\rho(x, x') = \int \frac{dy}{2\pi} e^{\frac{i}{4\pi k} x^2} \frac{1}{k} \frac{1}{2 \cosh \frac{x-y}{2k}} e^{-\frac{i}{4\pi k} y^2} \frac{1}{k} \frac{1}{2 \cosh \frac{y-x'}{2k}}. \tag{2.7}$$

If we use the Fourier transformation

$$\frac{1}{k} \frac{1}{2 \cosh \frac{x-y}{2k}} = \left\langle x \left| \frac{1}{2 \cosh \frac{\hat{p}}{2}} \right| y \right\rangle, \tag{2.8}$$

by introducing the coordinate operator \hat{q} and the momentum operator \hat{p} satisfying $[\hat{q}, \hat{p}] = i\hbar$ with $\hbar = 2\pi k$ and the coordinate eigenstate

$$\hat{q} |x\rangle = x|x\rangle, \tag{2.9}$$

normalized by $\langle x|x'\rangle = 2\pi\delta(x-x')$, we find that $\rho(x, x')$ can be expressed as

$$\rho(x, x') = \left\langle x \left| e^{\frac{i}{2\hbar} \hat{q}^2} \frac{1}{2 \cosh \frac{\hat{p}}{2}} e^{-\frac{i}{2\hbar} \hat{q}^2} \frac{1}{2 \cosh \frac{\hat{p}}{2}} \right| x' \right\rangle. \tag{2.10}$$

After some further canonical transformations with the help of the formula

$$e^{\frac{i}{2\hbar} \hat{q}^2} f(\hat{p}) e^{-\frac{i}{2\hbar} \hat{q}^2} = f(\hat{p} - \hat{q}), \quad e^{\frac{i}{2\hbar} \hat{p}^2} g(\hat{q}) e^{-\frac{i}{2\hbar} \hat{p}^2} = g(\hat{q} + \hat{p}), \tag{2.11}$$

and a similarity transformation, we finally find that $\rho(x, y)$ can be interpreted as a density matrix of a quantum-mechanical system

$$\rho(x, y) = \left\langle x \left| e^{-H} \right| y \right\rangle, \tag{2.12}$$

with (1.33), which leads to the expression

$$\rho(x, y) = \frac{1}{k} \frac{1}{\sqrt{2 \cosh \frac{x}{2}}} \frac{1}{2 \cosh \frac{x-y}{2k}} \frac{1}{\sqrt{2 \cosh \frac{y}{2}}}. \tag{2.13}$$

The formula (2.6) can be regarded as the $(-1)^\sigma$ -graded partition function of a particle propagating from a state labeled by the coordinate x_i to the permuted coordinate $x_{\sigma(i)}$. This is nothing but the partition function of a Fermi gas system. To avoid unnecessary complexities, hereafter we shall often drop the hats for the operators \hat{q} and \hat{p} .

It is convenient to define the grand partition function $\Xi_k(\mu)$ by introducing the chemical potential μ conjugate to N and summing over N . It turns out that $\Xi_k(\mu)$ can be written as a Fredholm

determinant of ρ ,

$$\Xi_k(\mu) = \sum_{N=0}^{\infty} e^{N\mu} Z_k(N) = \det(1 + e^\mu \rho), \tag{2.14}$$

and the grand potential $\tilde{J}_k(\mu) = \log \Xi_k(\mu)$ is given by³

$$\tilde{J}_k(\mu) = \sum_{\ell=1}^{\infty} \frac{(-1)^{\ell-1}}{\ell} e^{\ell\mu} \text{Tr} \rho^\ell. \tag{2.15}$$

The expression of the Fredholm determinant (2.14) and its expansion (2.15) can be shown by relating the coefficients of $\prod_{\ell=1}^{\infty} (\text{Tr} \rho^\ell)^{m_\ell}$ to the combinatorics of distributing N into an assembly of m_ℓ cycles of length ℓ satisfying $\sum_{\ell=1}^{\infty} m_\ell \ell = N$ [50].

In terms of the eigenvalues E_n ($n = 0, 1, \dots$) of the Hamiltonian H , the grand partition function is given by

$$\Xi_k(\mu) = \prod_{n=0}^{\infty} (1 + e^{\mu - E_n}); \tag{2.16}$$

hence, the problem boils down to the diagonalization of the Hamiltonian. The study of the spectrum of the ABJM Hamiltonian was initiated in Ref. [21] using the Hankel matrix representation of ρ , which was later generalized to other models.

Note that the partition function can also be rewritten as

$$Z_k(N) = \frac{1}{N!} \int \frac{d^N x}{(4\pi k)^N} \prod_{i < j} \tanh^2 \frac{x_i - x_j}{2k} \prod_{i=1}^N \frac{1}{2 \cosh \frac{x_i}{2}}, \tag{2.17}$$

which is useful for discussing the generalization to the ABJ model.

2.2. ABJ matrix model

In the previous subsection, we have given the Fermi gas formalism for the ABJM matrix model. Using this Fermi gas formalism, we shall explain how the exact instanton expansion was found in the next section. Before that, let us comment that the Fermi gas formalism is also applicable to the ABJ matrix model where the gauge group $U(N_1)_k \times U(N_2)_{-k}$ has different ranks. In generalizing the discussions in the previous subsection, there are two Fermi gas formalisms for the ABJ matrix model. One is to dress the density matrix in (2.13) by extra factors with the traces kept unchanged, while the other is to keep the density matrix and introduce the endpoints for the density matrix.

Since the extra physical ingredients of the ABJ theory compared with the ABJM theory are the fractional M2-branes, two formalisms amount to expressing the fractional branes by changing the closed-string backgrounds or from the open-string excitations. Hence, it is reasonable to call these two formalisms the closed-string formalism and open-string formalism, respectively.

After the perturbative coefficients C_k , $B_{k,M}$ (1.14) and A_k (1.15) and the worldsheet instantons were found from the consistency with the 't Hooft expansion [25], the topological string on local $\mathbb{P}^1 \times \mathbb{P}^1$, and the ABJM limit $M \rightarrow 0$, we can use either of these Fermi gas formalisms to further study the membrane instantons and the bound states.

³ Note that this grand potential $\tilde{J}_k(\mu)$ is different from the modified one defined in (1.8).

2.2.1. Closed-string formalism

Let us start the study of the ABJ matrix model (1.1) with different ranks. In what follows, we will set $N_1 = N, N_2 = N + M$ and assume $k, M > 0$ without loss of generality. First let us note that the partition function for $N = 0$ is nothing but the bosonic $U(M)$ Chern–Simons matrix model, which is given by [51]

$$Z_k(0, M) = k^{-\frac{M}{2}} \prod_{s=1}^{M-1} \left(2 \sin \frac{\pi s}{k}\right)^{M-s}. \tag{2.18}$$

As discussed in Refs. [36,37,39], if we define $\widehat{Z}_k(N, N + M)$ to be the absolute value of the partition function $Z_k(N, N + M)$ normalized by the Chern–Simons case ($N = 0$),

$$\widehat{Z}_k(N, N + M) = \left| \frac{Z_k(N, N + M)}{Z_k(0, M)} \right|, \tag{2.19}$$

we find that $\widehat{Z}_k(N, N + M)$ is given by an integration

$$\widehat{Z}_k(N, N + M) = \frac{1}{N!} \int \frac{d^N x}{(4\pi k)^N} \prod_{i < j} \tanh^2 \frac{x_i - x_j}{2k} \prod_{i=1}^N V(x_i), \tag{2.20}$$

with a dressing factor

$$V(x) = \frac{1}{e^{\frac{x}{2}} + (-1)^M e^{-\frac{x}{2}}} \prod_{s=-\frac{M-1}{2}}^{\frac{M-1}{2}} \tanh \frac{x + 2\pi i s}{2k}, \tag{2.21}$$

which is very similar to (2.17) in the ABJM matrix model. The Seiberg-like duality of the ABJ theory

$$U(N)_k \times U(N + M)_{-k} \Leftrightarrow U(N)_{-k} \times U(N + k - M)_k \tag{2.22}$$

implies that $\widehat{Z}_k(N, N + M)$ is invariant under the exchange $M \leftrightarrow k - M$ with k and N fixed. Also, note that the original partition function $Z_k(N, N + M)$ (as well as the Chern–Simons case $Z_k(0, M)$) vanishes when $k < M$, which is consistent with the conjecture that the supersymmetry of the ABJ theory is spontaneously broken for $k < M$.

Again, using the Cauchy identity, $\widehat{Z}_k(N, N + M)$ can be recast as a Fermi gas form

$$\widehat{Z}_k(N, N + M) = \frac{1}{N!} \sum_{\sigma \in S_N} (-1)^\sigma \int \frac{d^N x}{(2\pi)^N} \prod_{i=1}^N \rho(x_i, x_{\sigma(i)}), \tag{2.23}$$

with the density matrix (2.13) changed into

$$\rho(x, y) = \frac{1}{k} \sqrt{V(x)} \frac{1}{2 \cosh \frac{x-y}{2k}} \sqrt{V(y)}. \tag{2.24}$$

Using this dressed density matrix, the grand partition function $\Xi_{k,M}(\mu)$ in (1.7) is given by

$$\Xi_{k,M}(\mu) = |Z_k(0, M)| \det(1 + e^\mu \rho). \tag{2.25}$$

2.2.2. Open-string formalism

Having seen the description of the fractional branes by changing the closed-string background in Sect. 2.2.1, here let us turn to the description from the open string [38]. This formalism has a benefit

in the compatibility with the vacuum expectation values of the half-BPS Wilson loop, since we end up with an ultimate Fermi gas formalism that includes both the M2 fractional branes and the half-BPS Wilson loop. As a corollary of this formalism, we can reproduce, on one hand, the Fermi gas formalism for the vacuum expectation values of the half-BPS Wilson loop [52] and, on the other hand, that for the partition function of the M2 fractional branes [38].

To state the results we need some preparations. Instead of the partition function itself, since the formalism incorporates the insertion of the half-BPS Wilson loop as well, let us consider the vacuum expectation value of the half-BPS Wilson loop. After the localization of the supersymmetric theory, it was found [17,20] that the insertion means to include a loop operator in the matrix model:

$$\langle s_Y \rangle_k(N_1, N_2) = \frac{(-1)^{\frac{1}{2}N_1(N_1-1) + \frac{1}{2}N_2(N_2-1)}}{N_1!N_2!} \int \frac{d^{N_1}\mu}{(2\pi)^{N_1}} \frac{d^{N_2}\nu}{(2\pi)^{N_2}} s_Y(e^\mu | e^\nu) \times \left[\frac{\prod_{i < j}^{N_1} 2 \sinh \frac{\mu_i - \mu_j}{2} \prod_{i < j}^{N_2} 2 \sinh \frac{\nu_i - \nu_j}{2}}{\prod_{i=1}^{N_1} \prod_{j=1}^{N_2} 2 \cosh \frac{\mu_i - \nu_j}{2}} \right]^2 e^{\frac{ik}{4\pi} (\sum_{i=1}^{N_1} \mu_i^2 - \sum_{j=1}^{N_2} \nu_j^2)}. \quad (2.26)$$

