Modular invariance in superstring theory from $\mathcal{N} = 4$ super-Yang-Mills

Shai M. Chester,[a] **Michael B. Green,**[b,c] **Silviu S. Pufu,**[d] **Yifan Wang**[d,e] **and Congkao Wen**[b]

[a] *Department of Particle Physics and Astrophysics, Weizmann Institute of Science,*
Rehovot, Israel

[b] *School of Physics and Astronomy, Queen Mary University of London,*
London, E1 4NS, U.K.

[c] *DAMTP,*
Wilberforce Road, Cambridge CB3 0WA, U.K.

[d] *Joseph Henry Laboratories, Princeton University,*
Princeton, NJ 08544, U.S.A.

[e] *Center of Mathematical Sciences and Applications and Jefferson Physical Laboratory,*
Harvard University,
Cambridge, MA 02138, U.S.A.

E-mail: shai.chester@weizmann.ac.il, m.b.green@damtp.cam.ac.uk,
spufu@princeton.edu, yifanw@g.harvard.edu, c.wen@qmul.ac.uk

ABSTRACT: We study the four-point function of the lowest-lying half-BPS operators in the $\mathcal{N} = 4$ SU(N) super-Yang-Mills theory and its relation to the flat-space four-graviton amplitude in type IIB superstring theory. We work in a large-N expansion in which the complexified Yang-Mills coupling τ is fixed. In this expansion, non-perturbative instanton contributions are present, and the SL(2, $\mathbb{Z}$) duality invariance of correlation functions is manifest. Our results are based on a detailed analysis of the sphere partition function of the mass-deformed SYM theory, which was previously computed using supersymmetric localization. This partition function determines a certain integrated correlator in the undeformed $\mathcal{N} = 4$ SYM theory, which in turn constrains the four-point correlator at separated points. In a normalization where the two-point functions are proportional to $N^2 - 1$ and are independent of τ and $\bar{\tau}$, we find that the terms of order $\sqrt{N}$ and $1/\sqrt{N}$ in the large N expansion of the four-point correlator are proportional to the non-holomorphic Eisenstein series $E(\frac{3}{2}, \tau, \bar{\tau})$ and $E(\frac{5}{2}, \tau, \bar{\tau})$, respectively. In the flat space limit, these terms match the corresponding terms in the type IIB S-matrix arising from R^4 and $D^4 R^4$ contact interactions, which, for the R^4 case, represents a check of AdS/CFT at finite string coupling. Furthermore, we present striking evidence that these results generalize so that, at order

$N^{\frac{1}{2}-m}$ with integer $m \geq 0$, the expansion of the integrated correlator we study is a linear sum of non-holomorphic Eisenstein series with half-integer index, which are manifestly $\mathrm{SL}(2, \mathbb{Z})$ invariant.

Contents

1 Introduction

1.1 Overview

It has long been appreciated that string theory can be given a non-perturbative definition through the anti-de Sitter / conformal field theory (AdS/CFT) correspondence [1–3]. In principle, such a definition allows one to calculate properties of closed string theory or M-theory, and hence of quantum gravity, in terms of properties of certain CFTs such as the $SU(N)$ $\mathcal{N} = 4$ supersymmetric Yang-Mills (SYM) theory, without any explicit reference to standard stringy methods. For instance, as originally proposed in [4–6] and further explored in [7–9], one can obtain flat space scattering amplitudes of massless particles by taking certain kinematic limits of CFT correlation functions. While in general it is very difficult to determine the CFT correlation functions in the regime of interest, recent

progress combining various analytic bootstrap techniques [10–28] and supersymmetry has allowed the determination of CFT correlators, at strong coupling, in a variety of examples arising from top-down AdS/CFT constructions [17, 29–34]. From the CFT correlators in these examples, one can then extract the first few terms in the low-energy expansion of scattering amplitudes in superstring theory or M-theory. In the cases studied so far, these scattering amplitudes had been known from other string theory arguments, and therefore these calculations represent remarkable precision tests of AdS/CFT.

The main focus of this work is on the $\mathcal{N} = 4$ SYM theory, and, in particular, on how one can reproduce the $\mathrm{SL}(2, \mathbb{Z})$ modular properties of superstring amplitudes from CFT correlators. Before discussing these modular properties, however, let us briefly review the recent work on the four-point correlators of $\mathcal{N} = 4$ SYM in the 't Hooft limit, and then discuss how these computations should be modified in order to access $\mathrm{SL}(2, \mathbb{Z})$ invariants. As we will explain, in order to uncover the modular properties, we should consider a strong coupling limit that is different from the usual 't Hooft limit.

CFT correlation functions in 't Hooft limit and type IIB graviton amplitudes

Refs. [17, 33, 34] studied the four-point function of superconformal primary operators in the same multiplet as the stress-energy tensor of the $\mathcal{N} = 4$ SYM theory, considered in the large-N 't Hooft limit, where the 't Hooft coupling $\lambda \equiv g_{\mathrm{YM}}^2 N$ is fixed and in the regime $\lambda \gg 1$. In a normalization in which the disconnected term in the four-point function scales as $N^4 \lambda^0$ at large N and λ, the leading connected contribution is of order $N^2 \lambda^0$ and reproduces the tree amplitude of type IIB supergravity. The next correction, of order $N^2 \lambda^{-\frac{3}{2}}$, reproduces the "stringy" correction corresponding to the eight-derivative contact interaction of the form R^4 (a well-known contraction of four Riemann tensors) in the effective IIB superstring action [35, 36]. Several other terms of higher order in $1/\lambda$ and in $1/N$ have been matched with terms that arise in the low energy expansion of string perturbation theory [33, 34].

The procedure that allows the above comparisons between string perturbation theory and $\mathcal{N} = 4$ correlation functions is as follows. For the four-point function mentioned above, the analytic bootstrap consistency conditions [11, 33, 34] determine the position dependence of the tree-level Witten diagrams corresponding to higher derivative terms in the string effective action, up to a number of undetermined coefficients. At low orders in the derivative expansion, these coefficients can be fully determined, in principle, from constraints derived using supersymmetric localization [29–34]. Indeed, as shown in [33] in the case of the $\mathcal{N} = 4$ SYM theory, one can obtain integrated constraints on the 4-point functions of the operators in the stress-tensor multiplet by taking four derivatives of the partition function Z of the mass-deformed $\mathcal{N} = 4$ SYM theory placed on a round S^4. This mass deformation of the $\mathcal{N} = 4$ SYM theory is referred to as the $\mathcal{N} = 2^*$ theory, and its S^4 partition function was computed by Pestun using supersymmetric localization [37]. As shown in [37], this partition function takes the form of a finite-dimensional integral over constant values of one of the vector multiplet scalars. The integrand is a product of a classical contribution, a one-loop contribution, and contributions from instantons located at the north and south poles of S^4 [38–41].

The procedure implemented in [33] for analyzing the $\mathcal{N} = 4$ integrated correlation function made use of the constraint coming from the $m = 0$ limit of the mixed derivative (where m is the mass deformation parameter)

$$\tau_2^2 \partial_\tau \partial_{\bar{\tau}} \partial_m^2 \log Z \Big|_{m=0}, \tag{1.1}$$

where $\tau = \tau_1 + i\tau_2 \equiv \frac{\theta}{2\pi} + \frac{4\pi i}{g_{\mathrm{YM}}^2}$ is the complexified gauge coupling. As mentioned above, this leads to an explicit calculation of various terms in the double expansion in $1/N$ and $1/\lambda$ of the CFT correlation function [33, 34], and these terms reproduced the analogous terms in the low energy expansion of the type IIB superstring tree amplitude that are perturbative in g_s.

$SL(2, \mathbb{Z})$ duality of CFT correlation functions and type IIB graviton amplitudes

These connections between $\mathcal{N} = 4$ SYM and type IIB superstring perturbation theory are part of a much richer non-perturbative story that incorporates the constraints of S-duality. In particular, understanding how the $\mathrm{SL}(2, \mathbb{Z})$ S-duality of the type IIB superstring theory [42] arises as the image of Montonen-Olive $\mathrm{SL}(2, \mathbb{Z})$ duality of $\mathcal{N} = 4$ super-Yang-Mills theory [43–45] involves understanding the holographic connection between Yang-Mills instantons and D-instantons [46–49], which will be a crucial feature of this paper.

Before we go into more details about the $\mathrm{SL}(2, \mathbb{Z})$ invariance properties of the IIB amplitudes and corresponding SYM correlators, we note that neither $\mathcal{N} = 4$ SYM nor the IIB theory on $AdS_5 \times S^5$ are entirely $\mathrm{SL}(2, \mathbb{Z})$ invariant. As explained in [50], specifying the SYM theory requires knowing the global form of the gauge group, as well as a discrete theta angle, and general $\mathrm{SL}(2, \mathbb{Z})$ transformations change both the global structure of the gauge group and the discrete theta angle.[1] However, local correlators in $\mathcal{N} = 4$ SYM theory (and correspondingly the amplitudes in IIB string theory) are insensitive to such subtleties, and it is in this sense that we are exploring their $\mathrm{SL}(2, \mathbb{Z})$ invariance properties in this paper.

$SL(2, \mathbb{Z})$ *and the IIB superstring low energy expansion.* The exact coefficients of higher-derivative interactions in the low energy expansion of the four-graviton amplitude in IIB superstring theory are $\mathrm{SL}(2, \mathbb{Z})$-invariant functions of the complex string coupling $\tau_s = \chi_s + i/g_s$ that have been explicitly determined up to order $D^6 R^4$. These interactions preserve a fraction of the 32 supersymmetries (they are so-called F-terms), and their form is severely constrained by supersymmetry combined with S-duality. For example, the coefficient of the 1/2-BPS R^4 interaction is proportional to a non-holomorphic Eisenstein series $E(\frac{3}{2}, \tau, \bar{\tau})$ [53–55], which will be defined in appendix A. When expanded at

---

[1]For the SU(N) cases, these theories are labeled as $(\mathrm{SU}(N)/\mathbb{Z}_k)_n$ where k is a positive divisor of N and n a $\mathbb{Z}_k$-valued theta angle [50]. Under the S generator of $\mathrm{SL}(2, \mathbb{Z})$, the SU($N$) theory (no discrete theta angle for this case) is mapped to $(\mathrm{SU}(N)/\mathbb{Z}_n)_0$ and vice versa. On the bulk side, the type IIB string theory contains a nontrivial topological sector on AdS_5 described by a Chern-Simons-like theory involving the NS and RR 2-form fields [51]. The discrete data involved in specifying the boundary SYM theory translates into a choice of boundary conditions for the bulk topological theory, and such boundary conditions transform nontrivially under $\mathrm{SL}(2, \mathbb{Z})$ [51, 52]. These topological subtleties are important for understanding the $\mathrm{SL}(2, \mathbb{Z})$ properties of extended objects (such as line operators) or when considering topologically nontrivial backgrounds (such as nontrivial $H^2(M_4, \mathbb{Z}_N)$ on the 4d boundary manifold M_4).

small string coupling, this Eisenstein series has two terms that are power behaved in g_s, corresponding to genus-0 and genus-1 contributions in string perturbation theory. In addition, there is an infinite sequence of exponentially suppressed non-perturbative terms due to D-instanton effects, where the contribution of a charge-k D-instanton is proportional to $e^{-2\pi k/g_s}$. Similar comments apply to the coefficient of the 1/4-BPS D^4R^4 interaction, which has a coefficient proportional to $E(\frac{5}{2}, \tau, \bar{\tau})$ [56]. Whereas the Eisenstein series satisfy Laplace eigenvalue equations, the coefficient of the 1/8-BPS interaction, D^6R^4, is a novel modular function that satisfies an inhomogeneous Laplace eigenvalue equation [57] that is also reviewed in appendix A.[2]

$\mathrm{SL}(2, \mathbb{Z})$ *and correlation functions in* $\mathcal{N} = 4$ *SYM.* In the usual 't Hooft limit, the 't Hooft coupling λ is kept fixed as $N \to \infty$, which requires g_{YM} to be small. However, this is incompatible with $\mathrm{SL}(2, \mathbb{Z})$ duality, which has an action on the complex coupling $\tau \equiv \frac{\theta}{2\pi} + \frac{4\pi i}{g_{\mathrm{YM}}^2}$ given by

$$\tau \to \tau' = \frac{a\tau + b}{c\tau + d}, \tag{1.2}$$

where $a, b, c, d \in \mathbb{Z}$ and $ad - bc = 1$. In particular, this transformation mixes weak coupling and strong coupling effects. Therefore, in order to consider the action of $\mathrm{SL}(2, \mathbb{Z})$ on correlation functions in the large-N limit, it is necessary to consider the limit in which g_{YM} is fixed as $N \to \infty$. Such a limit had been considered in [33], where it was referred to as the "very strong coupling limit," and it was also considered in an analogous context in [61].[3] In particular, instanton effects of order $e^{-8\pi^2 k/g_{\mathrm{YM}}^2} = e^{-8\pi^2 kN/\lambda}$, where k (the instanton number) is a positive integer, survive this limit, whereas they are exponentially suppressed in the usual 't Hooft limit. In the very strong coupling limit, it is the terms of order $N^{1/2}$, $N^{-1/2}$, and N^{-1} that correspond to the R^4, D^4R^4, and D^6R^4 mentioned above.

Using a similar strategy to that outlined above in the 't Hooft limit, we find that the analytic bootstrap constraints combined with the supersymmetric localization constraints coming from (1.1) yield, in the very strong coupling limit, the same Eisenstein series that appear in the low energy expansion of the type IIB four-graviton amplitude. Indeed, the constraint from (1.1) is sufficient to determine the coefficient of $N^{1/2}$ in the large-N expansion of the CFT four-point function, and, as we will show, this coefficient ends up being proportional to $E(\frac{3}{2}, \tau, \bar{\tau})$, with precisely the right proportionality factor to match the corresponding term in the superstring amplitude.[4] This is a full non-perturbative precision test of AdS/CFT!

---

[2]These results were rederived in [58] using on-shell superamplitude methods [59]. Furthermore, more general F-terms involving higher-point interactions have also been determined [60].

[3]Non-'t Hooft limits have been considered before in other gauge theories. For instance, in the 3d $\mathrm{U}(N)_k \times \mathrm{U}(N)_{-k}$ ABJM theory [62], the limit in which k is fixed while $N \to \infty$ corresponds to M-theory on $AdS_4 \times S^7/\mathbb{Z}_k$. In the same theory, a different non-'t Hooft limit, namely that in which $N \to \infty$ with finite $\mu \equiv N/k^5$ considered in [32] is somewhat similar to the very strong coupling limit considered here.

[4]It is worth noting that the fact that the four-point functions of operators in the stress tensor multiplet of $\mathcal{N} = 4$ SYM, and in particular the quantity (1.1), are $\mathrm{SL}(2, \mathbb{Z})$ modular invariants is no surprise. Indeed, the operators belonging to the stress tensor multiplet transform with well-defined holomorphic and anti-holomorphic modular weights $(w, -w)$ under the action of $\mathrm{SL}(2, \mathbb{Z})$ [61, 63, 64]. The bottom component of the multiplet we consider has the weight $w = 0$, therefore the corresponding correlators are $\mathrm{SL}(2, \mathbb{Z})$ invariant.

Note that those contributions that are perturbative in $1/\lambda$ in the 't Hooft limit are also power-behaved in g_{YM} in the limit in which $N \to \infty$ and g_{YM} is fixed. Thus, the perturbative terms evaluated in earlier work reproduce the two terms that are power-behaved in g_{YM} in $E(\frac{3}{2}, \tau, \bar{\tau})$. Our main challenge is to show that the exponentially suppressed terms in the Eisenstein series can also be reproduced from the $\mathcal{N} = 4$ SYM theory. Using (1.1), we find that these exponentially suppressed terms come from considering the $m = 0$ limit of the instanton contributions to the $\mathcal{N} = 2^*$ partition function.[5]

At order $N^{-1/2}$ in the large-N expansion, the constraint (1.1) is no longer enough to fully determine the CFT correlation function, but we do find that the integrated correlator (1.1) is proportional to the $E(\frac{5}{2}, \tau, \bar{\tau})$ modular invariant that appears in the flat space limit of the IIB amplitude. Combining the integrated constraint with the flat space string theory answer, we obtain the complete CFT correlator up to order $N^{-1/2}$. While the $E(\frac{3}{2}, \tau, \bar{\tau})$ modular invariant appearing at order $N^{1/2}$ multiplies a single term of schematic form R^4 in the effective field theory in AdS, the $E(\frac{5}{2}, \tau, \bar{\tau})$ invariant multiplies a linear combination of the $D^4 R^4$ interaction and R^4/L^4, where L is the radius of AdS.

At higher orders in the $1/N$ expansion we do not have sufficient information to determine the four-point correlator in $\mathcal{N} = 4$ SYM. We will nevertheless argue, more conjecturally, that any given order in the large-N expansion of the integrated correlation function (1.1) is a finite linear sum of non-holomorphic Eisenstein series with rational coefficients.

1.2 Outline

The rest of this paper is organized as follows. In section 2 we give a brief review of the 4-point function of the stress tensor superconformal primary operator in $\mathcal{N} = 4$ SYM, its relation to the string theory scattering amplitude, and the supersymmetric localization constraint coming from the mixed derivative (1.1). Section 3 describes the main technical achievement of this paper, which is the evaluation of the instanton contributions to the integrated correlation function. These contributions are associated with factors of the Nekrasov partition function [40, 41] that enter into the localization result for the mass-deformed S^4 partition function and contribute to the $\mathcal{N} = 4$ integrated correlator described by (1.1). We will determine the k-instanton contributions to this quantity at orders $N^{\frac{1}{2}}$ and $N^{-\frac{1}{2}}$, and show that they match, respectively, the kth Fourier modes with respect to θ of $E(\frac{3}{2}, \tau, \bar{\tau})$ and $E(\frac{5}{2}, \tau, \bar{\tau})$. The analysis of terms that are higher order in $1/N$ in (1.1) is developed further in section 4. We end with a discussion of our results in section 5. Several technical details are relegated to the appendices.

2 Four-point function in $\mathcal{N} = 4$ SYM

Let us start with a brief review of the setup of the four-point function of the stress tensor superconformal primary operator in SYM, its relation to the 10d IIB flat space graviton S-matrix and constraints from supersymmetric localization. For more details, see ref. [33].

---

[5]While the strict $m = 0$ limit of the $\mathcal{N} = 2^*$ instanton partition function is trivial, the subleading m^2 order is not and contributes to (1.1).

This operator transforms in the **20′** of the $SO(6)_R$ R-symmetry, and it can be represented as a traceless symmetric tensor $S_{IJ}(\vec{x})$ with $I, J = 1, \ldots, 6$ as $SO(6)_R$ fundamental indices. In order to avoid a proliferation of indices, it is customary to contract them with null polarization vectors Y^I satisfying $Y \cdot Y \equiv \sum_{I=1}^{6} Y^I Y^I = 0$. Superconformal symmetry implies that the four-point function of the operator $S(\vec{x}, Y) \equiv S_{IJ}(\vec{x})Y^I Y^J$ takes the form [65, 66]

$$\langle S(\vec{x}_1, Y_1) \cdots S(\vec{x}_4, Y_4) \rangle = \frac{1}{\vec{x}_{12}^4 \vec{x}_{34}^4} \left[\vec{\mathcal{S}}_{\text{free}} + \mathcal{T}(U, V)\vec{\Theta} \right] \cdot \vec{\mathcal{B}}, \qquad (2.1)$$

where $\vec{x}_{ij} \equiv \vec{x}_i - \vec{x}_j$, and

$$\begin{aligned}
\vec{\mathcal{S}}_{\text{free}} &\equiv \left(1 \ U^2 \ \frac{U^2}{V^2} \ \frac{1}{c}\frac{U^2}{V} \ \frac{1}{c}\frac{U}{V} \ \frac{1}{c}U \right), \\
\vec{\Theta} &\equiv \left(V \ UV \ U \ U(U - V - 1) \ 1 - U - V \ V(V - U - 1) \right), \\
\mathcal{B} &\equiv \left(Y_{12}^2 Y_{34}^2 \ Y_{13}^2 Y_{24}^2 \ Y_{14}^2 Y_{23}^2 \ Y_{13}Y_{14}Y_{23}Y_{24} \ Y_{12}Y_{14}Y_{23}Y_{34} \ Y_{12}Y_{13}Y_{24}Y_{34} \right).
\end{aligned} \qquad (2.2)$$

Here, c is the conformal anomaly coefficient, which for an $SU(N)$ gauge group equals $c = (N^2 - 1)/4$; the quantities $U \equiv \frac{\vec{x}_{12}^2 \vec{x}_{34}^2}{\vec{x}_{13}^2 \vec{x}_{24}^2}$ and $V \equiv \frac{\vec{x}_{14}^2 \vec{x}_{23}^2}{\vec{x}_{13}^2 \vec{x}_{24}^2}$ are the usual conformal invariant cross-ratios; and $Y_{ij} \equiv Y_i \cdot Y_j$ are $SO(6)_R$ invariants. Importantly, the only non-trivial information in the correlator (2.1) is encoded in a single function of the conformal cross-ratios, $\mathcal{T}(U, V)$.

More generally, to describe the holographic correlators, it is simpler to work in Mellin space [67, 68]. Let us thus define the Mellin transform $\mathcal{M}$ of $\mathcal{T}$ via

$$\mathcal{T}(U, V) = \int_{-i\infty}^{i\infty} \frac{ds\, dt}{(4\pi i)^2} U^{\frac{s}{2}} V^{\frac{u}{2}-2} \Gamma\left[2 - \frac{s}{2}\right]^2 \Gamma\left[2 - \frac{t}{2}\right]^2 \Gamma\left[2 - \frac{u}{2}\right]^2 \mathcal{M}(s, t), \qquad (2.3)$$

where $u \equiv 4 - s - t$. Crossing symmetry $\mathcal{M}(s, t) = \mathcal{M}(t, s) = \mathcal{M}(s, u)$, as well as the analytic properties of the Mellin amplitude (for a detailed description see [33]), restrict $\mathcal{M}(s, t)$ to have the following $1/c$ expansion at fixed Yang-Mills coupling:

$$\mathcal{M}(s, t) = \frac{\alpha}{(s-2)(t-2)(u-2)} \frac{1}{c} + \frac{\beta}{c^{7/4}} + \frac{\mathcal{M}_{\text{1-loop}}(s, t)}{c^2} + \frac{\gamma_1(s^2 + t^2 + u^2) + \gamma_2}{c^{\frac{9}{4}}} + \cdots, \quad (2.4)$$

where the coefficients α, β, γ_i, etc. are potentially non-trivial functions of $(\tau, \bar{\tau})$. Here, $\mathcal{M}_{\text{1-loop}}$ is the regularized supergravity one-loop amplitude that can be found in [34] and will not be discussed here; in this work, we will instead focus mostly on the $1/c^{7/4}$ and $1/c^{9/4}$ terms, corresponding, respectively, to the R^4 and $D^4 R^4$ interaction vertices. As explained in [33], the constant α can be found as follows. The free theory contribution in (2.1), when expanded in conformal blocks, contains twist two operators of all spins. In the interacting $SU(N)$ gauge theory, however, one expects no operators of twist precisely two, except for those operators belonging to the stress tensor multiplet. Thus, the conformal block decomposition of the second term in (2.1) must cancel, in part, that of the first term. A careful analysis shows that this requirement implies [33, 69]

$$\alpha = 8. \qquad (2.5)$$

Note that this argument relies crucially on the gauge group being SU(N), and not U(N). A U(N) gauge theory contains a free U(1) sector, and the $S \times S$ OPE contains many operators of twist two beyond those in the stress tensor multiplet.

