5d partition functions with a twist

Crichigno, P.M.; Jain, D.; Willett, B.

Citation for published version (APA):

5d partition functions with a twist

P. Marcos Crichigno,[a] **Dharmesh Jain**[b] **and Brian Willett**[c]

[a]*Institute for Theoretical Physics, University of Amsterdam,*
Science Park 904, Postbus 94485, 1090 GL, Amsterdam, The Netherlands

[b]*Theory Division, Saha Institute of Nuclear Physics,*
1/AF Bidhan Nagar, Kolkata 700064, India

[c]*Kavli Institute for Theoretical Physics, University of California,*
Santa Barbara, CA 93106, U.S.A.

E-mail: p.m.crichigno@uva.nl, d.jain@saha.ac.in,
bwillett@kitp.ucsb.edu

ABSTRACT: We derive the partition function of 5d $\mathcal{N} = 1$ gauge theories on the manifold $S_b^3 \times \Sigma_{\mathfrak{g}}$ with a partial topological twist along the Riemann surface, $\Sigma_{\mathfrak{g}}$. This setup is a higher dimensional uplift of the two-dimensional A-twist, and the result can be expressed as a sum over solutions of Bethe-Ansatz-type equations, with the computation receiving nontrivial non-perturbative contributions. We study this partition function in the large N limit, where it is related to holographic RG flows between asymptotically locally AdS_6 and AdS_4 spacetimes, reproducing known holographic relations between the corresponding free energies on S^5 and S^3 and predicting new ones. We also consider cases where the 5d theory admits a UV completion as a 6d SCFT, such as the maximally supersymmetric $\mathcal{N} = 2$ Yang-Mills theory, in which case the partition function computes the 4d index of general class $\mathcal{S}$ theories, which we verify in certain simplifying limits. Finally, we comment on the generalization to $\mathcal{M}_3 \times \Sigma_{\mathfrak{g}}$ with more general three-manifolds $\mathcal{M}_3$ and focus in particular on $\mathcal{M}_3 = \Sigma_{\mathfrak{g}'} \times S^1$, in which case the partition function relates to the entropy of black holes in AdS_6.

Contents

1   Introduction and summary

There has been tremendous progress in obtaining exact results for supersymmetric gauge theories in various numbers of spacetime dimensions. These computations typically involve protected observables, which may be studied using non-renormalization theorems or, in the case of partition functions on compact manifolds, the localization technique [1, 2]. These results have opened new windows on quantum field theories and dualities between them.

Moreover, they have shown how theories in different dimensions are intricately interrelated. Often, subtle properties of a theory, such as its dual descriptions or space of marginal parameters, can become obvious when we embed this theory into a higher-dimensional framework.

In this paper, we extend the list of exact observables by computing the partition function of arbitrary 5d $\mathcal{N} = 1$ gauge theories on manifolds of the form $\mathcal{M}_3 \times \Sigma_{\mathfrak{g}}$, i.e., a product of a three-manifold and a genus-$\mathfrak{g}$ Riemann surface. We will focus mainly on the case $\mathcal{M}_3 = S_b^3$, the "squashed sphere," but we also discuss other cases. This is an interesting observable for a variety of reasons. First, five-dimensional quantum field theories are a fascinating arena with many counter-intuitive properties, and thus exact results can lead to valuable new insights. Although 5d gauge theories are infrared trivial, in many cases they are believed to arise as a relevant deformation of nontrivial UV SCFTs, which we may probe by computing suitable quantities protected under RG flows. In particular, there have been many such exact results for partition functions of these theories on closed manifolds; see [3, 4] for recent reviews and many examples and references. These results can be used to study interesting and subtle properties of these theories, such as the appearance of enhanced global symmetries at the SCFT point [5]; see, e.g., [6]. Then we expect the exact partition function on $S_b^3 \times \Sigma_{\mathfrak{g}}$ to lead to further probes into these 5d theories.

However, the observable $Z_{S_b^3 \times \Sigma_{\mathfrak{g}}}$ in particular is interesting because of its various connections to quantum field theories in other spacetime dimensions. In the remainder of this Introduction, we briefly review some of these features and summarize our main results.

The higher dimensional A-model and 2d TQFT. The first important connection is to two-dimensional QFT, and it is this connection which makes the computation of the partition function possible. Specifically, the background we consider involves a topological twist along $\Sigma_{\mathfrak{g}}$. Higher-dimensional theories compactified on $\Sigma_{\mathfrak{g}}$ with a partial topological twist have been well-studied in recent years [7–13]. These computations can be expressed as observables in a certain 2d topological quantum field theory (TQFT), namely the topological A-twist [14] of the effective theory obtained by compactification of the higher dimensional parent theory.

Here we take a similar approach, and study the effective 2d $\mathcal{N} = (2, 2)$ theory obtained by compactifying a 5d $\mathcal{N} = 1$ theory on $\mathcal{M}_3 \times \mathbb{R}^2$. As described in section 2, the result for $\mathcal{M}_3 = S_b^3$ takes the form of a sum over supersymmetric "Bethe vacua" of this 2d theory,[1]

$$Z_{S_b^3 \times \Sigma_{\mathfrak{g}}}(\nu)_{\mathfrak{n}} = \sum_{\hat{u} \in \mathcal{S}_{\mathrm{BE}}} \Pi_i(\hat{u}, \nu)^{\mathfrak{n}_i} \mathcal{H}(\hat{u}, \nu)^{\mathfrak{g}-1}, \tag{1.1}$$

where the objects

$$\Pi_i(u, \nu) = \exp\left(2\pi i \frac{\partial \mathcal{W}_{S_b^3 \times \mathbb{R}^2}}{\partial \nu_i}\right), \quad \mathcal{H}(u, \nu) = e^{2\pi i \Omega_{S_b^3 \times \mathbb{R}^2}} \det_{a,b} \frac{\partial^2 \mathcal{W}_{S_b^3 \times \mathbb{R}^2}}{\partial u_a \partial u_b}, \tag{1.2}$$

are refered to as the "flux operator" and "handle-gluing operator," respectively, and are built in terms of the *effective twisted superpotential*, $\mathcal{W}_{S_b^3 \times \mathbb{R}^2}$, and the *effective dilaton*,

---

[1]Here ν and $\mathfrak{n}$ are, respectively, supersymmetric mass parameters and gauge fluxes for background symmetries through $\Sigma_{\mathfrak{g}}$, as we describe in more detail in section 2 below.

$\Omega_{S_b^3 \times \mathbb{R}^2}$, controlling the low energy effective theory on the 2d Coulomb branch. Finally, the set of supersymmetric Bethe vacua of the theory, $\mathcal{S}_{\text{BE}}$, is defined as the solutions to certain Bethe-Ansatz-type equations, which may be written as

$$\mathcal{S}_{\text{BE}} = \left\{ \hat{u} \; \middle| \; \Pi_a(\hat{u}) \equiv \exp\left(2\pi i \frac{\partial \mathcal{W}_{S_b^3 \times \mathbb{R}^2}}{\partial u_a}(\hat{u}) \right) = 1, \quad a = 1, \ldots, r_G \right\} / W_G \,. \tag{1.3}$$

One novel feature of the five-dimensional computation, relative to lower-dimensional setups, is that the effective action controlling the 2d TQFT includes nontrivial non-perturbative corrections. That is, we may write the effective twisted superpotential as

$$\mathcal{W}_{S_b^3 \times \mathbb{R}^2}(u, \nu, \gamma) = \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\text{pert}}(u, \nu, \gamma) + \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\text{inst}}(u, \nu, \gamma) \,, \tag{1.4}$$

and similarly for the effective dilaton $\Omega_{S_b^3 \times \mathbb{R}^2}$. In particular, these depend on the gauge coupling, g_5, through a parameter $\gamma = -\frac{\pi(b+b^{-1})}{g_5^2}$, with the perturbative piece dominating for small g_5.

Let us first state our result for the perturbative contribution, for definiteness. Consider a 5d $\mathcal{N} = 1$ gauge theory with gauge group G, hypermultiplets in a representation $R = \bigoplus_i R_i$, and with a 5d CS term corresponding to the gauge-invariant functional $\text{Tr}_{CS}(\cdot)$ on the Lie algebra of G. Then we may write the various operators above to perturbative accuracy as:

$$\left(\Pi_a^{\text{pert}}\right)^{\mathfrak{m}_a} \left(\Pi_i^{\text{pert}}\right)^{\mathfrak{n}_i} e^{2\pi i(\mathfrak{g}-1)\Omega^{\text{pert}}} = e^{-\pi i \text{Tr}_{CS}(\mathfrak{m}u^2) + 2\pi i \gamma \text{Tr}(\mathfrak{m}u)} \prod_{\substack{\alpha \in Ad(G) \\ \alpha \neq 0}} s_b(\alpha(u) - iQ)^{-\alpha(\mathfrak{m})+1-\mathfrak{g}}$$

$$\times \prod_i \prod_{\rho \in R_i} s_b(\rho(u) + \nu_i)^{\rho(\mathfrak{m}) + \mathfrak{n}_i + (r_i-1)(\mathfrak{g}-1)} \,, \tag{1.5}$$

which are expressed in terms of the double sine function [15, 16], and we define $Q = \frac{1}{2}(b + b^{-1})$.[2] The above result is very reminiscent of the integrand of the S_b^3 partition function of a 3d $\mathcal{N} = 2$ theory [16], and we comment on this relation below.

However, as mentioned above, there are also non-perturbative contributions from instantons. To compute these, we take a detour into equivariant localization on the space $S_b^3 \times S_\epsilon^2$. This is an uplift of the $S_{\epsilon_1}^2 \times S_{\epsilon_2}^2$ partition function of [17], which receives contributions from point-like instantons. This allows us to extract the non-perturbative contributions to the effective action. We find that the full twisted superpotential and dilaton can be written in terms of the Nekrasov-Shatashvili (NS) limit of the 5d instanton partition function [18]. Namely, the latter has the following asymptotic behavior as we take one of the equivariant parameters to zero [19]:

$$Z_{\mathbb{R}^2_{\mathfrak{q}_1 = e^{2\pi i \epsilon_1}} \times \mathbb{R}^2_{\mathfrak{q}_2 = e^{2\pi i \epsilon_2}} \times S^1} \left(x = e^{2\pi i \tilde{u}}, y = e^{2\pi i \tilde{\nu}}, z = e^{2\pi i \tilde{\gamma}} \right)$$

$$\underset{\epsilon_2 \to 0}{\longrightarrow} \exp\left\{ 2\pi i \left(\frac{1}{\epsilon_2} \mathcal{W}_{NS}^{(5d)}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; \epsilon_1) - \Omega_{NS}^{(5d)}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; \epsilon_1) + O(\epsilon_2) \right) \right\} \,. \tag{1.6}$$

---

[2]To be precise, this result is expressed in the variables familiar from the S_b^3 partition function, while the variables more natural for the A-model on $S_b^3 \times \mathbb{R}^2$ turn out to be $\tilde{u} = iQ^{-1}u$, $\tilde{\nu} = iQ^{-1}\nu$, and $\tilde{\gamma} = iQ^{-1}\gamma$, as we explain below.

Then we may write the full, non-perturbative twisted superpotential and dilaton of the theory on $S_b^3 \times \mathbb{R}^2$ as:

$$\mathcal{W}_{S_b^3 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = \frac{1}{Qb} \mathcal{W}_{NS}^{(5d)}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; -b^2) + \frac{1}{Qb^{-1}} \mathcal{W}_{NS}^{(5d)}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; -b^{-2}), \qquad (1.7)$$

$$\Omega_{S_b^3 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = \Omega_{NS}^{(5d)}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; -b^2) + \Omega_{NS}^{(5d)}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; -b^{-2}). \qquad (1.8)$$

Although in principle this gives the complete answer, in practice it is difficult to compute the non-perturbative corrections in a useful form, and so we will mainly focus on various simplifying limits where the instanton contributions can be explicitly characterized.

Finally, we comment that the form of (1.7) as a sum of two contributions is closely related to the factorization of the S_b^3 partition function of 3d $\mathcal{N} = 2$ theories into two holomorphic blocks [20], which are associated to the solid tori in the Heegard decomposition of S^3. Our result then naturally generalizes to the case of $\mathcal{M}_3 \times \Sigma_{\mathfrak{g}}$ for arbitrary lens spaces, $\mathcal{M}_3 = L(p, q)$. In section 5 we briefly comment on this generalization in the case of $\mathcal{M}_3 = S^2 \times S^1$, with a topological twist on S^2, and point out an interesting relation to the 5d prepotential.