Here the loop operator $s_Y(e^\mu | e^\nu) = s_Y(e^{\mu_1}, e^{\mu_2}, \dots, e^{\mu_{N_1}} | e^{\nu_1}, e^{\nu_2}, \dots, e^{\nu_{N_2}})$ is the so-called supersymmetric Schur polynomial, which is nothing but the character of the hidden supergroup $U(N_1 | N_2)$. This is a symmetric polynomial with N_1 variables μ_i and N_2 variables ν_j , labeled by a Young tableau Y . It is probably simplest to find this supersymmetric Schur polynomial by expressing the original Schur polynomial in terms of a diagonal matrix $\text{diag}(e^{\mu_i})_{i=1}^{N_1}$ and the trace tr (in other words, the power sum symmetric polynomials) and replacing the diagonal matrix by a diagonal supermatrix $\text{diag}((e^{\mu_i})_{i=1}^{N_1} | (-e^{\nu_j})_{j=1}^{N_2})$ and the trace by a supertrace. For example, let us consider the supersymmetric Schur polynomial of the second symmetric tensor in the supergroup $U(N_1 | N_2)$ with N_1 and N_2 being large enough:

$$s_{\square\square}(e^\mu) = \frac{1}{2} (\text{tr} \text{diag}(e^\mu))^2 + \frac{1}{2} \text{tr} \text{diag}(e^{2\mu}). \quad (2.27)$$

Then, after the replacement, we find that the supersymmetric Schur polynomial can be decomposed into the original one by

$$s_{\square\square}(e^\mu | e^\nu) = s_{\square\square}(e^\mu) + s_{\square}(e^\mu) s_{\square}(e^\nu) + s_{\square}(e^\nu). \quad (2.28)$$

Also, let us define the vacuum expectation value in the grand canonical ensemble,

$$[s_Y]_{k,M}^{\text{GC}}(z) = \sum_{N=0}^{\infty} \langle s_Y \rangle_k(N, N + M) z^N, \quad (2.29)$$

as for the partition function. We further normalize it by the ABJM grand canonical partition function:

$$\langle s_Y \rangle_{k,M}^{\text{GC}}(z) = \frac{[s_Y]_{k,M}^{\text{GC}}(z)}{[1]_{k,0}^{\text{GC}}(z)}. \quad (2.30)$$

Then, we can prove

$$\langle s_Y \rangle_{k,M}^{\text{GC}}(z) = \det(H_{p,q}(z))_{\substack{1 \leq p \leq M+r \\ 1 \leq q \leq M+r}}, \quad (2.31)$$

where $H_{p,q}(z)$ is given by

$$H_{p,q}(z) = \begin{cases} E_{l_p} (1 + zQP)^{-1} E_{-M+q-1}, & \text{for } 1 \leq q \leq M, \\ zE_{l_p} (1 + zQP)^{-1} QE_{a_{q-M}}, & \text{for } 1 \leq q - M \leq r. \end{cases} \quad (2.32)$$

We have to explain the meaning of each expression. First of all, we regard $(Q)_{\nu,\mu}$ and $(P)_{\mu,\nu}$ as matrices and $(E_j)_\nu$ as a vector with continuous indices μ and ν . Again, as explained below (2.5), we multiply matrices with continuous indices by the integrations (1.2). So each expression is a scalar constructed by the inner product of two vectors with many matrices sandwiched in between. The explicit forms of $(Q)_{\nu,\mu}$, $(P)_{\mu,\nu}$, and $(E_j)_\nu$ are

$$(Q)_{\nu,\mu} = \frac{1}{2 \cosh \frac{\nu-\mu}{2}}, \quad (P)_{\mu,\nu} = \frac{1}{2 \cosh \frac{\mu-\nu}{2}}, \quad (E_j)_\nu = e^{(j+\frac{1}{2})\nu}. \quad (2.33)$$

Now let us turn to the integers l_p and a_q . In short, these integers are the p th leg length and q th arm length, which appear in the Frobenius symbol $(a_1 a_2 \cdots a_r; l_1 l_2 \cdots l_{r+M})$. The Frobenius symbol for the ABJM $U(N|N)$ case is found by drawing the diagonal line for the Young diagram and counting the number of boxes horizontally or vertically for the arm length or the leg length, respectively. For the general ABJ case with $U(N|N+M)$, all we have to do is to shift the diagonal line by M boxes. Note that, after shifting the diagonal line, the numbers of arm and leg lengths are no longer equal. Hence, we cannot consider the determinant of the square matrix. We can restore, however, the equality of the numbers by considering the ‘‘inverse’’ arm length as well. The subscript $-M+q-1$ appearing in (2.32) is nothing but this inverse arm length. Therefore, we can say that the lower expression in (2.32) comes from the usual arm length while the upper one is from the inverse arm length. (See, e.g., Refs. [38,52] for figures explaining the Frobenius symbol).

Before explaining how the expression is obtained, let us first see how this formula is beautiful and useful simultaneously. First let us consider the ABJM case $N_2 = N_1$ with a non-trivial insertion of the loop operator. Then, we find that $M = 0$ and the upper expression of $H_{p,q}$ is absent. If we further use the same formula for the hook Young tableau,

$$\langle s_{(a_q;l_p)} \rangle_{k,0}^{\text{GC}} = z E_{l_p} (1 + zQP)^{-1} Q E_{a_q}, \quad (2.34)$$

then, as a corollary of the above formula, we can express the vacuum expectation value of a general loop operator as the determinant of those of the hook loop operators constructed by taking all combinations of the arm and leg lengths from the original Young tableau [52]:

$$\langle s_Y \rangle_{k,0}^{\text{GC}}(z) = \det \left(\langle s_{(a_q;l_p)} \rangle_{k,0}^{\text{GC}}(z) \right)_{\substack{1 \leq p \leq r \\ 1 \leq q \leq r}}. \quad (2.35)$$

Note that, without taking the vacuum expectation value, this is a classical mathematical result stating that the character itself satisfies the same type of formula,

$$s_Y = \det \left(s_{(a_q;l_p)} \right)_{\substack{1 \leq p \leq r \\ 1 \leq q \leq r}}, \quad (2.36)$$

which is called the Giambelli formula. Hence, in other words, what we have shown is that we can put to each loop operator in the Giambelli formula (2.36) the normalized vacuum expectation values in the sense of the grand canonical ensemble. Namely, the normalized vacuum expectation values defined in (2.30) are Giambelli compatible.

Next, let us turn to another special case of the partition function with no insertions $s_Y = 1$. Then, we find

$$\langle 1 \rangle_{k,M}^{\text{GC}}(z) = \det \left(E_{M-p} (1 + zQP)^{-1} E_{-M+q-1} \right)_{\substack{1 \leq p \leq M \\ 1 \leq q \leq M}}, \quad (2.37)$$

Table 2. An explanation of how the $EQPE$ term in the expansion of $E(1 + zQP)^{-1}E$ in (2.32) appears.

$$\det \begin{pmatrix} \frac{1}{u_1+v_1} & \boxed{\frac{1}{u_1+v_2}} & \cdots & \frac{1}{u_1+v_{N_2}} \\ \vdots & \vdots & & \vdots \\ \frac{1}{u_{N_1}+v_1} & \frac{1}{u_{N_1}+v_2} & \cdots & \frac{1}{u_{N_1}+v_{N_2}} \\ \boxed{v_1^{N_2-N_1-1}} & v_2^{N_2-N_1-1} & \cdots & v_{N_2}^{N_2-N_1-1} \\ v_1^{N_2-N_1-2} & v_2^{N_2-N_1-2} & \cdots & v_{N_2}^{N_2-N_1-2} \\ \vdots & \vdots & & \vdots \\ v_1 & v_2 & \cdots & v_{N_2} \\ 1 & 1 & \cdots & 1 \end{pmatrix} \det \begin{pmatrix} \boxed{\frac{1}{v_1+u_1}} & \frac{1}{v_2+u_1} & \cdots & \frac{1}{v_{N_2}+u_1} \\ \vdots & \vdots & & \vdots \\ \frac{1}{v_1+u_{N_1}} & \frac{1}{v_2+u_{N_1}} & \cdots & \frac{1}{v_{N_2}+u_{N_1}} \\ v_1^{N_2-N_1-1} & v_2^{N_2-N_1-1} & \cdots & v_{N_2}^{N_2-N_1-1} \\ v_1^{N_2-N_1-2} & \boxed{v_2^{N_2-N_1-2}} & \cdots & v_{N_2}^{N_2-N_1-2} \\ \vdots & \vdots & & \vdots \\ v_1 & v_2 & \cdots & v_{N_2} \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

since $l_p = M - p$ in this case. If we regard the trace operators without endpoints as closed strings and the meson operators $EQPQP \cdots E$, which appear by expanding the expression in (2.32), as open strings, this formula seems to state that the fractional M2-branes can be constructed from the open strings. This is reminiscent of the fact that classical D-brane solutions can be studied from open-string field theory [53].

So far we have stressed the beauty of this formula. It is also interesting to note that this formula is actually very useful. In Sect. 3.2, we shall see that the special form of the density matrix is very useful for computing the powers of the density matrix. Our formula in (2.32) also shows the same computability. Namely, due to the expression of the meson operator, we can start with a vector E and multiply matrices Q and P to it one after another without difficulty.

Let us now turn to the explanation of how this expression is obtained. First let us explain the formula for the partition function $\langle 1 \rangle_{k,M}^{GC}(z)$. Again, our starting point is the matrix model (1.1). Instead of the Cauchy determinant formula (2.2), let us consider the following formula:

$$\det \begin{pmatrix} \left(\frac{1}{u_i + v_j} \right)_{\substack{1 \leq i \leq N_1 \\ 1 \leq j \leq N_2}} \\ \left(v_j^{N_2-N_1-i} \right)_{\substack{1 \leq i \leq N_2-N_1 \\ 1 \leq j \leq N_2}} \end{pmatrix} = \frac{\prod_{i < j}^{N_1} (u_i - u_j) \prod_{i < j}^{N_2} (v_i - v_j)}{\prod_{i,j}^{N_1, N_2} (u_i + v_j)}. \tag{2.38}$$

This is a natural combination of the Vandermonde determinant and the Cauchy determinant, which reduces to the Vandermonde determinant for the special case of $N_1 = 0$, while reducing to the Cauchy determinant (2.2) for $N_2 = N_1$. This formula can be proved by starting with the Cauchy determinant (2.2) with matrix size $N_2 \times N_2$ and sending $v_{N_2}, v_{N_2-1}, \dots, v_{N_1+1}$ to infinity one after another.

Then, after plugging in $u_i = e^{x_i}$ and $v_j = e^{y_j}$, we can reproduce the measure of the ABJ matrix model. Let us see how we can put the measure into a Fredholm determinantal expression as in the case of the ABJM matrix model. Since the substitution can be done at any stage, for simplicity, let us stick to the variables u_i and v_j .

Since in the partition function (1.1) we have two factors of this measure, we have two determinants (see Table 2). So there are a lot of terms in expanding the determinant. Suppose we pick up one element from the lower $(N_2 - N_1) \times N_2$ submatrix of the first determinant, say $v_1^{N_2-N_1-1}$. This means that, in this term, we are not allowed to pick up the other elements in the same row or in the

same column. This element $v_1^{N_2-N_1-1}$ is associated with the integration of the variable v_1 . To finish the integration of v_1 , we need to decide which element to pick up from the first column in the second determinant.

If we pick up one element from the lower $(N_2 - N_1) \times N_2$ submatrix of the second determinant, say $v_1^{N_2-N_1-2}$, this again means that we do not pick up the others in the same row or column. Then, we do not have other elements depending on the variable v_1 and we can complete the v_1 integration.

Instead, suppose that in the second determinant we pick up another element from the upper $N_1 \times N_2$ submatrix, say $(v_1 + u_1)^{-1}$. Again, there are no other elements associated with the variable v_1 and we can integrate v_1 . But now we are associated with the integration of another variable u_1 . To finish the u_1 integration, we need to pick up one element from the first row of the first determinant, say $(u_1 + v_2)^{-1}$. (Note that we can no longer pick up the element $(u_1 + v_1)^{-1}$ from the first determinant, because we already pick up $v_1^{N_2-N_1-1}$). Then, we are associated with the integration of the other variable v_2 .