At each order in $1/c$, one can impose constraints on the coefficients β, γ_i, etc. by either comparing with the (super)graviton four-point scattering amplitude in type IIB string theory in the flat space limit or using the quantity (1.1) (or other similar quantities) derived from supersymmetric localization. Let us first discuss the constraints from the flat space scattering amplitude, and then those from supersymmetric localization.

2.1 Constraints from the flat space limit

The IIB four-point scattering amplitude of 10d gravitons and superpartners are restricted by supersymmetry to be proportional to a single function $f(\mathbf{s}, \mathbf{t})$

$$\mathcal{A}(\mathbf{s}, \mathbf{t}) = \mathcal{A}_{\text{SG tree}}(\mathbf{s}, \mathbf{t}) f(\mathbf{s}, \mathbf{t}), \tag{2.6}$$

where $\mathcal{A}_{\text{SG tree}}$ is the tree-level four-point supergravity amplitude,[6] $\mathbf{s}$ and $\mathbf{t}$ are the Mandelstam invariants. We will also define $\mathbf{u} \equiv -\mathbf{s} - \mathbf{t}$. In turn, this function has an expansion at small momentum (more correctly, the expansion is for small values of the dimensionless product between momentum and the string length ℓ_s) of the form

$$f(\mathbf{s}, \mathbf{t}) \equiv 1 + f_{R^4}(\mathbf{s}, \mathbf{t})\ell_s^6 + f_{\text{1-loop}}(\mathbf{s}, \mathbf{t})\ell_s^8 + f_{D^4 R^4}(\mathbf{s}, \mathbf{t})\ell_s^{10} + \cdots . \tag{2.7}$$

Here, the coefficient function that appears at each order in the expansion may be a nontrivial function of the complexified string coupling $\tau_s = \chi_s + i/g_s$. In fact, the functions f_{R^4} and $f_{D^4 R^4}$ can be written in terms of non-holomorphic Eisenstein series as [54–57]

$$f_{R^4} = \frac{\mathbf{stu}}{64} g_s^{\frac{3}{2}} E(\tfrac{3}{2}, \tau_s, \bar{\tau}_s),$$
$$f_{D^4 R^4} = \frac{\mathbf{stu}(\mathbf{s}^2 + \mathbf{t}^2 + \mathbf{u}^2)}{2^{11}} g_s^{\frac{5}{2}} E(\tfrac{5}{2}, \tau_s, \bar{\tau}_s). \tag{2.8}$$

The Eisenstein series has the following expansion at small g_s (see appendix A for details)

$$E(r, \tau_s, \bar{\tau}_s) = \frac{2\zeta(2r)}{g_s^r} + 2\sqrt{\pi} g_s^{r-1} \frac{\Gamma(r - \frac{1}{2})}{\Gamma(r)} \zeta(2r - 1)$$
$$+ \frac{2\pi^r}{\Gamma(r)\sqrt{g_s}} \sum_{k \neq 0} |k|^{r-\frac{1}{2}} \sigma_{1-2r}(|k|) K_{r-\frac{1}{2}}(2\pi g_s^{-1} |k|) e^{2\pi i k \chi_s}, \tag{2.9}$$

where the divisor sum $\sigma_p(k)$ is defined as $\sigma_p(k) = \sum_{d>0,d|k} d^p$, and $K_{r-\frac{1}{2}}$ is the Bessel function of second kind of index $r - 1/2$.

---

[6]This is given by $\frac{\delta^{16}(Q)}{\mathbf{stu}}$ in the superamplitude notation where Q denotes the 16-component supermomentum variable. See, for instance, [58, 70]. In particular, the component corresponding to the four-graviton scattering is given by $\frac{R^4}{\mathbf{stu}}$, where R here denotes the linearized Riemann curvature tensor.

The relation between the function $f(\mathbf{s}, \mathbf{t})$ in (2.6) and the Mellin amplitude (2.4) is given by the flat space limit formula [33]

$$f(\mathbf{s}, \mathbf{t}) = \frac{\mathbf{stu}}{2^{11}\pi^2 g_s^2 \ell_s^8} \lim_{L/\ell_s \to \infty} L^{14} \int_{\kappa - i\infty}^{\kappa + i\infty} \frac{d\alpha}{2\pi i} e^\alpha \alpha^{-6} \mathcal{M}\left(\frac{L^2}{2\alpha}\mathbf{s}, \frac{L^2}{2\alpha}\mathbf{t}\right), \qquad (2.10)$$

where $\kappa > 0$.[7] This relation, as well as the AdS/CFT dictionary

$$\tau_s = \tau, \qquad \frac{L^4}{\ell_s^4} = \lambda = g_{\text{YM}}^2 \sqrt{4c+1} \qquad (2.11)$$

imply that

$$\beta(\tau, \bar{\tau}) = \frac{15}{4\sqrt{2\pi^3}} E(\tfrac{3}{2}, \tau, \bar{\tau}), \qquad \gamma_1(\tau, \bar{\tau}) = \frac{315}{128\sqrt{2\pi^5}} E(\tfrac{5}{2}, \tau, \bar{\tau}), \qquad \text{etc.} \qquad (2.12)$$

2.2 Constraints from supersymmetric localization

As explained in [33], supersymmetric localization also imposes constraints on the coefficients of the expansion in (2.4). While there are several possible supersymmetric localization constraints, the one studied in [33] came from the mixed derivative $\frac{\partial^4 \log Z}{\partial \tau \partial \bar{\tau} \partial m^2}\big|_{m=0}$ of the $\mathcal{N} = 2^*$ theory on a round S^4.

In a large c expansion, a careful analysis of the integrated constraints that follow from $\frac{\partial^4 \log Z}{\partial \tau \partial \bar{\tau} \partial m^2}\big|_{m=0}$ as well as the ansatz (2.4) gives [33]

$$\frac{\partial_\tau \partial_{\bar{\tau}} \partial_m^2 \log Z}{\partial_\tau \partial_{\bar{\tau}} \log Z}\bigg|_{m=0} = 2 - \frac{\beta(\tau, \bar{\tau})}{5c^{3/4}} + \frac{C_{\text{1-loop}}}{c} - \frac{16\gamma_1(\tau, \bar{\tau}) + 7\gamma_2(\tau, \bar{\tau})}{35c^{5/4}} + \cdots, \qquad (2.13)$$

where $C_{\text{1-loop}}$ is a constant that depends on the precise form of the $\mathcal{M}_{\text{1-loop}}$ amplitude that we will not study here. We have normalized the integrated four-point correlators by

$$\partial_\tau \partial_{\bar{\tau}} \log Z\bigg|_{m=0} = \frac{c}{2(\Im\tau)^2}. \qquad (2.14)$$

In the following section, using the matrix model for the $\mathcal{N} = 2^*$ partition function derived by Pestun [37], we will show that, up to an additive ambiguity that is a sum of holomorphic and anti-holomorphic terms in the complexified coupling,

$$\partial_m^2 \log Z\big|_{m=0} = -(4c+1)\log \Im\tau - \frac{\sqrt{2}}{\pi^{3/2}} E(\tfrac{3}{2}, \tau, \bar{\tau}) c^{1/4} + \frac{3}{16\sqrt{2\pi^5}} E(\tfrac{5}{2}, \tau, \bar{\tau}) \frac{1}{c^{1/4}} + \cdots. \quad (2.15)$$

(See also eq. (4.1) in section 4.) From comparing this expression to (2.13) and (2.14), we then derive constraints on the coefficients appearing in the Mellin amplitude (2.4). In particular, at order $c^{-3/4}$, eqs. (2.13)–(2.15) imply that β should take precisely the same value as derived in (2.12) using the flat space limit constraint! This can be viewed as a

---

[7]When evaluating this integral, it is useful to note that $\int_{\kappa - i\infty}^{\kappa + i\infty} \frac{d\alpha}{2\pi i} e^\alpha \alpha^{-n} = \frac{1}{\Gamma(n)}$.

derivation of the expression for f_{R^4} given in (2.8) or as a precision test of AdS/CFT at finite g_s. At order $c^{-5/4}$, we can combine (2.13)–(2.15) with the constraint (2.12) obtained from the flat space limit to deduce that

$$\gamma_2 = -3\gamma_1 \,. \tag{2.16}$$

What remains to be done is to derive eq. (2.15), which we carry out in the following section.

3 Eisenstein series from the mass-deformed S^4 partition function

3.1 Setup

As shown in [37], up to an overall normalization constant, the mass-deformed partition function of the $\mathcal{N} = 4$ SU(N) SYM theory (preserving an $\mathcal{N} = 2$ subalgebra) is

$$Z(m, \tau, \bar\tau) = \int d^{N-1}a \, \frac{\prod_{i<j} a_{ij}^2 H^2(a_{ij})}{H(m)^{N-1} \prod_{i\neq j} H(a_{ij} + m)} e^{-\frac{8\pi^2}{g_{\rm YM}^2} \sum_i a_i^2} |Z_{\rm inst}(m, \tau, a_{ij})|^2 \,, \tag{3.1}$$

where $a_{ij} \equiv a_i - a_j$, the integration is over N real variables a_i, $i = 1, \ldots, N$, subject to the constraint $\sum_i a_i = 0$, $H(z)$ is the product of two Barnes G-functions, and $Z_{\rm inst}$ is the contribution from instantons localized at the poles of S^4 that we will come to shortly. The constrained integral over the a_i can be implemented, for instance, by an integral over N unconstrained a_i's with a $\delta(\sum_i a_i)$ insertion. Note that the normalization constant that was dropped from (3.1) depends on the radius of the sphere, as required by the existence of a conformal anomaly, but it is independent of the coupling $(\tau, \bar\tau)$ and mass m. Consequently, it will drop out of the ratio (2.13) that is used to related the sphere partition function to the four-point correlator of the **20′** operator at separated points.

Evaluating $Z(m, \tau, \bar\tau)$ for all m seems to be a complicated task, and we will not pursue it here in full generality. Instead, what we need is to evaluate the second mass derivative at zero mass, $\partial_m^2 Z(m, \tau, \bar\tau)\big|_{m=0}$. Let us write the instanton partition function as

$$Z_{\rm inst}(m, \tau, a_{ij}) = \sum_{k=0}^{\infty} e^{2\pi i k\tau} Z_{\rm inst}^{(k)}(m, a_{ij}) \,, \tag{3.2}$$

with $Z_{\rm inst}^{(k)}(m, a_{ij})$ representing the contribution of the k-instanton sector and normalized such that $Z_{\rm inst}^{(0)}(m, a_{ij}) = 1$. Notably, $Z_{\rm inst}(0, \tau, a_{ij}) = 1$ [37] so the instantons do not contribute to the sphere partition function at the conformal point. Then one can argue that

$$
\begin{aligned}
\partial_m^2 \log Z\big|_{m=0} &= \partial_m^2 \log Z\big|_{m=0}^{\rm pert} + \partial_m^2 \log Z\big|_{m=0}^{\rm inst} \,, \\
\partial_m^2 \log Z\big|_{m=0}^{\rm pert} &\equiv \left\langle \partial_m^2 \prod_{i<j} \frac{H^2(a_{ij})}{H(a_{ij} - m)H(a_{ij} + m)} \right\rangle \bigg|_{m=0} \,, \\
\partial_m^2 \log Z\big|_{m=0}^{\rm inst} &\equiv \sum_{k=1}^{\infty} (e^{ik\theta} + e^{-ik\theta}) e^{-\frac{8\pi^2 k}{g_{\rm YM}^2}} \left\langle \partial_m^2 Z_{\rm inst}^{(k)}(m, a_{ij}) \right\rangle \bigg|_{m=0} \,,
\end{aligned}
\tag{3.3}
$$

where the expectation value is defined to be in the Hermitian matrix model at $m = 0$.[8]
The perturbative terms $\partial_m^2 \log Z|_{m=0}^{\text{pert}}$ were shown in [34] to take the form[9]

$$\partial_m^2 \log Z|_{m=0}^{\text{pert}} = 2N^2 \log g_{\text{YM}} + \sqrt{N} \left[\frac{16\zeta(3)}{g_{\text{YM}}^3} + \frac{g_{\text{YM}}}{3} \right] - \frac{1}{\sqrt{N}} \left[\frac{12\zeta(5)}{g_{\text{YM}}^5} + \frac{g_{\text{YM}}^3}{1440} \right] + \dots , \quad (3.5)$$

where further perturbative terms will be discussed in section 4. These perturbative terms match those of the expected Eisenstein series in (2.15) as defined in (2.9). To similarly match the instanton terms, we need to show that at large N we have

$$e^{-\frac{8\pi^2 k}{g_{\text{YM}}^2}} \left\langle \partial_m^2 Z_{\text{inst}}^{(k)}(m, a_{ij}) \right\rangle \bigg|_{m=0} = -\frac{16\sqrt{N}}{g_{\text{YM}}} k \, \sigma_{-2}(k) \, K_1(8\pi^2 k/g_{\text{YM}}^2)$$
$$+ \frac{2}{g_{\text{YM}}\sqrt{N}} k^2 \sigma_{-4}(k) \, K_2(8\pi^2 k/g_{\text{YM}}^2) + \cdots , \quad (3.6)$$

As a warm-up, let us start with the one-instanton case $k = 1$, and then continue with the case of multiple instantons.

3.2 One-instanton sector

The one-instanton contribution $Z_{\text{inst}}^{(1)}$ is [37, 40, 41] (see also [71])

$$Z_{\text{inst}}^{(1)}(m, a_{ij}) = -m^2 \sum_{l=1}^{N} \prod_{j \neq l} \frac{(a_l - a_j + i)^2 - m^2}{(a_l - a_j)(a_l - a_j + 2i)} . \quad (3.7)$$

Taking two derivatives with respect to the mass and evaluating the result at $m = 0$ gives

$$\partial_m^2 Z_{\text{inst}}^{(1)} \bigg|_{m=0} = I_1 , \qquad I_1 \equiv -2 \sum_{l=1}^{N} \prod_{j \neq l} \frac{(a_l - a_j + i)^2}{(a_l - a_j)(a_l - a_j + 2i)} . \quad (3.8)$$

The quantity I_1 can be written in terms of a contour integral as

$$I_1 = 4 \int \frac{dz}{2\pi} \left[\prod_j \frac{(z - a_j)^2}{(z - a_j)^2 + 1} - 1 \right] = 4 \int \frac{dz}{2\pi} \left[\exp\left(\sum_j \log \frac{(z - a_j)^2}{(z - a_j)^2 + 1} \right) - 1 \right] , \quad (3.9)$$

---

[8]We could consider the expectation values in (3.3) in either the SU(N) or the U(N) theories, whose partition functions are

$$Z^{\text{SU}(N)}\big|_{m=0} = \int d^N a \, \delta\left(\sum_i a_i \right) e^{-\frac{8\pi^2}{g_{\text{YM}}^2} \sum_i a_i^2} \prod_{i<j} a_{ij}^2 , \qquad Z^{\text{U}(N)}\big|_{m=0} = \int d^N a \, e^{-\frac{8\pi^2}{g_{\text{YM}}^2} \sum_i a_i^2} \prod_{i<j} a_{ij}^2 . \quad (3.4)$$

Indeed, for any function F that depends only on the differences a_{ij} one can show that

$$\int d^N a \, \delta\left(\sum_i a_i \right) e^{-\frac{8\pi^2}{g_{\text{YM}}^2} \sum_i a_i^2} F(a_{ij}) = \sqrt{\frac{8\pi}{g_{\text{YM}}^2 N}} \int d^N a \, e^{-\frac{8\pi^2}{g_{\text{YM}}^2} \sum_i a_i^2} F(a_{ij}) .$$

Thus, even in the presence of an insertion depending only on a_{ij} (as is the case in (3.3)), the partition functions for the SU(N) and U(N) theories differ by a multiplicative constant that is independent of the operator being inserted. It follows that normalized expectation values are equal in the SU(N) and U(N) theories.

[9]These expressions were given in eq. (3.1) and (3.20) of [34] in the strong coupling expansion, which we can simply convert to the very strong coupling expansion by replacing $\lambda \to g_{\text{YM}}^2 N$.

where the integration contour is a counter-clockwise contour surrounding the poles at $z = a_j + i$. Note that the subtraction of 1 from the integrand does not contribute to the final result, but it does make the integrand decay as $1/z^2$ at $|z| \to \infty$. Thus, the integration contour in (3.9) can be taken to be the real line.

We need to evaluate the expectation value of I_1 in the Hermitian matrix model (3.4). In the saddle point approximation, where at leading order in $1/N$ correlation functions factorize, we can approximately write the expectation value of the second expression in (3.9) as

$$\langle I_1 \rangle \approx 4 \int_{-\infty}^{\infty} \frac{dz}{2\pi} \left[\exp \left\langle \sum_j \log \frac{(z - a_j)^2}{(z - a_j)^2 + 1} \right\rangle - 1 \right]. \qquad (3.10)$$

We can change variables from z to x where $z = x\sqrt{\lambda}/(2\pi) = x\frac{\sqrt{N}g_{\text{YM}}}{2\pi}$, and rescale the a_j's similarly: $a_j = \sqrt{\frac{\lambda}{2\pi}}b_j$. It is a standard result on Hermitian matrix models that, at leading order in large N, the b_j become dense and their density is described by the Wigner semicircle law as

$$\rho(b) = \frac{2}{\pi}\sqrt{1 - b^2}, \qquad b \in [-1, 1], \qquad (3.11)$$

normalized so that $\int_{-1}^{1} db\, \rho(b) = 1$. Making the replacement $\sum_j (\cdots) \to N \int db\, (\cdots)$, we can then write (3.10) approximately as

$$\langle I_1 \rangle \approx 2\sqrt{N}\frac{g_{\text{YM}}}{\pi} \int_{-\infty}^{\infty} \frac{dx}{2\pi} \left[\exp\left(-N \int_{-1}^{1} db\, \rho(b) \log\left(1 + \frac{4\pi^2}{Ng_{\text{YM}}^2(x - b)^2} \right) \right) - 1 \right]. \quad (3.12)$$

The leading contribution to the integral is given by taking $N \to \infty$ in the integrand. In this limit, $N \log\left(1 + \frac{4\pi^2}{(x-b)^2 g_{\text{YM}}^2 N} \right) \to \frac{4\pi^2}{(x-b)^2 g_{\text{YM}}^2}$, so

$$\langle I_1 \rangle \approx 2\sqrt{N}\frac{g_{\text{YM}}}{\pi} \int_{-\infty}^{\infty} \frac{dx}{2\pi} \left[\exp\left(-\frac{8\pi}{g_{\text{YM}}^2} \int_{-1}^{1} db\, \frac{\sqrt{1 - b^2}}{(x - b)^2} \right) - 1 \right]. \qquad (3.13)$$

Performing the b integral:

$$\int_{-1}^{1} db\, \frac{\sqrt{1 - b^2}}{(x - b)^2} = \begin{cases} \pi\left(-1 + \frac{1}{\sqrt{1 - x^{-2}}} \right), & \text{if } |x| > 1, \\ \infty, & \text{if } |x| < 1, \end{cases} \qquad (3.14)$$

we obtain

$$\langle I_1 \rangle \approx 2\sqrt{N}\frac{g_{\text{YM}}}{\pi} e^{\frac{8\pi^2}{g_{\text{YM}}^2}} \int_{-\infty}^{\infty} \frac{dx}{2\pi} \left[e^{-\frac{8\pi^2}{g_{\text{YM}}^2}\frac{1}{\sqrt{1 - x^{-2}}}} \theta(|x| - 1) - e^{-\frac{8\pi^2}{g_{\text{YM}}^2}} \right], \qquad (3.15)$$

where $\theta(x)$ is the Heaviside theta function.

The integrand is an even function of x, so we can just integrate from $x = 0$ to $x = \infty$ and multiply the answer by a factor of 2. On this interval, we can further change variables from x to t, where

$$t = \frac{x}{\sqrt{x^2 - 1}} \qquad \Longleftrightarrow \qquad x = \frac{t}{\sqrt{t^2 - 1}}, \qquad (3.16)$$

and obtain

$$\langle I_1 \rangle \approx 2\sqrt{N} \frac{g_{\mathrm{YM}}}{\pi^2} e^{\frac{8\pi^2}{g_{\mathrm{YM}}^2}} \left[-e^{-\frac{8\pi^2}{g_{\mathrm{YM}}^2}} + \int_1^\infty \frac{dt}{(t^2-1)^{3/2}} \left(e^{-\frac{8\pi^2}{g_{\mathrm{YM}}^2}t} - e^{-\frac{8\pi^2}{g_{\mathrm{YM}}^2}} \right) \right] . \tag{3.17}$$

We can now use the integral representation of the Bessel K_1 function,

$$\int_1^\infty dt \frac{e^{-at} - e^{-a}}{(t^2-1)^{3/2}} = e^{-a} - a K_1(a) , \tag{3.18}$$

to finally write the $\sqrt{N}$ term in the large-N expansion of $\langle I_1 \rangle$ as

$$\langle I_1 \rangle \big|_{\sqrt{N}} = -\sqrt{N} \frac{16}{g_{\mathrm{YM}}} e^{\frac{8\pi^2}{g_{\mathrm{YM}}^2}} K_1(8\pi^2/g_{\mathrm{YM}}^2) . \tag{3.19}$$

Combining with (3.8), this expression implies

$$e^{-\frac{8\pi^2}{g_{\mathrm{YM}}^2}} \left\langle \partial_m^2 Z_{\mathrm{inst}}^{(1)}(m, a_{ij}) \right\rangle \bigg|_{m=0} \approx -\sqrt{N} \frac{16 K_1(8\pi^2/g_{\mathrm{YM}}^2)}{g_{\mathrm{YM}}} , \tag{3.20}$$

in agreement with the expansion of the Eisenstein series — See eq. (3.6) in the case $k = 1$.

Obtaining the term that scales as $1/\sqrt{N}$ in (3.6) is not any harder, because this term is suppressed only by a factor of $1/N$ relative to the term we just computed, while the corrections to the approximations made in writing (3.10) and (3.12) are suppressed by $1/N^2$. Thus, the next term in (3.19) can be obtained by simply expanding (3.12) to one more order in $1/N$ so that

$$N \log \left(1 + \frac{4\pi^2}{(x-b)^2 g_{\mathrm{YM}}^2 N} \right) = \frac{4\pi^2}{(x-b)^2 g_{\mathrm{YM}}^2} - \frac{8\pi^4}{(x-b)^4 g_{\mathrm{YM}}^4 N} + \cdots , \tag{3.21}$$

and evaluating the effect of the second term in (3.21) in the same way as above. Plugging (3.21) into (3.12), we obtain

$$\langle I_1 \rangle \big|_{1/\sqrt{N}} = \frac{1}{\sqrt{N}} \frac{32\pi^2}{g_{\mathrm{YM}}^3} \int_{-\infty}^\infty \frac{dx}{2\pi} \left(\int_{-1}^1 db \frac{\sqrt{1-b^2}}{(x-b)^4} \right) \exp \left(-\frac{8\pi}{g_{\mathrm{YM}}^2} \int_{-1}^1 db \frac{\sqrt{1-b^2}}{(x-b)^2} \right) . \tag{3.22}$$

Performing the b integrals and the $x \to t$ substitution (3.16), we find

$$\langle I_1 \rangle \big|_{1/\sqrt{N}} = \frac{1}{\sqrt{N}} \frac{16\pi^2}{g_{\mathrm{YM}}^3} \int_1^\infty dt\, t \sqrt{t^2-1}\, e^{-\frac{8\pi^2}{g_{\mathrm{YM}}^2}(t-1)} = \frac{1}{\sqrt{N}} e^{\frac{8\pi^2}{g_{\mathrm{YM}}^2}} \frac{2 K_2(8\pi^2/g_{\mathrm{YM}}^2)}{g_{\mathrm{YM}}} . \tag{3.23}$$

Adding this expression to (3.19) and using the definition of I_1 in (3.8), we conclude that

$$e^{-\frac{8\pi^2}{g_{\mathrm{YM}}^2}} \left\langle \partial_m^2 Z_{\mathrm{inst}}^{(1)}(m, a_{ij}) \right\rangle \bigg|_{m=0} = -\sqrt{N} \frac{16 K_1(8\pi^2/g_{\mathrm{YM}}^2)}{g_{\mathrm{YM}}} + \frac{1}{\sqrt{N}} \frac{2 K_2(8\pi^2/g_{\mathrm{YM}}^2)}{g_{\mathrm{YM}}} + \mathcal{O}(N^{-1}) , \tag{3.24}$$

which is again in agreement with the expectation (3.6).