Large N limit and universal RG flows to 3d. As we discuss in detail in section 3, the large N analysis of the matrix model computing the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function reveals an interesting structure. In particular, we will show that to leading order in N the twisted superpotential $\mathcal{W}$ described above as well as the free energy, $F = -\text{Re} \log Z$, are specified by the S^5 partition function of the same theory. In particular, we find the large N relation

$$F_{S^3 \times \Sigma_{\mathfrak{g}}} = -\frac{8}{9}(\mathfrak{g} - 1)Q^2 \, F_\kappa(\mathfrak{n}) F_{S^5}, \qquad (1.9)$$

where the function $F_\kappa(\mathfrak{n})$ is given explicitly in (3.45) and generically depends on the theory under consideration. The case of the universal twist, however, is special [21]. This corresponds to a topological twist along the exact superconformal R-symmetry in the UV, in this case the Cartan of $\text{SU}(2)_R$, which amounts to setting $\mathfrak{n} = 0$. In this case, and setting $\kappa = -1$, the relation above becomes universal:

$$F_{S^3 \times \Sigma_{\mathfrak{g}}}^{\text{univ}} = -\frac{8}{9}(\mathfrak{g} - 1)Q^2 F_{S^5}. \qquad (1.10)$$

These large N results have an interesting interplay with holography. Indeed, the relation above was predicted in [21] (for the round sphere, $Q = 1$) to hold for any 5d $\mathcal{N} = 1$ theory with an AdS$_6$ dual using properties of 6d supergravity. Our result is a field theory derivation of this relation.

For nonzero flavor fluxes, the relation is no longer universal and one must consider specific theories. The holographic description of the twisted compactification of the Seiberg theory with flavor flux was recently constructed in [22]. Specifying the above formula to this case we recover the results of this reference as well.

Similar universal relations hold for the case $\Sigma_{\mathfrak{g}_1} \times \Sigma_{\mathfrak{g}_2} \times S^1$, which is described holographically by black holes in asymptotically locally AdS$_6$. We discuss this case in section 5.

The interplay with holography goes both ways. While the twisted partition functions can be computed exactly by the localization methods we describe, these are not always well suited to answering dynamical questions, such as the existence of interacting fixed points in the IR. The explicit construction of holographic RG flows such as the ones described above then indicate the existence of interacting IR fixed points, at least at large N.

5d $\longrightarrow$ 6d $\longrightarrow$ 4d. One of the surprising features of QFTs in five dimensions is that there are examples of 5d effective theories whose ultraviolet completion is not itself a 5d QFT, but rather a 6d theory where one of the dimensions has been compactified on a circle, S_β^1, with radius β. Here one identifies, roughly,

$$\beta \sim g_5^2 \,. \tag{1.11}$$

In particular, the KK excitations of this compactified theory, with action proportional to β^{-1}, can be naturally identified with instanton configurations in the low energy 5d gauge theory. The prototypical example of this phenomenon is the maximally supersymmetric $\mathcal{N} = 2$ super Yang-Mills theory in 5d for a gauge group of ADE type, which is believed to be equivalent to the S^1 compactification of the $\mathcal{N} = (2,0)$ SCFT associated to the corresponding Lie algebra.

Given the above discussion, we expect that the 5d partition function on $S_b^3 \times \Sigma_{\mathfrak{g}}$ computes the partition function of the corresponding 6d UV theory on $S_b^3 \times \Sigma_{\mathfrak{g}} \times S_\beta^1$. Then, compactifying the 6d theory on $\Sigma_{\mathfrak{g}}$ with flavor fluxes $\mathfrak{n}$, we may obtain a 4d theory, $\mathcal{T}_{\Sigma_{\mathfrak{g}},\mathfrak{n}}^{(4d)}$, and then we may also interpret this partition function as computing its $S_b^3 \times S_\beta^1$ partition function, better known as the "supersymmetric index",[3] of this theory,

$$Z_{S_b^3 \times \Sigma_{\mathfrak{g}}}(\nu, \gamma)_{\mathfrak{n}}[\mathcal{T}^{(5d)}] = Z_{S_b^3 \times S_\beta^1}(p, q, \mu)[\mathcal{T}_{\Sigma_{\mathfrak{g}},\mathfrak{n}}^{(4d)}] \,. \tag{1.12}$$

Here the precise mapping of parameters is given in (4.7) below.

There has been much work on understanding the compactification of 6d SCFTs on a compact Riemann surface, and the 4d $\mathcal{N} = 1$ and $\mathcal{N} = 2$ SQFTs that one obtains as a result. This began with the work of Gaiotto on the compactification of the A-type $\mathcal{N} = (2,0)$ theory on a punctured Riemann surface, with a twist preserving $\mathcal{N} = 2$ supersymmetry in 4d, leading to the celebrated theories of class $\mathcal{S}$ [23]. Subsequently this has been generalized in many directions, including compactifications of D and E-type $\mathcal{N} = (2,0)$ SCFTs, compactifications preserving only 4d $\mathcal{N} = 1$ SUSY [24, 25], and, more recently, compactifications of 6d $\mathcal{N} = (1,0)$ theories, leading to new classes of 4d $\mathcal{N} = 1$ theories and dualities [26, 27]. Many of these 4d theories have the property that they do not have known Lagrangian descriptions, and so more indirect methods are required for studying their properties and computing observables, such as their supersymmetric index.

However, these 5d QFTs, which are believed to be low energy descriptions of the 6d SCFTs on a circle, typically do have Lagrangian descriptions, and so we may, in principle, directly compute their $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function by the methods outlined above. In

---

[3]More precisely, the partition function and index differ by an overall factor which can be identified with the Casimir energy of the vacuum state of the theory on $S_b^3 \times \mathbb{R}$.

practice, we will perform these computations in certain simplifying limits, namely, large N, and large gauge coupling. By (1.11), the latter limit corresponds to large radius β. In the 4d index, the large β limit is dominated by the "Casimir energy" of the vacuum state of the 4d theory quantized on S_b^3. In these limits the instanton contributions are under better control, and so analytic evaluation of the index is possible. In addition, in the case of the maximal $\mathcal{N} = 2$ SYM theory, there is a special limit of parameters with additional supersymmetry, and where the instanton contributions are greatly simplified, and we will also be able to evaluate the index analytically in this limit. This turns out to be closely related to the "Schur" [28] and "mixed Schur" [29] limit of the superconformal index.

In all of these cases, we will find strong consistency checks of our computations by comparing to expected properties of the 4d index of the above theories. We stress that our computation in principle gives a new approach to computing the index of 4d theories without a known Lagrangian description. The form of the computation is in terms of the 2d TQFT dual to the 4d index of these theories, and so manifestly exhibits various 4d dualities, such as S-duality.

Outline. The paper is organized as follows. In section 2 we provide the derivation of the exact partition function by localization methods, as described above. In section 3 we study the large N limit of the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function, its relation to the partition function on S^5 in this limit, and discuss universal RG flows and holography. In section 4 we consider theories with a 6d UV completion and compute the 4d superconformal index of theories obtained by reduction on the Riemann surface. Finally, in section 5 we discuss the partition function on more general manifolds, including $S^2 \times S^1 \times \Sigma_{\mathfrak{g}}$, and discuss the counting of black hole microstates in AdS_6.

Note added. In the final stages of this work we learned about the work [30], which has some overlap with our discussion in section 5.

2 Derivation of partition function on $S_b^3 \times \Sigma_{\mathfrak{g}}$

In this section, we derive the exact partition function of 5d $\mathcal{N} = 1$ gauge theories on $S_b^3 \times \Sigma_{\mathfrak{g}}$, with a partial topological twist on $\Sigma_{\mathfrak{g}}$. We discuss different ways of carrying out the computation, which are complementary to each other and offer different points of view on this observable.

2.1 5d $\mathcal{N} = 1$ theories on curved backgrounds

Let us first review some basic properties of 5d $\mathcal{N} = 1$ supersymmetric gauge theories in flat, Euclidean space, in preparation for studying them on curved backgrounds. The 5d $\mathcal{N} = 1$ Yang-Mills action can be obtained by dimensional reduction from 6d $\mathcal{N} = 1$ on $\mathbb{R}^{5,1}$, and is given by [3]

$$S = \frac{1}{g_5^2} \mathrm{Tr} \int d^5x \left(\frac{1}{2} F^{\mu\nu} F_{\mu\nu} + i\Lambda_I \slashed{D} \Lambda^I - D^\mu \sigma D_\mu \sigma - \Lambda_I [\sigma, \Lambda^I] - \frac{1}{2} D^{IJ} D_{IJ} \right). \quad (2.1)$$

Here $F = dA + i[A, A]$ is the gauge field strength, σ is a real scalar, Λ_I is the symplectic Majorana gaugino, and D_{IJ} is an auxiliary scalar field. The indices $I, J = 1, 2$ correspond to the SU(2)$_R$ symmetry, and we impose the symplectic Majorana condition,

$$\left(\Lambda_\alpha^I\right)^* = \epsilon_{IJ}\omega^{\alpha\beta}\Lambda_\beta^J, \tag{2.2}$$

where $\alpha = 1, \ldots, 4$ is the spinor index, which corresponds to the fundamental representation of USp(4) $\cong$ Spin(5).

As will be important below, these theories are closely related to 4d $\mathcal{N} = 2$ theories, which are obtained upon dimensional reduction, and as for 4d $\mathcal{N} = 2$ theories their actions can be written in terms of a holomorphic *prepotential*, $\mathcal{F}(\mathcal{A})$. In a notation adapted to reduction to 4d, this can be written as

$$S = \frac{1}{4\pi}\text{Im}\left(\frac{1}{2}\int d^2\theta d^5 x \frac{\partial^2\mathcal{F}}{\partial\mathcal{A}^i\partial\mathcal{A}^j}W^i W^j + \int d^4\theta d^5 x \bar{\mathcal{A}}^i\frac{\partial\mathcal{F}}{\partial\mathcal{A}^i}\right), \tag{2.3}$$

where W^i and $\mathcal{A}^i$ are the usual 4d $\mathcal{N} = 1$ chiral superfields making up the 4d $\mathcal{N} = 2$ vector multiplet, and the complex scalar in the 4d vector multiplet can be identified with

$$a = A_5 + i\sigma, \tag{2.4}$$

where A_5 is the fifth component of the 5d gauge field. Then the most general action in 5d can be written in terms of a cubic prepotential [5, 31],

$$\mathcal{F}(\mathcal{A}) = \frac{1}{2g_5{}^2}\text{Tr}\,\mathcal{A}^2 + \frac{1}{6}c\,\text{Tr}\,\mathcal{A}^3, \tag{2.5}$$

where g_5 is the 5d gauge coupling, and the cubic term defines a 5d Chern-Simons term, which may be included for certain choices of gauge group G.

In addition to the vector multiplet, we may include hypermultiplets. Their field content consists of complex scalars q_I^A and spinors ψ^A, where I is the SU(2)$_R$ index, as above, and A is a gauge or flavor index, for which we assume the matter is in a psueudoreal representation, $R = \bigoplus_i R_i$. These are taken to satisfy reality conditions,

$$\left(q_I^A\right)^* = \Omega_{AB}\epsilon^{IJ}q_J^B, \quad \left(\psi_\alpha^A\right)^* = \Omega_{AB}\omega^{\alpha\beta}\psi_\beta^B. \tag{2.6}$$

These give rise to 4d $\mathcal{N} = 2$ hypermultiplets upon dimensional reduction. We may also turn on background vector multiplets coupled to flavor symmetries, and including a background value, m_i, for the real scalar in these multiplets gives a supersymmetric real mass for hypermultiplets charged under this symmetry. Integrating out these massive hypermultiplets gives rise to a simple correction to the prepotential (2.5) [5, 31],

$$\mathcal{F}_{\text{eff}} = \frac{1}{2g_5{}^2}\text{Tr}\,\mathcal{A}^2 + \frac{1}{6}c\,\text{Tr}\,\mathcal{A}^3 - \frac{1}{12}\sum_i\sum_{\rho\in R_o}|\rho(\mathcal{A}) + m_i|^3 + \frac{1}{12}\sum_{\alpha\in Ad(G)}|\alpha(\mathcal{A})|^3$$
$$\approx \frac{1}{2}t_{ab}\mathcal{A}^a\mathcal{A}^b + \frac{1}{6}c_{abc}\mathcal{A}^a\mathcal{A}^b\mathcal{A}^c + \cdots \tag{2.7}$$

where the latter expansion may be made for different asymptotic directions on the Coulomb branch, and defines the effective gauge coupling and Chern-Simons terms which are generated there.