Now, in the second determinant, if we choose from the lower $(N_2 - N_1) \times N_2$ submatrix, then we finish the v_2 integration. But if we choose from the upper $N_1 \times N_2$ submatrix, we need to finish another integration of the variable u .

We can continue this process and find a term consisting of a series of matrix multiplications, with two monomials of v on the two ends connected by products of $(v + u)^{-1}(u + v)^{-1}$. After plugging in $u_i = e^{x_i}$ and $v_j = e^{y_j}$, essentially v on the two ends turns into E while $(v + u)^{-1}(u + v)^{-1}$ becomes QP . This can be regarded as an open string with Chan–Paton factors on the two ends. Of course, there are also contributions from the pure traces without ends. Now the only problem is to count the combinatorial factor correctly. The answer can be found more efficiently by preparing a mathematical formula for the matrix multiplications. This was done in Ref. [38].

For the vacuum expectation values of the loop operator, we can use another beautiful formula [54], stating that the supersymmetric Schur polynomial can be expressed as a ratio of two determinants, where the determinant in the denominator is nothing but the Vandermonde–Cauchy determinant that we encountered in (2.38). Hence, by combining the supersymmetric Schur polynomial with the measure, we find that one of the two determinants appearing in the measure of the partition function is replaced by another determinant. Hence, we can repeat the process explained in Table 2 and find the formula (2.31).

3. Non-perturbative corrections at large N

After presenting the Fermi gas formalism for the ABJ(M) matrix model in the previous section, we can now start our study of the large- μ expansion for the grand potential. For simplicity, we will mainly focus on the ABJM case. Since the expression as the partition function of a Fermi gas system already suggests a statistical-mechanical method, we first start with the WKB \hbar expansion. After that, we turn to the study of the exact values of the partition function at finite N and see how these exact values lead to the numerical study of the grand partition function. Finally, we compare with the free energy of the topological string theory and see how the exact expression is found.

3.1. WKB expansion

As emphasized in Refs. [32,55], the Fermi gas formalism enables us to study the M-theory regime (1.32). One useful way to study this regime is to perform the semiclassical (or WKB) expansion (1.34) of the grand potential $\tilde{J}_k(\mu) = \log \Xi_k(\mu)$ around $k = 0$ in the limit $\mu \rightarrow \infty$.

For example, the leading-order term $\tilde{J}^{(0)}(\mu)$ in the WKB expansion is easily found by replacing $\text{Tr}\rho^\ell$ in (2.15) by the classical phase space integral [32]:

$$\begin{aligned} \tilde{J}^{(0)}(\mu) &= \sum_{\ell=1}^{\infty} \frac{(-1)^{\ell-1}}{\ell} e^{\ell\mu} \int \frac{dqdp}{2\pi} \left(\frac{1}{2 \cosh \frac{q}{2}} \frac{1}{2 \cosh \frac{p}{2}} \right)^\ell \\ &= \frac{e^\mu}{4} {}_3F_2 \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}; 1, \frac{3}{2}; \frac{e^{2\mu}}{16} \right) - \frac{e^{2\mu}}{8\pi^2} {}_4F_3 \left(1, 1, 1, 1; \frac{3}{2}, \frac{3}{2}, 2; \frac{e^{2\mu}}{16} \right). \end{aligned} \quad (3.1)$$

From the large- μ behavior of the hypergeometric functions in the above expression,

$$\begin{aligned} \tilde{J}^{(0)}(\mu) &= \frac{2\mu^3}{3\pi^2} + \frac{\mu}{3} + \frac{2\zeta(3)}{\pi^2} + \frac{2}{3\pi^2} \left(-6\mu^2 + 6\mu + 6 - \pi^2 \right) e^{-2\mu} \\ &\quad + \frac{1}{2\pi^2} \left(-36\mu^2 - 66\mu + 25 - 6\pi^2 \right) e^{-4\mu} + \mathcal{O} \left(e^{-6\mu} \right), \end{aligned} \quad (3.2)$$

one can see that the perturbative part of $\tilde{J}^{(0)}(\mu)$ is indeed a cubic polynomial in μ . The higher-order correction of the WKB expansion can be systematically computed by the method of Wigner transformation [32], or the semiclassical analysis of the thermodynamic Bethe ansatz (TBA) equation [35]. It turns out that the coefficients C_k and $B_{k,0}$ in (1.13) do not receive corrections higher than $\mathcal{O}(k^1)$ and $\mathcal{O}(k^3)$, respectively. On the other hand, the constant A_k in (1.13) receives corrections of all orders in k .

We can continue this process to higher and higher \hbar corrections and reproduce the Taylor expansion of the functions $a_\ell(k)$ and $b_\ell(k)$ in (1.38) up to a certain large order of k . The coefficient of the membrane instanton was first obtained in this way [32,35].

3.2. Exact values of partition function at finite N

To fully study the instanton expansion, let us start the computations of the exact values of the partition function at finite N . The density matrix is of the form

$$\rho(q_1, q_2) = \frac{E(q_1) E(q_2)}{M(q_1) + M(q_2)}. \quad (3.3)$$

This type of integration kernel is quite ubiquitous and appears from time to time in modern physics. In the context of light-cone string field theory [56], the so-called Neumann coefficients, representing the overlap between the one-string state and the two-string state, take this form. From the integrability viewpoint, this form has a close relation to the TBA equation and was studied in Ref. [57].

Among other things, this expression of the density matrix implies that we can compute the powers of the density matrix without much difficulty. To explain the computation, we first rewrite (3.3) schematically as

$$\{M, \rho\} = E \otimes E, \quad (3.4)$$

if we regard ρ , M , and E as a symmetric matrix, a diagonal matrix, and a vector with continuous indices, respectively. Then, by studying the commutator $[M, \rho^n]$ for even n or the anti-commutator $\{M, \rho^n\}$ for odd n , we find

$$\rho^n(q_1, q_2) = \sum_{m=0}^{n-1} (-1)^m \frac{(\rho^m E)(q_1) (\rho^{n-1-m} E)(q_2)}{M(q_1) - (-1)^n M(q_2)}. \quad (3.5)$$

Table 3. The first few exact values of the partition function of the ABJM matrix model. See Ref. [22] for more exact values.

$$\begin{aligned}
 Z_1(1) &= \frac{1}{4}, & Z_1(2) &= \frac{1}{16\pi}, & Z_1(3) &= \frac{-3 + \pi}{64\pi}, & Z_1(4) &= \frac{10 - \pi^2}{1024\pi^2}, \\
 Z_2(1) &= \frac{1}{8}, & Z_2(2) &= \frac{1}{32\pi^2}, & Z_2(3) &= \frac{10 - \pi^2}{512\pi^2}, & Z_2(4) &= \frac{24 - 32\pi^2 + 3\pi^4}{49\,152\pi^4}, \\
 Z_3(1) &= \frac{1}{12}, & Z_3(2) &= \frac{-3 + \pi}{48\pi}, & Z_3(3) &= \frac{9 + 108\pi - 64\sqrt{3}\pi}{5184\pi}, \\
 Z_4(1) &= \frac{1}{16}, & Z_4(2) &= \frac{-8 + \pi^2}{512\pi^2}, & Z_4(3) &= \frac{-8 - 32\pi + 11\pi^2}{8192\pi^2}, \\
 Z_6(1) &= \frac{1}{24}, & Z_6(2) &= \frac{54 - 5\pi^2}{5184\pi^2}, & Z_6(3) &= \frac{189 + 192\sqrt{3}\pi - 125\pi^2}{186\,624\pi^2}.
 \end{aligned}$$

Table 4. The first few non-perturbative terms of the grand potential of the ABJM matrix model obtained from the numerical study. See Refs. [22,23] for more terms.

$$\begin{aligned}
 J_1^{\text{np}} &= \left[\frac{4\mu^2 + \mu + 1/4}{\pi^2} \right] e^{-4\mu} + \left[-\frac{52\mu^2 + \mu/2 + 9/16}{2\pi^2} + 2 \right] e^{-8\mu} + \mathcal{O}(e^{-12\mu}), \\
 J_2^{\text{np}} &= \left[\frac{4\mu^2 + 2\mu + 1}{\pi^2} \right] e^{-2\mu} + \left[-\frac{52\mu^2 + \mu + 9/4}{2\pi^2} + 2 \right] e^{-4\mu} + \mathcal{O}(e^{-6\mu}), \\
 J_3^{\text{np}} &= \frac{4}{3} e^{-\frac{4}{3}\mu} - 2e^{-\frac{8}{3}\mu} + \left[\frac{4\mu^2 + \mu + 1/4}{3\pi^2} + \frac{20}{9} \right] e^{-4\mu} - \frac{88}{9} e^{-\frac{16}{3}\mu} + \mathcal{O}(e^{-\frac{20}{3}\mu}), \\
 J_4^{\text{np}} &= e^{-\mu} + \left[-\frac{4\mu^2 + 2\mu + 1}{2\pi^2} \right] e^{-2\mu} + \frac{16}{3} e^{-3\mu} + \left[-\frac{52\mu^2 + \mu + 9/4}{4\pi^2} + 2 \right] e^{-4\mu} + \mathcal{O}(e^{-5\mu}), \\
 J_6^{\text{np}} &= \frac{4}{3} e^{-\frac{2}{3}\mu} - 2e^{-\frac{4}{3}\mu} + \left[\frac{4\mu^2 + 2\mu + 1}{3\pi^2} + \frac{20}{9} \right] e^{-2\mu} - \frac{88}{9} e^{-\frac{8}{3}\mu} + \mathcal{O}(e^{-\frac{10}{3}\mu}).
 \end{aligned}$$

Note that this formula shows that, instead of multiplying the matrices directly, we can pick up a vector E and multiply ρ by it one after another, which is much easier.

Using this method, in Ref. [22] we obtained the exact values of the partition function up to $(k, N_{\text{max}}) = (1, 44), (2, 20), (3, 18), (4, 16), (6, 14)$. See Table 3 for a few examples. By now, with some small technical progress, we believe that we can proceed to more exact values. See Ref. [39] for many exact values of the ABJ partition function.

3.3. Grand potential

After obtaining many exact values, we can now proceed to studying the exact coefficient of the total non-perturbative effects [22]. The results are given in Table 4. Since our analysis contains some guesswork, we shall spell out the full details, so that the reader can judge the accuracy for herself.