Note that in order to go beyond the first two orders in the $1/N$ expansion, one would have to take into account the $1/N^2$ corrections to the saddle point evaluation of expectation values, which we will do later in section 4.

3.3 The $k > 1$ instanton sector

We now consider the instanton sector with general $k > 1$. Recall the instanton partition function of SU(N) $\mathcal{N} = 2^*$ SYM on S^4 can be written as in eq. (3.2). The k-instanton partition $Z_{\text{inst}}^{(k)}(m, a_{ij})$ may be further expressed[10] as a sum of contour integrals around poles indicated by a vector $\vec{Y} = (Y_1, Y_2, \ldots, Y_N)$ of N Young diagrams,

$$Z_{\text{inst}}^{(k)}(m, a_{ij}) = \sum_{|\vec{Y}|=k} Z_{k,\vec{Y}}(m, a_{ij}), \tag{3.25}$$

where the Young diagrams Y_i in $\vec{Y}$ are such that the total number of boxes is k, namely $|\vec{Y}| \equiv \sum_i |Y_i| = k$. A Young diagram Y consists of columns of non-increasing height $\lambda_1 \geq \lambda_2 \geq \ldots$, and the transpose Young diagram Y^T has columns $\lambda_1^T \geq \lambda_2^T \geq \ldots$. We will often write them compactly as $Y = [\lambda_1, \lambda_2, \ldots]$ and $Y^T = [\lambda_1^T, \lambda_2^T, \ldots]$.

Explicitly, $Z_{k,\vec{Y}}(m, a_{ij})$ is given by [40, 41][11]

$$Z_{k,\vec{Y}}(m, a_{ij}) = \frac{1}{k!} \left(\frac{\epsilon_+(m^2 + \epsilon_-^2/4)}{\epsilon_1 \epsilon_2(m^2 + \epsilon_+^2/4)} \right)^k \oint \prod_{I=1}^k \frac{d\phi_I}{2\pi} \prod_{i=1}^N \frac{(\phi_I - a_j)^2 - m^2}{(\phi_I - a_j)^2 + \epsilon_+^2/4} \tag{3.26}$$

$$\times \prod_{I<J}^k \frac{\phi_{IJ}^2 [\phi_{IJ}^2 + \epsilon_+^2][\phi_{IJ}^2 + (im - \epsilon_-/2)^2][\phi_{IJ}^2 + (im + \epsilon_-/2)^2]}{[\phi_{IJ}^2 + \epsilon_1^2][\phi_{IJ}^2 + \epsilon_2^2][\phi_{IJ}^2 + (im + \epsilon_+/2)^2][\phi_{IJ}^2 + (im - \epsilon_+/2)^2]},$$

where $\phi_{IJ} \equiv \phi_I - \phi_J$ and $\epsilon_\pm \equiv \epsilon_1 \pm \epsilon_2$, with $\epsilon_{1,2}$ being the squashing parameters of S^4. (We have $\epsilon_1 = \epsilon_2 = 1$ for a round sphere.) The multi-dimensional integration contour is determined by the Jeffrey-Kirwan (JK) prescription [75], which selects the poles to encircle based on the choice of an auxiliary vector η in $\mathbb{R}^k$.[12] As mentioned above, the set of relevant poles are in one-to-one correspondence with the vector of Young diagrams $\vec{Y} = (Y_1, Y_2, \ldots, Y_N)$. Each box in the jth Young diagram Y_j is labelled by its position (α, β) for positive integers α and β denoting the column and the row, respectively, as measured from the bottom left corner (see figure 1).[13] The integral then consists of oriented contours surrounding the poles at

$$\{\phi_I | 1 \leq I \leq k\} = \{a_j + i\epsilon_+/2 + (\alpha - 1)i\epsilon_1 + (\beta - 1)i\epsilon_2 \,|(\alpha, \beta) \in Y_i, \ 1 \leq i \leq N\}. \tag{3.27}$$

---

[10]Note that the instanton partition function for $\mathcal{N} = 2^*$ SYM was originally obtained for U(N) gauge group [40, 41]. Later in [72, 73], the SU(N) instanton partition function was obtained by factorizing out U(1) contributions $Z_{U(1)}$ motivated by the AGT correspondence [72]. Since $Z_{U(1)}$ is holomorphic in τ and independent of the a_i, this is absorbed into the holomorphic (anti-holomorphic) ambiguity of (2.15), and does not affect the physical four-point-functions. For this reason, we will simply use the results for U(N) instanton partition functions in the following. We thank Yuji Tachikawa for pointing out the reference [73].

[11]Our notation here is related, for instance, to equation (3) of [74] by sending $\phi_I \to -i\phi_I, a_j \to -ia_j$ and $m_N \to im + \epsilon_+/2$.

[12]The physics derivation of this contour choice from localization was given in [76].

[13]We use the French (as opposed to the English) convention for drawing Young diagrams, where the diagrams extend up and to the right.

For the case of the round S^4, which is relevant for our consideration, we set $\epsilon_1 = \epsilon_2 = 1$. The k-instanton partition function then reduces to

$$Z_{k,\vec{Y}}(m, a_{ij}) = \frac{1}{k!} \left(\frac{2m^2}{m^2+1} \right)^k \oint \prod_{I=1}^{k} \frac{d\phi_I}{2\pi} \prod_{i=1}^{N} \frac{(\phi_I - a_i)^2 - m^2}{(\phi_I - a_i)^2 + 1}$$
$$\times \prod_{I<J}^{k} \frac{\phi_{IJ}^2 (\phi_{IJ}^2 + 4)(\phi_{IJ}^2 - m^2)^2}{(\phi_{IJ}^2 + 1)^2 [(\phi_{IJ} - m)^2 + 1][(\phi_{IJ} + m)^2 + 1]}, \qquad (3.28)$$

and the contours now are surrounding

$$\{\phi_I | 1 \leq I \leq k\} = \{a_j + (\alpha + \beta - 1)i \,|(\alpha, \beta) \in Y_i,\ 1 \leq i \leq N\}. \qquad (3.29)$$

Note some of the simple poles in (3.27) for certain $\vec{Y}$ are degenerate when $\epsilon_1 = \epsilon_2$, giving rise to higher order poles in (3.28). For certain purposes, (3.26) is more convenient to use because it contains only simple poles when $\epsilon_1 \neq \epsilon_2$, and one can set $\epsilon_1 = \epsilon_2 = 1$ after evaluating the residues.

In the case of $k = 1$, (3.28) reduces to eq. (3.7) that we studied in the previous sections. In this section, we are interested in the cases with $k > 1$. For instance, for $k = 2$, the relevant poles are enumerated by the Young diagrams N-tuples with a total of two boxes:

$$\vec{Y} = \{\dots, \square\square, \dots, \}, \quad \vec{Y} = \{\dots, \begin{smallmatrix}\square\\\square\end{smallmatrix}, \dots, \}, \quad \vec{Y} = \{\dots, \square, \dots, \square, \dots\}, \qquad (3.30)$$

and the 2-instanton partition function is then obtained by the sum (3.28) over the contributions from each vector of Young diagrams shown in (3.30). In general, such a k-instanton partition function is very complicated and difficult to analyze, especially when k is large. However, as we will show below, at the order m^2 in a small m expansion (relevant for the four-point correlation function we consider), the k-instanton partition function simplifies greatly even for arbitrary k.

3.3.1 Dominance of rectangular Young diagrams

In this section, we will study the instanton partition function on S^4 in the small-mass expansion. It turns out to be convenient to keep ϵ_1 and ϵ_2 general at first and send $\epsilon_1 \to \epsilon_2$ at the end.[14] We will argue that the k-instanton partition function at order m^2, i.e.

$$I_k \equiv \partial_m^2 Z_{\text{inst}}^{(k)}(m, a_{ij})\big|_{m=0}, \qquad (3.31)$$

only receives contributions from a vector of Young diagrams $\vec{Y}$ where all the N diagrams are empty except for one, say $Y_{\hat{i}}$, which is either rectangular or related by a sequence of *partial transpositions* to a rectangular diagram (we define what we mean by partial transpositions below). In other words, $\vec{Y}$ is related by partial transpositions to the following

$$\begin{aligned} Y_i &= \varnothing && \text{if} && i \neq \hat{i}, \\ Y_{\hat{i}} &= Y_{p \times q} && \text{for} && p \leq q \in \mathbb{Z}_+,\ \ pq = k, \end{aligned} \qquad (3.32)$$

---

[14]The same final result can be reached if we instead used (3.28) where $\epsilon_1 = \epsilon_2$ from the beginning.

where $Y_{p \times q}$ denotes the rectangular Young diagram with p rows and q columns. See appendix B and Theorem B.1 for the complete proof. In the following we present a brief summary of the proof.

Given a Young diagram $Y = [\lambda_1, \ldots, \lambda_l]$, the *partial transposition* at horizontal position α, denoted by PT_α, is a local operation on Y that replaces the rightmost block (subdiagram) $P = [\lambda_\alpha, \lambda_{\alpha+1}, \ldots, \lambda_l]$ by its transpose, while leaving the rest of the Young diagram unchanged, provided that the new diagram is still a Young diagram. Otherwise, the particular partial transposition PT_α is defined to act trivially on Y (for a more detailed discussion, see appendix B). In particular PT_1 generates the usual transposition of Young diagrams. In general $\vec{Y}$ can have many different partners from *partial transpositions*. For instance, for $k = 4$, we have two cases of vectors $\vec{Y}$ of the type (3.32):

$$\vec{Y} = \{\ldots, \begin{array}{c}\square\square\square\square\end{array}, \ldots\}, \qquad \vec{Y} = \{\ldots, \begin{array}{c}\square\square\\\square\square\end{array}, \ldots\}. \tag{3.33}$$

The other Young diagram vectors $\vec{Y}$ that contribute are either of the form

$$\vec{Y} = \{\ldots, \begin{array}{c}\square\\\square\\\square\\\square\end{array}, \ldots\}, \tag{3.34}$$

and related by a single transposition to the first case in (3.33), or are of the form

$$\vec{Y} = \{\ldots, \begin{array}{c}\blacksquare\\\blacksquare\\\square\square\end{array}, \ldots\}, \qquad \vec{Y} = \{\ldots, \begin{array}{c}\square\\\square\blacksquare\blacksquare\end{array}, \ldots\}. \tag{3.35}$$

The two $\vec{Y}$'s in (3.35) are related to each other by a transposition, and the second one is related to the second diagram in (3.33) via a single partial transposition involving the orange block.

Note that if we set $\epsilon_1 = \epsilon_2 = 1$ from the beginning (i.e. starting with (3.28)), the $\vec{Y}$'s and their partners related by transpositions and partial transpositions degenerate in the sense that they give exactly the same poles in the integrand of the contour integral (see Lemma B.2).

It is useful to introduce the arm-length and leg-length of a box at position (α, β) (see figure 1):

$$h(\alpha, \beta) \equiv \lambda_\beta^T - \alpha, \qquad v(\alpha, \beta) \equiv \lambda_\alpha - \beta, \tag{3.36}$$

where λ_β^T is as defined below (3.25). In other words, $h(\alpha, \beta)$ and $v(\alpha, \beta)$ measure the number of boxes from the given box (α, β) to the right edge and the top edge, respectively, of the Young diagram.

As we show in appendix B, in general the contribution of a given $\vec{Y}$ to the second mass-derivative of the instanton partition function behaves as

$$I_{k,\vec{Y}} \equiv \partial_m^2 Z_{k,\vec{Y}}(m, a_{ij})\big|_{m=0} \sim (\epsilon_-)^{-2 + \sum_{i=1}^N \mu(Y_i)}, \tag{3.37}$$
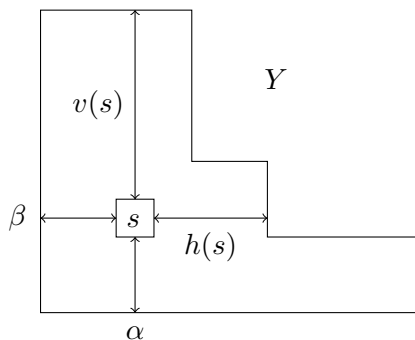
Figure 1. An example of Young diagram Y, the coordinates (α, β) for box s and the corresponding arm-length $h(s)$ and leg-length $v(s)$.

in the limit $\epsilon_- \to 0$, where

$$\mu(Y) = 2n_0(Y) - n_{-1}(Y) - n_1(Y) \tag{3.38}$$

with n_d representing the number of the boxes with $h - v = d$.

It is straightforward to see that all rectangular Young diagrams $Y_{p \times q}$ have

$$\mu(Y_{p \times q}) = \begin{cases} 1\,, & \text{if} \quad p \neq q\,, \\ 2\,, & \text{if} \quad p = q\,, \end{cases} \tag{3.39}$$

and, importantly, the operation of partial transposition does not change the value of $\mu(Y)$ (see Lemma B.4). Therefore, from (3.37), we see that each $\vec{Y}$ of the type (3.32) leads to a finite contribution to the instanton partition function in the small-mass expansion.[15]

Furthermore, all other types of Young diagram N-tuples either have $\sum_{i=1}^{N} \mu(Y_i) > 2$, and therefore their contributions manifestly vanish, or $\sum_{i=1}^{N} \mu(Y_i) = 2$, but for such a $\vec{Y}$ vector (where every non-empty Y_i is not related to a rectangular diagram by partial transpositions) there is always a cancellation between $\vec{Y}$ and a partner related by a certain involution that arises from a sequence of partial positions (see appendix B).[16] For instance, consider the following two $\vec{Y}$'s that are related by partial transposition to each another (transposing the orange block):

$$\vec{Y} = \{\ldots, \;\; , \ldots\}\,, \qquad \vec{Y} = \{\ldots, \;\; , \ldots\}\,. \tag{3.40}$$

The above diagrams in $\vec{Y}$'s are not related to any rectangular Young diagrams by partial transpositions, and each has $\mu = 2$ (therefore naively would lead to a finite contribution to

---

[15]One may worry about the divergence arising from $I_{k,(\ldots,Y_{p \times q},\ldots)} \sim (\epsilon_-)^{-1}$ for $p \neq q$, but the divergence is canceled out in the sum $I_{k,(\ldots,Y_{p \times q},\ldots)} + I_{k,(\ldots,Y_{q \times p},\ldots)}$ and leads to a finite result.

[16]Such a cancellation does not happen for (3.32) with a square Young diagram $Y_{p \times p}$, which also have $\mu = 2$ as shown in (3.39).

the instanton partition function). But the finite contributions from these two $\vec{Y}$'s in fact cancel out. This is a very general phenomenon. Again, we refer a general proof of all these statements to the appendix B.

3.3.2 Instanton partition function at order m^2

After showing the dominance of the rectangular Young diagrams for the leading m^2 order of the instanton partition function, let us now compute the residues necessary for evaluating (3.26) or (3.28). As mentioned above, while in general the contribution of any given Young diagram to the k-instanton partition function is rather complicated, the quadratic term in the small mass expansion will end up being quite simple.

For instance, the instanton partition function for $k = 2$ can be computed either from (3.26), in which case we have non-trivial contributions only from

$$\vec{Y} = \{\ldots, \boxed{}, \ldots\}, \qquad \vec{Y} = \{\ldots, \boxed{}, \ldots\}. \tag{3.41}$$

Alternatively, we can also use (3.28), in which case the above two $\vec{Y}$'s give the same set of poles. For the case of $k = 2$, for each Young diagram, there are two ways of distributing ϕ_I to the boxes of the Young diagram, which lead to two contributions to the partition function. For instance, using the formula (3.26), the two residues that we should evaluate for the first Young diagram in (3.41) are given by

$$R_1 = \text{Res}_{\phi_1 = a_j + i\epsilon_+/2} \text{Res}_{\phi_2 = \phi_1 + i\epsilon_1}, \qquad R_2 = \text{Res}_{\phi_2 = a_j + i\epsilon_+/2} \text{Res}_{\phi_1 = \phi_2 + i\epsilon_1}. \tag{3.42}$$

For the second Young diagram (which is the conjugate of the first one), the residues are computed in the same way, but with $\epsilon_1 \leftrightarrow \epsilon_2$. Furthermore, R_1 and R_2 give the same contributions since they simply exchange $\phi_1 \leftrightarrow \phi_2$. It is thus convenient to only evaluate the residue $\text{Res}_{\phi_2 = \phi_1 + i\epsilon_1}$ in R_1, and $\text{Res}_{\phi_1 = \phi_2 + i\epsilon_1}$ in R_2, and leave the remaining variable (namely ϕ_1 in the first case, and ϕ_2 in the second case) unintegrated. (The variable that remains is the one corresponding to the box of the Young diagram that sits in the bottom left corner.) Computing the residues explicitly and summing up the contributions from both Young diagrams, we obtain

$$I_{1\times2} = \oint \frac{dz}{2\pi} \prod_{k_a} \prod_{j=1}^{N} \frac{(z - a_j + k_a i)^2}{(z - a_j + k_a i)^2 + 1}$$

$$\times \left[5 + \sum_{j=1}^{N} \frac{3i}{(z - a_j + 2i)(z - a_j + i)(z - a_j)} \right], \tag{3.43}$$

where $k_a = \{0, 1\}$, and we have set $\epsilon_1 = \epsilon_2 = 1$ at the end of the computation and denoted the unintegrated variable as z in (3.43). The contour for the remaining z integral is around poles at $z = a_j + i$, with $j = 1, 2, \ldots, N$. Note the same result (3.43) can be obtained by using (3.28) (i.e. setting $\epsilon_1 = \epsilon_2 = 1$ from the beginning).

In general, for the case of k instantons, there are $k!$ ways of assigning ϕ_I's to a given Young diagram,[17] and we will integrate out all the $(k-1)$ ϕ_I's, but again leave the one that is assigned to the bottom left corner box unintegrated (just as the $k=2$ case), and denote it by z. The contour for the remaining z-integration is then a counter-clockwise contour surrounding the poles at $z = a_j + i$, with $j = 1, 2, \ldots, N$.

Let us consider another example before presenting a general k-instanton formula. For instance, the 4-instanton partition function, for which there are two types of Young diagrams that contribute. The first type is given by

$$\vec{Y} = \{\ldots, \square, \ldots\}, \qquad \vec{Y} = \{\ldots, \square\square\square\square, \ldots\}, \tag{3.44}$$

and the second kind is

$$\vec{Y} = \{\ldots, \square\square, \ldots\}, \quad \vec{Y} = \{\ldots, \square\square, \ldots\}, \quad \vec{Y} = \{\ldots, \square\square\square, \ldots\}. \tag{3.45}$$

Computing each contribution explicitly, we again find very compact results with similar structures as those of (3.43) from the two-instanton case. For the Young diagrams in (3.44), we find

$$I_{1\times 4} = \oint \frac{dz}{2\pi} \prod_{k_a} \prod_{j=1}^{N} \frac{(z - a_j + k_a i)^2}{(z - a_j + k_a i)^2 + 1} \tag{3.46}$$
$$\times \left[\frac{17}{4} + \sum_{j=1}^{N} \frac{45i}{2(z - a_j + 4i)(z - a_j + 3i)(z - a_j)} \right],$$

with $k_a = \{0, 1, 2, 3\}$, while for (3.45) we obtain

$$I_{2\times 2} = \oint \frac{dz}{2\pi} \prod_{k_a} \prod_{j=1}^{N} \frac{(z - a_j + k_a i)^2}{(z - a_j + k_a i)^2 + 1}, \tag{3.47}$$

with $k_a = \{0, 1, 1, 2\}$.

The above simple structures present in the examples we have studied generalize. Indeed, we find that the contribution to the k-instanton partition function coming from Young diagram vectors of the form (3.32) as well as its partial transpositions is given by

$$I_{p\times q} = \oint \frac{dz}{2\pi} \prod_{k_a} \prod_{j=1}^{N} \frac{(z - a_j + k_a i)^2}{(z - a_j + k_a i)^2 + 1} \times \left[\frac{4}{1 + \delta_{pq}} \left(\frac{1}{p^2} + \frac{1}{q^2} \right) \right.$$
$$\left. + \sum_{j=1}^{N} \frac{i f(p, q)}{(z - a_j + (p + q - 1)i)(z - a_j + (q - 1)i)(z - a_j + (p - 1)i)} \right], \tag{3.48}$$

---

[17]If we use (3.28) for the computation, one should also divide the symmetry factors due to the degeneracy of the poles (i.e. according to (3.29), the ϕ_I's corresponding to the boxes at positions (α, β) of a Young diagram with the same values of $\alpha + \beta$ surround the same poles.).

where the integration contour of the left-over z is a counter-clockwise contour surrounding the poles at $z = a_j + i$ (with $j = 1, 2, \ldots, N$). The k_a's (k of them) are read off from the vector of Young diagrams $\vec{Y}$ as in (3.32), and they are given by

$$k_a = \{0, 1, \cdots, p-1; 1, 2, \cdots, p; \cdots; q-1, q, \cdots, p+q-2\}. \tag{3.49}$$

Finally, the function $f(p, q)$ is

$$f(p, q) = \frac{2(q+p)(q-p)^2}{pq}. \tag{3.50}$$

This function is symmetric in $p \leftrightarrow q$ and vanishes at $p = q$. The formula (3.48), which is one of our main results, was obtained by studying the pattern of many non-trivial examples. We will study its large-N expansion in the next section. For the special case where the non-trivial rectangular Young diagram in $\vec{Y}$ is $Y_{1 \times k}$, we provide a proof in appendix C using a recursion relation satisfied by the instanton partition function [77, 78]. However, the same recursion relation proof for general $Y_{p \times q}$ becomes a bit cumbersome. Nevertheless, we have verified the formula (3.48) explicitly up to $k = 20$ instantons. Furthermore, the result (3.48) clearly has a structure that is consistent with (3.28): the constant part $\frac{4}{1+\delta_{pq}}\left(\frac{1}{p^2} + \frac{1}{q^2}\right)$ arises only from the second line of (3.28), whereas the term depending on z and a_j involves expanding the first line of (3.28) while taking the residues around higher order poles.