5d $\mathcal{N} = 1$ supersymmetry on curved spacetimes. There has been much work on placing 5d supersymmetric theories on curved spacetimes and computing their partition functions by localization. Some examples include S^5 [32–34], $\mathbb{CP}^2 \times S^1$ [35], $Y^{p,q}$ [36], and $S^4 \times S^1$ [6]; see [3, 4] for recent reviews, and many additional references. In addition, maximally supersymmetric 5d Yang-Mills theory was studied on $S^3 \times \Sigma_{\mathfrak{g}}$ in [37–39], although our result appears to differ from theirs.[4]

To write supersymmetric actions on such manifolds, we can employ the philosophy of [40] and consider a rigid limit of 5d $\mathcal{N} = 1$ supergravity. Such an approach was studied in [41–43]. In particular, in [43] it was shown that we may define a supersymmetric background for a 5d $\mathcal{N} = 1$ theory on any manifold, $\mathcal{M}_5$, admitting a transversally Hermitian structure. This means we may write a metric on $\mathcal{M}_5$ of the form

$$ds^2_{M_5} = S^2(d\psi + \rho)^2 + ds^2_4 \,. \tag{2.8}$$

Here ψ is a coordinate generating an isometry of the metric, i.e., $K = \partial_\psi$ is a Killing vector. The transverse directions admit a complex structure, and ds^2_4 is a Hermitian metric. With this structure, we may find solutions to the Killing spinor equations, which allow us to write actions on $\mathcal{M}_5$ preserving some supersymmetry.

Let us note that the above classification is very similar to that of [44] for coupling 3d $\mathcal{N} = 2$ supersymmetric theories to three-manifolds. Namely, in that case the necessary structure is a transversally holomorphic foliation. Then it is straightforward to check that, given a three-manifold $\mathcal{M}_3$ with background gravity fields, we may define a supersymmetric background on $\mathcal{M}_3 \times \mathbb{R}^2$ with

$$\zeta^{(5d)} = \zeta^{(3d)} \otimes \epsilon \,, \tag{2.9}$$

for any constant ϵ. In particular, in the case where $\mathcal{M}_3$ is a Seifert manifold [44], we may preserve one supercharge, $\zeta^{(3d)}$, of R-charge 1 and another, $\tilde{\zeta}^{(3d)}$, of R-charge -1, and so on $\mathcal{M}_3 \times \mathbb{R}^2$ we may preserve four supercharges.

Let us consider the algebra satisfied by the supercharges. First, in 3d, we have [44]

$$\{\delta_\zeta, \delta_\zeta\} = \{\delta_{\tilde\zeta}, \delta_{\tilde\zeta}\} = 0 \,,$$
$$\{\delta_\zeta, \delta_{\tilde\zeta}\} = -2i\mathcal{L}'_K + \zeta\tilde\zeta(z - \Delta H) \,, \tag{2.10}$$

where z is the central charge of the field being acted on, Δ is its R-charge, H is a scalar in the background supergravity multiplet, and $\mathcal{L}'_K$ is the modified Lie derivative, which is covariant under R- and central-charge symmetry gauge transformations, acting along the Killing vector,

$$K^\mu = \zeta\gamma^\mu\tilde\zeta \,. \tag{2.11}$$

Returning to the 5d superalgebra, let us define $\delta_\pm$ and $\tilde\delta_\pm$ as the transformations associated to $\zeta \otimes \epsilon_\pm$ and $\tilde\zeta \otimes \epsilon_\pm$, respectively, where we define $\epsilon_\pm$ as the ± 1 eigenvectors of σ_3. Then,

---

[4]Specifically, we find additional contributions from instantons and fermionic zero modes relative to their computation.

for example, we may compute

$$\{\delta_+, \delta_-\} = (\epsilon_+\epsilon_-)(-2i\mathcal{L}'_K + \zeta\tilde{\zeta}(z - \Delta H)) + (\zeta\zeta)(\epsilon_+\sigma_i\epsilon_-) = 0, \tag{2.12}$$

$$\{\delta_+, \tilde{\delta}_+\} = (\zeta\tilde{\zeta})(\epsilon_+\sigma_i\epsilon_+) = 2i\partial_z, \tag{2.13}$$

$$\{\delta_+, \tilde{\delta}_-\} = (\epsilon_+\epsilon_-)(-2i\mathcal{L}'_K + \zeta\tilde{\zeta}(z - \Delta H)) + (\zeta\tilde{\zeta})(\epsilon_+\sigma_i\epsilon_+) = -2i\mathcal{L}'_K + (z - \Delta H). \tag{2.14}$$

We can repeat this for the remaining supercharges. Then, if we identify $\mathcal{Q}_+ = \delta_+$, $\mathcal{Q}_- = \tilde{\delta}_-$, $\tilde{\mathcal{Q}}_+ = \tilde{\delta}_+$, and $\tilde{\mathcal{Q}}_- = \delta_-$, one can check that $\mathcal{Q}_\alpha$ and $\tilde{\mathcal{Q}}_\alpha$ generate the $\mathcal{N} = (2,2)$ superalgebra, with a central charge given by the operator

$$\tilde{Z} = -2i\mathcal{L}'_K + (z - \Delta H). \tag{2.15}$$

In other words, the 5d $\mathcal{N} = 1$ theory compactified on $\mathcal{M}_3$ gives rise to an effective 2d $\mathcal{N} = (2,2)$ theory, whose central charge is determined by the operator $\tilde{Z}$ above, which depends on the KK momentum along $\mathcal{M}_3$, R-charge, and real mass parameters, which contribute to the central charge, z.

Finally, to construct a background on the compact manifold, $\mathcal{M}_3 \times \Sigma_{\mathfrak{g}}$, we may perform a partial topological twist along $\Sigma_{\mathfrak{g}}$. From the 5d point of view, this amounts to turning on a flux, $\mathfrak{g} - 1$, on $\Sigma_{\mathfrak{g}}$ for the background gauge field coupled to the $U(1)_R \subset SU(2)_R$ symmetry. From the point of view of the effective 2d $\mathcal{N} = (2,2)$ theory, this is simply the topological A-twist [14]. We then expect the partition function to be an observable in an appropriate 2d topological quantum field theory (TQFT), which constrains the form of the answer. We review this structure in the next subsection.

In the rest of the paper, we will mostly focus on the case $\mathcal{M}_3 = S_b^3$.[5] Specifically, we will take the supersymmetric background of [45], which exhibits S_b^3 as an S^1 fiber bundle over the round metric on S^2, i.e.,

$$ds_3^2 = \tilde{\beta}^2(d\psi + a)^2 + d\Omega_{S^2}^2. \tag{2.16}$$

This preserves an $SU(2) \times U(1)$ isometry, and the Killing vector appearing in the superalgebra is $K = \partial_\psi$. Then, in this background we have $H = \frac{i}{2}(b + b^{-1}) \equiv iQ$ [44], and so the central charge in the $\mathcal{N} = (2,2)$ algebra (2.15) is

$$\begin{aligned}
\tilde{Z} &= -2i\mathcal{L}_{\partial_\psi} + (z + A_\psi - i\Delta Q) \\
&\equiv -2i\mathcal{L}_{\partial_\psi} + i\sigma,
\end{aligned} \tag{2.17}$$

which defines a variable, σ, valued in the complexification of the Cartan subalgebra of the flavor symmetry, which is a holomorphic combination of the background real scalar, z, and component, A_ψ, of the background gauge field along the direction K, as discussed in more detail in [11].

2.2 The 2d A-twist and Bethe Ansatz equations

As observed above, the supersymmetric background on $S_b^3 \times \Sigma_{\mathfrak{g}}$, when considered along the $\Sigma_{\mathfrak{g}}$ directions, takes the form of a 2d topological twist. Such setups, where a d-dimensional

---

[5]We will return to consider more general lens spaces in section 5.

gauge theory is placed on a manifold of the form $\mathcal{M}_{d-2} \times \Sigma_{\mathfrak{g}}$, and subjected to a topological twist along the Riemann surface, have been studied in many examples recently — see, e.g., [7–13]. These works are related to the gauge-Bethe correspondence of [46], and can be described as a "higher dimensional A-twist" [11].

On general grounds, we expect such an observable to be computed by an appropriate 2d TQFT, which tightly constrains the form the result may take. In fact, by studying the effective action of the low energy theory compactified to 2d, we may express the full answer in terms of two functions, the "effective twisted superpotential," $\mathcal{W}(\sigma)$ and the "effective dilaton," $\Omega(\sigma)$, which are functions of the twisted chiral field strength multiplet, Σ, associated to the 2d vector multiplet, $\mathcal{V}$,

$$\Sigma = -iD_-\tilde{D}_+\mathcal{V} = \sigma + i\sqrt{2}(\theta^+\tilde{\Lambda}_+ - \tilde{\theta}^-\Lambda_-) + \sqrt{2}\theta^+\tilde{\theta}^-(D - iF_{12}) + \cdots. \tag{2.18}$$

Then the effective action can be written in terms of these objects as [7]

$$S = \int d^2x \sqrt{g}\left(-2F_{12}^a\frac{\partial\mathcal{W}(\sigma)}{\partial\sigma_a} + \tilde{\Lambda}^a\Lambda^b\frac{\partial^2\mathcal{W}(\sigma)}{\partial\sigma_a\partial\sigma_b} + \frac{i}{2}\Omega(\sigma)R\right) + \mathcal{Q}(\cdots). \tag{2.19}$$

Here, σ may denote both dynamical and background vector multiplets associated to the gauge group, G, and flavor group, G_F, respectively, expanded in a Cartan basis. For the rest of the paper, we set the notation

$$\sigma \rightarrow \{u_a, \quad a = 1, \ldots, r_G, \qquad \nu_i, \quad i = 1, \ldots, r_{G_F}\}, \tag{2.20}$$

defining the gauge and flavor symmetry parameters, u and ν, respectively.

Using this low energy action, we may construct the partition function of this effective theory on a Riemann surface $\Sigma_{\mathfrak{g}}$ with the topological A-twist background, where we include twisted masses ν and background magnetic fluxes $\mathfrak{n}$ for the flavor symmetry. We find [7, 11]

$$Z_{\mathcal{M}_{d-2}\times\Sigma_{\mathfrak{g}}}(\nu)_{\mathfrak{n}} = \sum_{\hat{u}\in\mathcal{S}_{\mathrm{BE}}} \Pi_i(\hat{u}, \nu)^{\mathfrak{n}_i}\mathcal{H}(\hat{u}, \nu)^{\mathfrak{g}-1}, \tag{2.21}$$

where

$$\Pi_i(u, \nu) = \exp\left(2\pi i\frac{\partial\mathcal{W}}{\partial\nu_i}(u, \nu)\right), \quad \mathcal{H}(u, \nu) = e^{2\pi i\Omega(u,\nu)}\det_{a,b}\frac{\partial^2\mathcal{W}(u,\nu)}{\partial u_a\partial u_b}, \tag{2.22}$$

which we refer to as the "flux operator" and "handle-gluing operator," respectively, and $\mathcal{S}_{\mathrm{BE}}$ is the set of supersymmetric "Bethe vacua" of the theory, defined by

$$\mathcal{S}_{\mathrm{BE}} = \left\{\hat{u} \mid \Pi_a(\hat{u}, \nu) \equiv \exp\left(2\pi i\frac{\partial\mathcal{W}}{\partial u_a}(\hat{u}, \nu)\right) = 1, \quad a = 1, \ldots, r_G\right\}/W_G, \tag{2.23}$$

where we quotient by the action of the Weyl group, W_G, of G and discard any solutions on which it does not act freely.