Table 5. Mathematica code for studying the coefficients of non-perturbative effects.

```
Z={1/16,1/512-1/(64*Pi^2),11/8192-1/(1024*Pi^2)-1/(256*Pi)};
(*Add as many exact values as possible*)
$MaxExtraPrecision=300;
Xi[w_]:=Exp[(ga1) w+(a12 mu^2+be2 mu+ga2) w^2];
(*Add higher-order terms w^3, w^4, ... similarly when necessary*)
k = 4; c[k_]:= 2/(Pi^2 k); b[k_]:= 1/(3 k) + k/24;
a[k_]:= If[EvenQ[k], -(Zeta[3]/(Pi^2 k)) \
- 2/k Sum[m (k/2 - m) Log[2 Sin[(2 Pi m)/k]], {m,1,k/2-1}], \
-(Zeta[3]/(8 Pi^2 k)) + k/4 Log[2] \
- 1/k Sum[(k + (-1)^m (2 m - k))/4 (3 k - (-1)^m (2 m - k))/4 \
Log[2 Sin[(Pi m)/k]], {m,1,k-1}]];
J0 = c[k]^(-1/3) AiryAi[c[k]^(-1/3) (n-b[k])];
J1 = CoefficientList[CoefficientList[Series[Xi[w],{w,0,2}],w][[2]],mu]. \
Table[(-1)^m D[c[k]^(-1/3) AiryAi[c[k]^(-1/3) (n + 4/k - b[k])],{n,m}], \
{m,0,0}] // Simplify;
J2 = CoefficientList[CoefficientList[Series[Xi[w],{w,0,2}],w][[3]],mu]. \
Table[(-1)^m D[c[k]^(-1/3) AiryAi[c[k]^(-1/3) (n + 8/k - b[k])],{n,m}], \
{m,0,2}] // Simplify;
(*Define higher instanton terms J3, J4, ... similarly when necessary*)
```

Let us pick up the case of $k = 4$ for concreteness. Suppose that we already know the expansion structure of the non-perturbative effects:

$$J_4^{\text{np}}(\mu) = \gamma_1 e^{-\mu} + (\alpha_2 \mu^2 + \beta_2 \mu + \gamma_2) e^{-2\mu} + \gamma_3 e^{-3\mu} + \dots \tag{3.6}$$

By rewriting the large- μ expansion of the grand potential in terms of the large- N expansion of the partition function, we find

$$Z_4(N) = Z_4^{(0)}(N) + Z_4^{(1)}(N) + Z_4^{(2)}(N) + \dots, \tag{3.7}$$

where each term is given by

$$\begin{aligned} Z_4^{(0)}(N) &= e^{A_4} C_4^{-\frac{1}{3}} \text{Ai} \left[C_4^{-\frac{1}{3}} (N - B_4) \right], \\ Z_4^{(1)}(N) &= \gamma_1 e^{A_4} C_4^{-\frac{1}{3}} \text{Ai} \left[C_4^{-\frac{1}{3}} (N + 1 - B_4) \right], \\ Z_4^{(2)}(N) &= \left(\alpha_2 \partial_N^2 - \beta_2 \partial_N + \gamma_2 + \frac{\gamma_1^2}{2} \right) e^{A_4} C_4^{-\frac{1}{3}} \text{Ai} \left[C_4^{-\frac{1}{3}} (N + 2 - B_4) \right], \end{aligned} \tag{3.8}$$

with $B_4 = B_{4,0}$ (1.14). Note that the argument of the first non-perturbative effect $Z_4^{(1)}(N)$ is shifted by 1 because of the exponent factor $e^{-\mu}$ in $J_4^{\text{np}}(\mu)$ (3.6). In the second non-perturbative effect $Z_4^{(2)}(N)$, instead of the polynomial of μ , we replace μ by $-\partial_N$. Note also that the lower non-perturbative effects come from the expansion of the exponential function in the grand canonical partition function.

Now let us fit the exact values against this function form. We first prepare as many exact values as possible. In the following analysis, we shall utilize all the values with $N_{\text{max}} = 16$ in Ref. [22]. As long as we know the command “FindFit” in Mathematica, it is by no means difficult to obtain the fitted result. Still, the reader may find it useful to directly apply the given program, so we also list the program in Table 5. Note that, since we are only interested in integral k , we have used the expression for A_k in Ref. [41]. Then, if we ask Mathematica what is the best fit for γ_1 by

```
FindFit[Log[N[Z, 50]], a[k]+Log[Abs[J0+J1]], {ga1}, n],
```

we find

$$\gamma_1 = 0.987\,217\,181\,038\,943\,116\,089\,720\,984\,763\,427\,159\,164\,382\,260\,644\,01.$$

So we have 2% accuracy for γ_1 . We are sure that the reader is not satisfied with this. For example, it is not possible to distinguish π from $\sqrt{10}$ with 2% accuracy. But it is not difficult to roughly estimate the errors of this value. Although we have exact values for the partition function, since we neglect higher orders of the non-perturbative corrections in the grand potential, at each order we always have an error of $e^{-4\mu/k}$. After plugging in the stationary condition $N = \partial J / \partial \mu = C_k \mu^2$ (1.35), we find the errors

$$e^{-4\mu/k} = e^{-2\pi\sqrt{2N/k}} \simeq 10^{-6}, \quad (3.9)$$

with $k = 4$ and $N \simeq 10$. Since we have quadratic coefficients at the next order, the errors are actually much bigger than we expect. But anyway, we can improve our results by including higher instantons:

```
FindFit[Log[N[Z, 50]], a[k] + Log[Abs[J0+J1+J2]], {ga1, a12, be2, ga2}, n].
```

Now the result is much improved:

$$\gamma_1 = 1.000\,000\,031\,595\,205\,703\,112\,645\,819\,864\,097\,712\,750\,672\,969\,024\,25. \quad (3.10)$$

We can continue this fitting to higher and higher orders. For example, for

```
FindFit[Log[N[Z, 50]], a[k] + Log[Abs[J0+J1+J2+J3+J4+J5]], \
{ga1, a12, be2, ga2, ga3, a14, be4, ga4, ga5}, n]
```

we find

$$\gamma_1 = 1.000\,000\,000\,000\,000\,000\,000\,000\,121\,199\,386\,553\,425\,131\,382\,322\,92. \quad (3.11)$$

However, if we go to the sixth non-perturbative corrections, the accuracy suddenly drops:

$$\gamma_1 = 1.000\,000\,000\,000\,000\,003\,438\,828\,680\,298\,020\,743\,302\,548\,140\,694\,23. \quad (3.12)$$

The reason is that, although we only find roughly 25-digit accuracy, we have already used up the 50 digits in the exact values of the partition function. Therefore, we improve by preparing 100 digits from the beginning:

```
FindFit[Log[N[Z, 100]], a[k] + Log[Abs[J0+J1+J2+J3+J4+J5+J6]], \
{ga1, a12, be2, ga2, ga3, a14, be4, ga4, ga5, a16, be6, ga6}, n].
```

Then we find

$$\gamma_1 = 0.999\,999\,999\,999\,999\,999\,999\,999\,999\,998\,548\,662\,756\,071\,218\,409\,269\,666\,\dots \quad (3.13)$$

When we extend to the eighth non-perturbative corrections, an error message appears again:

```
Failed to converge to the requested accuracy or precision within
100 iterations.
```

But we can increase to 150 digits:

```
FindFit[Log[N[Z, 150]], a[k] + Log[Abs[J0+J1+J2+J3+J4+J5+J6+J7+J8]], \
{ga1, a12, be2, ga2, ga3, a14, be4, ga4, ga5, a16, be6, ga6, ga7, a18,
be8, ga8}, n]
```

to avoid the error message and find

$$\gamma_1 = 0.999\ 999\ 999\ 999\ 999\ 999\ 999\ 999\ 999\ 999\ 999\ 999\ 999\ 999\ 999\ 994\ 165\ 937\ \dots . \tag{3.14}$$

Note that we should end our fitting at this point. We have only 16 pieces of data and if we fit up to the eighth non-perturbative corrections we use up all the data and it does not make sense to fit with more unknowns. We believe that we can go a little further to find more data with good machines. However, at present, let us be satisfied with more than 50-digit accuracy. After we are satisfied with the first coefficient, we can simply plug in the correct value and proceed to finding coefficients of higher orders.

Here we pick up the $k = 4$ case and explain our numerical study and the guesswork in detail. We hope we have convinced the reader with the high accuracy that we can reach at this point.

3.4. Comparison with topological strings

Note that each term of the grand potential obtained in Table 4 has exactly the same exponent as the worldsheet instanton $e^{-\frac{4m}{k}\mu}$. Hence we can follow the proposal from Ref. [42] to compare the result with the free energy of the topological string theory on local $\mathbb{P}^1 \times \mathbb{P}^1$, which describes the worldsheet instantons. We find that the numerical values of the first instanton for $k = 3, 4, 6$ and the values of the second instanton for $k = 6$ match the prediction in (1.37) coming from F^{top} in (1.16). However, if we trust this expression and extrapolate to smaller k , we find that the match does not hold any more. Besides, the expression is divergent at $k = 1, 2$ for the first instanton and $k = 1, 2, 4$ for the second instanton. In fact, this divergence gives us an important clue to understanding the whole instanton expansion. The exponents of the terms containing divergences are $e^{-2\ell\mu}$, which are nothing but the exponents of the membrane instantons. Besides, if we look at the behavior of the worldsheet instantons around the divergence, we find

$$\begin{aligned} \lim_{k \rightarrow 2} d_1(k)e^{-\frac{4\mu}{k}} &= \left[\frac{4}{\pi^2(k-2)^2} + \frac{4(\mu+1)}{\pi^2(k-2)} + \frac{2\mu^2+2\mu+1}{\pi^2} + \frac{1}{3} \right] e^{-2\mu}, \\ \lim_{k \rightarrow 4} d_2(k)e^{-\frac{8\mu}{k}} &= \left[-\frac{8}{\pi^2(k-4)^2} - \frac{4(\mu+1)}{\pi^2(k-4)} - \frac{2\mu^2+2\mu+1}{2\pi^2} - \frac{7}{6} \right] e^{-2\mu}, \end{aligned} \tag{3.15}$$

whose finite part already looks close to the expression of the numerical values. In this way, we can expect that the membrane is also divergent at these points so that, in total, the divergences cancel among themselves. From Table 4 we also observe that the divergence of the membrane instanton $e^{-2\ell\mu}$ for an odd integer ℓ comes from even integers k , while the divergence for an even integer ℓ comes from all integers k . Also, the coefficients of the membrane instanton $e^{-2\ell\mu}$ for odd ℓ vanish at odd integers k . Analytically continuing to non-integral k , from this observation, we can expect that the ℓ th membrane instanton is given by a periodic function. In particular, if we assume that it is expressed by using the trigonometric function $\tan \frac{\pi k\ell}{4}$ and simple rational functions, we can determine the function form of the membrane instantons by combining the results of the WKB expansion. In fact, this is how (1.38) was first determined in Refs. [22,35].

Note that there are no other exponents $e^{-\left(\frac{4m}{k}+2\ell\right)\mu}$ labeled by (m, ℓ) in (1.24) that have the same contribution as the first and second membrane instantons $e^{-2\mu}$ and $e^{-4\mu}$, except the pure worldsheet instantons. For the higher instanton effects, however, the situation is more difficult. The reason is that besides the worldsheet instantons and the membrane instantons we need to consider the bound states

between them, $e^{-\left(\frac{4m}{k}+2\ell\right)\mu}$ ($m \neq 0, \ell \neq 0$) [23,35]. We do not have clues to them from any systematic expansions because the bound state effects are detected neither from the 't Hooft expansions nor from the WKB expansions. The only clue is the pole cancellation mechanism, as we have explained in Sect. 1.5 and used above to fix the coefficients of the membrane instantons. Luckily, we find that, if we assume that the bound state effects $f_{m,\ell}(\mu)$ are expressed in terms of the corresponding worldsheet instantons $d_m(k)$ and the corresponding membrane instantons $a_\ell(k)$ as ($A_{m,\ell}(k) = -2m\pi^2 a_\ell(k)$) [23]

$$f_{m,\ell}(\mu) = d_m(k) \sum_{(p_1, \dots, p_\ell)} \frac{A_{m,1}^{p_1}(k) A_{m,2}^{p_2}(k) \cdots A_{m,\ell}^{p_\ell}(k)}{p_1! p_2! \cdots p_\ell!}, \tag{3.16}$$

with the sum running over all of the allowed partitions (p_1, \dots, p_ℓ) of ℓ satisfying $p_1 + 2p_2 + \cdots + \ell p_\ell = \ell$, we can observe a general cancellation among the worldsheet instantons, the membrane instantons, and their bound states, whose finite results coincide beautifully with the numerical studies. This observation implies that, if we redefine the chemical potential μ as

$$\mu_{\text{eff}} = \mu + \frac{1}{C_k} \sum_{\ell=1}^{\infty} a_\ell(k) e^{-2\ell\mu}, \tag{3.17}$$

or, more explicitly, as in (1.23) for integral k ,

- the bound state effects are absorbed into the worldsheet instanton effects.