Finally, we remark that given the structure of the k-instanton contribution to the non-holomorphic Eisenstein series, especially the appearance of the divisor sum (see (2.9)), it is not surprising that the relevant $\vec{Y}$'s are only the rectangular ones. As we will show, each Young diagram $Y_{p \times q}$ contributes a term in the divisor sum for a non-holomorphic Eisenstein series (proportional to $p^{1-2r} + q^{1-2r}$ for $E(r, \tau, \bar{\tau})$).[18]

3.3.3 Large-N expansion

We will now compute the expectation value of $I_{p \times q}$ in the Hermitian matrix model (3.4), in the large-N expansion. The computation is similar to that in the one-instanton case presented in section 3.2, so we will therefore be brief here. In the large-N limit, we have

$$\langle I_{p \times q} \rangle \approx \frac{\sqrt{N} g_{\mathrm{YM}}}{2\pi} \int_{-\infty}^{\infty} \frac{dx}{2\pi} \left(\exp\left[-N \int_{-1}^{1} db \rho(b) \sum_{k_a} \log\left(1 + \frac{4\pi^2}{N g_{\mathrm{YM}}^2 (x - b + \frac{2\pi i}{\sqrt{N} g_{\mathrm{YM}}} k_a)^2} \right) \right] \right.$$
$$\left. \times \left(\frac{4}{1+\delta_{pq}} \frac{p^2+q^2}{p^2 q^2} + i \frac{(\frac{2\pi}{g_{\mathrm{YM}}})^3}{\sqrt{N}} \int_{-1}^{1} db \rho(b) \frac{f(p,q)}{g(x,b)} \right) - \frac{4}{1+\delta_{pq}} \frac{p^2+q^2}{p^2 q^2} \right), \tag{3.51}$$

with k_a given in (3.49). In the above, we have approximated the sums as integrals, and we have deformed the contour by subtracting an appropriate constant from the integrand. Again, the density measure $\rho(b)$ obeys Wigner's semi-circle law, given by $\rho(b) = \frac{2}{\pi}\sqrt{1 - b^2}$,

---

[18]In particular, the term $\frac{4}{1+\delta_{pq}}\left(\frac{1}{p^2} + \frac{1}{q^2}\right)$, which dominates in the large N limit, gives directly the divisor sum $\sigma_{-2}(|k|)$ in $E(\frac{3}{2}, \tau, \bar{\tau})$.

and

$$g(x,b) = \left[x - b + \frac{2\pi i}{\sqrt{N}g_{YM}}(p+q-1)\right]\left[x - b + \frac{2\pi i}{\sqrt{N}g_{YM}}(q-1)\right]\left[x - b + \frac{2\pi i}{\sqrt{N}g_{YM}}(p-1)\right].$$
(3.52)

Just as the one-instanton case, we expand the integrand in $1/N$, and the integration over x can be separated into different regions: $x \in \{-1, 1\}$, $x \in \{-\infty, -1\}$, and $x \in \{1, +\infty\}$. We find that the leading term in the $1/N$ expansion is given by

$$\langle I_{p\times q}\rangle\big|_{\sqrt{N}} = \frac{\sqrt{N}g_{YM}}{2\pi^2}\left(-\frac{4}{1+\delta_{pq}}\frac{p^2+q^2}{p^2q^2}\right) + \frac{\sqrt{N}g_{YM}}{\pi}\exp\left[k\frac{8\pi^2}{g_{YM}^2}\right]\left(\frac{4}{1+\delta_{pq}}\frac{p^2+q^2}{p^2q^2}\right)$$
$$\times \int_1^\infty \frac{dx}{2\pi}\left(\exp\left[-\frac{8k\pi^2 x}{g_{YM}^2(x^2-1)^{\frac{1}{2}}}\right] - \exp\left[-k\frac{8\pi^2}{g_{YM}^2}\right]\right).$$
(3.53)

After a change of integration variable identical to (3.16), it is straightforward to show that $\langle I_{p\times q}\rangle\big|_{\sqrt{N}}$ can be expressed in terms of a Bessel K_1 function:

$$\langle I_{p\times q}\rangle\big|_{\sqrt{N}} = -\frac{k}{1+\delta_{pq}}\left(\frac{1}{p^2}+\frac{1}{q^2}\right)\frac{16\sqrt{N}}{g_{YM}}\exp\left[k\frac{8\pi^2}{g_{YM}^2}\right]K_1\left(k\frac{8\pi^2}{g_{YM}^2}\right).$$
(3.54)

Summing over all possible rectangular $\vec{Y}$'s for a given k-instanton sector (namely all the divisors of k), we have

$$\sum_{pq=k,\,0<p\leq q}\langle I_{p\times q}\rangle\big|_{\sqrt{N}} = -\frac{16\sqrt{N}}{g_{YM}}k\,\sigma_{-2}(k)\exp\left[k\frac{8\pi^2}{g_{YM}^2}\right]K_1\left(k\frac{8\pi^2}{g_{YM}^2}\right),$$
(3.55)

where we have used

$$\sum_{pq=k,\,0<p\leq q}\frac{1}{1+\delta_{pq}}\left(\frac{1}{p^2}+\frac{1}{q^2}\right) = \sigma_{-2}(k).$$
(3.56)

The N^0 order term vanishes due to the fact that integrand is odd in x. Then, at the next order we have a $1/\sqrt{N}$ term that takes the following form

$$\langle I_{p\times q}\rangle\big|_{1/\sqrt{N}} = \frac{1}{1+\delta_{pq}}\frac{16\pi^2}{\sqrt{N}g_{YM}^5 k^2}\exp\left[k\frac{8\pi^2}{g_{YM}^2}\right]\int_1^\infty dx\left(\exp\left[-\frac{8k\pi^2 x}{g_{YM}^2(x^2-1)^{\frac{1}{2}}}\right]\right.$$
$$\times\left.\left[\frac{c_1 g_{YM}^2 x}{(x^2-1)^{\frac{5}{2}}} - \frac{c_2\pi^2}{(x^2-1)^3}\right]\right),$$
(3.57)

where c_1, c_2 are given by

$$c_1 = k(p-q)^2(p+q)(2p+2q-3) + \left(k+6\sum_a k_a^2\right)(p^2+q^2),$$

$$c_2 = 8k(p-q)^2(p+q)\sum k_a + 16(p^2+q^2)\left(\sum_a k_a\right)^2,$$
(3.58)

with k_a given in (3.49). Again, by a change of integration variable, the integral of $\langle I_{p \times q} \rangle \big|_{1/\sqrt{N}}$ reduces to a standard Bessel K_2 function,

$$
\begin{aligned}
\langle I_{p \times q} \rangle \big|_{1/\sqrt{N}} &= \frac{1}{1 + \delta_{pq}} \frac{16\pi^2}{\sqrt{N} g_{\mathrm{YM}}^5 k^2} \exp\left[k \frac{8\pi^2}{g_{\mathrm{YM}}^2} \right] \left[\frac{c_1 g_{\mathrm{YM}}^2}{\left(k \frac{8\pi^2}{g_{\mathrm{YM}}^2} \right)} - \frac{3 c_2 \pi^2}{\left(k \frac{8\pi^2}{g_{\mathrm{YM}}^2} \right)^2} \right] K_2\left(k \frac{8\pi^2}{g_{\mathrm{YM}}^2} \right) \\
&= \frac{k^2}{1 + \delta_{pq}} \left(\frac{1}{p^4} + \frac{1}{q^4} \right) \frac{2}{\sqrt{N} g_{\mathrm{YM}}} \exp\left[k \frac{8\pi^2}{g_{\mathrm{YM}}^2} \right] K_2\left(k \frac{8\pi^2}{g_{\mathrm{YM}}^2} \right) .
\end{aligned}
\tag{3.59}
$$

Again, taking into account all the relevant contributions from rectangular $\vec{Y}$'s with k boxes, the prefactor in the above formula becomes the divisor sum $\sigma_{-4}(k)$, namely,

$$
\sum_{pq=k,\, 0 < p \leq q} \frac{1}{1 + \delta_{pq}} \left(\frac{1}{p^4} + \frac{1}{q^4} \right) = \sigma_{-4}(k) .
\tag{3.60}
$$

Combining (3.59) with the result of the leading large-N term in (3.55), we obtain

$$
-\frac{16\sqrt{N}}{g_{\mathrm{YM}}} k \, \sigma_{-2}(k) \exp\left[k \frac{8\pi^2}{g_{\mathrm{YM}}^2} \right] K_1\left(k \frac{8\pi^2}{g_{\mathrm{YM}}^2} \right) + \frac{2}{\sqrt{N} g_{\mathrm{YM}}} k^2 \, \sigma_{-4}(k) \exp\left[k \frac{8\pi^2}{g_{\mathrm{YM}}^2} \right] K_2\left(k \frac{8\pi^2}{g_{\mathrm{YM}}^2} \right) .
\tag{3.61}
$$

Therefore we have proven (3.6). In the next section, we will study the higher order terms in the $1/N$ expansion, and show that in fact they are also given by non-holomorphic Eisenstein series.

4 Eisenstein series at higher orders in $1/N$

In this section we will provide additional evidence that the coefficients in the large-N expansion of $\partial_m^2 \log Z \big|_{m=0}$, which was derived in the previous sections to the first couple of orders in $1/N$ in terms of the Eisenstein series shown in (2.15), takes the form of Eisenstein series to all orders in $1/N$. In particular, we propose that, through order $N^{-7/2}$, $\partial_m^2 \log Z \big|_{m=0}$ is

$$
\begin{aligned}
\partial_m^2 \log Z \big|_{m=0} ={}& 2N^2 \log g_{\mathrm{YM}} - \frac{\sqrt{N}}{\pi^{\frac{3}{2}}} E(\tfrac{3}{2}, \tau, \bar{\tau}) + \frac{3}{16\sqrt{N} \pi^{\frac{5}{2}}} E(\tfrac{5}{2}, \tau, \bar{\tau}) \\
&+ \frac{1}{N^{\frac{3}{2}}} \left[-\frac{13}{2^9 \pi^{\frac{3}{2}}} E(\tfrac{3}{2}, \tau, \bar{\tau}) + \frac{135}{2^{11} \pi^{\frac{7}{2}}} E(\tfrac{7}{2}, \tau, \bar{\tau}) \right] \\
&+ \frac{1}{N^{\frac{5}{2}}} \left[-\frac{75}{2^{12} \pi^{\frac{5}{2}}} E(\tfrac{5}{2}, \tau, \bar{\tau}) + \frac{1575}{2^{14} \pi^{\frac{9}{2}}} E(\tfrac{9}{2}, \tau, \bar{\tau}) \right] \\
&+ \frac{1}{N^{\frac{7}{2}}} \left[\frac{1533}{2^{18} \pi^{\frac{3}{2}}} E(\tfrac{3}{2}, \tau, \bar{\tau}) - \frac{80325}{2^{21} \pi^{\frac{7}{2}}} E(\tfrac{7}{2}, \tau, \bar{\tau}) + \frac{2480625}{2^{23} \pi^{\frac{11}{2}}} E(\tfrac{11}{2}, \tau, \bar{\tau}) \right] \\
&+ O(N^{-\frac{9}{2}}) + (\text{anti})\text{holomorphic ambiguity} .
\end{aligned}
\tag{4.1}
$$

We can then take derivatives in τ and $\bar{\tau}$ to obtain the $\mathrm{SL}(2,\mathbb{Z})$ invariant quantity

$$
\begin{aligned}
\tau_2^2 \partial_\tau \partial_{\bar{\tau}} \partial_m^2 \log Z\big|_{m=0} = {} & \frac{N^2}{4} - \frac{3\sqrt{N}}{2^4 \pi^{\frac{3}{2}}} E(\tfrac{3}{2}, \tau, \bar{\tau}) + \frac{45}{2^8 \sqrt{N} \pi^{\frac{5}{2}}} E(\tfrac{5}{2}, \tau, \bar{\tau}) \\
& + \frac{1}{N^{\frac{3}{2}}} \left[-\frac{39}{2^{13} \pi^{\frac{3}{2}}} E(\tfrac{3}{2}, \tau, \bar{\tau}) + \frac{4725}{2^{15} \pi^{\frac{7}{2}}} E(\tfrac{7}{2}, \tau, \bar{\tau}) \right] \\
& + \frac{1}{N^{\frac{5}{2}}} \left[-\frac{1125}{2^{16} \pi^{\frac{5}{2}}} E(\tfrac{5}{2}, \tau, \bar{\tau}) + \frac{99225}{2^{18} \pi^{\frac{9}{2}}} E(\tfrac{9}{2}, \tau, \bar{\tau}) \right] \\
& + \frac{1}{N^{\frac{7}{2}}} \left[\frac{4599}{2^{22} \pi^{\frac{3}{2}}} E(\tfrac{3}{2}, \tau, \bar{\tau}) - \frac{2811375}{2^{25} \pi^{\frac{7}{2}}} E(\tfrac{7}{2}, \tau, \bar{\tau}) + \frac{245581875}{2^{27} \pi^{\frac{11}{2}}} E(\tfrac{11}{2}, \tau, \bar{\tau}) \right] \\
& + O(N^{-\frac{9}{2}}) \, .
\end{aligned}
\tag{4.2}
$$

The first piece of evidence for (4.1) comes from considering the terms that are perturbative in $1/\lambda = 1/(g_{\mathrm{YM}}^2 N)$, which as discussed above were computed in [34] and take the form[19]

$$
\begin{aligned}
\partial_m^2 \log Z\big|_{m=0}^{\mathrm{pert}} = {} & 2N^2 \log g_{\mathrm{YM}} + \sqrt{N} \left[-\frac{16\zeta(3)}{g_{\mathrm{YM}}^3} - \frac{g_{\mathrm{YM}}}{3} \right] + \frac{1}{\sqrt{N}} \left[\frac{12\zeta(5)}{g_{\mathrm{YM}}^5} + \frac{g_{\mathrm{YM}}^3}{1440} \right] \\
& + \frac{1}{N^{\frac{3}{2}}} \left[\frac{135\zeta(7)}{8 g_{\mathrm{YM}}^7} + \frac{g_{\mathrm{YM}}^5}{215040} - \frac{13\zeta(3)}{32 g_{\mathrm{YM}}^3} - \frac{13 g_{\mathrm{YM}}}{1536} \right] \\
& + \frac{1}{N^{\frac{5}{2}}} \left[\frac{1575\zeta(9)}{16 g_{\mathrm{YM}}^9} + \frac{g_{\mathrm{YM}}^7}{6881280} - \frac{75\zeta(5)}{64 g_{\mathrm{YM}}^5} - \frac{5 g_{\mathrm{YM}}^3}{73728} \right] \\
& + \frac{1}{N^{\frac{7}{2}}} \left[\frac{2480625\zeta(11)}{2048 g_{\mathrm{YM}}^{11}} + \frac{25 g_{\mathrm{YM}}^9}{2491416576} - \frac{80325\zeta(7)}{8192 g_{\mathrm{YM}}^7} - \frac{17 g_{\mathrm{YM}}^5}{6291456} \right. \\
& \left. \qquad + \frac{1533\zeta(3)}{16384 g_{\mathrm{YM}}^3} + \frac{511 g_{\mathrm{YM}}}{262144} \right] + O(N^{-\frac{9}{2}}) \, .
\end{aligned}
\tag{4.3}
$$

These terms match the perturbative part of the Eisenstein series in (4.1), as defined in (2.9). This match motivates the conjecture that the finite g_{YM} expression for $\partial_m^2 \log Z\big|_{m=0}$ can be derived to any order in $1/N$ by computing the perturbative terms as described in [34], and then simply replacing those by their Eisenstein completions using (2.9).

Further evidence for (4.1) comes from considering the instanton terms $\partial_m^2 \log Z\big|_{m=0}^{\mathrm{inst}}$, which are written as expectation values of sums and products of eigenvalues. In the previous sections, these quantities were computed using the saddle-point expansion, which is valid to leading order in $1/N^2$ (including the subleading in $1/N$ term). Subleading corrections in $1/N^2$ can be computed using topological recursion [79, 80]. This method naturally applies to the resolvent $W(y_1, \dots, y_n)$, which is defined as the connected expectation value

$$
W^n(y_1, \dots, y_n) \equiv N^{n-2} \left\langle \sum_{i_1} \frac{1}{y_1 - a_{i_1}} \cdots \sum_{i_n} \frac{1}{y_n - a_{i_n}} \right\rangle_{\mathrm{conn.}} ,
\tag{4.4}
$$

---

[19]The expression here includes a further order in $1/N$ and several more orders in $1/\lambda = 1/(g_{\mathrm{YM}}^2 N)$ relative to eqs. (3.1) and (3.20) of [34], which can be easily computed using the same methods explained in that work.

with the $1/N^2$ expansion

$$W^n(y_1, \ldots, y_n) \equiv \sum_{m=0}^{\infty} \frac{1}{N^{2m}} W_m^n(y_1, \ldots, y_n). \tag{4.5}$$

The coefficients W_m^n can be computed for finite λ for any n, m in a Gaussian matrix model using a recursion formula in n, m, starting with the base case W_0^1, as reviewed for the Gaussian U(N) SYM matrix model in [34]. (See Footnote 8.) Topological recursion can then be applied to any expectation value that can be written in terms of the resolvents $W^n(y_1, \ldots, y_n)$, for instance by taking derivatives or integrals in terms of y_i. Unfortunately, the operators that appear in the instanton terms (3.48) are written as products over a restricted set of eigenvalues a_i, which cannot be easily related to $W^n(y_1, \ldots, y_n)$. However, by expanding these products for small a_i, which is equivalent to a small g_{YM} expansion, they can be expressed as an infinite sum of polynomials in a_i, whose expectation values can then be easily related to $W^n(y_1, \ldots, y_n)$.

Let us begin by discussing the one-instanton case. By explicitly performing the sums and products in I_1 (eq. (3.8)) for many small values of N, we find that I_1 can be expanded for small a_i as

$$I_1 = -\frac{4\Gamma(N+\frac{1}{2})}{\sqrt{\pi}\Gamma(N)} - \frac{3\Gamma(N-\frac{1}{2})}{2\sqrt{\pi}\Gamma(N+2)}C_2 + \frac{315\Gamma(N-\frac{3}{2})}{32\sqrt{\pi}\Gamma(N+4)}C_2^2 - \frac{15(3-N+4N^2)\Gamma(N-\frac{3}{2})}{16\sqrt{\pi}\Gamma(N+4)}C_4 + \cdots \tag{4.6}$$

where we defined the invariants

$$C_p = \sum_{j,k}(a_j - a_k)^p. \tag{4.7}$$

The expectation values of these C_p can be related to coefficients of the large y_i expansion of n-body resolvents $W^n(y_1, \ldots, y_n)$ with $n \leq p$. Since the C_p are degree p polynomials in a_i, they must be proportional to $\lambda^{p/2}$ and their $1/N^2$ expansion truncates. For instance, for the C_p that are shown in (4.6), using the explicit expressions for W_m^n in appendix B of [34] we find

$$\langle C_2 \rangle = \lambda \left[\frac{N^2}{8\pi^2} - \frac{1}{8\pi^2} \right], \quad \langle C_2^2 \rangle = \lambda^2 \left[\frac{N^4}{64\pi^4} - \frac{1}{64\pi^4} \right], \quad \langle C_4 \rangle = 5\lambda^2 \left[\frac{N^2}{128\pi^4} - \frac{1}{128\pi^4} \right]. \tag{4.8}$$

We can then insert these expressions into the expectation value (4.6), set $\lambda = g_{\text{YM}}^2 N$, and expand in $1/N$ to get

$$\begin{aligned}
\langle I_1 \rangle = \sqrt{N} &\left[-\frac{4}{\sqrt{\pi}} - \frac{3g_{\text{YM}}^2}{16\pi^{5/2}} + \frac{15g_{\text{YM}}^4}{2048\pi^{9/2}} + O\left(g_{\text{YM}}^6\right) \right] \\
&+ \frac{1}{\sqrt{N}} \left[\frac{1}{2\sqrt{\pi}} + \frac{15g_{\text{YM}}^2}{128\pi^{5/2}} + \frac{105g_{\text{YM}}^4}{16384\pi^{9/2}} + O\left(g_{\text{YM}}^6\right) \right] \\
&+ \frac{1}{N^{\frac{3}{2}}} \left[-\frac{1}{32\sqrt{\pi}} + \frac{69g_{\text{YM}}^2}{2048\pi^{5/2}} + \frac{2175g_{\text{YM}}^4}{262144\pi^{9/2}} + O\left(g_{\text{YM}}^6\right) \right] \\
&+ \frac{1}{N^{\frac{5}{2}}} \left[-\frac{5}{256\sqrt{\pi}} + \frac{285g_{\text{YM}}^2}{16384\pi^{5/2}} + \frac{24675g_{\text{YM}}^4}{2097152\pi^{9/2}} + O\left(g_{\text{YM}}^6\right) \right] \\
&+ \frac{1}{N^{\frac{7}{2}}} \left[\frac{21}{8192\sqrt{\pi}} + \frac{5103g_{\text{YM}}^2}{524288\pi^{5/2}} + \frac{1158885g_{\text{YM}}^4}{67108864\pi^{9/2}} + O\left(g_{\text{YM}}^6\right) \right] + O(N^{-\frac{9}{2}}).
\end{aligned} \tag{4.9}$$

This is consistent according to (3.8) with the small g_{YM} expansion of

$$
\begin{aligned}
\left\langle \partial_m^2 Z_{\mathrm{inst}}^{(1)}(m, a_{ij}) \right\rangle \Big|_{m=0} = e^{\frac{8\pi^2}{g_{\mathrm{YM}}^2}} & \left[-\sqrt{N} \frac{16 K_1(8\pi^2/g_{\mathrm{YM}}^2)}{g_{\mathrm{YM}}} + \frac{2 K_2(8\pi^2/g_{\mathrm{YM}}^2)}{\sqrt{N} g_{\mathrm{YM}}} \right. \\
& + \frac{1}{32 g_{\mathrm{YM}} N^{\frac{3}{2}}} \left[-13 K_1(8\pi^2/g_{\mathrm{YM}}^2) + 9 K_3(8\pi^2/g_{\mathrm{YM}}^2) \right] \\
& + \frac{1}{128 g_{\mathrm{YM}} N^{\frac{5}{2}}} \left[-25 K_2(8\pi^2/g_{\mathrm{YM}}^2) + 15 K_4(8\pi^2/g_{\mathrm{YM}}^2) \right] \\
& + \frac{1}{g_{\mathrm{YM}} N^{\frac{7}{2}}} \left[\frac{1533 K_1\left(\frac{8\pi^2}{g_{\mathrm{YM}}^2}\right)}{16384} - \frac{5355 K_3\left(\frac{8\pi^2}{g_{\mathrm{YM}}^2}\right)}{32768} + \frac{2625 K_5\left(\frac{8\pi^2}{g_{\mathrm{YM}}^2}\right)}{32768} \right] + O(N^{-\frac{9}{2}}) \right],
\end{aligned}
\tag{4.10}
$$

which generalizes (3.24) to higher orders in $1/N$, and describes the one-instanton contribution to the Eisenstein series in (4.3). In appendix D, we similarly match all instantons up to $k = 12$ to order $O(g_{\mathrm{YM}}^4)$.

We should end by pointing out that the $\mathrm{SL}(2, \mathbb{Z})$-invariant expression (4.2) has the property that the coefficients multiplying $\pi^{-r} E(r, \tau, \bar{\tau})$ are all rational numbers. It would be interesting to understand the significance of this fact in relation to the set of modular functions that appear in the expression for the $\mathcal{N} = 4$ SYM correlator and superstring scattering amplitudes.