This result may often be alternatively derived by direct UV localization. There, one arrives at an expression of the form

$$Z_{\mathcal{M}_{d-2}\times\Sigma_{\mathfrak{g}}}(\nu)_{\mathfrak{n}} = \frac{1}{|W_G|}\sum_{\mathfrak{m}\in\Lambda_G}\oint_{\mathcal{C}_{JK}} du\, \Pi_i(u, \nu)^{\mathfrak{n}_i}\Pi_a(u, \nu)^{\mathfrak{m}_a}\mathcal{H}(u, \nu)^{\mathfrak{g}}e^{-2\pi i\Omega(u,\nu)}, \tag{2.24}$$

where we sum over gauge fluxes, $\mathfrak{m}$, in the lattice, Λ_G, of coweights of G, and $\mathcal{C}_{JK}$ is the so-called "Jeffrey-Kirwan contour" [47, 48]. Roughly speaking, we may define this contour by removing a small neighborhood of any singularities or boundaries at infinity in the integrand, leaving a non-singular region, $\hat{\mathcal{M}}$, of the complexified Cartan of G, and then the contour runs over those portions of $\partial\hat{\mathcal{M}}$ with

$$\text{sgn Im}(\partial_a \mathcal{W}(u,\nu)) = -\,\text{sgn}(\eta_a)\,, \qquad (2.25)$$

where η_a is an auxiliary parameter valued in the Cartan of the gauge group. In addition, we discard contributions from the neighborhood of points with enhanced Weyl symmetry. Then, if we choose $\eta_a \propto \mathfrak{m}_a$ with a positive constant, one can check the sum over $\mathfrak{m}_a$ is a convergent geometric series on the JK contour, and after performing it we find

$$Z_{\mathcal{M}_{d-2}\times\Sigma_{\mathfrak{g}}}(\nu)_{\mathfrak{n}} = \frac{1}{|W_G|}\oint_{\partial\hat{\mathcal{M}}} du\,\Pi_i(u,\nu)^{\mathfrak{n}_i}\prod_a \frac{1}{1-\Pi_a(u,\nu)}\mathcal{H}(u,\nu)^{\mathfrak{g}}e^{-2\pi i\Omega(u,\nu)}\,. \qquad (2.26)$$

The poles in this integrand are precisely at the solutions to (2.23) (appearing with multiplicity $|W_G|$, which cancels the prefactor), and taking their residues we recover the formula (2.21), demonstrating the equivalence of these two approaches. We refer to [11] for more details on this argument and the JK contour.

Thus, to derive the partition function on $\mathcal{M}_{d-2}\times\Sigma_{\mathfrak{g}}$ for general $\Sigma_{\mathfrak{g}}$, it suffices to compute the effective objects, $\mathcal{W}(u,\nu)$ and $\Omega(u,\nu)$, describing the low energy theory obtained after compactifying on $\mathcal{M}_{d-2}$, which in the case of present interest is S_b^3. In the next subsection we attempt to carry out this procedure directly by expanding the 5d fields in KK modes on S_b^3. However, the result we obtain in this way turns out to encode only the perturbative contribution to $Z_{S_b^3\times\Sigma_{\mathfrak{g}}}$. We then describe, in section 2.4, another approach which captures the full non-perturbative twisted superpotential and dilaton.

2.3 Reduction on S_b^3 and the perturbative partition function

Let us consider a 5d $\mathcal{N}=1$ theory on $S_b^3\times\mathbb{R}^2$. As discussed above, this may be described by an effective 2d $\mathcal{N}=(2,2)$ theory with infinitely many fields, arising from the KK modes on S_b^3. We begin by describing this in the case of a free 5d $\mathcal{N}=1$ hypermultiplet.

Hypermultiplet. To write the effective twisted superpotential and dilaton generated by a hypermultiplet, we will first need to describe how its modes on $S_b^3\times\mathbb{R}^2$ decompose into 2d fields on $\mathbb{R}^2$. Recall that the dimensional reduction of a 5d hypermultiplet to three dimensions is a 3d $\mathcal{N}=4$ hypermultiplet, or equivalently, a pair of 3d $\mathcal{N}=2$ chiral multiplets in conjugate representations. Thus, if we restrict to a point, x, on $\mathbb{R}^2$, the field content of the 5d hypermultiplet on the three transverse directions is that of a 3d $\mathcal{N}=4$ hypermultiplet. We may then understand the 2d field content by expanding this 3d hypermultiplet in modes on S_b^3.

Round sphere. For simplicity, let us first consider the round sphere, $b=1$. Following [49], we may decompose a chiral multiplet on S^3 into modes via

$$\phi = \sum_{\ell,m,n}\phi_{\ell,m,n}Y_{\ell,m,n},\; \psi = \sum_{\ell,m',n}\psi_{\ell,m',n}^+\chi_{\ell,m',n}^+ + \sum_{\ell,m,n'}\psi_{\ell,m,n'}^-\chi_{\ell,m,n'}^-,\; F = \sum_{\ell,m,n}F_{\ell,m,n}Y_{\ell,m,n},$$

$$(2.27)$$

where $Y_{\ell,m,n}$ run over the spherical harmonics on S^3, with $\ell \in \mathbb{Z}_{\geq 0}$ and $m, n \in \{-\frac{\ell}{2}, \ldots, \frac{\ell}{2}\}$ the angular momenta under the $\mathrm{SU}(2)_L \times \mathrm{SU}(2)_R \cong \mathrm{SO}(4)$ isometry group, and similarly for the spinor spherical harmonics, $\chi^{\pm}$.[6] Then these modes organize into the following short and long 0d $\mathcal{N} = 1$ multiplets of the action of the supercharges δ and $\tilde{\delta}$:

$$\text{short:} \qquad \left\{\phi_{\ell,\frac{\ell}{2},n}, \ \psi^+_{\ell,\frac{\ell+1}{2},n}\right\}, \quad \left\{\psi^-_{\ell,-\frac{\ell+1}{2},n}, \ F_{\ell,-\frac{\ell}{2},n}\right\}, \tag{2.28}$$

$$\text{long:} \qquad \left\{\phi_{\ell,m,n}, \ \psi^+_{\ell,m+\frac{1}{2},n}, \ \psi^-_{\ell,m+\frac{1}{2},n}, \ F_{\ell,m+1,n}\right\}, \qquad -\frac{\ell}{2} \leq m < \frac{\ell}{2}. \tag{2.29}$$

We expect contribution from long multiplets cancels out of protected observables, and so we will henceforth focus on the short multiplets.

For a 3d $\mathcal{N} = 4$ hypermultiplet, which consists of two 3d $\mathcal{N} = 2$ multiplets, Φ and $\tilde{\Phi}$ in conjugate gauge and flavor representations, we may organize the short $\mathcal{N} = 1$ multiplets into $\mathcal{N} = 2$ chiral multiplets, for $\ell \in \mathbb{Z}_{\geq 0}$ and $n \in \{-\frac{\ell}{2}, \ldots, \frac{\ell}{2}\}$,

$$\Phi^+_{\ell,n} \equiv \left\{\phi_{\ell,\frac{\ell}{2},n}, \ \psi^+_{\ell,\frac{\ell+1}{2},n}, \tilde{\psi}^-_{\ell,-\frac{\ell+1}{2},n}, \ \tilde{F}_{\ell,-\frac{\ell}{2},n}\right\}, \tag{2.30}$$

$$\Phi^-_{\ell,n} \equiv \left\{\tilde{\phi}_{\ell,\frac{\ell}{2},n}, \ \tilde{\psi}^+_{\ell,\frac{\ell+1}{2},n}, \psi^-_{\ell,-\frac{\ell+1}{2},n}, \ F_{\ell,-\frac{\ell}{2},n}\right\}. \tag{2.31}$$

Then one can check that, for a hypermultiplet with R-charge 1 coupled to a background gauge field with scalar $\tilde{\sigma}$, one has

$$[\delta, \tilde{\delta}]\Phi^{\pm}_{\ell,n} = (\ell + 1 \pm \tilde{\sigma})\, \Phi^{\pm}_{\ell,n}. \tag{2.32}$$

The above analysis was carried out at a single point, $x \in \mathbb{R}^2$. In general, we obtain 2d superfields $\Phi^{\pm}_{\ell,n}(x)$ which, as discussed above, form multiplets of the 2d $\mathcal{N} = (2,2)$ superalgebra. Specifically, it is easy to see that each mode gives rise to an independent 2d chiral multiplet, and from (2.32) and (2.17), we can see that these chiral multiplets have twisted masses

$$\tilde{m}^{\pm}_{\ell,n} = \ell + 1 \pm \tilde{\sigma}. \tag{2.33}$$

In addition, we have modes which contribute long multiplets in 2d, which have unprotected masses, and which do not contribute to protected observables.

We may now compute the effective twisted superpotential contributed by integrating out this compactified 5d hypermultiplet. Recall that the contribution of a single 2d chiral multiplet of twisted mass $\tilde{m}$ is [50][7]

$$\mathcal{W}_{\mathbb{R}^2}(\tilde{m}) = -\frac{1}{2\pi i}\tilde{m}(\log \tilde{m} - 1). \tag{2.34}$$

The full twisted superpotential from summing over all the short multiplets appearing in (2.33) is

$$\mathcal{W}^{\mathrm{hyp,U(1)}}_{S^3_{b=1} \times \mathbb{R}^2}(\tilde{\sigma}) = \sum_{\ell=0}^{\infty} \sum_{\pm} (\ell + 1) \mathcal{W}^{\Phi}_{\mathbb{R}^2}\left((\ell + 1) \pm \tilde{\sigma}\right) \equiv g_{b=1}(\tilde{\sigma}). \tag{2.35}$$

---

[6]Specifically, there m' and n' take values in $\{-\frac{\ell+1}{2}, \ldots, \frac{\ell+1}{2}\}$, representing the decomposition of the spinors into representations $\bigoplus_{\ell \geq 0}(\ell, \ell + \frac{1}{2}) \oplus (\ell + \frac{1}{2}, \ell)$ of $\mathrm{SU}(2) \times \mathrm{SU}(2)$.

[7]More precisely we should introduce a dynamical scale, μ, and write the logarithm as $\log\left(\frac{\tilde{m}}{\mu}\right)$, but this would drop out of the computation below.

After suitable regularization (see (2.42) below), this infinite sum may be explicitly evaluated to give

$$g_{b=1}(\tilde{\sigma}) = -2\frac{1}{(2\pi i)^3}\text{Li}_3(e^{2\pi i\tilde{\sigma}}) + \tilde{\sigma}\frac{1}{(2\pi i)^2}\text{Li}_2(e^{2\pi i\tilde{\sigma}}) + \frac{1}{24}\tilde{\sigma}(2\tilde{\sigma}^2 - 1)\,. \quad (2.36)$$

Squashed sphere. The above computation is generalized in a straightforward manner to S_b^3 for general b by expanding the fields in spherical harmonics, following [45].[8] One finds,

$$[\delta, \tilde{\delta}]\Phi_{\ell,n}^{\pm} = \left\{\ell + 1 + \frac{2(b-b^{-1})}{b+b^{-1}}n \pm \tilde{\sigma}\right\}\Phi_{\ell,n}^{\pm}\,. \quad (2.37)$$

A more general R-charge, Δ, can be obtained by shifting, in our convention, $\tilde{\sigma} \to \tilde{\sigma}+1-\Delta$.