Besides, we have a bonus from this redefinition:

- the quadratic polynomial coefficients of the membrane instantons are reduced into linear polynomials,

$$J(\mu_{\text{eff}}) = \frac{C}{3} \mu_{\text{eff}}^3 + B \mu_{\text{eff}} + A + \sum_{m=1}^{\infty} d_m(k) e^{-\frac{4m\mu_{\text{eff}}}{k}} + \sum_{\ell=1}^{\infty} (\tilde{b}_\ell(k) \mu_{\text{eff}} + \tilde{c}_\ell(k)) e^{-2\ell\mu_{\text{eff}}}, \tag{3.18}$$

and the coefficients $\tilde{c}_\ell(k)$ are expressed in terms of $\tilde{b}_\ell(k)$ as

$$\tilde{c}_\ell(k) = -k^2 \frac{\partial}{\partial k} \frac{\tilde{b}_\ell(k)}{2\ell k}. \tag{3.19}$$

3.5. Topological string theory

In the above several subsections, we have seen that the WKB expansion, computation of the partition function, and the pole cancellation mechanism can determine all of the worldsheet instantons, the membrane instantons, and their bound states order by order. It is not clear, however, what the coefficient functions of higher orders generally look like. In this subsection, we shall explain a further relation between the ABJ(M) matrix model and topological string theory and determine the coefficient functions systematically.

We shall sketch a relation between the membrane instanton and the refined topological string theory in the NS limit after recapitulating the relation between the worldsheet instanton and the (unrefined) topological strings. The large- N expansion of the free energy in the ABJ(M) theory is formally equivalent to the free energy of pure Chern–Simons theory on $\mathbb{R}\mathbb{P}^3$, which in turn is related to the topological string theory on a non-compact Calabi–Yau manifold, local $\mathbb{P}^1 \times \mathbb{P}^1$ [42]. It turns out that the chemical potential in the ABJ(M) Fermi gas is a natural variable in the large-radius expansion on the topological string side. The Kähler parameters of two \mathbb{P}^1 in local $\mathbb{P}^1 \times \mathbb{P}^1$ are identified with

the chemical potential as in (1.21), and they are exchanged under the Seiberg-like duality of the ABJ theory. Surprisingly, it was found in Ref. [24] that the non-perturbative corrections in the ABJ(M) matrix model are completely determined by the refined topological string on local $\mathbb{P}^1 \times \mathbb{P}^1$, and the precise relation is given by (1.16). The first term in (1.16) represents the worldsheet instanton, which corresponds to the standard topological string. This part of the correspondence (1.16) is the usual story for the genus expansion of matrix models, and it can be checked, in principle, to any order of the string coupling $g_s = 2/k$. The non-trivial part of the proposal in Ref. [24] is the second term in (1.16), namely, the correspondence between the membrane instantons in the ABJ(M) matrix model and the refined topological string in the NS limit. Note that in (1.16) the pole cancellation mechanism works within every BPS multiplet. Namely, we can check that the combination of F_{top} and F_{NS} in (1.16) with common BPS indices $N_{j_L, j_R}^{\mathbf{d}}$ does not contain any divergences. Also, we stress that not only the singular part but also the finite part are correctly reproduced from (1.16).

As discussed in Refs. [58–60], the refined topological string theory in the NS limit is closely related to the quantization of mirror curves of local Calabi–Yau manifolds. One can easily see that the Fermi surface $H(x, p) = 0$ of the ABJM Fermi gas corresponds to the mirror curve of the diagonal local $\mathbb{P}^1 \times \mathbb{P}^1$. More generally, the mirror curve of local $\mathbb{P}^1 \times \mathbb{P}^1$ is given by

$$H(x, p) = -1 + e^x + e^p + z_1 e^{-x} + z_2 e^{-p} = 0. \tag{3.20}$$

We can “quantize” this curve with $[x, p] = i\hbar$, where the Planck constant \hbar of this quantum-mechanical system is unrelated to the previous Planck constant in the Fermi gas system and is determined later. The corresponding difference equation for the wave function $\Psi(x)$ is

$$(-1 + e^x + z_1 e^{-x}) \Psi(x) + \Psi(x + \hbar) + z_2 \Psi(x - \hbar) = 0, \tag{3.21}$$

whose formal solution

$$\Psi(x) = \exp\left[\frac{1}{\hbar} S(x, \hbar)\right], \quad S(x, \hbar) = \sum_{n=0}^{\infty} S_n(x) \hbar^{2n}, \tag{3.22}$$

gives the “quantum periods”

$$\Pi_A = \oint_A dS, \quad \Pi_B = \oint_B dS. \tag{3.23}$$

Using the quantum periods, we can compute the free energy in the NS limit [60].

After identifying the two moduli z_I ($I = 1, 2$) in (3.20) and \hbar as

$$z_I = e^{-\frac{T_I}{g_s}}, \quad \hbar = \frac{2\pi i}{g_s} = \pi i k, \tag{3.24}$$

we find that the membrane instanton coefficients $a_\ell(k)$ and $b_\ell(k)$ appear as the coefficients of the quantum A-period Π_A and the quantum B-period Π_B , respectively, in the small- z_I expansion. Let us see explicitly how this works in the following.

We shall explain how to compute the quantum periods systematically up to any desired order in the small- z_I expansion. As discussed in Refs. [24,60], it is convenient to introduce $V(x)$ by

$$V(x) = \frac{\Psi(x + \hbar)}{\Psi(x)}. \tag{3.25}$$

Then (3.21) is rewritten as

$$V(X) = 1 - X - \frac{z_1}{X} - \frac{z_2}{V(Xq^{-1})}, \tag{3.26}$$

where we have also introduced $X = e^x$ and $q = e^{\hbar}$. One can easily solve (3.26) in the small- z_I expansion as

$$V(X) = 1 - X - \frac{z_1}{X} + \frac{qz_2}{X - q} - \frac{q^3 z_1 z_2}{X(X - q)^2} + \frac{q^4 z_2^2}{(X - q)^2(X - q^2)} + \mathcal{O}(z^3). \tag{3.27}$$

Once we know $V(X)$, the quantum periods are obtained by rewriting the contour integral as a sum of residues:⁴

$$\begin{aligned} \Pi_A(z_1, z_2; q) &= -\text{Res}_{X=0} \frac{1}{X} \log \left[\frac{V(X)}{1 - X} \right], \\ \Pi_B(z_1, z_2; q) &= -\sum_{j \geq 0} \text{Res}_{X=q^j} \frac{\log(X)}{X} \log \left[\frac{V(X)}{1 - X} \right]. \end{aligned} \tag{3.28}$$

The quantum A-period is given by

$$\Pi_A(z_1, z_2, q) = z_1 + z_2 + (4 + q + q^{-1}) z_1 z_2 + \frac{3}{2} (z_1^2 + z_2^2) + \mathcal{O}(z^3), \tag{3.29}$$

which defines a quantum version of the mirror map

$$\frac{1}{2} \log \frac{Q_I}{z_I} = \Pi_A(z_1, z_2, q). \tag{3.30}$$

Here Q_I is related to the ‘‘effective’’ Kähler parameter T_I^{eff} by

$$Q_I = e^{-\frac{T_I^{\text{eff}}}{gs}}, \tag{3.31}$$

which takes care of the effect of the bound states. Inverting the quantum mirror map, we can express z_I in terms of Q_I :

$$-\frac{1}{2} \log \frac{z_I}{Q_I} = Q_1 - \frac{Q_1^2}{2} + Q_2 - \frac{Q_2^2}{2} + \frac{q^2 - q^{-2}}{q - q^{-1}} Q_1 Q_2 + \mathcal{O}(Q^3). \tag{3.32}$$

On the other hand, the quantum B-period is related to the derivative of the free energy in the NS limit as

$$\begin{aligned} Q_2 \frac{\partial}{\partial Q_2} F_{\text{NS}}(Q_1, Q_2) &= -\frac{2}{\hbar} \left[2\Pi_B^{\text{even}}(z_1, z_2; q) - \Pi_A(z_1, z_2, q)^2 \right], \\ Q_1 \frac{\partial}{\partial Q_1} F_{\text{NS}}(Q_1, Q_2) &= -\frac{2}{\hbar} \left[2\Pi_B^{\text{even}}(z_2, z_1; q) - \Pi_A(z_1, z_2, q)^2 \right], \end{aligned} \tag{3.33}$$

where Π_B^{even} denotes the even-power part of the quantum B-period in the \hbar expansion. Calculating the residue in (3.28), the right-hand side of (3.33) is easily found to be

$$\begin{aligned} \frac{1}{\hbar} \left[2\Pi_B^{\text{even}}(z_1, z_2; q) - \Pi_A(z_1, z_2, q)^2 \right] &= \frac{q + 1}{q - 1} z_2 + \frac{(q + 1)^3}{(q - 1)q} z_1 z_2 \\ &\quad + \frac{5 + 8q + 5q^2}{2(q^2 - 1)} z_2^2 + \mathcal{O}(z^3). \end{aligned} \tag{3.34}$$

⁴ Here we will use a slightly different normalization of quantum periods as compared to Ref. [24].

Rewriting the above expression in terms of Q_I using the quantum mirror map (3.32) and integrating (3.33), we finally arrive at

$$F_{\text{NS}}(Q_1, Q_2, q) = -\frac{q+1}{q-1}(Q_1 + Q_2) - \frac{(q^2+1)}{4(q^2-1)}(Q_1^2 + Q_2^2) - \frac{(q^2+1)(q+1)}{q(q-1)}Q_1Q_2 + \mathcal{O}(Q^3). \tag{3.35}$$

Plugging this expression into (1.16), we can show that our formula (1.16) reproduces all the known results of the membrane instantons and the bound states in the ABJ(M) matrix model [24,39].

From this non-trivial correspondence in the ABJ(M) theory between the membrane instantons and the refined topological string theory in the NS limit, in Ref. [24] it is further conjectured that this correspondence holds for general local Calabi–Yau manifolds, and the combination of F_{top} and F_{NS} appearing in (1.16) gives a non-perturbative completion of the topological string theory. However, there is one subtle point: we have to turn on an extra discrete B -field along the worldsheet instanton part in order for the pole cancellation to work for the general case. See Ref. [24] for details.

In Refs. [61–65], it is argued that this correspondence between the membrane instantons and the refined topological string theory in the NS limit in the ABJ(M) theory can be naturally understood from the “dual” perspective, by studying the spectral problem (2.16) of the Hamiltonian in the Fermi gas system.

4. Related topics

In the previous section, we have reviewed the progress in understanding the partition functions of the ABJM theory and the ABJ theory. Here we shall explain some progress after finding the exact instanton expansion of the ABJ(M) partition function.

4.1. Exact results with supersymmetry enhancement

As seen before, the large- μ expansion of the grand potential is completely determined by the refined topological string theory on local $\mathbb{P}^1 \times \mathbb{P}^1$ [24]. However, it is not technically easy to compute the higher-order instanton corrections in practice. It is well known that, at $k = 1, 2$, the supersymmetry of the ABJM theory is enhanced from $\mathcal{N} = 6$ to $\mathcal{N} = 8$. It is natural to expect that for these levels some simplifications happen. In fact, as we will see below, the large- μ expansion of the grand potential drastically simplifies at $k = 1, 2$ [66]. We can write this down in a closed form in terms of the topological string free energy.