5 Conclusion

In this paper, we studied the four-point correlator $\langle SSSS \rangle$ of the superconformal primary operator S transforming in the $\mathbf{20'}$ of $\mathrm{SO}(6)_R$ in the $\mathcal{N} = 4$ SYM theory, in the "very strong coupling" limit in which N is sent to infinity at fixed g_{YM}. In this limit, the action of $\mathrm{SL}(2, \mathbb{Z})$ modular transformations on the $\langle SSSS \rangle$ correlator is manifest. In particular, we studied the constraints on $\langle SSSS \rangle$ coming from the flat space limit of the IIB string theory amplitudes, and those coming from the integrated four-point function $\tau_2^2 \partial_\tau \partial_{\bar{\tau}} \partial_m^2 \log Z \big|_{m=0}$. The latter can be computed using supersymmetric localization. Starting from Pestun's localization expression [37] for the partition function Z, we argued that when $\tau_2^2 \partial_\tau \partial_{\bar{\tau}} \partial_m^2 \log Z \big|_{m=0}$ is expanded in $1/N$, the first two sub-leading terms (of orders $N^{\frac{1}{2}}$ and $N^{-\frac{1}{2}}$, respectively) can be written as the Eisenstein series $E(\frac{3}{2}, \tau, \bar{\tau})$ and $E(\frac{5}{2}, \tau, \bar{\tau})$, respectively. Our argument is not completely rigorous because it relies on studying the k-instanton contribution for many values of k and deducing the general pattern, but we hope that it should be possible to provide a more rigorous argument in future work. Using solely the relation between the integrated $\langle SSSS \rangle$ correlator and $\tau_2^2 \partial_\tau \partial_{\bar{\tau}} \partial_m^2 \log Z \big|_{m=0}$ from [33], we completely determined the $N^{\frac{1}{2}}$ term in the large N expansion of $\langle SSSS \rangle$. This term corresponds to an effective $\ell_s^{-2} R^4$ coupling in AdS_5, which, in the flat space limit, matches the $\ell_s^{-2} R^4$ contribution to the Type IIB graviton S-matrix as computed at finite string coupling g_s in [53–55]. This is a precision test of AdS/CFT at finite g_s! We then used the $\ell_s^2 D^4 R^4$ term in the Type IIB S-matrix, which is also known at finite g_s, as well as the $N^{-\frac{1}{2}}$ term in $\tau_2^2 \partial_\tau \partial_{\bar{\tau}} \partial_m^2 \log Z \big|_{m=0}$ to completely determine $\langle SSSS \rangle$ at order $N^{-\frac{1}{2}}$. In Mellin

space, this expression contains two polynomial terms, both proportional to $E(\frac{5}{2}, \tau, \bar{\tau})$, one corresponding to an $\ell_s^2 D^4 R^4$ contact term in AdS_5 and one to an $\frac{\ell_s^2}{L^4} R^4$ term. Finally, using a small g_{YM} expansion, we gave non-trivial evidence that each of the terms in the $1/N$ expansion of $\tau_2^2 \partial_\tau \partial_{\bar{\tau}} \partial_m^2 \log Z\big|_{m=0}$ is a finite linear combination of non-holomorphic Eisenstein series.

The fact that we can derive the full CFT correlator at order $N^{\frac{1}{2}}$ generalizes the analysis in [48, 81] of the k-instanton measure in the large-N limit of SU(N) $\mathcal{N} = 4$ SYM theory at lowest order in g_{YM}. That analysis, which was based on a large-N saddle point solution of the ADHM constraints, demonstrated a number of general features of the holographic relationship between Yang-Mills instantons and D-instantons. In particular, in the large-N saddle point approximation the k-instanton moduli space (where $k \ll N$) is dominated by the region in which the positions of the instantons and their scale sizes coincide, so they are represented by a single point in $AdS_5 \times S^5$, with all the instantons in commuting SU(2) subgroups of SU(N). This is interpreted holographically as a single D-instanton of charge k in the dual type IIB theory. Furthermore, the divisor sum $\sigma_{-2}(|k|)$ arises as the partition function of the D-instanton matrix model, i.e. the partition function of 10d $\mathcal{N} = 1$ SU(k) Yang-Mills theory reduced to zero space-time dimensions.

The result of this large-N ADHM analysis is reproduced in our procedure by the first term in the small-g_{YM} expansion of the Bessel function in the kth Fourier mode of the Eisenstein series $E(\frac{3}{2}, \tau, \bar{\tau})$ (the function $\mathcal{F}_k$ defined in (A.7)). The fact that the dominant contribution to the Nekrasov partition function in the $m \to 0$ limit has a single cluster of boxes should correspond to properties of the large-N analysis of the ADHM construction. However, this correspondence is difficult to make precise since our analysis is based on taking a limit of the non-conformal $\mathcal{N} = 2^*$ whereas conformal invariance is explicit in the large-N ADHM construction. The connection of the D-instanton measure with the SU(k) D-instanton matrix model partition function is also not obvious in our procedure. Nevertheless, the fact that our procedure packages an infinite number of perturbative corrections to the k-instanton contribution into a K-Bessel function is a most significant generalization of [48, 81] and an essential requirement of SL(2, $\mathbb{Z}$) covariance.

As shown in section 4 our integrated constraint $\tau_2^2 \partial_\tau \partial_{\bar{\tau}} \partial_m^2 \log Z\big|_{m=0}$ has an expansion in half-integer powers of $1/N$ (apart from the first term). However, it is well known that the low energy expansion of the string amplitude does also contain even powers of ℓ_s^2, which lead to integer powers of $1/N$, the most relevant one being the 1/8-BPS interaction $D^6 R^4$. A more complete analysis of the holographic correspondence should therefore also include terms with integer powers of $1/N$. We expect that such terms will appear in other quantities that can be computed using supersymmetric localization, such as $\partial_m^4 \log Z\big|_{m=0}$ or $\partial_b^2 \partial_m^2 \log Z\big|_{m=0,b=1}$, where $b = \epsilon_1/\epsilon_2$ is a parameter that defines the squashing deformation of S^4 that appear in (3.26) (recall that up to now $\epsilon_1 = \epsilon_2 = 1$). We expect that determining these three distinct integrated four-point correlation functions should eliminate the ambiguities in determining the expansion of the $AdS_5 \times S^5$ type IIB string theory amplitudes up to order $D^6 R^4$. In other words, this procedure should uniquely determine the BPS protected interactions without the need to input known results from flat-space type IIB superstring theory.

We have so far only considered four-point correlators. It was argued in [63, 64] that all four-point functions of short operators are invariant under both the bonus U(1) symmetry as well as SL(2, $\mathbb{Z}$). As we have seen, the four-point integrated correlators are linear combinations of non-holomorphic Eisenstein series, which are manifestly SL(2, $\mathbb{Z}$) invariant. Such a statement would not hold for ($n \geq 5$)-point functions, which transform as modular forms with non-trivial modular weights.[20] These correlation functions should correspond to higher-point superstring amplitudes that violate the U(1) R-symmetry of type IIB supergravity. Such U(1)-violating superstring amplitudes (especially those that violate U(1) maximally) are identified in [60, 82], and more importantly the F-terms (terms up to the same number of derivatives as $D^6 R^4$) have also been determined using maximal supersymmetry and SL(2, $\mathbb{Z}$) symmetry [60]. The coefficients of these U(1)-violating interactions are modular forms with non-zero modular weights, and it would be of interest to understand how they arise from the supersymmetric localization computation.

Lastly, it is interesting to compare the calculation presented in this paper to calculations done in the 3d ABJM theory [62] with gauge group U(N)$_k$ × U(N)$_{-k}$ and $\mathcal{N} = 6$ supersymmetry. An analogous computation that includes non-perturbative contributions can also be performed in that case [32], in the very strong coupling limit in which N is taken to infinity while N/k^5 is held fixed. In this limit, the ABJM theory is dual to type IIA string theory on $AdS_4 \times \mathbb{CP}^3$ at finite string coupling g_s. However, in this case, all the non-perturbative contributions to the type IIA scattering amplitudes of the lowest closed string states vanish. Thus, in order to obtain non-trivial non-perturbative contributions to CFT correlators that can be matched to string scattering amplitudes, one is led to consider the case of the 4d $\mathcal{N} = 4$ theory that was studied in the present paper. Nevertheless, it is worth pointing out that the closest 3d analog of the formulas presented in section 4 that include resummed instanton contributions would be the mass-deformed partition function of ABJM theory that can be computed [83] to all orders in the $1/N$ expansion using the Fermi gas method developed in [84].

Acknowledgments

---

[20]A modular form with holomorphic and anti-holomorphic weights (w, w') transforms as $f^{(w,w')}(\tau', \bar{\tau}') \to (c\tau + d)^w (c\bar{\tau} + d)^w f^{(w,w')}(\tau, \bar{\tau})$ under the action of SL(2, $\mathbb{Z}$).

tality during the initial stages of this work. The ACP is supported by National Science Foundation Grant No. PHY-1607611.

A Non-holomorphic Eisenstein series

The first four terms in the low energy expansion of the four-graviton amplitude in the type IIB superstring theory correspond to BPS protected effective interactions that take the form

$$
\mathcal{A}(\mathbf{s},\mathbf{t}) = \frac{R^4}{\ell_s^8 g_s^2} \left(\frac{1}{\mathbf{stu}} + \frac{\ell_s^6 g_s^{\frac{3}{2}}}{2^6} E(\tfrac{3}{2},\tau,\bar\tau) + \frac{\ell_s^{10} g_s^{\frac{5}{2}}}{2^{11}} (\mathbf{s}^2 + \mathbf{t}^2 + \mathbf{u}^2) E(\tfrac{5}{2},\tau,\bar\tau) \right.
$$
$$
\left. + \frac{\ell_s^{12} g_s^3}{2^{12}} (\mathbf{s}^3 + \mathbf{t}^3 + \mathbf{u}^3) \mathcal{E}(\tau,\bar\tau) + \dots \right),
\tag{A.1}
$$

where R signifies the linearised Weyl curvature tensor, which has the form $R_{\mu\nu\rho\sigma} = k_{[\mu}\epsilon_{\nu]}k_{[\rho}\epsilon_{\sigma]}$ (where $[\cdots]$ denotes anti-symmetrization of the indices), where k_μ is a null ten-dimensional momentum and $\epsilon_\nu\epsilon_\sigma$ is a graviton polarization. The symbol R^4 denotes the particular contraction of four curvature tensors that is implied by ten-dimensional $\mathcal{N} = 2$ supersymmetry.

The function $E(r,\tau,\bar\tau)$ is a non-holomorphic Eisenstein series, which is a modular function of $(\tau,\bar\tau)$ that satisfies the Laplace eigenvalue equation

$$
(\Delta_\tau - r(r-1)) E(r,\tau,\bar\tau) = 0,
\tag{A.2}
$$

where the hyperbolic laplacian is defined by $\Delta_\tau = 4\tau_2^2 \partial_\tau \partial_{\bar\tau} = \tau_2^2 (\partial_{\tau_1}^2 + \partial_{\tau_2}^2)$. Assuming moderate growth as $\tau_2 \to \infty$ (in other words, assuming $E(r,\tau,\bar\tau)$ grows no faster than τ_2^a for any a, so it is no more singular than perturbative string theory in the $g_s \to 0$ limit) the unique SL$(2,\mathbb{Z})$-invariant solution to this equation is[21]

$$
E(r,\tau,\bar\tau) = \sum_{(m,n)\neq(0,0)} \frac{\tau_2^r}{|m\tau+n|^{2r}}.
\tag{A.3}
$$

It is straightforward to show that $E(r,\tau,\bar\tau)$ is invariant under a SL$(2,\mathbb{Z})$ transformation,

$$
E(r,\tau,\bar\tau) \to E(r,\tau',\bar\tau'), \quad \text{with} \quad \tau \to \tau' = \frac{a\tau+b}{c\tau+d},
\tag{A.4}
$$

for $a,b,c,d \in \mathbb{Z}$ and $ad - bc = 1$.

---

[21]It is often convenient to define the non-holomorphic Eisenstein series in terms of a Poincaré series $\hat{E}(r,\tau,\bar\tau) = \frac{1}{2}\pi^{-r}\Gamma(r)E(r,\tau,\bar\tau)$, where

$$
\hat{E}(r,\tau,\bar\tau) = \pi^{-r}\Gamma(r)\zeta(2r)\frac{1}{2}\sum_{\gamma\in\Gamma_\infty\backslash\mathrm{SL}(2,\mathbb{Z})} \Im(\gamma(\tau))^r,
$$

and $\Gamma_\infty = \left\{ \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \,\middle|\, n\in\mathbb{Z} \right\}$. The expression $\hat{E}(r,\tau,\bar\tau)$ is manifestly invariant under SL$(2,\mathbb{Z})$ and satisfies the important functional relation $\hat{E}(r,\tau,\bar\tau) = \hat{E}(1-r,\tau,\bar\tau)$.

A non-holomorphic Eisenstein series has an expansion in Fourier modes of the form

$$E(r, \tau, \bar{\tau}) = \sum_{k \in \mathbb{Z}} \mathcal{F}_k(r, \tau_2) \, e^{2\pi i k \tau_1} \,, \tag{A.5}$$

where the zero mode consists of two power behaved terms,

$$\mathcal{F}_0(r, \tau_2) = 2\zeta(2r) \, \tau_2^r + \frac{2\sqrt{\pi} \, \Gamma(r - \frac{1}{2}) \zeta(2r - 1)}{\Gamma(r)} \, \tau_2^{1-r} \,, \tag{A.6}$$

and the non-zero modes are proportional to K-Bessel functions,

$$\mathcal{F}_k(r, \tau_2) = \frac{2 \, \pi^r}{\Gamma(r)} \, |k|^{r-\frac{1}{2}} \, \sigma_{1-2r}(|k|) \sqrt{\tau_2} \, K_{r-\frac{1}{2}}(2\pi |k| \tau_2) \,, \qquad k \neq 0 \,, \tag{A.7}$$

where the divisor sum is defined by

$$\sigma_p(k) = \sum_{d > 0, d|k} d^p \tag{A.8}$$

for $k > 0$, and $\sigma_{-p}(k) = k^{-p} \, \sigma_p(k)$.

The two power-behaved terms in $\mathcal{F}_0(r, \tau_2)$ in (A.6) correspond to tree-level and $(r - \frac{1}{2})$-loop contributions in string perturbation theory. Using the asymptotic behavior of the K-Bessel function

$$K_\nu(z) \sim \sqrt{\frac{\pi}{2z}} \, e^{-z} \left(1 + O\left(\frac{1}{z}\right) \right) \,, \tag{A.9}$$

we see that the non-zero mode $\mathcal{F}_k(r, \tau_2)$ behaves as $e^{-2\pi |k| \tau_2}$ and has the form of a k D-instanton contribution.

The terms proportional to $E(\frac{3}{2}, \tau, \bar{\tau})$ and $E(\frac{5}{2}, \tau, \bar{\tau})$ in (A.1) are coefficients of R^4 and $D^4 R^4$ interactions in the type IIB low energy effective action. These are, respectively, 1/2-BPS and 1/4-BPS interactions. The last term in (A.1) corresponds to a term proportional to the 1/8-BPS interaction $D^6 R^4$, with a coefficient $\mathcal{E}(\tau, \bar{\tau})$ that satisfies the inhomogeneous Laplace eigenvalue equation [57, 58]

$$(\Delta_\tau - 12) \, \mathcal{E}(\tau, \bar{\tau}) = -E^2(\tfrac{3}{2}, \tau, \bar{\tau}) \,. \tag{A.10}$$

The solution to this equation [85] is qualitatively different from an Eisenstein series. The zero mode is of the form in the large-τ_2 limit,

$$\mathcal{F}_{0, \mathcal{E}}(\tau_2) := \int_{-\frac{1}{2}}^{\frac{1}{2}} d\tau_1 \mathcal{E}(\tau, \bar{\tau}) = \frac{2 \, \zeta(3)^2}{3} \, \tau_2^3 + \frac{4 \, \zeta(2) \, \zeta(3)}{3} \, \tau_2 + \frac{4 \, \zeta(4)}{\tau_2} + O(e^{-4\pi \tau_2}) \,. \tag{A.11}$$

The power-behaved contributions correspond to string perturbation theory up to genus three. The symbol $O(e^{-4\pi |k| \tau_2})$ denotes a specific infinite sum of D-instanton-anti D-instanton contributions with zero total instanton number (details of which are in [85]). Similarly, each mode of non-zero mode number k has the form of a sum of D-instanton-anti D-instanton contributions with instanton numbers k_1 and k_2 satisfying $k_1 + k_2 = k \neq 0$. We see in the main text that such a $(\mathbf{s}^3 + \mathbf{t}^3 + \mathbf{u}^3) \, R^4$ contribution does not arise from our

analysis of the flat-space limit of $\tau_2^2 \partial_\tau \partial_{\bar{\tau}} \partial_m^2 \log Z$ since its contribution to the integrated correlation function vanishes.

The four terms in the low energy expansion of the four-graviton amplitude explicitly shown in (A.1) correspond to local BPS interactions that are fully determined by supersymmetry, while higher derivative terms are not expected to be protected and have not been fully determined.

B Rectangular dominance

As discussed in the main text, the full Nekrasov partition function for the mass-deformed $\mathcal{N} = 2^*$ SU(N) SYM with squashing parameters $\epsilon_{1,2}$ at instanton number k is given by a sum over N-tuples of Young diagrams $\vec{Y} = (Y_1, Y_2, \ldots, Y_N)$ with k boxes in total,

$$
\begin{aligned}
Z_{\text{inst}}^{(k)}(m, a_{ij}, \epsilon_{1,2}) &= \sum_{|\vec{Y}|=k} Z_{\vec{Y}} \\
Z_{\vec{Y}} &\equiv \frac{\prod_{i,j=1}^{N} \prod_{s \in Y_i} (E(a_{ij}, Y_i, Y_j, s) - im - \epsilon_+/2) \prod_{t \in Y_j} (-E(a_{ji}, Y_j, Y_i, t) - im + \epsilon_+/2)}{\prod_{s \in Y_i} E(a_{ij}, Y_i, Y_j, s) \prod_{t \in Y_j} (\epsilon_+ - E(a_{ji}, Y_j, Y_i, t))},
\end{aligned}
\tag{B.1}
$$

where $\epsilon_\pm \equiv \epsilon_1 \pm \epsilon_2$ as in the main text, and

$$
E(a_{ij}, Y_i, Y_j, s) \equiv i a_{ji} - \epsilon_1 h_j(s) + \epsilon_2 (v_i(s) + 1). \tag{B.2}
$$

Here, s labels a box (α, β) (α-th column and β-th row) in a given Young diagram as figure 1, and $h_i(s)$ and $v_i(s)$ denote the arm-length and leg-length, respectively, of the box s in the diagram Y_i. (Each individual Young diagram Y consists of columns of non-increasing heights $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l$ with $\lambda_l \geq 1$. The transpose (conjugate) diagram Y^T has columns of heights $\lambda_1^T \geq \lambda_2^T \geq \cdots \geq \lambda_m^T$ with $\lambda_m^T \geq 1$. Then the arm-length h and leg-length v of the box s in Y are given by (see figure 1)

$$
h(s) = \lambda_\beta^T - \alpha, \qquad v(s) = \lambda_\alpha - \beta. \tag{B.3}
$$

Note that the definitions of h and v extend beyond boxes in Y to the entire quadrant $(\alpha, \beta) \in \mathbb{Z}_+^2$ in the obvious way. In particular they can be negative (e.g. when Y is empty).)

In the rest of this appendix, we prove the following theorem:

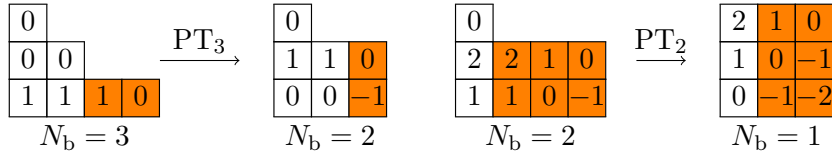**Theorem B.1.** *For the second mass derivative of the instanton partition function at $m = 0$*

$$
I_k(a_{ij}) \equiv \partial_m^2 \left(\lim_{\epsilon_{1,2} \to 1} Z_{\text{inst}}^{(k)}(m, a_{ij}, \epsilon_{1,2}) \right) \bigg|_{m=0} \tag{B.4}
$$

only the summands in (B.1) *with $\vec{Y}$ of the following type*

$$
\begin{aligned}
Y_i &= \varnothing \text{ if } i \neq \hat{i} \\
Y_{\hat{i}} &= \underbrace{[p, p, \ldots, p]}_{q} \text{ with } pq = k \text{ for } p, q \in \mathbb{Z}_+,
\end{aligned}
\tag{B.5}
$$

and those with $Y_{\hat{i}}$ replaced by its partial transpositions (which we define next) contribute.

Figure 2. Example of partial transpositions and the corresponding changes in $N_{\rm b}(Y)$. The subdiagram P is colored in orange. The numbers in the individual boxes s is the value $d(s)$.

Given a Young diagram $Y = [\lambda_1, \lambda_2, \ldots, \lambda_l]$ (with $\lambda_l \geq 1$), we define its *partial transposition* at position α, with $1 \leq \alpha \leq l$, to be

$$\mathrm{PT}_\alpha(Y) = [\lambda_1, \lambda_2, \ldots, \lambda_{\alpha-1}, \lambda'_1, \ldots, \lambda'_{\lambda_\alpha}]\,, \tag{B.6}$$

where λ'_β are column lengths of the transpose of the subdiagram $P = [\lambda_\alpha, \lambda_{\alpha+1}, \ldots, \lambda_l]$, if $\lambda_{\alpha-1} \geq \lambda'_1$. Equivalently the partial transposition at α-th column is given by the disjoint union

$$\mathrm{PT}_\alpha(Y) = (Y \backslash P) \sqcup P^T\,, \tag{B.7}$$

where P denotes the block (Young subdiagram) to the right of the $(\alpha - 1)$-st column, and $Y \backslash P$ is the complement Young subdiagram (see examples in figure 2). In particular, the usual transposition is a partial transposition at $\alpha = 1$. For notational simplicity, we will often suppress the subscript α when the context is clear.

We will need the following useful properties of the map $\mathrm{PT}_\alpha(\cdot)$. We start with the obvious lemma:

Lemma B.2. *For a Young diagram Y, the multiset $\Delta(Y) = \{\alpha + \beta \,|\, (\alpha, \beta) \in Y\}$ is invariant under partial transpositions.*

Consequently the partial transpositions preserve the set of poles (3.29) in the contour integral for the instanton partition function. A useful corollary that follows is the following:

Corollary B.2.1. *Among all the Young diagrams related by (a sequence of) partial transpositions to a rectangular diagram $Y_{p \times q}$ with p columns of height q each, the maximum width (or height) is $p + q - 1$.*

Proof. A single partial transposition at the right-most column gives a Young diagram with width $p + q - 1$. The fact that this is the maximal value that can be achieved by any sequence of partial transpositions follows from the previous lemma by noting that

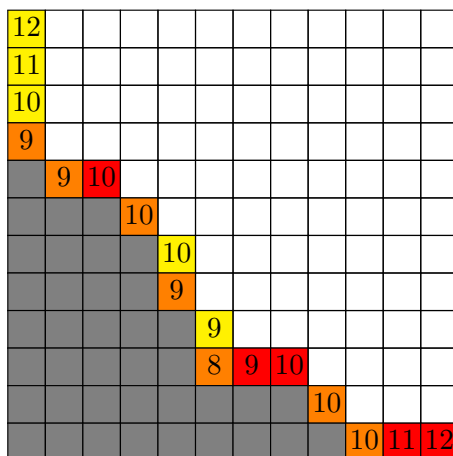$$\max[\Delta(Y_{p \times q})] = p + q. \tag{B.8}$$

$\square$

Lemma B.3. *For Young diagram $Y = [\lambda_1, \ldots, \lambda_l]$ with $\lambda_l \geq 1$, the multiset*

$$\Delta_B(Y) \equiv \{\alpha + \lambda_\alpha \,|\, 1 \leq \alpha \leq M\} \tag{B.9}$$

is invariant under partial transpositions, where M is any integer equal to or larger than the maximum width of Young diagrams related to Y by partial transpositions. Note that λ_α for $\alpha > l$ is defined to be 0.

Figure 3. An example with $Y = [8,7,7,6,4,2,2,2,1]$ (gray boxes) in a square of size $M = 12$. The boxes at $(\lambda_\alpha^T + 1, \alpha)$ are colored in yellow; the boxes at $(\alpha, \lambda_\alpha + 1)$ are colored in red; the boxes common to both are colored in orange. The numbers in the yellow and orange boxes are part of $\Delta_B(Y^T)$ while those in the red and orange boxes are part of $\Delta_B(Y)$. Thus, in this example, $\Delta_B(Y) = \{9,9,10,10,9,8,9,10,10,10,11,12\}$ and $\Delta_B(Y^T) = \{10,10,8,9,9,10,10,9,9,10,11,12\}$.