As above, each of these modes corresponds to a 2d chiral multiplet, with twisted masses

$$\tilde{m}_{\ell,n}^{\pm} = \ell + 1 + \frac{2(b-b^{-1})}{b+b^{-1}}n \pm \tilde{\sigma}\,. \quad (2.38)$$

Then, the twisted superpotential obtained by integrating out these chirals reads

$$\mathcal{W}_{S_b^3\times\mathbb{R}^2}^{\text{hyp,U}(1)}(\tilde{\sigma}) = \sum_{\ell=0}^{\infty}\sum_{n=-\frac{\ell-1}{2}}^{\frac{\ell-1}{2}}\sum_{\pm}\mathcal{W}_{\mathbb{R}^2}^{\Phi}\left(\ell + 1 + \frac{2(b-b^{-1})}{b+b^{-1}}n \pm \tilde{\sigma}\right) \equiv g_b(\tilde{\sigma})\,, \quad (2.39)$$

which, after suitable regularization, defines a function $g_b(\tilde{\sigma})$. We regularize this sum as follows. Note that, using $e^{2\pi i\partial_{\tilde{\sigma}}\mathcal{W}_{\mathbb{R}^2}(\tilde{\sigma})} = \tilde{\sigma}^{-1}$, we may formally write

$$\exp\left(2\pi i\frac{\partial g_b(\tilde{\sigma})}{\partial\tilde{\sigma}}\right) = \prod_{\ell=0}^{\infty}\prod_{n=-\frac{\ell-1}{2}}^{\frac{\ell-1}{2}}\frac{\ell + 1 + \frac{2(b-b^{-1})}{b+b^{-1}}n - \tilde{\sigma}}{\ell + 1 + \frac{2(b-b^{-1})}{b+b^{-1}}n + \tilde{\sigma}}$$

$$= \prod_{j,k\geq 0}\frac{(j+\frac{1}{2})b + (k+\frac{1}{2})b^{-1} - Q\tilde{\sigma}}{(j+\frac{1}{2})b + (k+\frac{1}{2})b^{-1} + Q\tilde{\sigma}} \equiv s_b(-iQ\tilde{\sigma})\,, \quad (2.40)$$

where we have defined

$$Q = \frac{1}{2}(b + b^{-1})\,, \quad (2.41)$$

and $s_b(x)$ is the double sine function [15, 16], which can be rigorously defined as a meromorphic function of σ. We may then set

$$g_b(\tilde{\sigma}) = \frac{1}{2\pi i}\int d\tilde{\sigma}\log s_b(-iQ\tilde{\sigma})\,. \quad (2.42)$$

Next consider the effective dilaton. This depends on the choice of R-symmetry used to twist the theory on the Riemann surface, $\Sigma_{\mathfrak{g}}$, and the contribution of a hypermultiplet will depend on its charge, r, under this R-symmetry. We stress that r, which is in general

---

[8]Here we find it convenient to use a slightly non-standard normalization of the scalar $\tilde{\sigma}$. Our normalization is related to the usual scalar, σ, on S_b^3 by $\sigma = -iQ\tilde{\sigma}$, where Q is as in (2.41). We will comment more on this relation below.

integer-quantized due to the nontrivial flux, $\mathfrak{g} - 1$, through $\Sigma_{\mathfrak{g}}$ for the R-symmetry background gauge field, is distinct from the R-charge appearing in the coupling to S_b^3, which we denote Δ, and which in general is a real number. Then the contribution of a single 2d chiral multiplet to the effective dilaton is:

$$\Omega_{\mathbb{R}^2}(\tilde{m}) = -(r-1)\log\tilde{m}, \tag{2.43}$$

and proceeding as above, we find:

$$\Omega_{S_b^3\times\mathbb{R}^2}^{\text{hyp},\text{U}(1)}(\tilde{\sigma}) = (r-1)\ell_b(\tilde{\sigma}), \tag{2.44}$$

where we have defined

$$\ell_b(\tilde{\sigma}) \equiv \frac{1}{2\pi i}\log s_b(-iQ\tilde{\sigma}). \tag{2.45}$$

Below it will often be natural to take $r = 1$, as this is the R-charge obtained by twisting by the $\text{U}(1)_R \subset \text{SU}(2)_R$ symmetry appearing in the $\mathcal{N} = 1$ superalgebra, and in this case the hypermultiplet has vanishing contribution to the effective dilaton.

We now have all the elements we need to compactify the $\mathbb{R}^2$ factor to $\Sigma_{\mathfrak{g}}$ using the 2d A-twist. Namely, from (2.21) above we have, for a flux $\mathfrak{m}$ on $\Sigma_{\mathfrak{g}}$,

$$Z_{S_b^3\times\Sigma_{\mathfrak{g}}}^{\text{hyp},\text{U}(1)}(\tilde{\sigma})_{\mathfrak{m}} = \Pi(\tilde{\sigma})^{\mathfrak{m}}e^{2\pi i(\mathfrak{g}-1)\Omega(\tilde{\sigma})} = s_b(-iQ\tilde{\sigma})^{\mathfrak{m}+(r-1)(\mathfrak{g}-1)}. \tag{2.46}$$

General gauge theory. It is straightforward to generalize this result to a set of hypermultiplets living in a representation $R = \oplus_i R_i$ of a gauge group G. As above we take $\tilde{\sigma} \to \tilde{u}$ and $\tilde{\nu}$ for the gauge and flavor symmetry (real mass) parameters, respectively. Then we have

$$\mathcal{W}_{S_b^3\times\mathbb{R}^2}^{hyp}(\tilde{u},\tilde{\nu}) = \sum_i\sum_{\rho\in R_i}g_b(\rho(\tilde{u}) + \tilde{\nu}_i), \tag{2.47}$$

$$\Omega_{S_b^3\times\mathbb{R}^2}^{hyp}(\tilde{u},\tilde{\nu}) = \sum_i\sum_{\rho\in R_i}(r_i - 1)\ell_b(\rho(\tilde{u}) + \tilde{\nu}_i), \tag{2.48}$$

where we have introduced masses $\tilde{\nu}_i$ and R-charges r_i for the hypermultiplets. In addition to the hypermultiplets, we expect a contribution to the 2d A-model from the vector multiplets and the classical action. For the former, we write

$$\mathcal{W}_{S_b^3\times\mathbb{R}^2}^{vec}(\tilde{u}) = -\sum_{\alpha\in Ad(G)'}g_b(\alpha(\tilde{u}) + 1), \tag{2.49}$$

$$\Omega_{S_b^3\times\mathbb{R}^2}^{vec}(\tilde{u}) = -\sum_{\alpha\in Ad(G)'}\ell_b(\alpha(\tilde{u}) + 1)$$

$$= -\sum_{\alpha>0}\frac{1}{2\pi i}\log\left[2\sin\left(\pi b Q\alpha(\tilde{u})\right)2\sin\left(\pi b^{-1}Q\alpha(\tilde{u})\right)\right]. \tag{2.50}$$

where the sums are over the set, $Ad(G)$, of roots of G, and the primes denote that we include only the non-zero roots in the sum. This can be motivated as follows. Given a $\mathcal{N} = 1$ 5d vector multiplet and an adjoint hypermultiplet of R-charge zero, we may give

the latter an expectation value, without breaking R-symmetry, and completely Higgsing the gauge group. At low energies this leaves no degrees of freedom for either the vector or hypermultiplet and thus their contributions should cancel out in the partition function. This leads to the expression above. We will also derive this by an alternate method in section 2.4 below.

Finally, the classical contribution to $\mathcal{W}(u)$ and $\Omega(u)$ can in principle be computed directly, but we can also take a more indirect approach and compute them on the background $S_b^3 \times \Sigma_{\mathfrak{g}}$. The contribution of the Yang-Mills term was computed in [39], and we may similarly evaluate the CS contribution and find:

$$Z_{S_b^3 \times \Sigma_{\mathfrak{g}}} = e^{\pi i Q^2 \mathrm{Tr}_{\mathrm{CS}}(\mathfrak{m}\tilde{u}^2) - 2\pi i Q^2 \tilde{\gamma}\mathfrak{m}\tilde{u}}, \tag{2.51}$$

where we defined[9]

$$\tilde{\gamma} = -\frac{2\pi i}{g_5{}^2}. \tag{2.52}$$

From this and the general form of (2.21), we can read off

$$\mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathrm{classical}}(\tilde{u}) = -\frac{1}{2}Q^2 \tilde{\gamma}\tilde{u}^2 - \frac{1}{6}Q^2 \mathrm{Tr}_{\mathrm{CS}}(\tilde{u}^3), \qquad \Omega_{S_b^3 \times \mathbb{R}^2}^{\mathrm{classical}} = 0. \tag{2.53}$$

Putting these pieces together, we arrive at the following result for the twisted superpotential of a 5d $\mathcal{N} = 1$ theory

$$\mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathrm{pert}}(\tilde{u}, \tilde{\nu}) = \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathrm{hyp}}(\tilde{u}, \tilde{\nu}) + \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathrm{vec}}(\tilde{u}) + \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathrm{classical}}(\tilde{u}), \tag{2.54}$$

and similarly for $\Omega_{S_b^3 \times \mathbb{R}^2}^{\mathrm{pert}}$, where, with foresight, we have labeled these as the perturbative contributions, for reasons that will be clear below.

We may nevertheless proceed as in the discussion above in section 2.2 and write the answer for the (perturbative) partition function on $S_b^3 \times \Sigma_{\mathfrak{g}}$. Before doing this, it will be convenient to define rescaled variables, which are normalized in a way which will look more natural in the context of S_b^3 partition function, namely,

$$u = -iQ\tilde{u}, \qquad \nu = -iQ\tilde{\nu}, \qquad \gamma = -iQ\tilde{\gamma}. \tag{2.55}$$

Then we can write the perturbative contribution to the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function as

$$Z_{S_b^3 \times \Sigma_{\mathfrak{g}}}^{\mathrm{pert}}(\nu)_{\mathfrak{n}} = \sum_{\hat{u} \in \mathcal{S}_{\mathrm{BE}}} \Pi_i^{\mathrm{pert}}(\hat{u}, \nu)^{\mathfrak{n}_i} \mathcal{H}^{\mathrm{pert}}(\hat{u}, \nu)^{\mathfrak{g}-1}, \tag{2.56}$$

where

$$\Pi_i^{\mathrm{pert}}(u, \nu) \equiv \prod_{\rho \in R_i} s_b(\rho(u) + \nu_i), \quad \Pi_a^{\mathrm{pert}}(\hat{u}, \nu) \equiv \prod_i \prod_{\rho \in R_i} s_b(\rho(\hat{u}) + \nu_i)^{\rho_a} \prod_{\alpha \in Ad(G)'} s_b(\alpha(\hat{u}) - iQ)^{-\alpha_a},$$

$$\mathcal{H}^{\mathrm{pert}}(u) \equiv \prod_i \prod_{\rho \in R_i} (s_b(\rho(u) + \nu_i))^{r_i - 1} \prod_{\alpha > 0} \left[2\sin\left(\pi b\alpha(u)\right) 2\sin\left(\pi b^{-1}\alpha(u)\right)\right]^{-1} \det_{a,b} \frac{-Q}{2\pi} \frac{\partial \log \Pi_a^{\mathrm{pert}}}{\partial u_b}, \tag{2.57}$$

---

[9]The 5d Yang-Mills coupling naturally appears in a complex combination with the 5d theta angle, θ_{5d}, and we may naturally extend the definition of $\tilde{\gamma}$ as $\tilde{\gamma} = \frac{2\pi i}{g_5{}^2} + \frac{\theta_{5d}}{2\pi}$. However, we will mostly restrict our attention to $\theta_{5d} = 0$ in this paper, and do not write the theta angle explicitly below.

and

$$\mathcal{S}_{\text{BE}}^{\text{pert}} = \{\, \hat{u} \,\mid\, \Pi_a(\hat{u}, \nu) = 1, \quad a = 1, \ldots, r_G \,\}/W_G \,.\tag{2.58}$$

Alternatively, we may write this in terms of an integral over the JK contour, as in (2.24),

$$Z_{S_b^3 \times \Sigma_{\mathfrak{g}}}^{\text{pert}}(\nu)_{\mathfrak{n}} = (-iQ)^{-r_G} \frac{1}{|W_G|} \sum_{\mathfrak{m} \in \Lambda_G} \oint du\, \Pi_i^{\text{pert}}(u, \nu)^{\mathfrak{n}_i} \Pi_a^{\text{pert}}(u, \nu)^{\mathfrak{m}_a} \mathcal{H}^{\text{pert}}(u, \nu)^{\mathfrak{g}} e^{-2\pi i \Omega^{\text{pert}}(u, \nu)} \,.\tag{2.59}$$

Alternative perspective: a direct sum of 3d theories. Before moving on to evaluate the validity of this formula more closely, let us mention an alternative perspective on the above result. Let us again return to the case of a single hypermultiplet on $S_b^3 \times \Sigma_{\mathfrak{g}}$. Then, rather than expanding in modes on S_b^3, we may expand in modes on $\Sigma_{\mathfrak{g}}$. After the topological twist we expect many of the modes to cancel, with the number that remain determined by the total flux felt by the chiral multiplet.