Let us first write down the results in Ref. [66]. The complete large- μ expansion of $J_k(\mu)$ at $k = 1, 2$ is exactly given by

$$J_1(\mu) = \frac{1}{16\pi^2} \left(F_0(t) - tF_0'(t) + \frac{1}{2}t^2F_0''(t) \right) + \frac{3\mu}{8} + \frac{\log 2}{4} + F_1(t) + F_1^{\text{NS}}(t),$$

$$J_2(\mu) = \frac{1}{4\pi^2} \left(F_0(t) - tF_0'(t) + \frac{1}{2}t^2F_0''(t) \right) + \frac{\mu}{4} + F_1(t) + F_1^{\text{NS}}(t). \tag{4.1}$$

We need to explain the notations in these equations.⁵ The functions $F_0(t)$ and $F_1(t)$ are the standard genus zero and genus one free energies on local $\mathbb{P}^1 \times \mathbb{P}^1$, respectively. These are computed in the

⁵ We denote the effective Kähler modulus T^{eff} by t in the current case with fixed k . Since there are two Kähler moduli t_1 and t_2 in local $\mathbb{P}^1 \times \mathbb{P}^1$, the free energy is in general a function of these two parameters (t_1, t_2) . Here we denote the free energy in the diagonal slice by $F_g(t, t) = F_g(t)$.

standard way for the special geometry. The function $F_1^{\text{NS}}(t)$ is the first correction to the refined topological string free energy in the NS limit. The Kähler modulus t is related to the complex modulus z by the mirror map

$$t = -\log z + 4z {}_4F_3\left(1, 1, \frac{3}{2}, \frac{3}{2}; 2, 2, 2; -16z\right). \tag{4.2}$$

The only difference between $k = 1$ and $k = 2$ is the identification of z and μ :

$$z = e^{-4\mu} \quad \text{for } k = 1, \quad z = e^{-2\mu} \quad \text{for } k = 2. \tag{4.3}$$

At large radius point ($t \rightarrow \infty$), the genus zero free energy is expanded as

$$F_0(t) = \frac{t^3}{6} - 2\zeta(3) + 4e^{-t} - \frac{9}{2}e^{-2t} + \mathcal{O}(e^{-3t}). \tag{4.4}$$

Eliminating t by (4.2), one easily finds

$$\begin{aligned} F_0(t) - tF_0'(t) + \frac{1}{2}t^2F_0''(t) &= -\frac{1}{6}\log^3 z - 2\zeta(3) + 4\left(\log^2 z - \log z + 1\right)z \\ &\quad - \frac{1}{2}\left(52\log^2 z - 2\log z + 9\right)z^2 + \mathcal{O}(z^3). \end{aligned} \tag{4.5}$$

The free energies $F_1(t)$ and $F_1^{\text{NS}}(t)$ are exactly given by

$$\begin{aligned} F_1(t) &= -\frac{1}{12}\log[64z(1+16z)] - \frac{1}{2}\log\left[\frac{K(-16z)}{\pi}\right] = -\frac{\log z}{12} + \frac{2}{3}z - \frac{10}{3}z^2 + \mathcal{O}(z^3), \\ F_1^{\text{NS}}(t) &= \frac{1}{12}\log z - \frac{1}{24}\log(1+16z) = \frac{\log z}{12} - \frac{2}{3}z + \frac{16}{3}z^2 + \mathcal{O}(z^3), \end{aligned} \tag{4.6}$$

where $K(z)$ is the complete elliptic integral of the first kind. Plugging these results into (4.1), one can check that the large- μ expansion in Table 4 is correctly reproduced.

In the following, we briefly explain how to obtain these results. Here we focus on the case of $k = 2$. What we should do is to take the limit $k \rightarrow 2$ in the general formula (1.16). As mentioned before, both the worldsheet instanton correction and the membrane instanton correction are divergent in this limit, but the sum of them is finite. Since we already know that the divergences are canceled, we can focus on the finite parts. Let us see the worldsheet instanton part in (1.16). It is easy to see that the finite part comes only from $g = 0, 1$. The $g = 1$ part is trivial. In the limit $k \rightarrow 2$, each term of the $g = 0$ part is

$$\lim_{k \rightarrow 2} \left(2 \sin \frac{2\pi w}{k}\right)^{-2} e^{-\frac{4dw\mu}{k}} \sim \frac{e^{-2dw\mu}}{12\pi^2 w^2} \left(3 + \pi^2 w^2 + 6dw\mu + 6d^2 w^2 \mu^2\right), \tag{4.7}$$

where \sim means that we extract only the finite part. Therefore, the finite part of the worldsheet instanton correction is finally given by

$$\frac{1}{4\pi^2} F_0^{\text{inst}}(t) - \frac{t}{4\pi^2} \partial_t F_0^{\text{inst}}(t) + \frac{t^2}{8\pi^2} \partial_t^2 F_0^{\text{inst}}(t) + F_1^{\text{inst}}(t), \tag{4.8}$$

where $F_g^{\text{inst}}(t)$ is the instanton part in $F_g(t)$. We have used the relation $t = 2\mu_{\text{eff}}$ at $k = 2$. Similarly, as shown in Ref. [66], the finite part of the membrane instanton correction is finally given by

$$\frac{1}{8} (\partial_{t_1} - \partial_{t_2})^2 F_0(t_1, t_2) |_{t_1=t_2=t} + F_1^{\text{NS,inst}}(t) = -\frac{t}{8} + \frac{\mu}{4} + F_1^{\text{NS,inst}}(t). \tag{4.9}$$

From these results, one obtains the exact expressions (4.1).

Once the modified grand potential is known, one can compute the grand partition function. As seen before, the grand potential is constructed from $J_k(\mu)$ by (1.8), with the subscript M dropped here. Let us plug the result (4.1) into (1.8). For $k = 2$, using

$$e^{J_2(\mu+2\pi in)} = e^{J_2(\mu)} \exp \left[\pi in^2 \tau + 2\pi in \left(\xi - \frac{1}{12} \right) \right], \tag{4.10}$$

with

$$\tau = \frac{2i}{\pi} F_0''(t), \quad \xi = \frac{1}{2\pi^2} (t F_0''(t) - F_0'(t)), \tag{4.11}$$

we find that the grand partition function is expressed in terms of the elliptic theta function by

$$\Xi_2(\mu) = e^{J_2(\mu)} \vartheta_3 \left(\xi - \frac{1}{12}, \tau \right). \tag{4.12}$$

Similarly, the grand partition function at $k = 1$ is given by

$$\Xi_1(\mu) = e^{J_1(\mu)} \vartheta_3 \left(\frac{\xi}{2} - \frac{7}{24}, \tau \right). \tag{4.13}$$

We stress that these expressions are *exact*. All the instanton corrections are encoded through the topological string free energies $F_0(t)$, $F_1(t)$, and $F_1^{\text{NS}}(t)$.

There are two important consequences of the exact expressions (4.12) and (4.13). One is that one can know the eigenvalues of the density matrix ρ as the zeros of $\Xi_k(\mu)$. In fact, it is well known that the theta function has an infinite number of zeros. It was confirmed in Ref. [66] that the zeros in (4.12) and (4.13) precisely reproduce the eigenvalues of ρ at $k = 1, 2$, numerically computed by the exact quantization condition in Ref. [61]. The other important point is that it is possible to analytically continue $\Xi_k(\mu)$ from $\mu \rightarrow \infty$ to $\mu \rightarrow -\infty$. The continuation is essentially the modular S-transformation $\tau \rightarrow \bar{\tau} = -1/\tau$. In Ref. [66], this analytic continuation was indeed performed, and then the small- e^μ expansion was computed. The results perfectly agree with the exact values of the partition function in Table 3 (see Refs. [21,22] for more values)!

Furthermore, in Refs. [66,67], the exact results for ABJM at $k = 4$ and ABJ at $(k, M) = (2, 1)$ were presented. These also reproduce the known results perfectly. Note that at $k = 4$ the supersymmetry is no longer enhanced. Nevertheless, we can write down the grand potential in a closed form. Moreover, in Ref. [67], exact functional equations among the grand partition functions in the ABJ matrix model were found. These relations are quite similar to the so-called quantum Wronskian relations [68]. This seems to imply an interesting connection to integrable models.

4.2. 't Hooft expansion and Borel resummation

Throughout this paper, we focus on the large- N expansion in the M-theory limit (1.32). As already seen in Sect. 1.4, it is also interesting to consider the 't Hooft limit (1.26). In this limit, the ABJM theory is dual to type IIA string theory on $AdS_4 \times \mathbb{CP}^3$, and the large- N expansion is related to the genus expansion in this string theory. In the 't Hooft limit, the free energy admits the large- N expansion (1.27). The genus g free energy $F_g(\lambda)$ can be exactly computed by using the holomorphic anomaly equation order by order [25]. Then, the genus expansion (1.27) turns out to be an asymptotic series for a given finite λ . The standard way to resum such formal asymptotic series is known as Borel

resummations. For the genus expansion (1.27), let us define its Borel transform by

$$\mathcal{B}F(\zeta) = \sum_{g=2}^{\infty} \frac{F_g(\lambda)}{(2g-3)!} \zeta^{2g-2}. \tag{4.14}$$

Remarkably, as observed in Refs. [29,69], this Borel transform has no singularities on the positive real axis. Therefore, the following Borel summation is well defined:

$$\mathcal{S}F(N, \lambda) = N^{-2}F_0(\lambda) + F_1(\lambda) + \int_0^{\infty} \frac{d\zeta}{\zeta} e^{-\zeta} \mathcal{B}F(N\zeta). \tag{4.15}$$

The crucial conclusion in Ref. [69] is that this Borel summation does not reproduce the exact values of $F(N, \lambda)$. There is an exponentially suppressed correction between them:

$$F^{\text{exact}}(N, \lambda) - \mathcal{S}F(N, \lambda) = \mathcal{O}\left(e^{-\pi N \sqrt{2/\lambda}}\right). \tag{4.16}$$

This correction is nothing but the membrane instanton correction! It is not captured by the perturbative $1/N$ expansion. Thus, in the 't Hooft limit, the perturbative genus expansion and its Borel resummation are insufficient to reconstruct the exact free energy. We need to take into account the membrane instanton corrections in addition to the perturbative expansion. The incorporation of such membrane instantons in the Borel analysis is an open problem. As pointed out in Ref. [29], such corrections should be related to complex singularities of the Borel transform (4.14).

As reviewed in this paper, we already know the complete M-theoretic expansion of the free energy with the help of the topological string. It should, in principle, be possible to use this result to understand the membrane instanton corrections to the genus expansion. Interestingly, it was observed in Ref. [70] that the Borel resummation (4.15) already contains a membrane instanton-like correction in the M-theoretic expansion from the Fermi gas result. It would be interesting to reveal a quantitative relation between the 't Hooft expansion and the M-theoretic expansion.

4.3. Other theories

In Sect. 4.1, we have seen special cases with supersymmetry enhancements. Here, let us see how the analysis can be generalized to other theories of M2-branes. In particular, when we are interested in gaining the universal property of the M2-branes, it is very important to generalize to as many theories as possible. We shall list some results obtained recently.

First, let us give a viewpoint from quiver diagrams, where the vertices represent the gauge group factor $U(N)$ and the edges represent the bifundamental matters between the two connected factors. Then, the quiver diagram of the ABJM theory is nothing but the Dynkin diagram of affine A_1 or \widehat{A}_1 . Then, it is natural to ask about a general matrix model associated with \widehat{A}_r .