Proof. It suffices to prove that $\Delta_B(Y) = \Delta_B(Y^T)$, because for any partial transposition with respect to a subdiagram $P \subset Y$, $\Delta_B(Y \backslash P)$ is clearly invariant. The entries in $\Delta_B(Y)$ consists of $\alpha + \beta$ for boxes (α, β) located on the North-East boundary of the Young diagram.

We proceed by induction. If Y consists of a single box, $\Delta_B(Y) = \Delta_B(Y^T)$ trivially for any $M \geq 1$. Suppose $\Delta_B(Y) = \Delta_B(Y^T)$ for some large enough M as specified in the statement of the lemma. We show it continues to hold upon adding a box to Y. First of all, in order to generate a Young diagram, the additional box must be added to an *inward* corner of the North-East border of Y, i.e. a coordinate (α, β) that satisfies

$$(\alpha, \beta) = (\alpha, \lambda_\alpha + 1) = (\lambda_\beta^T + 1, \beta). \tag{B.10}$$

We denote the new Young diagram obtained in this fashion by Y'. Then $\Delta_B(Y')$ is given by $\Delta_B(Y)$ with the element $\alpha + \lambda_\alpha$ (an orange entry in figure 3) replaced by $\alpha + \lambda_\alpha + 1$. Similarly $\Delta_B(Y'^T)$ is given by $\Delta_B(Y^T) = \Delta_B(Y)$ with the element $\lambda_\beta^T + \beta$ replaced by $\lambda_\beta^T + \beta + 1$. Due to (B.10), we have $\alpha + \lambda_\alpha = \lambda_\beta^T + \beta = \alpha + \beta - 1$, as well as $\alpha + \lambda_\alpha + 1 = \lambda_\beta^T + \beta + 1 = \alpha + \beta$. Thus, $\Delta_B(Y')$ and $\Delta_B(Y'^T)$ are each obtained from $\Delta_B(Y)$ and $\Delta_B(Y^T)$, respectively, by replacing an occurrence of $\alpha + \beta - 1$ with $\alpha + \beta$. Since by assumption $\Delta_B(Y) = \Delta_B(Y^T)$, we conclude that $\Delta_B(Y') = \Delta_B(Y'^T)$. See figure 3 for an example. □

For a box with coordinates $s = (\alpha, \beta)$ in a Young diagram Y, we define $d(s) = h(s) - v(s)$. Then we also define $n_a(Y)$ to be number of coordinates s with $d(s) = a$.

Lemma B.4. *For a Young diagram Y,*

$$\mu(Y) = 2n_0(Y) - (n_1(Y) + n_{-1}(Y)) \tag{B.11}$$

is invariant under partial transpositions. In fact both $n_0(Y)$ and $n_1(Y) + n_{-1}(Y)$ are separately invariant.

Proof. Under a partial transposition that involves transposing the subdiagram P of Y, the box at $(\alpha, \beta) \in P$ gets mapped to the box at $(\beta, \alpha) \in P^T$. Thus, focusing on boxes in P, the boxes with $d(s) = 0$ are preserved, whereas the boxes with $d(s) = \pm 1$ are mapped to boxes with $d(s) = \mp 1$ — See figure 2. Consequently, the contribution to $\mu(Y)$ from P does not change under the transposition. Under the partial transposition, the values of $d(s)$ for the boxes in $s \in Y \backslash P$ get permuted within each column. To see this let's consider for example a partial transposition $\text{PT}_{\hat{\alpha}}$ with respect to the sub-diagram $P = [\rho_1, \rho_2, \dots]$ (namely $\rho_a = \lambda_{\hat{\alpha}+a-1}$) with $\lambda_{\hat{\alpha}-1} \geq \rho_1^T$ (so that it's a nontrivial operation). Focusing on the boxes in the α-th column with $1 \leq \alpha \leq \hat{\alpha} - 1$, their $d = h - v$ values before the partial transposition are given by

$$\{\rho_\beta^T + \beta + (\hat{\alpha} - \lambda_\alpha - \alpha) | 1 \leq \beta \leq \lambda_\alpha\} \tag{B.12}$$

whereas after the partial transposition, we have

$$\{\rho_\beta + \beta + (\hat{\alpha} - \lambda_\alpha - \alpha) | 1 \leq \beta \leq \lambda_\alpha\}. \tag{B.13}$$

From Lemma B.3, we see the above two multisets are the same. Hence the contribution to n_d coming from $Y \backslash P$ does not change under the transposition of P (see examples in figure 2). Therefore we conclude $\mu(Y)$ is a partial transposition invariant, and clearly so is $n_0(Y)$. $\qquad\square$

In particular, a rectangular Young diagram $Y_{p \times q}$ with p columns and q rows has

$$n_0(Y) = \min(p, q), \qquad n_1(Y) = \min(p-1, q), \qquad n_{-1}(Y) = \min(p, q-1), \tag{B.14}$$

which implies that

$$\mu(Y_{p \times q}) = \begin{cases} 2 & \text{if } p = q, \\ 1 & \text{if } p \neq q. \end{cases} \tag{B.15}$$

To proceed, let us make two more definitions. We define $N_{\text{b}}(Y) \geq 1$ to be the minimal number of rectangular blocks (in a horizontal decomposition) in a Young diagram Y (see figure 2 for examples). Note that while $N_{\text{b}}(Y)$ is invariant under (full) transposition, it typically changes under partial transpositions (see figure 2 for example). We give the set of all Young diagram Y with given size $|Y| = k$ a lexicographic (total) ordering as follows. For two Young diagrams $Y = [\lambda_1, \lambda_2, \dots, \lambda_l]$ and $Y' = [\rho_1, \rho_2, \dots, \rho_m]$, we have $Y \geq Y'$ when $\lambda_1 > \rho_1$ or when $\lambda_1 = \rho_1$ and $\lambda_2 > \rho_2$ or when $\lambda_1 = \rho_1$, $\lambda_2 = \rho_2$, and $\lambda_3 > \rho_3$, etc., or when $\lambda_i = \rho_i$ for all i.

Given a Young diagram Y, we consider all possible partial transpositions, and among the ones that minimize $N_{\text{b}}(Y)$, we take the smallest Young diagram according to the lexicographic ordering, and call it $Y_{\min}$. For example if Y is related by (possibly several successive) partial transpositions to a rectangular Young diagram $Y_{p \times q}$ for some $p \leq q$, then clearly

$$Y_{\min} = Y_{p \times q}, \tag{B.16}$$

because $N_{\text{b}}(Y_{p \times q}) = 1$.
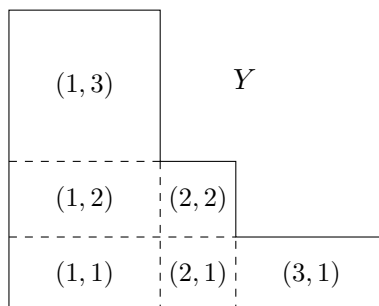
Figure 4. Decomposition of the Young diagram Y into rectangular blocks by perpendicular lines from the $N_b = 3$ corners. The individual blocks are labelled as (i,j) with $i + j \leq N_b + 1 = 4$ here.

Lemma B.5. *If $Y_{\min}$ has $N_b(Y_{\min}) = c$ rectangular blocks, then we must have*

$$\mu(Y) \geq c \,. \tag{B.17}$$

In particular $\mu(Y) = 1$ if and only if Y is related by partial transpositions to a diagram of the type $Y_{p \times q}$, for some p and q.

Proof. Since μ is invariant under partial transposition, we can take $Y = Y_{\min}$. By assumption, the Young diagram Y has a minimum of c rectangular blocks in a horizontal decomposition. Correspondingly, Y has c outward-pointing corners along the North-East boundary. We obtain a finer decomposition of Y into $1 + 2 + \cdots + c = \frac{c(c+1)}{2}$ smaller rectangular blocks by drawing perpendicular lines from the c corners in an obvious fashion (see figure 4). Each of the c blocks at the corners contributes 2 or 1 to μ depending on whether it has equal or non-equal sides. Therefore, it suffices to show that the $\frac{c(c-1)}{2}$ interior blocks each contribute non-negatively to $\mu(Y)$.

We label these rectangular blocks as (i,j) for $1 \leq i, j \leq c$ and $i + j \leq c + 1$, and let their sizes be $p_j \times q_i$. The corner blocks are given by those with $i + j = c + 1$. Each interior rectangular block labeled by (i,j) contributes at least -1 to $\mu(Y)$, and this happens precisely when either $p_j + p_{j+1} = q_{i+1}$, or $q_i + q_{i+1} = p_{j+1}$ (see figure 5). Let's assume this is the case for the interior rectangular block (i,j). But this means N_b can be reduced by 1 after a transposition of the subdiagram involving the rectangular blocks (k,l) with $k > i, l \geq j$ if $p_j + p_{j+1} = q_{i+1}$, and similarly from transposing the subdiagram involving the rectangular blocks (k,l) with $k \geq i, l > j$ if $q_i + q_{i+1} = p_{j+1}$. Since such operations are all achievable by a sequence of partial transpositions, this contradicts with the fact that $Y = Y_{\min}$ minimizes N_b. Thus each interior rectangular block can only contribute non-negatively to $\mu(Y)$. Therefore $\mu(Y)$ is bounded from below by the contributions of the $N_b = c$ corner blocks and the lemma follows. □

We will also need the following lemma concerning the general properties of terms that appear in the summand of (B.1).

Lemma B.6. *For a pair of Young diagrams Y_1 and Y_2, the following function*

$$G_{Y_1,Y_2}(a) \equiv \prod_{s \in Y_1} (a - (h_{Y_1}(s) - v_{Y_2}(s))) \prod_{t \in Y_2} (-a - (h_{Y_2}(t) - v_{Y_1}(t)) \tag{B.18}$$
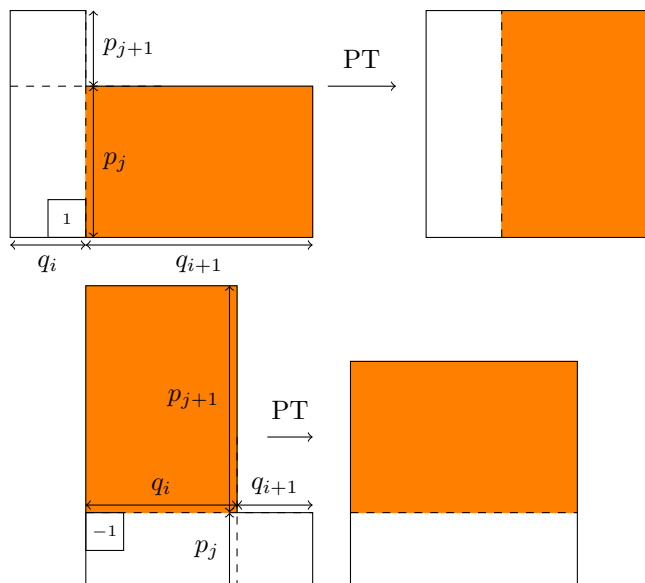
Figure 5. Reduction of N_b (or corners) by partial transpositions. Note that the PT in the right diagram involves a sequence of three partial transpositions.

satisfies

$$G_{Y_1,Y_2}(a) = G_{Y_1,Y_2^T}(a) \tag{B.19}$$

Proof. We use the following identity from appendix A.1 of [77],

$$\begin{aligned}
G_{Y_1,Y_2}(a) = &\frac{\prod_{(\alpha,\beta)\in Y_1}(a+1+M-\alpha-\beta)\prod_{(\alpha,\beta)\in Y_2}(-a+1+M-\alpha-\beta)}{\prod_{\alpha=\beta=1}^{M}(a+\alpha-\beta)} \\
&\times \prod_{\alpha=\beta=1}^{M}(a-(h_{Y_1}((\alpha,\beta)))-v_{Y_2}((\alpha,\beta)))
\end{aligned} \tag{B.20}$$

which holds when M is any positive integer larger than the widths and heights of Y_1 and Y_2. Clearly, the term in the first line of (B.20) is invariant under $Y_2 \to Y_2^T$, so it suffices to show that the last factor is also invariant. This in turn follows from the simple identity (see Lemma B.3 and figure 3)

$$\prod_{\alpha=1}^{M}(\alpha+\rho_\alpha) = \prod_{\alpha=1}^{M}(\alpha+\rho_\alpha^T) \tag{B.21}$$

where ρ_α are column lengths for $Y_2 = [\rho_1,\ldots,\rho_m]$ (and zero for α larger than the width of Y_2). □

The strategy of the proof for the main theorem B.1 is as follows. We first focus on the case where $\vec{Y}$ contains a single nonempty Young diagram $Y_{\hat{i}} = Y$. By studying the structure of the summand $Z_{\vec{Y}}$ in (B.1) in the limit $\epsilon_{1,2} \to 1$, we argue that the contribution

to (B.4) vanishes if $\mu(Y) > 2$. Furthermore, if $\mu(Y) = 2$ and Y is not related by partial transpositions to a rectangular Young diagram, its contribution to (B.4) cancels that of its partner from an involution with respect to a particular subdiagram which we will define. Consequently, only the cases with Y related to a rectangular Young diagram by partial transpositions (either with $\mu(Y) = 2$ or $\mu(Y) = 1$, since in the latter case all Young diagrams are related by partial transpositions to a rectangular one) matter for (B.4). Next, we extend the analysis to general $\vec{Y}$. Once again, the structure of the summand in (B.1) implies that the only other relevant case corresponds to $\vec{Y}$ having two non-empty Young diagrams. Up to a reordering of the Y_i, let us consider $\vec{Y} = (Y_1, Y_2, \varnothing, \ldots, \varnothing)$. Moreover, one needs $\mu(Y_1) = \mu(Y_2) = 1$ for a nonvanishing contribution to (B.4). We then show that the contribution from $\vec{Y} = (Y_1, Y_2, \varnothing, \ldots, \varnothing)$ to (B.4) (though non-vanishing) cancels with that from $\vec{Y} = (Y_1, Y_2^T, \varnothing, \ldots, \varnothing)$. This completes the proof of Theorem B.1. Below we explain each step in the above order in more detail.

We first take $\vec{Y}$ to be such that the only nonempty Young diagram is $Y_{\hat{i}} = Y$. The quantity $Z_{\vec{Y}}$ defined in (B.1) takes the form

$$Z_{\vec{Y}} = F_1(Y)F_2(Y) \tag{B.22}$$

where

$$
\begin{aligned}
F_1(Y) &= \prod_{s \in Y} \frac{(E(0,Y,Y,s) - im - \epsilon_+/2)(-E(0,Y,Y,s) - im + \epsilon_+/2)}{E(0,Y,Y,s)(\epsilon_+ - E(0,Y,Y,s))}, \\
F_2(Y) &= \prod_{j=1,j\neq\hat{i}}^{N} \prod_{s \in Y} \frac{(E(a_{\hat{i}j},Y,\varnothing,s) - im - \epsilon_+/2)}{E(a_{\hat{i}j},Y,\varnothing,s)} \frac{(E(a_{\hat{i}j},Y,\varnothing,s) + im - \epsilon_+/2)}{E(a_{\hat{i}j},Y,\varnothing,s) - \epsilon_+}.
\end{aligned}
\tag{B.23}
$$

Our goal here is to understand properties of $Z_{\vec{Y}}$ in the limit $\epsilon_- \to 0$ (and $\epsilon_+ \to 2$) in relation to the shape of the Young diagram Y.

The quantity $F_2(Y)$ is manifestly finite and nonzero (for generic a_i) for all $m, \epsilon_\pm$, while $F_1(Y)$ has a subtler behavior in the limit $\epsilon_- \to 0$. To understand how $F_1(Y)$ behaves in this limit, we write it as

$$F_1(Y) = F_1^0(Y)F_1^+(Y)F_1^-(Y)F_1^r(Y), \tag{B.24}$$

where we have decomposed the product over $s \in Y$ as $Y = Y_0 \sqcup Y_+ \sqcup Y_- \sqcup Y_r$ defined by

$$
\begin{aligned}
Y_0 &= \{s \in Y | h(s) - v(s) = 0\}, \\
Y_\pm &= \{s \in Y | h(s) - v(s) = \pm 1\}, \\
Y_r &= Y \backslash (Y_0 \sqcup Y_+ \sqcup Y_-),
\end{aligned}
\tag{B.25}
$$

and $F_1^0, F_1^\pm, F_1^r$ correspond to products over these disjoint subsets, respectively. Note that

according to our previous definition, $n_0 = |Y_0|$ and $n_{\pm 1} = |Y_\pm|$. More explicitly, we have

$$F_1^0 = \prod_{s \in Y_0} \frac{(h\epsilon_- + \frac{\epsilon_-}{2} + m)(h\epsilon_- + \frac{\epsilon_-}{2} - m)}{(h\epsilon_- - \epsilon_2)(h\epsilon_- + \epsilon_1)},$$

$$F_1^+ = \prod_{s \in Y_+} \frac{(h\epsilon_- + \frac{\epsilon_+}{2} + m)(h\epsilon_- + \frac{\epsilon_+}{2} - m)}{h\epsilon_-(h\epsilon_- + \epsilon_+)}, \tag{B.26}$$

$$F_1^- = \prod_{s \in Y_-} \frac{(h\epsilon_- + \frac{\epsilon_+}{2} + m - 2\epsilon_2)(h\epsilon_- + \frac{\epsilon_+}{2} - m - 2\epsilon_2)}{(h\epsilon_- - 2\epsilon_2)(h + 1)\epsilon_-},$$

where, for notational simplicity, we suppressed the s dependence in $h(s)$. For the quantity (B.4) of interest, we need to take the limit $\epsilon_- \to 0$. In this limit, F_1^0 could potentially vanish, $F_1^\pm$ could potentially blow up, and F_1^r is finite and nonzero.

Expanding in small ϵ_- (and taking $\epsilon_+ = 2$), we have

$$F_1^0 = \prod_{s \in Y_0} \left(-m^2 - (h + 1/2)^2 \epsilon_-^2\right) (1 + \mathcal{O}(\epsilon_-^2)),$$

$$F_1^+ = \prod_{s \in Y_+} \frac{1}{2h\epsilon_-} \left(1 + \frac{3}{2}h\epsilon_- + \mathcal{O}(\epsilon_-^2)\right), \tag{B.27}$$

$$F_1^- = \prod_{s \in Y_-} \frac{1}{-2(h + 1)\epsilon_-} \left(1 - \frac{3}{2}(h + 1)\epsilon_- + \mathcal{O}(\epsilon_-^2)\right).$$

From lemma B.5, $\mu(Y) = 2n_0 - n_1 - n_{-1} \geq 1$, which implies that there can be at most a simple pole in ϵ_- from F_1 at order m^2. For the purpose of extracting the order m^2 term from Z_Y, we can take the truncation (which we denote by $\simeq$)

$$F_1^0 \simeq \prod_{s \in Y_0} \left(-m^2 - (h + 1/2)^2 \epsilon_-^2\right),$$

$$F_1^+ \simeq \prod_{s \in Y_+} \frac{1}{2h\epsilon_-} \left(1 + \frac{3}{2}h\epsilon_-\right),$$

$$F_1^- \simeq \prod_{s \in Y_-} \frac{1}{-2(h + 1)\epsilon_-} \left(1 - \frac{3}{2}(h + 1)\epsilon_-\right), \tag{B.28}$$

$$F_1^r \simeq F_1^r|_{m=0}.$$

We recollect the relevant contributions to F_1 as

$$F_1(Y) \simeq m^2 \epsilon_-^{\mu(Y)-2}(F_1^{(0)} + \epsilon_- F_1^{(1)})F_1^r \tag{B.29}$$

where

$$F_1^{(0)}(Y) = \frac{1}{\prod_{s \in Y_+}(2h) \prod_{s \in Y_-}(-2v)} \prod_{s \in Y_0} \left(h + \frac{1}{2}\right)^2 \sum_{s \in Y_0} \frac{1}{\left(h + \frac{1}{2}\right)^2} \tag{B.30}$$

and

$$F_1^{(1)}(Y) = \frac{3}{2} F_1^{(0)}(Y) \left(\sum_{s \in Y_+} h - \sum_{s \in Y_-} v\right). \tag{B.31}$$

In addition,

$$F_1^r = F_1^{r(0)} + \epsilon_- F_1^{r(1)} \,, \tag{B.32}$$

where

$$F_1^{r(0)}(Y) = \prod_{s \in Y} \frac{(h-v)^2}{(h-v)^2 - 1} \,. \tag{B.33}$$

Similarly, it suffices to set $m = 0$ for F_2,

$$F_2(Y) \simeq F_2^{(0)}(Y) + \epsilon_- F_2^{(1)}(Y) \,, \tag{B.34}$$

where we only kept the terms up to first order in ϵ_-. Here

$$F_2^{(0)}(Y) = \prod_{j=1, j \neq \hat{i}}^{N} \prod_{s \in Y} \frac{(ia_{j\hat{i}} - (h_\varnothing(s) - v_Y(s))^2}{(ia_{j\hat{i}} - (h_\varnothing(s) - v_Y(s))^2 - 1} \tag{B.35}$$

and

$$F_2^{(1)}(Y) \tag{B.36}$$

$$= F_2^{(0)}(Y) \sum_{j=1, j \neq \hat{i}}^{N} \sum_{s \in Y} \frac{h_\varnothing(s) + v_Y(s) + 1}{(ia_{j\hat{i}} - (h_\varnothing(s) - v_Y(s))(ia_{j\hat{i}} - (h_\varnothing(s) - v_Y(s) + 1))(ia_{j\hat{i}} - (h_\varnothing(s) - v_Y(s) - 1))} \,.$$

We have thus spelled out which parts of (B.22) and (B.23) matter for evaluating the limit $\epsilon_{1,2} \to 1$ needed in (B.4), and how they depend on the combinatoric properties of the Young diagram Y. We are ready to see which Young diagrams can potentially contribute to (B.4). From Lemma B.5, any Young diagram Y that is not related by partial transposition to a rectangular one has $\mu(Y) \geq 2$. From (B.29), the contribution to (B.4) from this particular Young diagram vector $\vec{Y}$ (with the single non-empty Young diagram Y) goes as,

$$I_{k,\vec{Y}} \sim \lim_{\epsilon_- \to 0} \epsilon_-^{\mu(Y)-2} \tag{B.37}$$

Thus if $\mu(Y) > 2$, it vanishes identically. We then move onto cases of Y with $\mu(Y) = 2$, and its contribution to (B.4) is then deduced from

$$F_1(Y) \simeq m^2 F_1^{(0)}(Y) \,, \qquad F_2(Y) \simeq F_2^{(0)}(Y) \tag{B.38}$$

to be

$$I_{k,\vec{Y}} = 2 F_1^{(0)}(Y) F_2^{(0)}(Y) \,. \tag{B.39}$$

It is easy to see that the corresponding $Y_{\min}$ takes the form as in figure 6 from stacking 3 rectangular boxes of sizes $p_1 \times q_1$, $p_2 \times q_1$, and $p_1 \times q_2$ respectively where the sides satisfy (see proof of Lemma B.5)

$$q_2 \geq p_1 + p_2 + 1 \,, \qquad q_1 > p_2 \,. \tag{B.40}$$

Under partial transpositions, we will show that the $p_2 \times q_1$ block of $Y_{\min}$ gets mapped to a Young subdiagram of the resulting Young diagram Y. We define W^Y to be this particular
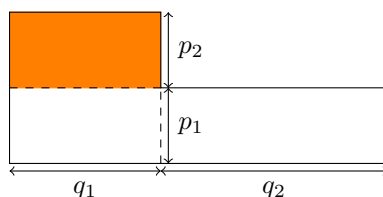
Figure 6. $Y_{\min}$ for a Young diagram Y with $\mu(Y) = 2$ and not related to a rectangular diagram by partial transpositions.