The 2d index theorem states that for a Dirac fermion, χ, the difference in the number of right- and left-moving fermionic zero modes on $\Sigma_{\mathfrak{g}}$ is given by

$$n_R^\chi - n_L^\chi = -t_\chi\, \eta_\Sigma \,, \qquad \eta_\Sigma = \begin{cases} 2|\mathfrak{g} - 1| & \text{for } \mathfrak{g} \neq 1 \\ 1 & \text{for } \mathfrak{g} = 1 \end{cases} \,,\tag{2.60}$$

where t_χ is the charge of the corresponding fermion in the background $F = T\text{dvol}(\Sigma_{\mathfrak{g}})$, $T = \frac{\kappa}{2}(T_R + \mathfrak{m}_I T_I)$, where $T_{R,I}$ are the R-symmetry and other possible background symmetry generators, respectively, and $\kappa = \{1, 0, -1\}$ for $\mathfrak{g} = 0, 1$, and $\mathfrak{g} > 1$, respectively. Applying this to a fermion with R-charge $(r - 1)$ and gauge charge 1, and denoting $\hat{\mathfrak{m}} \equiv \mathfrak{m} + (r - 1)(\mathfrak{g} - 1)$, with $\mathfrak{m}$ an integer-quantized gauge flux, gives the reduction rule

$$\text{5d hyper of charge 1} \rightarrow \begin{cases} |\hat{\mathfrak{m}}| & \text{3d chirals of charge 1} & \text{for } \hat{\mathfrak{m}} > 0 \\ |\hat{\mathfrak{m}}| & \text{3d chirals of charge } -1 & \text{for } \hat{\mathfrak{m}} < 0 \end{cases} \,.\tag{2.61}$$

Thus, since each chiral multiplet contributes a factor of $s_b(u)$ to the S_b^3 partition function, the latter may be written as (using also $s_b(-u) = s_b(u)^{-1}$)

$$Z_{S_b^3 \times \Sigma_{\mathfrak{g}}}^{\text{hyp,U(1)}}(u) = s_b(u)^{\mathfrak{m} + (r-1)(\mathfrak{g}-1)} \,,\tag{2.62}$$

which is precisely what we found in (2.46).

More generally, given a gauge theory, we may consider the effective 3d field content obtained for each choice of flux, $\mathfrak{m} \in \Lambda_G$, on the Riemann surface, $\Sigma_{\mathfrak{g}}$. This matter content describes an effective 3d $\mathcal{N} = 2$ theory, which we may denote $\mathcal{T}_{\Sigma_{\mathfrak{g}}, \mathfrak{m}}^{(3d)}$. Then we expect that, schematically, the full 5d theory compactified on $\Sigma_{\mathfrak{g}} \times \mathbb{R}^3$ can be written as a direct sum of 3d theories

$$\mathcal{T}_{\Sigma_{\mathfrak{g}}}^{(5d)} = \bigoplus_{\mathfrak{m} \in \Lambda_G} \mathcal{T}_{\Sigma_{\mathfrak{g}}, \mathfrak{m}}^{(3d)} \,.\tag{2.63}$$

If we take this relation literally, then we may also compute the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function as

$$Z_{S_b^3 \times \Sigma_{\mathfrak{g}}} = \sum_{\mathfrak{m} \in \Lambda_G} Z_{S_b^3} \left[\mathcal{T}_{\Sigma_{\mathfrak{g}}, \mathfrak{m}}^{(3d)} \right] \,.\tag{2.64}$$

We claim this is precisely the interpretation of the formula (2.59). Namely, each summand in that formula takes the form of the integral of a product of double sine functions, and such an integral describes the S_b^3 partition function of a particular 3d $\mathcal{N} = 2$ theory, which we claim is precisely $\mathcal{T}_{\Sigma_{\mathfrak{g}},\mathfrak{m}}^{(3d)}$.

2.4 Reduction to 4d and the instanton partition function

While the above two perspectives nicely complement each other, they also suffer from an important shortcoming. In both approaches we considered the compactification of the manifold from $d = 5$ to $d' = 2$ (in the first case) or $d' = 3$ (in the second). This compactification can be viewed as studying the theory in a limit where the compactified manifold is very small, and deriving the effective dynamics on the directions that remain large. While this is valid for any finite relative size of the S_b^3 and $\Sigma_{\mathfrak{g}}$ factors, owing to the topological invariance along $\Sigma_{\mathfrak{g}}$, the strict limit relevant to the compactification may not commute with the inclusion of non-perturbative effects, since in 5d these are associated to codimension-4 field configurations, namely, instantons. As we will see below, instantons do indeed modify the dynamics of the compactified theory, and are important for describing the full effective twisted superpotential and dilaton which compute the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function. In this section, we describe an alternative method, based on reduction to 4d, which captures these non-perturbative contributions. We obtain a perturbative piece which matches the result above but which is supplemented by an instanton contribution, which we write below.

To start, let us consider a 5d $\mathcal{N} = 1$ theory on $\mathbb{R}^4 \times S^1$. This gives an effective 4d $\mathcal{N} = 2$ theory on $\mathbb{R}^4$, with towers of KK modes corresponding to the Fourier modes of S^1. For example, a 5d $\mathcal{N} = 1$ hypermultiplet, Φ, with real mass, m, gives rise to a tower of 4d hypermultiplets, Φ_n, with complex masses

$$m_n = im + A_5 + \frac{n}{r}, \tag{2.65}$$

where r is the radius of S^1, and A_5 is the holonomy of the background gauge field along the S^1. Below we will mostly work in units with $r = 1$, but it will sometimes be important to keep it as a free parameter.

For the purpose of computing the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function, we may realize the S_b^3 factor as an S^1 fibration over S^2. Then the partition function on this five-manifold is equivalent to a partition function on $S^2 \times \Sigma_{\mathfrak{g}}$, where we insert one unit of flux on S^2 for the $U(1)_{KK}$ symmetry, corresponding to translations along the S^1 fiber.

For the rest of this subsection, let us restrict our attention to the case of $\Sigma_{\mathfrak{g}=0} = S^2$, so that the underlying four-manifold is $S^2 \times S^2$. The partition function of 4d $\mathcal{N} = 2$ gauge theories on this space was considered in [17], and so we may apply their results to gain another perspective on our computation.

The 4d computation. Let us briefly review the computation of the partition function on $S^2 \times S^2$, following [17]. Consider the following vector field on $S^2 \times S^2$, which generates an infinitesimal isometry

$$v = 2\epsilon_1(iz_1\partial_{z_1} - i\bar{z}_1\partial_{\bar{z}_1}) + 2\epsilon_2(iz_2\partial_{z_2} - i\bar{z}_2\partial_{\bar{z}_2}), \tag{2.66}$$

where z_i are the complex coordinates on each S^2 factor. Then we may consider the equivariant deformation of the Donaldson-Witten topological twist on $S^2 \times S^2$ [1]. The partition function can then be computed by equivariant localization, and reduces to a contribution from the four fixed points of the $U(1)_1 \times U(1)_2$ isometry above, namely, at the products of the poles of the two S^2 factors. More precisely, one finds that the path integral localizes to the locus

$$[\Phi, \bar{\Phi}] = [F, \Phi] = [F, \bar{\Phi}] = 0, \quad \iota_v F = id\Phi, \quad \iota_v D\bar{\Phi} = 0, \qquad (2.67)$$

where Φ and F are the complex scalar and field strength, respectively, in the $\mathcal{N} = 2$ vector multiplet. The first conditions imply we may take $\Phi, \bar{\Phi}$, and F in the same Cartan subalgebra of the gauge group, G, with rank r_G. The second and third then imply that we may write, in this Cartan basis,

$$\frac{F_a}{2\pi i} = \mathfrak{m}_{1,a}\omega_1 + \mathfrak{m}_{2,a}\omega_2, \quad \Phi_a = \tilde{u}_a + \frac{1}{2}\mathfrak{m}_{1,a}\epsilon_1 h_1 + \frac{1}{2}\mathfrak{m}_{2,a}\epsilon_2 h_2, \quad a = 1, \ldots, r_G, \quad (2.68)$$

where $\tilde{u}_a \in \mathbb{C}$ parameterizes the allowed profiles of Φ, $\mathfrak{m}_{i,a} \in \mathbb{Z}$ label the GNO fluxes of the gauge field on the two S^2 factors, and ω_i and h_i are the volume forms and height functions, respectively, on the two S^2 factors, i.e., $\omega_i = \frac{1}{4\pi}\sin\theta_i d\theta_i d\phi_i$ and $h_i = \cos\theta_i$ in the usual coordinates on S^2. In addition, the path integral receives contributions from point-like instantons localized at each of the fixed points.

Let us consider a 4d $\mathcal{N} = 2$ theory with gauge group G and hypermultiplets in the representation $R = \bigoplus_{i=1}^{s} R_i$. Then the partition function is given by an integral over the BPS locus above

$$Z_{S^2_{\epsilon_1} \times S^2_{\epsilon_2}}(\tilde{\nu}, \tau)_{\mathfrak{n}_1, \mathfrak{n}_2} = \sum_{\mathfrak{m}_1, \mathfrak{m}_2} \oint d\tilde{u} \prod_\ell Z_{\mathbb{R}^2_{\epsilon_1^{(\ell)}} \times \mathbb{R}^2_{\epsilon_2^{(\ell)}}}(\tilde{u}^{(\ell)}, \tilde{\nu}^{(\ell)}, \tau) \qquad (2.69)$$

Here τ is the complexified gauge coupling, $\tilde{u} = (\tilde{u}_a)$ and $\tilde{\nu} = (\tilde{\nu}_i)$ are complex scalar parameters associated to the gauge and flavor symmetry, respectively, $\mathfrak{m}_{1,2}$ and $\mathfrak{n}_{1,2}$ are the corresponding fluxes through the two S^2 factors, the index ℓ runs over the four fixed points, $\ell \in \{nn, ns, sn, ss\}$, sitting at the north ($n$) and south ($s$) poles of the two spheres, and we have defined

$$\epsilon_1^{(\ell)} = \begin{cases} \epsilon_1, & \ell = nn \text{ or } ns \\ -\epsilon_1, & \ell = sn \text{ or } ss \end{cases} \qquad \epsilon_2^{(\ell)} = \begin{cases} \epsilon_2, & \ell = nn \text{ or } sn \\ -\epsilon_2, & \ell = ns \text{ or } ss \end{cases}$$

$$\tilde{u}_a^{(\ell)} = \tilde{u}_a + \frac{1}{2}\left(\mathfrak{m}_{1,a}\epsilon_1^{(\ell)} + \mathfrak{m}_{2,a}\epsilon_2^{(\ell)}\right) \qquad \tilde{\nu}_i^{(\ell)} = \tilde{\nu}_i + \frac{1}{2}\left(\mathfrak{n}_{1,i}\epsilon_1^{(\ell)} + \mathfrak{n}_{2,i}\epsilon_2^{(\ell)}\right). \qquad (2.70)$$

Finally, $Z_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}}$ is the partition function on $\mathbb{R}^4$ in the Ω-background [18], which describes the contribution in the local neighborhood of each fixed point. It can be decomposed as

$$Z_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}}(\tilde{u}, \tilde{\nu}, \tau) = Z_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}}^{\text{classical}}(\tilde{u}, \tau) Z_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}}^{\text{1-loop}}(\tilde{u}, \tilde{\nu}) Z_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}}^{\text{inst}}(\tilde{u}, \tilde{\nu}, \tau). \qquad (2.71)$$

Let us analyze each of these pieces in turn. First, the classical contribution comes from the Yang-Mills term, and is given by

$$Z_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}}^{\text{classical}}(\tilde{u}, \tau) = \exp\left(-\pi i \tau \text{Tr}\frac{\tilde{u}^2}{\epsilon_1 \epsilon_2}\right), \qquad (2.72)$$

where we define $\mathrm{Tr}(\tilde{u}^2) = K^{ab}\tilde{u}_a\tilde{u}_b$, with K^{ab} the Killing form. Taking the contribution from the four fixed points using (2.70), we find

$$Z^{\text{classical}}_{S^2_{\epsilon_1} \times S^2_{\epsilon_2}}(\tilde{u}, \tau)_{\mathfrak{m}_1, \mathfrak{m}_2} \equiv \prod_\ell Z^{\text{classical}}_{\mathbb{R}^2_{\epsilon_1^{(\ell)}} \times \mathbb{R}^2_{\epsilon_2^{(\ell)}}}(\tilde{u}^{(\ell)}) = e^{-8\pi i \tau K^{ab}\mathfrak{m}_{1,a}\mathfrak{m}_{2,b}} . \tag{2.73}$$