It has already been found in Ref. [32] that, if the density matrix is Hermitian, the perturbative part of any \widehat{A}_r matrix model is summed up to the Airy function with the coefficient C given explicitly in Refs. [32,71]. The other coefficients are not easy to obtain in general. In Ref. [40] it was found that, as long as the quiver diagram associated with the levels is a repetition of an original quiver diagram, we can express the grand potential of the new quiver diagram completely in terms of that of the original diagram. In Ref. [72] we restricted ourselves to the $\mathcal{N} = 4$ case [73], where the Chern–Simons levels are given by

$$k_a = \frac{k}{2} (s_a - s_{a-1}), \quad s_a = \pm 1, \tag{4.17}$$

and found that the explicit form of the coefficient B is given by

$$B = \frac{B^{(0)}}{k} + kB^{(2)}, \tag{4.18}$$

with

$$\begin{aligned} B^{(0)} &= -\frac{1}{6} \left[\frac{\Sigma(p)}{\Sigma(q)} + \frac{\Sigma(q)}{\Sigma(p)} - \frac{4}{\Sigma(q)\Sigma(p)} \right], \\ B^{(2)} &= \frac{1}{24} \left[\Sigma(q)\Sigma(p) - 12 \left(\frac{\Sigma(q, p, q)}{\Sigma(q)} + \frac{\Sigma(p, q, p)}{\Sigma(p)} - \frac{\Sigma(q, p)\Sigma(p, q)}{\Sigma(q)\Sigma(p)} \right) \right]. \end{aligned} \tag{4.19}$$

Here we adopt the notation of $\Sigma(L)$, with L denoting an alternating sequence of q and p , whose definition is given by

$$\begin{aligned} \Sigma(q) &= \sum_{a=1}^m q_a, & \Sigma(p) &= \sum_{a=1}^m p_a, \\ \Sigma(q, p) &= \sum_{1 \leq a \leq b \leq m} q_a p_b, & \Sigma(p, q) &= \sum_{1 \leq a \leq b \leq m} p_a q_b, \\ \Sigma(q, p, q) &= \sum_{1 \leq a \leq b < c \leq m} q_a p_b q_c, & \Sigma(p, q, p) &= \sum_{1 \leq a < b \leq c \leq m} p_a q_b p_c. \end{aligned} \tag{4.20}$$

Note that the condition in each sum can be stated as the requirement that we choose q_a and p_a out of a sequence $q_1, p_1, q_2, p_2, \dots, q_m, p_m$ by respecting its ordering.

Furthermore, if we restrict ourselves to the special $\mathcal{N} = 4$ (q, p) model where the q edges with $s_a = +1$ and the p edges with $s_a = -1$ are separated, we find that the coefficient A is given in terms of that of the ABJM matrix model by [72]

$$A = \frac{1}{2} \left(p^2 A_{\text{ABJM}}(qk) + q^2 A_{\text{ABJM}}(pk) \right). \tag{4.21}$$

A similar proposal was made in Ref. [41] for the so-called M-theoretic matrix model [55,74], whose Fermi gas Hamiltonian is related to that of the $(q, 1)$ model by a canonical transformation [75]. In Ref. [76], the membrane instanton of the (q, p) model was studied and a systematic expansion in terms of the derivative operators acting on the classical expression was found. Using the results, we were able to study the $(2, 2)$ model in full detail [77] and find that the non-perturbative effects are given by the refined topological string theory, as in the ABJM matrix model (1.12). It still awaits to identify the Calabi–Yau manifold whose BPS index describes the $(2, 2)$ model, but we observed that the diagonal Gopakumar–Vafa invariant is nothing but that of the D_5 del Pezzo surface. It is interesting to find that the correspondence between the topological string theory and the matrix model coming from the superconformal theory is not restricted to the case of the ABJM theory. In Ref. [78], the non-perturbative effects were further explored. Based on the systematic WKB analysis and exact computations in various special cases, analytic expressions of the first few membrane instanton corrections and worldsheet instanton corrections were conjectured for general (q, p) and k .

The generalization, however, is not restricted to the $\widehat{A}\widehat{D}\widehat{E}$ quiver. In Ref. [79], an $\widehat{A}\widehat{D}\widehat{E}$ sector was specified by requiring that the long-range force among the eigenvalues vanishes. It is then interesting to ask whether there is a Fermi gas formalism for these series and whether the perturbative corrections are still summed up to the Airy function. In Refs. [80,81], we continue to study the \widehat{D} quiver. We find that, after using the Cauchy determinant (2.2) in a different way, we can still give a Fermi gas

formalism for this type of theory and prove the previous proposal [82]:

$$C = \frac{1}{\pi^2 k} \left(\frac{1}{\sigma_0 \sigma_1} + \sum_{m=1}^{r-1} \frac{s_m - s_{m+1}}{\sigma_m \sigma_{m+1}} + \frac{s_r}{\sigma_r \sigma_{r+1}} \right), \quad (4.22)$$

with the variables σ_m given as

$$\sigma_m = \sum_{n=1}^r (|s_m - s_n| + |s_m + s_n|) - 4|s_m|, \quad \sigma_0 = 2(r-2), \quad \sigma_{r+1} = 2 \sum_{n=1}^r |s_n|, \quad (4.23)$$

by the reordered s ,

$$0 \leq |s_r| \leq |s_{r-1}| \leq \dots \leq |s_1|. \quad (4.24)$$

As in the $\mathcal{N} = 4$ case of the \widehat{A} quiver, here we can also restrict ourselves to the case of uniform s_a , $s_a = 1$, and find that the coefficient A is consistent with

$$A = \frac{1}{2} \left(A_{\text{ABJM}}(2rk) + r^2 A_{\text{ABJM}}(2(r-2)k) \right), \quad (4.25)$$

up to $\mathcal{O}(k^5)$, which is very similar to (4.21) for the (q, p) model. We hope that all these findings will finally be completed to a more systematic understanding.

5. Summary and further directions

In this review article, we have explained carefully the quantum effects of the ABJM theory and the ABJ theory, where the partition function and the vacuum expectation values of the half-BPS Wilson loop can be studied in full detail, including both the perturbative and non-perturbative corrections.

We have seen many interesting properties. For example, the leading large- N behavior of the free energy is given by $\sim k^{\frac{1}{2}} N^{\frac{3}{2}}$. The perturbative $1/N$ corrections are summed up to an Airy function. The non-perturbative corrections include the effects interpreted as worldsheet instantons and those interpreted as membrane instantons. In general, the coefficients of both instantons and their bound states can be divergent at certain values of k , though as a whole the divergences in these instanton effects cancel among themselves and the final result is completely finite.

Also, in Sect. 4.3, we have seen that many generalizations enjoy these interesting properties. Hence, one further direction is to generalize our analysis to as many theories as possible and identify this class clearly. It is surprising to find that the superconformal Chern–Simons theory of the \widehat{D} quiver can also be formulated with the Fermi gas system. It is then natural to ask whether we can apply the same method even for the \widehat{E} quiver or its cousins with orthogonal or symplectic groups [74,83]. We hope that through these studies we will end up with a universal understanding of the M2-branes.

Next, let us also stress that, after a large number of numerical analyses, our large- μ expansion of the grand potential (or large- N expansion of the partition function) remains a conjecture. It would, of course, be interesting to see how this conjecture is proved with a mathematical method. The study explained in Sect. 4.1 would be a first attempt in this direction.

Finally, let us come back to a more physical question. Our original motivation in studying the ABJM partition function was to understand the M2-branes. After seeing all of the perturbative and non-perturbative corrections, it is reasonable to come back to ask what we have learned about M2-brane interactions. Unfortunately, we do not have much to say about this at this stage, except that some of the results can be rederived from the supergravity [84–86]. However, we believe that the beautiful structure that we find will be a guide towards a clearer understanding of the $N^{\frac{3}{2}}$ law or even a better description of the mysterious M-theory.

Acknowledgements

We are grateful to all of our collaborators at various stages of this project for valuable discussions, including Hiroyuki Fuji, Alba Grassi, Shinji Hirano, Masazumi Honda, Marcos Marino, Tomoki Nosaka, and Masaki Shigemori. The work of S.M. was supported in part by JSPS Grant-in-Aid for Scientific Research (C) # 26400245. The work of K.O. was supported in part by JSPS Grant-in-Aid for Young Scientists (B) # 23740178.

Funding

Open Access funding: SCOAP³.

References

- [1] E. Witten, Nucl. Phys. B **268**, 253 (1986).
- [2] E. Witten, Prog. Math. **133**, 637 (1995) [[arXiv:hep-th/9207094](#)] [[Search INSPIRE](#)].
- [3] B. M. Zupnik and D. V. Khetselius, Sov. J. Nucl. Phys. **47**, 730 (1988) [*Yad. Fiz.* **47**, 1147 (1988)] [[Search INSPIRE](#)].
- [4] H.-C. Kao and K.-M. Lee, Phys. Rev. D **46**, 4691 (1992) [[arXiv:hep-th/9205115](#)] [[Search INSPIRE](#)].
- [5] H.-C. Kao, K.-M. Lee, and T. Lee, Phys. Lett. B **373**, 94 (1996) [[arXiv:hep-th/9506170](#)] [[Search INSPIRE](#)].
- [6] D. Gaiotto and X. Yin, J. High Energy Phys. **0708**, 056 (2007) [[arXiv:0704.3740](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [7] D. L. Jafferis and A. Tomasiello, J. High Energy Phys. **0810**, 101 (2008) [[arXiv:0808.0864](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [8] D. Gaiotto and E. Witten, J. High Energy Phys. **1006**, 097 (2010) [[arXiv:0804.2907](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [9] K. Hosomichi, K.-M. Lee, S. Lee, S. Lee, and J. Park, J. High Energy Phys. **0807**, 091 (2008) [[arXiv:0805.3662](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [10] O. Aharony, O. Bergman, D. L. Jafferis, and J. Maldacena, J. High Energy Phys. **0810**, 091 (2008) [[arXiv:0806.1218](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [11] K. Hosomichi, K.-M. Lee, S. Lee, S. Lee, and J. Park, J. High Energy Phys. **0809**, 002 (2008) [[arXiv:0806.4977](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [12] O. Aharony, O. Bergman, and D. L. Jafferis, J. High Energy Phys. **0811**, 043 (2008) [[arXiv:0807.4924](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [13] E. Witten, Nucl. Phys. B **443**, 85 (1995) [[arXiv:hep-th/9503124](#)] [[Search INSPIRE](#)].
- [14] I. R. Klebanov and A. A. Tseytlin, Nucl. Phys. B **475**, 164 (1996) [[arXiv:hep-th/9604089](#)] [[Search INSPIRE](#)].
- [15] M. Benna, I. Klebanov, T. Klose, and M. Smedback, J. High Energy Phys. **0809**, 072 (2008) [[arXiv:0806.1519](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [16] E. Witten, Nucl. Phys. B **460**, 335 (1996) [[arXiv:hep-th/9510135](#)] [[Search INSPIRE](#)].
- [17] N. Drukker and D. Trancanelli, J. High Energy Phys. **1002**, 058 (2010) [[arXiv:0912.3006](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [18] V. Pestun, Commun. Math. Phys. **313**, 71 (2012) [[arXiv:0712.2824](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [19] K. Hosomichi, [[arXiv:1502.04543](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [20] A. Kapustin, B. Willett, and I. Yaakov, J. High Energy Phys. **1003**, 089 (2010) [[arXiv:0909.4559](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [21] Y. Hatsuda, S. Moriyama, and K. Okuyama, J. High Energy Phys. **1210**, 020 (2012) [[arXiv:1207.4283](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [22] Y. Hatsuda, S. Moriyama, and K. Okuyama, J. High Energy Phys. **1301**, 158 (2013) [[arXiv:1211.1251](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [23] Y. Hatsuda, S. Moriyama, and K. Okuyama, J. High Energy Phys. **1305**, 054 (2013) [[arXiv:1301.5184](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [24] Y. Hatsuda, M. Marino, S. Moriyama, and K. Okuyama, J. High Energy Phys. **1409**, 168 (2014) [[arXiv:1306.1734](#)] [[hep-th](#)] [[Search INSPIRE](#)].
- [25] N. Drukker, M. Marino, and P. Putrov, Commun. Math. Phys. **306**, 511 (2011) [[arXiv:1007.3837](#)] [[hep-th](#)] [[Search INSPIRE](#)].