Figure 7. An example with $Y_{\min} = [4,4,4,2,2,2,2,2]$ (first diagram above). The other Y's are related by partial transpositions to $Y_{\min}$ and the subdiagram W^Y is marked in orange.

subdiagram in Y (see figure 7 for an example). Then we define the Young diagram Y' (correspondingly the Young diagram vector $\vec{Y}'$ with a single non-empty Young diagram $Y'_{\hat{i}} = Y'$) to be obtained from Y by a sequence of partial transpositions that send W^Y to its transpose while keeping the rest of the diagram fixed.(It is always possible to find such a sequence of partial transpositions.)[22] We denote this involution by ι_Y.

We show below that

$$I_{k,\vec{Y}} + I_{k,\vec{Y}'} = 0. \tag{B.41}$$

This is a consequence of the properties of $F_1^{(0)}(Y)$ and $F_2^{(0)}(Y)$ under partial transpositions: the latter is even whereas the former is odd for the particular type of Y and involution ι_Y we consider here.

We first show that $F_2^{(0)}(Y)$ is invariant under partial transposition with respect to an arbitrary subdiagram P. We have the decomposition

$$F_2^{(0)}(Y) = \prod_{j=1,j\neq\hat{i}}^{N} \prod_{s\in Y\backslash P} \frac{(ia_{j\hat{i}} - (h_\varnothing(s) - v_Y(s))^2}{(ia_{j\hat{i}} - (h_\varnothing(s) - v_Y(s))^2 - 1} \prod_{s\in P} \frac{(ia_{j\hat{i}} - (h_\varnothing(s) - v_Y(s))^2}{(ia_{j\hat{i}} - (h_\varnothing(s) - v_Y(s))^2 - 1}. \tag{B.42}$$

Since for general $s = (\alpha, \beta) \in Y$ and $Y = [\lambda_1, \ldots, \lambda_l]$

$$h_\varnothing(s) - v_Y(s) = -\alpha - (\lambda_\alpha - \beta), \tag{B.43}$$

---

[22]The existence of such a sequence of partial transpositions is guaranteed for Y of the form in figure 8 (with the orange block W^Y possibly replaced by its partial transpositions). In the later part of the section, we will prove that for Y with $\mu(Y) = 2$ and not related to rectangular diagrams by partial transpositions, it takes the form as in figure 8. Here we assume this is the case. Note that $\mu(Y) = 2$ demands the bottom-left $Y_{t\times r}$ sub-diagram (gray block) to contribute -1 to $\mu(Y)$ since each of the three exterior colored blocks along the North-East boarder (orange or gray) contribute at least 1 to $\mu(Y)$. This means we have either $\lambda_1^T - r = \lambda_r$ or $\lambda_t^T = \lambda_1 - t$(see proof of Lemma B.5). In the former case, ι_Y is achieved by PT_{r+1} followed by PT_{r+t+1}; in the latter case, $\iota_Y = \mathrm{PT}_1 \cdot \mathrm{PT}_{r+t+1} \cdot \mathrm{PT}_{t+1} \cdot \mathrm{PT}_1$.

where λ_β^T does not appear, the first factor in (B.42) from $Y \backslash P$ does not change under $P \to P^T$. It is also easy to see that the second factor is also invariant. For $(\alpha, \beta) \in P$ and $(\beta, \alpha) \in P^T$ with $P = [\rho_1, \ldots, \rho_m]$,

$$
\begin{aligned}
h_\varnothing(\alpha, \beta) - v_Y(\alpha, \beta) &= -\alpha - (l - m) - \rho_\alpha + \beta \, , \\
h_\varnothing(\beta, \alpha) - v_{\mathrm{PT}(Y)}(\beta, \alpha) &= -\beta - (l - m) - \rho_\beta^T + \alpha
\end{aligned}
\tag{B.44}
$$

are mapped to each another through

$$
\alpha \to 1 + \rho_\beta^T - \alpha \, , \qquad \beta \to 1 + \rho_\alpha - \beta \, .
\tag{B.45}
$$

Thus we've shown

$$
F_2^{(0)}(Y) = F_2^{(0)}(\mathrm{PT}(Y)) \, .
\tag{B.46}
$$

Let's now show that $F_1^{(0)}(Y)$ is odd under the involution ι_Y if $\mu(Y) = 2$. We rewrite $F_1^{(0)}(Y)$ as

$$
\begin{aligned}
F_1^{(0)}(Y) &= f_1(Y) f_2(Y) f_3(Y) \\
f_1 &= \frac{1}{\prod_{s \in Y_+}(2h) \prod_{s \in Y_-}(-2v)} \\
f_2 &= \prod_{s \in Y_0}\left(h + \frac{1}{2}\right)^2 \sum_{s \in Y_0} \frac{1}{\left(h + \frac{1}{2}\right)^2} \\
f_3 &= \prod_{s \in Y_r} \frac{(h - v)^2}{(h - v)^2 - 1} \, .
\end{aligned}
\tag{B.47}
$$

Once again, we separate the products above as

$$
f_i(Y) = \prod_{s \in Y \backslash W^Y} \cdots \prod_{s \in W^Y} \cdots \equiv f_i^1(Y) f_i^2(Y)
\tag{B.48}
$$

where f_i^1 and f_i^2 only receives contributions from entries in $Y \backslash W^Y$ and W^Y respectively. Since ι_Y simply transposes W^Y, which exchanges h and v, we have

$$
f_1^2(Y) = -f_1^2(\iota_Y(Y)) \, , \qquad f_2^2(Y) = f_2^2(\iota_Y(Y)) \, , \qquad f_3^2(Y) = f_3^2(\iota_Y(Y)) \, .
\tag{B.49}
$$

Now if Y takes the form as in figure 8 with W^Y a rectangular block (or type either $p_2 \times q_1$ or its transpose as in figure 6), then clearly f_i^1 are invariant under the involution ι_Y since the values of h (or v) do not change for boxes in the gray parts of the Young diagram as in figure 8. Furthermore, if W^Y is replaced its partial transpositions (leaving the rest of Y fixed), the contributions to f_i^1 still come from the same gray blocks as in figure 8, and consequently invariance of f_i^1 under ι_Y continues to hold thanks to corollary B.2.1 (see also proof of Lemma B.4). Since Y is related to $Y_{\min}$ in figure 6 by a sequence of partial transpositions, it is sufficient to show that any partial transposition preserves the form of Y up to partial transpositions restricted to the subdiagram W^Y in figure 8 (in particular W^Y remains a Young subdiagram under any partial transposition of Y).

We proceed by induction. First note $Y_{\min}$ is of the form in figure 8 with $r = 0, t = p_1$ and $p = p_2, q = q_1$. Now suppose we start with a general Y as in figure 8, we show that any partial transposition PT_α preserves its form, namely only the gray blocks contain boxes with $h - v = 0$ or ± 1 among all boxes in $Y \backslash W^Y$. Clearly, the only nontrivial partial transpositions are PT_α with $\alpha = 1$ which corresponds to the usual transposition, or $1 < \alpha \leq r$ and $\lambda_1^T - \alpha + 1 \leq \lambda_{\alpha-1}$, or $\lambda_1^T \geq \alpha \geq r + q + 1$ and $\lambda_1^T - \alpha + 1 \leq t$. The first case clearly preserves the form of Y as in figure 8. In the second case, the induction hypothesis implies the values $h - v$ for the boxes just below the top left gray block in figure 8 satisfy

$$r - \alpha + (q + p + t) - \lambda_\alpha \leq -2 \tag{B.50}$$

for $1 \leq \alpha \leq r$, and this is sufficient to ensure that the left white block in figure 8 does not contain entries with $h - v$ equal to 0 or ± 1. Suppose we perform a partial transposition $\mathrm{PT}_x(Y)$ for some $1 < x \leq r$, then the resulting Young diagram is again of the form in figure 8 with the change in the parameters,

$$(p, q, t, r) \to (q, p, r - x + 1, t + x - 1) \tag{B.51}$$

and to ensure that the (new) left white block as in figure 8 does not contain entries with $h - v$ equal to 0 or ± 1, we need

$$t + x - 1 - \alpha + (q + p + t - x + 1) - \lambda_\alpha \leq -2 \,, \tag{B.52}$$

which follows from (B.50). A similar argument applies to the third case, focusing on the right bottom white block in figure 8.

We have thus proven that

$$F_1^{(0)}(Y) = -F_1^{(0)}(\iota_Y(Y)) \tag{B.53}$$

for the special type of Young diagram Y and involution ι_Y considered here.

Putting together the facts that $F_1^{(0)}(Y)$ and $F_2^{(0)}(Y)$ are respectively odd and even under the involution ι_Y, it follows that the contributions to (B.4) from Y and $\iota_Y(Y)$ cancel pairwise in the $\epsilon_- \to 0$ limit as a consequence of (B.53) and (B.46). We have now shown that in the case where $\vec{Y}$ contains a single non-empty Young diagram, if this Young diagram is not related to a rectangular diagram by partial transpositions, it does not contribute to (B.4).

We now explain why only $\vec{Y}$ with a single non-empty Young diagram contributes to (B.4). The form of F_1 and the behavior (B.29) implies the only other relevant case involves two non-empty Young-diagram in $\vec{Y}$ which we can take to be

$$\vec{Y} = \{Y_1, Y_2, \varnothing, \dots, \varnothing\} \,, \tag{B.54}$$

up to an S_N permutation, and we must have $\mu(Y_1) = \mu(Y_2) = 1$ (i.e. related to non-square rectangular Young diagram by partial transpositions). In this case, we have pair-wise cancellations between such $\vec{Y}$ and the partner $\vec{Y}'$ with Y_2 (or Y_1) replaced by its transpose

$$\vec{Y}' = \{Y_1, Y_2^T, \varnothing, \dots, \varnothing\} \,. \tag{B.55}$$
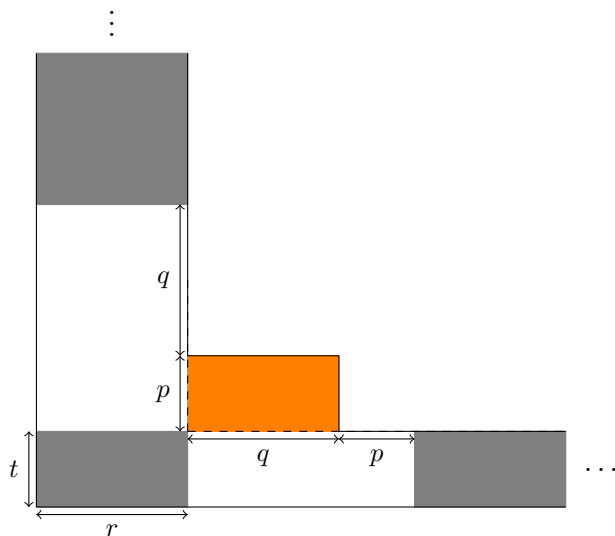
Figure 8. Young diagram Y related to $Y_{\min}$ in figure 6 by partial transpositions. The orange block denotes W^Y. The gray blocks contain all the boxes with $h - v = 0$ or ± 1 in $Y \backslash W^Y$. Note that r, t are non-negative integers. If $t = 0$, the only gray block is the one on the top left; if $r = 0$, the only gray block is the one on the bottom right. In particular $Y_{\min}$ in figure 6 corresponds to the special case $r = 0, t = p_1$ and $p = p_2, q = q_1$.

To see this, recall that the summand $Z_{\vec{Y}}$ of (B.1) in this case decomposes as the following product

$$Z_{\vec{Y}} = Z_{Y_1} Z_{Y_2} Z_{Y_1, Y_2} \,, \tag{B.56}$$

where each factor only depends on the Young diagram(s) in the subscript (e.g. Z_{Y_1} only involves the product in (B.1) with $i = 1$ and $j \neq 2$ or $j = 1$ and $i \neq 2$). From the argument we presented for the case $\vec{Y} = \{Y, \varnothing, \ldots, \varnothing\}$, it's easy to see that Z_{Y_1} is invariant under $Y_2 \to Y_2^T$ while Z_{Y_2} is odd since $\mu(Y_2) = 1$ (see equation (B.28) and (B.30)). Thus it suffices to prove that in the limit (B.4)

$$Z_{Y_1, Y_2} \simeq \prod_{s \in Y_1} \frac{(ia_{21} - (h_{Y_1}(s) - v_{Y_2}(s))^2}{(ia_{21} - (h_{Y_1}(s) - v_{Y_2}(s))^2 - 1} \prod_{t \in Y_2} \frac{(ia_{12} - (h_{Y_2}(t) - v_{Y_1}(t))^2}{(ia_{12} - (h_{Y_2}(t) - v_{Y_1}(t))^2 - 1} \tag{B.57}$$

satisfies

$$Z_{Y_1, Y_2} = Z_{Y_1, Y_2^T} \,, \tag{B.58}$$

which indeed follows from Lemma B.6.

We have thus finished the proof of theorem B.1: the limit (B.4) of double mass derivatives of the Nekrasov partition function (B.1) is dominated by $\vec{Y}$ with a single non-empty Young diagram of rectangular shape (and its partial transpositions).

C Recursion relations

In this appendix we will give a proof of (3.48) for the special cases in which the nontrivial Young diagram in the vector $\vec{Y}$ in (3.32) is of the form $Y_{1 \times k}$ or $Y_{k \times 1}$. Recall a Young

diagram vector $\vec{Y}$'s contribution to the k-instanton partition function is given by

$$Z_{k,\vec{Y}}(m, a_{ij}, \epsilon_i) = \frac{1}{k!}\left(\frac{\epsilon_+(m^2+\epsilon_-^2/4)}{\epsilon_1\epsilon_2(m^2+\epsilon_+^2/4)}\right)^k \oint \prod_{I=1}^{k}\frac{d\phi_I}{2\pi}\prod_{j=1}^{N}\frac{(\phi_I-a_j)^2-m^2}{(\phi_I-a_j)^2+\epsilon_+^2/4} \tag{C.1}$$

$$\times \prod_{I<J}^{k}\frac{\phi_{IJ}^2[\phi_{IJ}^2+\epsilon_+^2][\phi_{IJ}^2+(im-\epsilon_-/2)^2][\phi_{IJ}^2+(im+\epsilon_-/2)^2]}{[\phi_{IJ}^2+\epsilon_1^2][\phi_{IJ}^2+\epsilon_2^2][\phi_{IJ}^2+(im+\epsilon_+/2)^2][\phi_{IJ}^2+(im-\epsilon_+/2)^2]}.$$

For the round S^4, by setting $\epsilon_1 = \epsilon_2 = 1$ we have

$$Z_{k,\vec{Y}}(m, a_{ij}) = \frac{1}{k!}\left(\frac{2m^2}{m^2+1}\right)^k \oint \prod_{I=1}^{k}\frac{d\phi_I}{2\pi}\prod_{i=1}^{N}\frac{(\phi_I-a_i)^2-m^2}{(\phi_I-a_i)^2+1}$$

$$\times \prod_{I<J}^{k}\frac{\phi_{IJ}^2(\phi_{IJ}^2+4)(\phi_{IJ}^2-m^2)^2}{(\phi_{IJ}^2+1)^2[(\phi_{IJ}-m)^2+1][(\phi_{IJ}+m)^2+1]}. \tag{C.2}$$

In this appendix, we will first keep ϵ_1, ϵ_2 arbitrary, and set them equal to 1 at the end. That is because with general ϵ_1, ϵ_2, the instanton partition function only has simple poles and obeys a simple recursion relation [77, 78]. Indeed, it is straightforward to see that $Z_{k,\vec{Y}}(m, a_{ij}, \epsilon_i)$ satisfies the following recursion relation

$$Z_{k+1,\vec{Y}_+}(m, a_{ij}, \epsilon_i) = Z_{k,\vec{Y}}(m, a_{ij}, \epsilon_i)\frac{1}{k+1}\frac{\epsilon_+(m^2+\epsilon_-^2/4)}{\epsilon_1\epsilon_2(m^2+\epsilon_+^2/4)}\oint\frac{d\phi}{2\pi}\prod_{j=1}^{N}\frac{(\phi-a_j)^2-m^2}{(\phi-a_j)^2+\epsilon_+^2/4} \tag{C.3}$$

$$\times \prod_{J=1}^{k}\frac{(\phi-\hat{\phi}_J)^2[(\phi-\hat{\phi}_J)^2+\epsilon_+^2][(\phi-\hat{\phi}_J)^2+(im-\epsilon_-/2)^2][(\phi-\hat{\phi}_J)^2+(im+\epsilon_-/2)^2]}{[(\phi-\hat{\phi}_J)^2+\epsilon_1^2][(\phi-\hat{\phi}_J)^2+\epsilon_2^2][(\phi-\hat{\phi}_J)^2+(im+\epsilon_+/2)^2][(\phi-\hat{\phi}_J)^2+(im-\epsilon_+/2)^2]},$$

where $\vec{Y}_+$ (with $k+1$ boxes) is a Young diagram vector that is obtained from $\vec{Y}$ by adding one more box. The contour for the integration over ϕ is determined by the position of the box that we add to $\vec{Y}$ during the recursion procedure, and $\hat{\phi}_J$ are the poles for evaluating $Z_{k,\vec{Y}}(m, a_{ij}, \epsilon_i)$, which are determined by $\vec{Y}$ according to (3.27).

We will here consider the contributions from Young diagram vector $\vec{Y}$ as in (3.32) where the nontrivial Young diagram is a single row Young diagram $Y_{1\times k}$ or a single column Young diagram $Y_{k\times 1}$. These Young diagrams can be constructed recursively by adding one box at a time, and one can solve the recursion relation straightforwardly for the instanton partition function at order m^2.

We define

$$I_{1\times k}(\epsilon_1, \epsilon_2) = \partial_m^2 Z_{k,(\ldots,Y_{1\times k},\ldots)}(m, a_{ij}, \epsilon_i)\big|_{m=0}. \tag{C.4}$$

The recursion relation for $I_{1\times k}(\epsilon_1, \epsilon_2)$ is solved by beginning with the initial data, which we take to be the contribution from the $\vec{Y}$ with nontrivial Young diagram $Y_{1\times 2}$. For simplicity, we will begin by setting $N = 1$ since the argument for general N is more complicated, in which case the initial data is

$$I_{1\times 2}(\epsilon_1, \epsilon_2) = \oint\frac{dz}{2\pi}\left[\frac{3(z-a_1)(z-a_1+i\epsilon_1)}{4\epsilon_1^2(z-a_1-i\epsilon_1)(z-a_1+2i\epsilon_1)}\frac{1}{\epsilon_1-\epsilon_2}\right.$$

$$\left.+ \frac{3\epsilon_1^4+(z-a_1)(z-a_1+i\epsilon_1)\left(22\epsilon_1^2+8(z-a_1)(z-a_1+i\epsilon_1)\right)}{4\epsilon_1^3(z-a_1-i\epsilon_1)^2(z-a_1+2i\epsilon_1)^2}\right], \tag{C.5}$$

where the integration is around the poles at $z = a_1 + i\epsilon_+/2$, but the integrand is our focus here. From this we find that, in general, the $\vec{Y}$ with nontrivial Young diagram $Y_{1 \times k}$ gives the following contribution

$$I_{1 \times k}(\epsilon_1, \epsilon_2) = \oint \frac{dz}{2\pi} \left[\frac{(2k^2-2)(z-a_1)((z-a_1)+i(k-1)\epsilon_1)}{\epsilon_1^2 \, k^2 k!((z-a_1)-i\epsilon_1)((z-a_1)+ik\epsilon_1)} \frac{1}{\epsilon_1 - \epsilon_2} \right. \tag{C.6}$$
$$\left. + \frac{(k-1)^2(k+1)\epsilon_1^4 + (z-a_1)(z-a_1+i(k-1)\epsilon_1)\left((6k^2-2)\epsilon_1^2 + 4k(z-a_1)(z-a_1+i(k-1)\epsilon_1)\right)}{k \, k! \, \epsilon_1^3 (z-a_1-i\epsilon_1)^2(z-a_1+ik\epsilon_1)^2} \right].$$

Here, we only keep the terms of order $O((\epsilon_1 - \epsilon_2)^0)$. The contribution from the $\vec{Y}$ with nontrivial Young diagram $Y_{k \times 1}$ is obtained by the transposition, which simply exchanges ϵ_1 with ϵ_2,

$$I_{k \times 1}(\epsilon_1, \epsilon_2) = I_{1 \times k}(\epsilon_2, \epsilon_1). \tag{C.7}$$

The singular terms cancel in the sum of these two contributions, and the final result is given by

$$\begin{aligned} I_{1 \times k} &= \left[I_{1 \times k}(\epsilon_1, \epsilon_2) + I_{k \times 1}(\epsilon_1, \epsilon_2) \right]\big|_{\epsilon_1 = \epsilon_2 = 1} \\ &= \oint \frac{dz}{2\pi} \prod_{k_a=0}^{k-1} \frac{(z-a_1+k_a i)^2}{(z-a_1+k_a i)^2 + 1} \times \frac{1}{k!} \left[\frac{4}{1+\delta_{1k}} \left(1 + \frac{1}{k^2} \right) \right. \\ &\quad \left. + \frac{2i(k+1)(k-1)^2}{k(z-a_1+ki)(z-a_1+(k-1)i)(z-a_1)} \right], \end{aligned} \tag{C.8}$$

where we have set $\epsilon_1 = \epsilon_2 = 1$ at the end of computation. Having determined the $N = 1$ case, the generalization of the above integrand to arbitrary N is straightforward. This can be seen by using (C.2) as follows.

If $I_{1 \times k}$ were computed from (C.2), we would take consecutive residues surrounding the poles at $\phi_{IJ}^2 + 1 = 0$. Since they are higher order poles, one needs to expand the integrand when taking the residues. The constant term $\frac{4}{1+\delta_{1k}}\left(1 + \frac{1}{k^2}\right)$ of $I_{1 \times k}$ clearly comes from the second line in (C.2), which is independent of N, whereas the term containing $z - a_1$ requires expanding the first line in (C.2), and its generalization to general N is obvious. Therefore, for general N we have

$$\begin{aligned} I_{1 \times k} &= \oint \frac{dz}{2\pi} \prod_{k_a=0}^{k-1} \prod_{j=1}^{N} \frac{(z-a_j+k_a i)^2}{(z-a_j+k_a i)^2 + 1} \times \frac{1}{k!} \left[\frac{4}{1+\delta_{1k}} \left(1 + \frac{1}{k^2} \right) \right. \\ &\quad \left. + \sum_{j=1}^{N} \frac{2i(k+1)(k-1)^2}{k(z-a_j+ki)(z-a_j+(k-1)i)(z-a_j)} \right], \end{aligned} \tag{C.9}$$

and the expression (after summing over $k!$ identical contributions) agrees with the general formula given in (3.48).