For the perturbative and instanton part, it is useful to first write the contribution to the corresponding equivariant indices [34, 51]. The appropriate index for the vector multiplet is that for the so-called self-dual complex, and for the hypermultiplet it is the Dirac complex, where both are twisted by the vector bundle, V, in which the fields take values. These can both be related to the index of the Dolbeault complex for V, $\mathrm{ind}(\bar{\partial}_V)$, namely,

$$\mathrm{ind}_{\text{self-dual}} = \frac{1 + e^{i(\epsilon_1 + \epsilon_2)}}{2}\mathrm{ind}(\bar{\partial}_V), \quad \mathrm{ind}_{\text{Dirac}} = -e^{i\frac{\epsilon_1 + \epsilon_2}{2}}\frac{e^{i\tilde{\nu}} + e^{-i\tilde{\nu}}}{2}\mathrm{ind}(\bar{\partial}_V) . \tag{2.74}$$

Thus, we first consider the Dolbeault index, $\mathrm{ind}(\bar{\partial}_V)$. This may be decomposed into a perturbative piece and an instanton contribution,

$$\mathrm{ind}(\bar{\partial}_V) \equiv \mathrm{ind}_{\text{pert}}(\bar{\partial}_V) + \mathrm{ind}_{\text{inst}}(\bar{\partial}_V) . \tag{2.75}$$

The former can be written for an arbitrary representation R of an arbitrary gauge group G as

$$\mathrm{ind}_{\text{pert}}(\bar{\partial}_V) = \frac{\mathrm{tr}_R e^{i\tilde{u}}}{(e^{i\epsilon_1} - 1)(e^{i\epsilon_2} - 1)} . \tag{2.76}$$

The form of the instanton contribution is more subtle, and depends in a detailed way on the representation and gauge group, with explicit expressions typically known only for the classical gauge groups. For example, for the adjoint representation of $G = \mathrm{U}(N)$, and working in the k instanton sector, we have [34, 52]

$$\mathrm{ind}_{\text{inst}}(\bar{\partial}_V) = -e^{-i\frac{\epsilon_1 + \epsilon_2}{2}}\left(\mathrm{tr}_N(e^{i\tilde{u}})\,\mathrm{tr}_{\bar{k}}(e^{i\phi}) + c.c.\right) + (1 - e^{-i\epsilon_1})(1 - e^{-i\epsilon_2})\,\mathrm{tr}_{\mathrm{adj}_k}(e^{i\phi}), \tag{2.77}$$

where ϕ are the equivariant parameters for the $\mathrm{U}(k)$ symmetry acting on the k-instanton moduli space.

The 1-loop determinant in a given instanton background is given by making the formal replacement,

$$\mathrm{ind} = \sum_\lambda w_\lambda e^\lambda \quad \rightarrow \quad \prod_\lambda \lambda^{w_\lambda} . \tag{2.78}$$

Let us focus on the perturbative contribution first. Let us choose $\delta_i \in \{\pm 1\}$ so that $|e^{i\epsilon_i\delta_i}| < 1$ (we always assume $\epsilon_i \in \mathbb{C}\backslash\mathbb{R}$). Then the natural expansion of the index of the Dolbeault complex is (taking a charge 1 representation of $\mathrm{U}(1)$ for simplicity)

$$\mathrm{ind}^{\mathrm{U}(1)}_{\text{pert}}(\bar{\partial}_V) = \delta_1\delta_2 \sum_{k_1, k_2 \geq 0} e^{i\tilde{u} + i\delta_1\epsilon_1\left(k_1 + \frac{1-\delta_1}{2}\right) + i\delta_2\epsilon_2\left(k_2 + \frac{1-\delta_2}{2}\right)} , \tag{2.79}$$

and so the corresponding 1-loop determinant is, using (2.78),

$$Z^{\text{pert}, \mathrm{U}(1), \bar{\partial}_V}_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}} = \prod_{k_1, k_2 \geq 0} \left(\tilde{u} + \delta_1\epsilon_1\left(k_1 + \frac{1-\delta_1}{2}\right) + \delta_2\epsilon_2\left(k_2 + \frac{1-\delta_2}{2}\right)\right)^{\delta_1\delta_2} . \tag{2.80}$$

If we consider instead the Dirac complex, corresponding to a hypermultiplet, we find, using (2.74)

$$Z_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}}^{\mathrm{pert},\mathrm{U}(1),\mathrm{hyp}} = \prod_{k_1,k_2 \geq 0} \left[\left(\tilde{u} + \delta_1 \epsilon_1 \left(k_1 + \frac{1}{2} \right) + \delta_2 \epsilon_2 \left(k_2 + \frac{1}{2} \right) \right) \right.$$
$$\left. \times \left(-\tilde{u} + \delta_1 \epsilon_1 \left(k_1 + \frac{1}{2} \right) + \delta_2 \epsilon_2 \left(k_2 + \frac{1}{2} \right) \right) \right]^{-\delta_1 \delta_2 / 2} \tag{2.81}$$

Returning to $S^2_{\epsilon_1} \times S^2_{\epsilon_2}$, we may use the above result to compute the contributions from the four fixed points, identifying parameters as in (2.70). Let us assume now for concreteness that $|e^{i\epsilon_i}| < 1$. Then one finds many of terms in the infinite products above cancel, and we are left with

$$Z_{S^2_{\epsilon_1} \times S^2_{\epsilon_2}}^{\mathrm{pert},\mathrm{U}(1),\mathrm{hyp}}(\tilde{u})_{\mathfrak{m}_1,\mathfrak{m}_2} = \left([\tilde{u}; \epsilon_1, \epsilon_2]_{\mathfrak{m}_1,\mathfrak{m}_2} [-\tilde{u}; \epsilon_1, \epsilon_2]_{-\mathfrak{m}_1,-\mathfrak{m}_2} \right)^{-1/2} \tag{2.82}$$

where we defined

$$[\tilde{u}; \epsilon_1, \epsilon_2]_{\mathfrak{m}_1,\mathfrak{m}_2} \equiv \prod_{k_1 = -\frac{1}{2}(|\mathfrak{m}_1|-1)}^{\frac{1}{2}(|\mathfrak{m}_1|-1)} \prod_{k_2 = -\frac{1}{2}(|\mathfrak{m}_2|-1)}^{\frac{1}{2}(|\mathfrak{m}_2|-1)} (\tilde{u} + \epsilon_1 k_1 + \epsilon_2 k_2)^{\mathrm{sgn}(\mathfrak{m}_1)\,\mathrm{sgn}(\mathfrak{m}_2)} . \tag{2.83}$$

We can simplify this using

$$[\tilde{u}; \epsilon_1, \epsilon_2]_{\mathfrak{m}_1,\mathfrak{m}_2} = [\tilde{u}; \epsilon_1, \epsilon_2]_{-\mathfrak{m}_1,-\mathfrak{m}_2} = (-1)^{\mathfrak{m}_1 \mathfrak{m}_2} [-\tilde{u}; \epsilon_1, \epsilon_2]_{\mathfrak{m}_1,\mathfrak{m}_2} . \tag{2.84}$$

Then we find[10]

$$Z_{S^2_{\epsilon_1} \times S^2_{\epsilon_2}}^{\mathrm{pert},\mathrm{U}(1),\mathrm{hyp}}(\tilde{u})_{\mathfrak{m}_1,\mathfrak{m}_2} = [\tilde{u}; \epsilon_1, \epsilon_2]_{\mathfrak{m}_1,\mathfrak{m}_2}^{-1} . \tag{2.85}$$

We note for later convenience that in the limit of vanishing equivariant parameters, $\epsilon_{1,2} \to 0$, we have

$$[\tilde{u}; \epsilon_1, \epsilon_2]_{\mathfrak{m}_1,\mathfrak{m}_2} \xrightarrow{\epsilon_2 \to 0} \prod_{k_1 = -(|\mathfrak{m}_1|-1)/2}^{(|\mathfrak{m}_1|-1)/2} (\tilde{u} + \epsilon_1 k_1)^{\mathrm{sgn}(\mathfrak{m}_1)\mathfrak{m}_2} \xrightarrow{\epsilon_1 \to 0} \tilde{u}^{\mathfrak{m}_1 \mathfrak{m}_2} . \tag{2.86}$$

The perturbative contribution for a more general hypermultiplet is then computed by

$$Z_{S^2_{\epsilon_1} \times S^2_{\epsilon_2}}^{\mathrm{1\text{-}loop},\mathrm{hyp}}(\tilde{u}, \tilde{\nu})_{\mathfrak{m}_1,\mathfrak{n}_1;\mathfrak{m}_2,\mathfrak{n}_2} = \prod_i \prod_{\rho \in R_i} [\rho(\tilde{u}) + \tilde{\nu}_i; \epsilon_1, \epsilon_2]_{\rho(\mathfrak{m}_1)+\mathfrak{n}_{1i},\rho(\mathfrak{m}_2)+\mathfrak{n}_{2i}}^{-1} \tag{2.87}$$

For the vector multiplet, a similar argument, using (2.74), gives

$$Z_{S^2_{\epsilon_1} \times S^2_{\epsilon_2}}^{\mathrm{1\text{-}loop},\mathrm{vec}}(\tilde{u})_{\mathfrak{m}_1,\mathfrak{m}_2} = \prod_{\alpha \in Ad(G)'} \left([\alpha(\tilde{u}); \epsilon_1, \epsilon_2]_{\alpha(\mathfrak{m}_1)-1,\alpha(\mathfrak{m}_2)-1} [\alpha(\tilde{u}); \epsilon_1, \epsilon_2]_{\alpha(\mathfrak{m}_1)+1,\alpha(\mathfrak{m}_2)+1} \right)^{1/2}$$
$$= \prod_{\alpha > 0} [\alpha(\tilde{u}); \epsilon_1, \epsilon_2]_{\alpha(\mathfrak{m}_1)-1,\alpha(\mathfrak{m}_2)-1} [\alpha(\tilde{u}); \epsilon_1, \epsilon_2]_{\alpha(\mathfrak{m}_1)+1,\alpha(\mathfrak{m}_2)+1} \tag{2.88}$$

---

[10]Here we have chosen a convenient overall phase, which was ambiguous in the above infinite products. We similarly fix several such phases below. It would be desirable to fix these phases more rigorously from first principles, perhaps by careful consideration of the 4d 't Hooft anomalies and the 5d parity anomaly, as discussed in [53, 54].

Then the full perturbative contribution is

$$Z^{\text{pert}}_{S^2_{\epsilon_1} \times S^2_{\epsilon_2}}(\tilde{u}, \tilde{\nu}, \tau)_{\mathfrak{m}_1, \mathfrak{n}_1; \mathfrak{m}_2, \mathfrak{n}_2} = Z^{\text{classical}}_{S^2_{\epsilon_1} \times S^2_{\epsilon_2}}(\tilde{u}, \tau)_{\mathfrak{m}_1, \mathfrak{m}_2} Z^{\text{1-loop,vec}}_{S^2_{\epsilon_1} \times S^2_{\epsilon_2}}(\tilde{u})_{\mathfrak{m}_1, \mathfrak{m}_2} Z^{\text{1-loop,hyp}}_{S^2_{\epsilon_1} \times S^2_{\epsilon_2}}(\tilde{u}, \tilde{\nu})_{\mathfrak{m}_1, \mathfrak{n}_1; \mathfrak{m}_2, \mathfrak{n}_2}$$

(2.89)

The full nonperturbative expression is obtained by including the contribution from the remaining terms in (2.77), which correspond to the instanton contribution. The detailed form of the instanton contribution will depend on the gauge group and matter representation, but has the general form

$$Z^{\text{inst}}_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}}(\tilde{u}, \tilde{\nu}, \tau) = \sum_{k=0}^{\infty} z^k Z^{(k)}_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}}(\tilde{u}, \tilde{\nu}),$$

(2.90)

where we defined $z = e^{2\pi i \tau}$, the classical contribution from the instanton action, and $Z^{(k)}_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}}$ is the contribution to the 1-loop determinant in the k-instanton background. For example, in the U(N) case, it is given by integration over the U(k) equivariant parameters appearing in (2.77), and can be expressed in some cases as a sum over N-colored Young diagrams [18]. We will discuss an explicit example when we consider the 5d uplift below.