- [26] J. Bagger and N. Lambert, Phys. Rev. D **77**, 065008 (2008) [arXiv:0711.0955 [hep-th]] [Search INSPIRE].
- [27] A. Gustavsson, Nucl. Phys. B **811**, 66 (2009) [arXiv:0709.1260 [hep-th]] [Search INSPIRE].
- [28] C. P. Herzog, I. R. Klebanov, S. S. Pufu, and T. Tesileanu, Phys. Rev. D **83**, 046001 (2011) [arXiv:1011.5487 [hep-th]] [Search INSPIRE].
- [29] N. Drukker, M. Marino, and P. Putrov, J. High Energy Phys. **1111**, 141 (2011) [arXiv:1103.4844 [hep-th]] [Search INSPIRE].
- [30] H. Fuji, S. Hirano, and S. Moriyama, J. High Energy Phys. **1108**, 001 (2011) [arXiv:1106.4631 [hep-th]] [Search INSPIRE].
- [31] K. Okuyama, Prog. Theor. Phys. **127**, 229 (2012) [arXiv:1110.3555 [hep-th]] [Search INSPIRE].
- [32] M. Marino and P. Putrov, J. Stat. Mech. **1203**, P03001 (2012) [arXiv:1110.4066 [hep-th]] [Search INSPIRE].
- [33] M. Hanada, M. Honda, Y. Honma, J. Nishimura, S. Shiba, and Y. Yoshida, J. High Energy Phys. **1205**, 121 (2012) [arXiv:1202.5300 [hep-th]] [Search INSPIRE].
- [34] P. Putrov and M. Yamazaki, Mod. Phys. Lett. A **27**, 1250200 (2012) [arXiv:1207.5066 [hep-th]] [Search INSPIRE].
- [35] F. Calvo and M. Marino, J. High Energy Phys. **1305**, 006 (2013) [arXiv:1212.5118 [hep-th]] [Search INSPIRE].
- [36] H. Awata, S. Hirano, and M. Shigemori, Prog. Theor. Exp. Phys. **2013**, 053B04 (2013) [arXiv:1212.2966] [Search INSPIRE].
- [37] M. Honda, J. High Energy Phys. **1312**, 046 (2013) [arXiv:1310.3126 [hep-th]] [Search INSPIRE].
- [38] S. Matsumoto and S. Moriyama, J. High Energy Phys. **1403**, 079 (2014) [arXiv:1310.8051 [hep-th]] [Search INSPIRE].
- [39] M. Honda and K. Okuyama, J. High Energy Phys. **1408**, 148 (2014) [arXiv:1405.3653 [hep-th]] [Search INSPIRE].
- [40] M. Honda and S. Moriyama, [arXiv:1404.0676 [hep-th]] [Search INSPIRE].
- [41] Y. Hatsuda and K. Okuyama, J. High Energy Phys. **1410**, 158 (2014) [arXiv:1407.3786 [hep-th]] [Search INSPIRE].
- [42] M. Marino and P. Putrov, J. High Energy Phys. **1006**, 011 (2010) [arXiv:0912.3074 [hep-th]] [Search INSPIRE].
- [43] A. Cagnazzo, D. Sorokin, and L. Wulff, J. High Energy Phys. **1005**, 009 (2010) [arXiv:0911.5228 [hep-th]] [Search INSPIRE].
- [44] D. L. Jafferis, J. High Energy Phys. **1205**, 159 (2012) [arXiv:1012.3210 [hep-th]] [Search INSPIRE].
- [45] N. Hama, K. Hosomichi, and S. Lee, J. High Energy Phys. **1103**, 127 (2011) [arXiv:1012.3512 [hep-th]] [Search INSPIRE].
- [46] M. Bershadsky, S. Cecotti, H. Ooguri, and C. Vafa, Commun. Math. Phys. **165**, 311 (1994) [arXiv:hep-th/9309140] [Search INSPIRE].
- [47] T. Azeyanagi, M. Fujita, and M. Hanada, Phys. Rev. Lett. **110**, 121601 (2013) [arXiv:1210.3601 [hep-th]] [Search INSPIRE].
- [48] M. Alim, S. T. Yau, and J. Zhou, [arXiv:1506.01375 [hep-th]] [Search INSPIRE].
- [49] K. Becker, M. Becker, and A. Strominger, Nucl. Phys. B **456**, 130 (1995) [arXiv:hep-th/9507158] [Search INSPIRE].
- [50] R. P. Feynman, Statistical Mechanics: A Set of Lectures (Frontiers in Physics) (Perseus, New York, 1972).
- [51] E. Witten, Commun. Math. Phys. **121**, 351 (1989).
- [52] Y. Hatsuda, M. Honda, S. Moriyama, and K. Okuyama, J. High Energy Phys. **1310**, 168 (2013) [arXiv:1306.4297 [hep-th]] [Search INSPIRE].
- [53] A. Sen and B. Zwiebach, J. High Energy Phys. **0003**, 002 (2000) [arXiv:hep-th/9912249] [Search INSPIRE].
- [54] E. M. Moens and Joris Van der Jeugt, J. Algebraic Combinatorics **17**, 283 (2003).
- [55] A. Grassi and M. Marino, [arXiv:1403.4276 [hep-th]] [Search INSPIRE].
- [56] M. B. Green, J. H. Schwarz, and L. Brink, Nucl. Phys. B **219**, 437 (1983).
- [57] C. A. Tracy and H. Widom, Commun. Math. Phys. **179**, 667 (1996) [arXiv:solv-int/9509003] [Search INSPIRE].
- [58] A. Mironov and A. Morozov, J. High Energy Phys. **1004**, 040 (2010) [arXiv:0910.5670 [hep-th]] [Search INSPIRE].

- [59] A. Mironov and A. Morozov, *J. Phys. A* **43**, 195401 (2010) [[arXiv:0911.2396](#) [hep-th]] [[Search INSPIRE](#)].
- [60] M. Aganagic, M. C. N. Cheng, R. Dijkgraaf, D. Krefl, and C. Vafa, *J. High Energy Phys.* **1211**, 019 (2012) [[arXiv:1105.0630](#) [hep-th]] [[Search INSPIRE](#)].
- [61] J. Kallen and M. Marino, [[arXiv:1308.6485](#) [hep-th]] [[Search INSPIRE](#)].
- [62] J. Kallen, [[arXiv:1407.0625](#) [hep-th]] [[Search INSPIRE](#)].
- [63] M.-x. Huang and X.-f. Wang, [[arXiv:1406.6178](#) [hep-th]] [[Search INSPIRE](#)].
- [64] A. Grassi, Y. Hatsuda, and M. Marino, [[arXiv:1410.3382](#) [hep-th]] [[Search INSPIRE](#)].
- [65] X. Wang, G. Zhang, and M. X. Huang, [[arXiv:1505.05360](#) [hep-th]] [[Search INSPIRE](#)].
- [66] S. Codesido, A. Grassi, and M. Marino, [[arXiv:1409.1799](#) [hep-th]] [[Search INSPIRE](#)].
- [67] A. Grassi, Y. Hatsuda, and M. Marino, [[arXiv:1410.7658](#) [hep-th]] [[Search INSPIRE](#)].
- [68] V. V. Bazhanov, S. L. Lukyanov, and A. B. Zamolodchikov, *Commun. Math. Phys.* **190**, 247 (1997) [[arXiv:hep-th/9604044](#)] [[Search INSPIRE](#)].
- [69] A. Grassi, M. Marino, and S. Zakany, *J. High Energy Phys.* **1505**, 038 (2015) [[arXiv:1405.4214](#) [hep-th]] [[Search INSPIRE](#)].
- [70] Y. Hatsuda and K. Okuyama, [[arXiv:1505.07460](#) [hep-th]] [[Search INSPIRE](#)].
- [71] D. R. Gulotta, C. P. Herzog, and S. S. Pufu, *J. High Energy Phys.* **1112**, 077 (2011) [[arXiv:1105.2817](#) [hep-th]] [[Search INSPIRE](#)].
- [72] S. Moriyama and T. Nosaka, *J. High Energy Phys.* **1411**, 164 (2014) [[arXiv:1407.4268](#) [hep-th]] [[Search INSPIRE](#)].
- [73] Y. Imamura and K. Kimura, *J. High Energy Phys.* **0810**, 040 (2008) [[arXiv:0807.2144](#) [hep-th]] [[Search INSPIRE](#)].
- [74] M. Mezei and S. S. Pufu, *J. High Energy Phys.* **1402**, 037 (2014) [[arXiv:1312.0920](#) [hep-th]] [[Search INSPIRE](#)].
- [75] B. Assel, N. Drukker, and J. Felix, [[arXiv:1504.07636](#) [hep-th]] [[Search INSPIRE](#)].
- [76] S. Moriyama and T. Nosaka, [[arXiv:1410.4918](#) [hep-th]] [[Search INSPIRE](#)].
- [77] S. Moriyama and T. Nosaka, *J. High Energy Phys.* **1505**, 022 (2015) [[arXiv:1412.6243](#) [hep-th]] [[Search INSPIRE](#)].
- [78] Y. Hatsuda, M. Honda, and K. Okuyama, [[arXiv:1505.07120](#) [hep-th]] [[Search INSPIRE](#)].
- [79] D. R. Gulotta, J. P. Ang, and C. P. Herzog, *J. High Energy Phys.* **1201**, 132 (2012) [[arXiv:1111.1744](#) [hep-th]] [[Search INSPIRE](#)].
- [80] S. Moriyama and T. Nosaka, [[arXiv:1504.07710](#) [hep-th]] [[Search INSPIRE](#)].
- [81] N. Drukker and J. Felix, [[arXiv:1501.02268](#) [hep-th]] [[Search INSPIRE](#)].
- [82] P. M. Crichigno, C. P. Herzog, and D. Jain, *J. High Energy Phys.* **1303**, 039 (2013) [[arXiv:1211.1388](#) [hep-th]] [[Search INSPIRE](#)].
- [83] D. R. Gulotta, C. P. Herzog, and T. Nishioka, *J. High Energy Phys.* **1204**, 138 (2012) [[arXiv:1201.6360](#) [hep-th]] [[Search INSPIRE](#)].
- [84] O. Aharony, A. Hashimoto, S. Hirano, and P. Ouyang, *J. High Energy Phys.* **1001**, 072 (2010) [[arXiv:0906.2390](#) [hep-th]] [[Search INSPIRE](#)].
- [85] S. Bhattacharyya, A. Grassi, M. Marino, and A. Sen, *Classical Quantum Gravity* **31**, 015012 (2014) [[arXiv:1210.6057](#) [hep-th]] [[Search INSPIRE](#)].
- [86] A. Dabholkar, N. Drukker, and J. Gomes, [[arXiv:1406.0505](#) [hep-th]] [[Search INSPIRE](#)].