D   Instantons at higher order in $1/N$

In this appendix we compute the instanton contributions to $\partial_m^2 \log Z\big|_{m=0}$ to $O(N^{-\frac{9}{2}})$ in a small $g_{\rm YM}$ expansion to subleading order for instantons $k = 2, \ldots 12$. The result matches

the conjecture in (4.1) for $\partial_m^2 \log Z\big|_{m=0}$ at finite g_{YM} in terms of Eisenstein series, which generalizes the match found for the perturbative terms and the $k = 1$ instanton in the main text. By explicitly performing the sums and products in $I_{p \times q}$ (3.48) for many small values of N, we find that $I_{p \times q}$ can be expanded for small a_i as

$$I_{p \times q}(N, a_{ij}) = I_{p \times q}^{(0)}(N) + I_{p \times q}^{(2)}(N) C_2(a_{ij}) + \cdots , \tag{D.1}$$

where recall that C_2 is defined in (4.7). When $p = q$, which includes the one-instanton case (4.6), we found closed form expressions for $I_{p \times q}^{(0)}(N)$ and $I_{p \times q}^{(2)}(N)$, but for $p \neq q$ we could only find recursion relations in N. In either case, these formulae can be expanded explicitly at large N. For instance, for $k = 2$ we find the recursion relations:

$$\left(-1600N^2 - 7310N - 8256\right) I_{2 \times 1}^{(0)}(N+2) + \left(1440N^2 + 5859N + 5958\right) I_{2 \times 1}^{(0)}(N+3)$$
$$+ (N+2)(160N + 491) I_{2 \times 1}^{(0)}(N+1) = 0 , \qquad I_{2 \times 1}^{(0)}(2) = -\frac{134}{27} , \qquad I_{2 \times 1}^{(0)}(3) = -\frac{517}{81} , \tag{D.2}$$

and

$$\left(-320N^2 - 734N - 456\right) I_{2 \times 1}^{(2)}(N+2) + \left(288N^2 + 1323N + 684\right) I_{2 \times 1}^{(2)}(N+3)$$
$$+ N(32N + 51) I_{2 \times 1}^{(2)}(N+1) = 0 , \qquad I_{2 \times 1}^{(2)}(2) = -\frac{10}{243} , \qquad I_{2 \times 1}^{(2)}(3) = -\frac{20}{729} , \tag{D.3}$$

which we can solve in a large-N expansion to get

$$I_{2 \times 1}^{(0)} = -5\sqrt{\frac{2}{\pi}}\sqrt{N} + \frac{17\sqrt{\frac{1}{N}}}{8\sqrt{2\pi}} + \frac{325\left(\frac{1}{N}\right)^{3/2}}{1024\sqrt{2\pi}} + \frac{2155\left(\frac{1}{N}\right)^{5/2}}{8192\sqrt{2\pi}} + \frac{1543605\left(\frac{1}{N}\right)^{7/2}}{4194304\sqrt{2\pi}} + O(N^{-\frac{9}{2}}) ,$$
$$I_{2 \times 1}^{(2)} = -\frac{15}{8\sqrt{2\pi}N^{5/2}} + \frac{255}{128\sqrt{2\pi}N^{7/2}} - \frac{11025}{16384\sqrt{2\pi}N^{9/2}} + \frac{478485}{131072\sqrt{2\pi}N^{11/2}} + O(N^{-\frac{13}{2}}) . \tag{D.4}$$

The cases $k = 3, \ldots, 12$ are increasingly more complicated so we put them in an attached Mathematica file.

We can then use the expectation value $\langle C_2 \rangle$ in (4.8) with $\lambda = g_{\mathrm{YM}}^2 N$ to compute $\langle I_{p \times q} \rangle$ to $O(N^{-\frac{9}{2}})$. The result in each case is consistent with the small g_{YM} expansion to $O(g_{\mathrm{YM}}^2)$ of

$$\left\langle \partial_m^2 Z_{\mathrm{inst}}^{p \times q}(m, a_{ij}) \right\rangle \bigg|_{m=0} = \frac{e^{\frac{8pq\pi^2}{g_{\mathrm{YM}}^2}}}{1 + \delta_{p,q}} \left[-\sqrt{N}\frac{16K_1\left(\frac{8pq\pi^2}{g_{\mathrm{YM}}^2}\right)}{g_{\mathrm{YM}}}\left(\frac{p}{q} + \frac{q}{p}\right) + \frac{2K_2\left(\frac{8pq\pi^2}{g_{\mathrm{YM}}^2}\right)}{g_{\mathrm{YM}}\sqrt{N}}\left(\frac{p^2}{q^2} + \frac{q^2}{p^2}\right) \right.$$

$$+ \frac{1}{32g_{\mathrm{YM}}N^{\frac{3}{2}}}\left[-13K_1\left(\frac{8pq\pi^2}{g_{\mathrm{YM}}^2}\right)\left(\frac{p}{q} + \frac{q}{p}\right) + 9K_3\left(\frac{8pq\pi^2}{g_{\mathrm{YM}}^2}\right)\left(\frac{p^3}{q^3} + \frac{q^3}{p^3}\right) \right]$$

$$+ \frac{1}{128g_{\mathrm{YM}}N^{\frac{5}{2}}}\left[-25K_2\left(\frac{8pq\pi^2}{g_{\mathrm{YM}}^2}\right)\left(\frac{p^2}{q^2} + \frac{q^2}{p^2}\right) + 15K_4\left(\frac{8pq\pi^2}{g_{\mathrm{YM}}^2}\right)\left(\frac{p^4}{q^4} + \frac{q^4}{p^4}\right) \right]$$

$$+ \frac{1}{g_{\mathrm{YM}}N^{\frac{7}{2}}}\left[\frac{1533K_1\left(\frac{8pq\pi^2}{g_{\mathrm{YM}}^2}\right)}{16384}\left(\frac{p}{q} + \frac{q}{p}\right) - \frac{5355K_3\left(\frac{8pq\pi^2}{g_{\mathrm{YM}}^2}\right)}{32768}\left(\frac{p^3}{q^3} + \frac{q^3}{p^3}\right) + \frac{2625K_5\left(\frac{8pq\pi^2}{g_{\mathrm{YM}}^2}\right)}{32768}\left(\frac{p^5}{q^5} + \frac{q^5}{p^5}\right) \right]$$

$$+ O(N^{-\frac{9}{2}}) \right] , \tag{D.5}$$

which describes the $p \times q$ instanton terms in (4.1). This is a very nontrivial check of the conjectured finite g_{YM} expression for $\partial_m^2 \log Z \big|_{m=0}$.

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

References

[1] J.M. Maldacena, *The large N limit of superconformal field theories and supergravity*, [*Int. J. Theor. Phys.* **38** (1999) 1113](#) [*Adv. Theor. Math. Phys.* **2** (1998) 231] [[hep-th/9711200](#)] [[INSPIRE](#)].

[2] E. Witten, *Anti-de Sitter space and holography*, [*Adv. Theor. Math. Phys.* **2** (1998) 253](#) [[hep-th/9802150](#)] [[INSPIRE](#)].

[3] S.S. Gubser, I.R. Klebanov and A.M. Polyakov, *Gauge theory correlators from noncritical string theory*, [*Phys. Lett. B* **428** (1998) 105](#) [[hep-th/9802109](#)] [[INSPIRE](#)].

[4] J. Polchinski, *S matrices from AdS space-time*, [hep-th/9901076](#) [[INSPIRE](#)].

[5] L. Susskind, *Holography in the flat space limit*, [*AIP Conf. Proc.* **493** (1999) 98](#) [[hep-th/9901079](#)] [[INSPIRE](#)].

[6] S.B. Giddings, *Flat space scattering and bulk locality in the AdS/CFT correspondence*, [*Phys. Rev. D* **61** (2000) 106008](#) [[hep-th/9907129](#)] [[INSPIRE](#)].

[7] J. Penedones, *Writing CFT correlation functions as AdS scattering amplitudes*, [*JHEP* **03** (2011) 025](#) [[arXiv:1011.1485](#)] [[INSPIRE](#)].

[8] A. Fitzpatrick and J. Kaplan, *Analyticity and the Holographic S-matrix*, [*JHEP* **10** (2012) 127](#) [[arXiv:1111.6972](#)] [[INSPIRE](#)].

[9] A. Fitzpatrick and J. Kaplan, *Scattering States in AdS/CFT*, [arXiv:1104.2597](#) [[INSPIRE](#)].

[10] L.F. Alday and A. Bissi, *The superconformal bootstrap for structure constants*, [*JHEP* **09** (2014) 144](#) [[arXiv:1310.3757](#)] [[INSPIRE](#)].

[11] L. Rastelli and X. Zhou, *Mellin amplitudes for $AdS_5 \times S^5$*, [*Phys. Rev. Lett.* **118** (2017) 091602](#) [[arXiv:1608.06624](#)] [[INSPIRE](#)].

[12] O. Aharony, L.F. Alday, A. Bissi and E. Perlmutter, *Loops in AdS from Conformal Field Theory*, [*JHEP* **07** (2017) 036](#) [[arXiv:1612.03891](#)] [[INSPIRE](#)].

[13] L. Rastelli and X. Zhou, *Holographic Four-Point Functions in the (2, 0) Theory*, [*JHEP* **06** (2018) 087](#) [[arXiv:1712.02788](#)] [[INSPIRE](#)].

[14] L. Rastelli and X. Zhou, *How to Succeed at Holographic Correlators Without Really Trying*, [*JHEP* **04** (2018) 014](#) [[arXiv:1710.05923](#)] [[INSPIRE](#)].

[15] S. Caron-Huot and A.-K. Trinh, *All tree-level correlators in $AdS_5 \times S_5$ supergravity: hidden ten-dimensional conformal symmetry*, [*JHEP* **01** (2019) 196](#) [[arXiv:1809.09173](#)] [[INSPIRE](#)].

[16] L.F. Alday and A. Bissi, *Generalized bootstrap equations for $\mathcal{N} = 4$ SCFT*, [*JHEP* **02** (2015) 101](#) [[arXiv:1404.5864](#)] [[INSPIRE](#)].

[17] L.F. Alday, A. Bissi and E. Perlmutter, *Genus-One String Amplitudes from Conformal Field Theory*, [*JHEP* **06** (2019) 010](#) [[arXiv:1809.10670](#)] [[INSPIRE](#)].

[18] L.F. Alday and A. Bissi, *Loop Corrections to Supergravity on $AdS_5 \times S^5$*, *Phys. Rev. Lett.* **119** (2017) 171601 [arXiv:1706.02388] [INSPIRE].

[19] F. Aprile, J.M. Drummond, P. Heslop and H. Paul, *Quantum Gravity from Conformal Field Theory*, *JHEP* **01** (2018) 035 [arXiv:1706.02822] [INSPIRE].

[20] F. Aprile, J.M. Drummond, P. Heslop and H. Paul, *Loop corrections for Kaluza-Klein AdS amplitudes*, *JHEP* **05** (2018) 056 [arXiv:1711.03903] [INSPIRE].

[21] L.F. Alday and S. Caron-Huot, *Gravitational S-matrix from CFT dispersion relations*, *JHEP* **12** (2018) 017 [arXiv:1711.02031] [INSPIRE].

[22] F. Aprile, J. Drummond, P. Heslop and H. Paul, *Double-trace spectrum of $N = 4$ supersymmetric Yang-Mills theory at strong coupling*, *Phys. Rev. D* **98** (2018) 126008 [arXiv:1802.06889] [INSPIRE].

[23] F. Aprile, J. Drummond, P. Heslop and H. Paul, *One-loop amplitudes in $AdS_5 \times S^5$ supergravity from $\mathcal{N} = 4$ SYM at strong coupling*, *JHEP* **03** (2020) 190 [arXiv:1912.01047] [INSPIRE].

[24] L.F. Alday and X. Zhou, *Simplicity of AdS Supergravity at One Loop*, *JHEP* **09** (2020) 008 [arXiv:1912.02663] [INSPIRE].

[25] J.M. Drummond and H. Paul, *One-loop string corrections to AdS amplitudes from CFT*, arXiv:1912.07632 [INSPIRE].

[26] S. Giusto, R. Russo and C. Wen, *Holographic correlators in AdS_3*, *JHEP* **03** (2019) 096 [arXiv:1812.06479] [INSPIRE].

[27] L. Rastelli, K. Roumpedakis and X. Zhou, *$AdS_3 \times S^3$ Tree-Level Correlators: Hidden Six-Dimensional Conformal Symmetry*, *JHEP* **10** (2019) 140 [arXiv:1905.11983] [INSPIRE].

[28] S. Giusto, R. Russo, A. Tyukov and C. Wen, *Holographic correlators in AdS_3 without Witten diagrams*, *JHEP* **09** (2019) 030 [arXiv:1905.12314] [INSPIRE].

[29] S.M. Chester, S.S. Pufu and X. Yin, *The M-theory S-matrix From ABJM: Beyond 11D Supergravity*, *JHEP* **08** (2018) 115 [arXiv:1804.00949] [INSPIRE].

[30] S.M. Chester and E. Perlmutter, *M-Theory Reconstruction from (2,0) CFT and the Chiral Algebra Conjecture*, *JHEP* **08** (2018) 116 [arXiv:1805.00892] [INSPIRE].

[31] D.J. Binder, S.M. Chester and S.S. Pufu, *Absence of $D^4 R^4$ in M-theory From ABJM*, *JHEP* **04** (2020) 052 [arXiv:1808.10554] [INSPIRE].

[32] D.J. Binder, S.M. Chester and S.S. Pufu, *AdS_4/CFT_3 from weak to strong string coupling*, *JHEP* **01** (2020) 034 [arXiv:1906.07195] [INSPIRE].

[33] D.J. Binder, S.M. Chester, S.S. Pufu and Y. Wang, *$\mathcal{N} = 4$ Super-Yang-Mills correlators at strong coupling from string theory and localization*, *JHEP* **12** (2019) 119 [arXiv:1902.06263] [INSPIRE].

[34] S.M. Chester, *Genus-2 holographic correlator on $AdS_5 \times S^5$ from localization*, *JHEP* **04** (2020) 193 [arXiv:1908.05247] [INSPIRE].

[35] D.J. Gross and E. Witten, *Superstring Modifications of Einstein's Equations*, *Nucl. Phys. B* **277** (1986) 1 [INSPIRE].

[36] M.T. Grisaru, A.E.M. van de Ven and D. Zanon, *Two-Dimensional Supersymmetric σ-models on Ricci Flat Kähler Manifolds Are Not Finite*, *Nucl. Phys. B* **277** (1986) 388 [INSPIRE].

[37] V. Pestun, *Localization of gauge theory on a four-sphere and supersymmetric Wilson loops*, *Commun. Math. Phys.* **313** (2012) 71 [`arXiv:0712.2824`] [InSPIRE].

[38] G.W. Moore, N. Nekrasov and S. Shatashvili, *Integrating over Higgs branches*, *Commun. Math. Phys.* **209** (2000) 97 [`hep-th/9712241`] [InSPIRE].

[39] A. Losev, N. Nekrasov and S.L. Shatashvili, *Issues in topological gauge theory*, *Nucl. Phys. B* **534** (1998) 549 [`hep-th/9711108`] [InSPIRE].

[40] N.A. Nekrasov, *Seiberg-Witten prepotential from instanton counting*, *Adv. Theor. Math. Phys.* **7** (2003) 831 [`hep-th/0206161`] [InSPIRE].

[41] N. Nekrasov and A. Okounkov, *Seiberg-Witten theory and random partitions*, *Prog. Math.* **244** (2006) 525 [`hep-th/0306238`] [InSPIRE].

[42] C.M. Hull and P.K. Townsend, *Unity of superstring dualities*, *Nucl. Phys. B* **438** (1995) 109 [`hep-th/9410167`] [InSPIRE].

[43] C. Montonen and D.I. Olive, *Magnetic Monopoles as Gauge Particles?*, *Phys. Lett. B* **72** (1977) 117 [InSPIRE].

[44] E. Witten and D.I. Olive, *Supersymmetry Algebras That Include Topological Charges*, *Phys. Lett. B* **78** (1978) 97 [InSPIRE].

[45] H. Osborn, *Topological Charges for N = 4 Supersymmetric Gauge Theories and Monopoles of Spin 1*, *Phys. Lett. B* **83** (1979) 321 [InSPIRE].

[46] T. Banks and M.B. Green, *Nonperturbative effects in AdS in five-dimensions $\times S^5$ string theory and d = 4 SUSY Yang-Mills*, *JHEP* **05** (1998) 002 [`hep-th/9804170`] [InSPIRE].

[47] M. Bianchi, M.B. Green, S. Kovacs and G. Rossi, *Instantons in supersymmetric Yang-Mills and D instantons in IIB superstring theory*, *JHEP* **08** (1998) 013 [`hep-th/9807033`] [InSPIRE].

[48] N. Dorey, T.J. Hollowood, V.V. Khoze, M.P. Mattis and S. Vandoren, *Multi-instanton calculus and the AdS/CFT correspondence in N = 4 superconformal field theory*, *Nucl. Phys. B* **552** (1999) 88 [`hep-th/9901128`] [InSPIRE].

[49] N. Dorey, T.J. Hollowood, V.V. Khoze and M.P. Mattis, *The calculus of many instantons*, *Phys. Rept.* **371** (2002) 231 [`hep-th/0206063`] [InSPIRE].

[50] O. Aharony, N. Seiberg and Y. Tachikawa, *Reading between the lines of four-dimensional gauge theories*, *JHEP* **08** (2013) 115 [`arXiv:1305.0318`] [InSPIRE].

[51] E. Witten, *AdS/CFT correspondence and topological field theory*, *JHEP* **12** (1998) 012 [`hep-th/9812012`] [InSPIRE].

[52] D. Gaiotto, D. Mazac and M.F. Paulos, *Bootstrapping the 3d Ising twist defect*, *JHEP* **03** (2014) 100 [`arXiv:1310.5078`] [InSPIRE].

[53] M.B. Green and M. Gutperle, *Effects of D instantons*, *Nucl. Phys. B* **498** (1997) 195 [`hep-th/9701093`] [InSPIRE].

[54] M.B. Green, M. Gutperle and P. Vanhove, *One loop in eleven-dimensions*, *Phys. Lett. B* **409** (1997) 177 [`hep-th/9706175`] [InSPIRE].

[55] M.B. Green and S. Sethi, *Supersymmetry constraints on type IIB supergravity*, *Phys. Rev. D* **59** (1999) 046006 [`hep-th/9808061`] [InSPIRE].

[56] M.B. Green, H.-h. Kwon and P. Vanhove, *Two loops in eleven-dimensions*, *Phys. Rev. D* **61** (2000) 104010 [hep-th/9910055] [INSPIRE].

[57] M.B. Green and P. Vanhove, *Duality and higher derivative terms in M-theory*, *JHEP* **01** (2006) 093 [hep-th/0510027] [INSPIRE].

[58] Y. Wang and X. Yin, *Constraining Higher Derivative Supergravity with Scattering Amplitudes*, *Phys. Rev. D* **92** (2015) 041701 [arXiv:1502.03810] [INSPIRE].

[59] H. Elvang, D.Z. Freedman and M. Kiermaier, *A simple approach to counterterms in N = 8 supergravity*, *JHEP* **11** (2010) 016 [arXiv:1003.5018] [INSPIRE].

[60] M.B. Green and C. Wen, *Modular Forms and* $SL(2,\mathbb{Z})$-*covariance of type IIB superstring theory*, *JHEP* **06** (2019) 087 [arXiv:1904.13394] [INSPIRE].

[61] A. Basu, M.B. Green and S. Sethi, *A curious truncation of N = 4 Yang-Mills*, *Phys. Rev. Lett.* **93** (2004) 261601 [hep-th/0406267] [INSPIRE].

[62] O. Aharony, O. Bergman, D.L. Jafferis and J. Maldacena, *N = 6 superconformal Chern-Simons-matter theories, M2-branes and their gravity duals*, *JHEP* **10** (2008) 091 [arXiv:0806.1218] [INSPIRE].

[63] K.A. Intriligator, *Bonus symmetries of N = 4 superYang-Mills correlation functions via AdS duality*, *Nucl. Phys. B* **551** (1999) 575 [hep-th/9811047] [INSPIRE].

[64] K.A. Intriligator and W. Skiba, *Bonus symmetry and the operator product expansion of N = 4 SuperYang-Mills*, *Nucl. Phys. B* **559** (1999) 165 [hep-th/9905020] [INSPIRE].

[65] B. Eden, A.C. Petkou, C. Schubert and E. Sokatchev, *Partial nonrenormalization of the stress tensor four point function in N = 4 SYM and AdS/CFT*, *Nucl. Phys. B* **607** (2001) 191 [hep-th/0009106] [INSPIRE].

[66] M. Nirschl and H. Osborn, *Superconformal Ward identities and their solution*, *Nucl. Phys. B* **711** (2005) 409 [hep-th/0407060] [INSPIRE].

[67] G. Mack, *D-dimensional Conformal Field Theories with anomalous dimensions as Dual Resonance Models*, *Bulg. J. Phys.* **36** (2009) 214 [arXiv:0909.1024] [INSPIRE].

[68] G. Mack, *D-independent representation of Conformal Field Theories in D dimensions via transformation to auxiliary Dual Resonance Models. Scalar amplitudes*, arXiv:0907.2407 [INSPIRE].

[69] F. Aprile, J.M. Drummond, P. Heslop and H. Paul, *Unmixing Supergravity*, *JHEP* **02** (2018) 133 [arXiv:1706.08456] [INSPIRE].

[70] R.H. Boels and D. O'Connell, *Simple superamplitudes in higher dimensions*, *JHEP* **06** (2012) 163 [arXiv:1201.2653] [INSPIRE].

[71] J.G. Russo and K. Zarembo, *Massive N = 2 Gauge Theories at Large N*, *JHEP* **11** (2013) 130 [arXiv:1309.1004] [INSPIRE].

[72] L.F. Alday, D. Gaiotto and Y. Tachikawa, *Liouville Correlation Functions from Four-dimensional Gauge Theories*, *Lett. Math. Phys.* **91** (2010) 167 [arXiv:0906.3219] [INSPIRE].

[73] L.F. Alday and Y. Tachikawa, *Affine* SL(2) *conformal blocks from 4d gauge theories*, *Lett. Math. Phys.* **94** (2010) 87 [arXiv:1005.4469] [INSPIRE].

[74] T. Okuda and V. Pestun, *On the instantons and the hypermultiplet mass of N = 2∗ super Yang-Mills on* S^4, *JHEP* **03** (2012) 017 [arXiv:1004.1222] [INSPIRE].

[75] L.C. Jeffrey and F.C. Kirwan, *Localization for nonabelian group actions*, `alg-geom/9307001`.

[76] C. Hwang, J. Kim, S. Kim and J. Park, *General instanton counting and 5d SCFT*, *JHEP* **07** (2015) 063 [*Addendum ibid.* **04** (2016) 094] [`arXiv:1406.6793`] [INSPIRE].

[77] S. Kanno, Y. Matsuo and H. Zhang, *Extended Conformal Symmetry and Recursion Formulae for Nekrasov Partition Function*, *JHEP* **08** (2013) 028 [`arXiv:1306.1523`] [INSPIRE].

[78] S. Nakamura, F. Okazawa and Y. Matsuo, *Recursive method for the Nekrasov partition function for classical Lie groups*, *PTEP* **2015** (2015) 033B01 [`arXiv:1411.4222`] [INSPIRE].

[79] B. Eynard, *Topological expansion for the 1-Hermitian matrix model correlation functions*, *JHEP* **11** (2004) 031 [`hep-th/0407261`] [INSPIRE].

[80] B. Eynard and N. Orantin, *Algebraic methods in random matrices and enumerative geometry*, `arXiv:0811.3531` [INSPIRE].

[81] N. Dorey, V.V. Khoze, M.P. Mattis and S. Vandoren, *Yang-Mills instantons in the large N limit and the AdS/CFT correspondence*, *Phys. Lett. B* **442** (1998) 145 [`hep-th/9808157`] [INSPIRE].

[82] R.H. Boels, *Maximal R-symmetry violating amplitudes in type IIB superstring theory*, *Phys. Rev. Lett.* **109** (2012) 081602 [`arXiv:1204.4208`] [INSPIRE].

[83] T. Nosaka, *Instanton effects in ABJM theory with general R-charge assignments*, *JHEP* **03** (2016) 059 [`arXiv:1512.02862`] [INSPIRE].

[84] M. Mariño and P. Putrov, *ABJM theory as a Fermi gas*, *J. Stat. Mech.* **1203** (2012) P03001 [`arXiv:1110.4066`] [INSPIRE].

[85] M.B. Green, S.D. Miller and P. Vanhove, *$SL(2,\mathbb{Z})$-invariance and D-instanton contributions to the D^6R^4 interaction*, *Commun. Num. Theor. Phys.* **09** (2015) 307 [`arXiv:1404.2192`] [INSPIRE].