Putting these ingredients together, we may write

$$Z_{S^2_{\epsilon_1} \times S^2_{\epsilon_2}}(\tilde{\nu}, \tau)_{\mathfrak{n}_1, \mathfrak{n}_2} = \frac{1}{|W_G|} \sum_{\mathfrak{m}_1, \mathfrak{m}_2} \oint d\tilde{u} \, Z^{\text{pert}}_{S^2_{\epsilon_1} \times S^2_{\epsilon_2}}(\tilde{u}, \tilde{\nu}, \tau)_{\mathfrak{m}_1, \mathfrak{n}_1; \mathfrak{m}_2, \mathfrak{n}_2} \prod_\ell Z^{\text{inst}}_{\mathbb{R}^2_{\epsilon_1^{(\ell)}} \times \mathbb{R}^2_{\epsilon_2^{(\ell)}}}(\tilde{u}^{(\ell)}, \tilde{\nu}^{(\ell)}, \tau).$$

(2.91)

Uplifting to 5d. We may now uplift this result to our desired background, $S^3_b \times S^2$, by applying this computation to the effective 4d $\mathcal{N} = 2$ theory, $\tilde{\mathcal{T}}^{(4d)}$, obtained by dimensional reduction of our chosen 5d theory, $\mathcal{T}^{(5d)}$, as outlined above.

More precisely, our strategy will be to exhibit S^3_b as an S^1 fibration over the topologically twisted S^2. Let us first describe the case $b = 1$. Then it was shown in [11] that the usual supersymmetric background on the round sphere is an S^1 fibration over the topological A-twist background on the 2-sphere. Here we must include one unit of flux for the connection fibering the S^1. Then this background may be equivalently obtained by considering the effective 2d theory obtained by dimensional reduction, which has a U(1)$_{KK}$ global symmetry corresponding to translations along the S^1 direction, and inserting a unit flux on S^2 for this global symmetry. We will employ the same strategy here, this time with an additional S^2 factor in the geometry. In other words, we consider the effective 4d $\mathcal{N} = 2$ theory obtained by dimensional reduction along the S^1 fiber. Here we perform an ordinary topological twist along each S^2 factor, which corresponds to the $\epsilon_1, \epsilon_2 \to 0$ limit of the equivariant background discussed above. Finally, we must turn on a unit flux for the U(1)$_{KK}$ symmetry along one of the S^2 factors.

In the case of a non-round sphere, $b \neq 1$, we will argue below that this can also be exhibited as an S^1 fibration over S^2, but now with a nonzero Omega-background on the S^2 [55]. This will again fit into the framework described above, but now rather than taking both ϵ_1 and ϵ_2 to zero, we will keep a nonzero ϵ_1, specifically, we find that the appropriate value is

$$\epsilon_1 = \epsilon_b \equiv -\frac{2}{r} \frac{b - b^{-1}}{b + b^{-1}},$$

(2.92)

where r is the S^1 radius, which we will often set to 1 below. Note this vanishes when we set $b = 1$. Thus, we arrive at the following schematic relation

$$Z_{S_b^3 \times S^2}^{\mathcal{T}^{(5d)}}(\tilde{\nu})_{\mathfrak{n}_1, \mathfrak{n}_2} \cong \lim_{\epsilon \to 0} Z_{S_{\epsilon_b}^2 \times S_\epsilon^2}^{\tilde{\mathcal{T}}^{(4d)}} \left(\tilde{\nu}, \tilde{\nu}_{KK} = \frac{1}{r} \right)_{\mathfrak{n}_1, \mathfrak{n}_{1,KK}=1; \mathfrak{n}_2, \mathfrak{n}_{2,KK}=0}, \qquad (2.93)$$

where $\tilde{\mathcal{T}}^{(4d)}$ is the effective 4d theory. Here the arguments of the partition functions correspond to the mass and fluxes for flavor symmetries. On the r.h.s., there is an additional flavor symmetry, the $U(1)_{KK}$ symmetry corresponding to translations along S^1, which we have assigned a mass $\frac{1}{r}$ and a unit flux on the first S^2 factor, which gives rise to the Hopf fibration of S^3 described above. We will also see in a moment that the flux $\mathfrak{n}_1$ can be absorbed into a shift of $\tilde{\nu}$, as expected since there is no nontrivial 2-cycle on S^3 on which to support a flux. We will also demonstrate that the partition function can be written in terms of the 5d Nekrasov partition function, or more precisely, of the Nekrasov-Shatashvili limit of this partition function.

Let us now make these statements more precise. As above, the integrand of the partition function factorizes into contributions from the perturbative (classical and 1-loop) piece, and the instanton contribution. First, the classical contributions on $S_b^3 \times \Sigma_\mathfrak{g}$ were already presented above in (2.51), and we may use that result in the present case, $\mathfrak{g} = 0$, as well.

Next we have the perturbative and instanton contributions. To compute these, let us return to the contribution to the Dolbeault index in 4d from a single fixed point, given by (2.75),

$$\text{ind}(\bar{\partial}_V) \equiv \text{ind}_{\text{pert}}(\bar{\partial}_V) + \text{ind}_{\text{inst}}(\bar{\partial}_V). \qquad (2.94)$$

Then when we add a tower of KK modes with masses $\frac{n}{r}$, $n \in \mathbb{Z}$, the index is modified to

$$\text{ind}^{(5d)}(\bar{\partial}_V) \equiv \sum_{n \in \mathbb{Z}} e^{\frac{in}{r}} \text{ind}(\bar{\partial}_V), \qquad (2.95)$$

and similarly for the vector and hypermultiplet indices. Then we expect that when we construct the 1-loop determinants on $\mathbb{R}^4 \times S^1$, these will be related to those in 4d by, schematically,

$$Z_{\mathbb{R}_{\epsilon_1}^2 \times \mathbb{R}_{\epsilon_2}^2}^{\text{1-loop}} = \prod_\lambda \lambda \quad \Rightarrow \quad Z_{\mathbb{R}_{\epsilon_1}^2 \times \mathbb{R}_{\epsilon_2}^2 \times S^1}^{\text{1-loop}} = \prod_{n \in \mathbb{Z}} \prod_\lambda \left(\lambda + \frac{n}{r} \right) \quad \propto \quad \prod_\lambda 2 \sin \pi r \lambda. \qquad (2.96)$$

This holds for both the perturbative contribution and the 1-loop contribution in an instanton background, as discussed in [34] in the context of the S^5 partition function.

To see how this works in more detail, let us consider the perturbative contribution for a single hypermultiplet. Recall that the perturbative contribution of the hypermultiplet on $\mathbb{R}_{\epsilon_1}^2 \times \mathbb{R}_{\epsilon_2}^2$ is given by (2.81), namely,

$$Z_{\mathbb{R}_{\epsilon_1}^2 \times \mathbb{R}_{\epsilon_2}^2}^{\text{pert,U(1),hyp}} = \prod_{k_1, k_2 \geq 0} \left[\left(\tilde{u} + \delta_1 \epsilon_1 \left(k_1 + \frac{1}{2} \right) + \delta_2 \epsilon_2 \left(k_2 + \frac{1}{2} \right) \right) \right.$$
$$\left. \times \left(-\tilde{u} + \delta_1 \epsilon_1 \left(k_1 + \frac{1}{2} \right) + \delta_2 \epsilon_2 \left(k_2 + \frac{1}{2} \right) \right) \right]^{-\delta_1 \delta_2/2}. \qquad (2.97)$$

The uplift to $\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2} \times S^1$ is then given by[11]

$$
\begin{aligned}
Z^{\text{pert},\text{U}(1),\text{hyp}}_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2} \times S^1}(\tilde{u}) = \prod_{n \in \mathbb{Z}} \prod_{k_1,k_2 \geq 0} &\left[\left(\tilde{u} + \frac{n}{r} + \delta_1 \epsilon_1 \left(k_1 + \frac{1}{2} \right) + \delta_2 \epsilon_2 \left(k_2 + \frac{1}{2} \right) \right) \right. \\
&\left. \times \left(-\tilde{u} - \frac{n}{r} + \delta_1 \epsilon_1 \left(k_1 + \frac{1}{2} \right) + \delta_2 \epsilon_2 \left(k_2 + \frac{1}{2} \right) \right) \right]^{-\delta_1 \delta_2/2} \\
= \prod_{k_1,k_2 \geq 0} &\left[2i \sin\left[\pi r \left(\tilde{u} + \delta_1 \epsilon_1 \left(k_1 + \frac{1}{2} \right) + \delta_2 \epsilon_2 \left(k_2 + \frac{1}{2} \right) \right) \right] \right. \\
&\left. \times 2i \sin\left[\pi r \left(-\tilde{u} + \delta_1 \epsilon_1 \left(k_1 + \frac{1}{2} \right) + \delta_2 \epsilon_2 \left(k_2 + \frac{1}{2} \right) \right) \right] \right]^{-\delta_1 \delta_2/2}.
\end{aligned}
\tag{2.98}
$$

Note this is naturally a function of the parameters

$$
x = e^{2\pi i r \tilde{u}}, \qquad \mathfrak{q}_i = e^{2\pi i r \epsilon_i}. \tag{2.99}
$$

We will sometimes emphasize this by writing the 5d partition function as

$$
Z_{\mathbb{R}^2_{\mathfrak{q}_1} \times \mathbb{R}^2_{\mathfrak{q}_2} \times S^1}(x), \tag{2.100}
$$

where the meaning of the subscripts and arguments should be clear from context. Now we may glue four copies of this function with parameters identified as in (2.70), and one finds that the infinite product simplifies to

$$
Z^{\text{pert},\text{U}(1),\text{hyp}}_{S^2_{\epsilon_1} \times S^2_{\epsilon_2} \times S^1}(\tilde{u})_{\mathfrak{m}_1,\mathfrak{m}_2} = (x; \mathfrak{q}_1, \mathfrak{q}_2)^{-1}_{\mathfrak{m}_1,\mathfrak{m}_2}, \tag{2.101}
$$

where the parameters are as in (2.99), and we have defined

$$
(x; \mathfrak{q}_1, \mathfrak{q}_2)_{\mathfrak{m}_1,\mathfrak{m}_2} = \prod_{\substack{k_1 = \\ -\frac{1}{2}(|\mathfrak{m}_1|-1)}}^{\frac{1}{2}(|\mathfrak{m}_1|-1)} \prod_{\substack{k_2 = \\ -\frac{1}{2}(|\mathfrak{m}_2|-1)}}^{\frac{1}{2}(|\mathfrak{m}_2|-1)} \left(x^{1/2} \mathfrak{q}_1^{k_1/2} \mathfrak{q}_2^{k_2/2} - x^{-1/2} \mathfrak{q}_1^{-k_1/2} \mathfrak{q}_2^{-k_2/2} \right)^{\text{sgn}(\mathfrak{m}_1)\text{sgn}(\mathfrak{m}_2)}.
\tag{2.102}
$$

This result may also be obtained by directly uplifting the perturbative contribution,

$$
Z^{\text{pert},\text{U}(1),\text{hyp}}_{S^2_{\epsilon_1} \times S^2_{\epsilon_2} \times S^1}(\tilde{u})_{\mathfrak{m}_1,\mathfrak{m}_2} = \prod_{n \in \mathbb{Z}} Z^{\text{pert},\text{U}(1),\text{hyp}}_{S^2_{\epsilon_1} \times S^2_{\epsilon_2}} \left(\tilde{u} + \frac{n}{r} \right)_{\mathfrak{m}_1,\mathfrak{m}_2} = \prod_{n \in \mathbb{Z}} \left[\tilde{u} + \frac{n}{r}; \epsilon_1, \epsilon_2 \right]^{-1}_{\mathfrak{m}_1,\mathfrak{m}_2} \\
= (x; \mathfrak{q}_1, \mathfrak{q}_2)^{-1}_{\mathfrak{m}_1,\mathfrak{m}_2}. \tag{2.103}
$$

The above computation represents the 1-loop determinant on the background $S^2_{\epsilon_1} \times S^2_{\epsilon_2} \times S^1$, where the S^1 appears as trivial product. We will return to this case in section 5 below, but our present interest is in the partition function on $S^3_b \times S^2$. Then we must include a magnetic flux for fields charged under the $\text{U}(1)_{KK}$ symmetry. Recall that the magnetic

---

[11]Here we fix the overall normalization arising from regularizing the product over n for later convenience. As mentioned above, it would be interesting to fix this from first principles.

flux enters the fixed point contributions on $S^2_{\epsilon_1} \times S^2_{\epsilon_2}$ by shifting the eigenvalues by $\frac{\mathfrak{m}_i \epsilon_i}{2}$. Thus, including a unit $\mathrm{U}(1)_{KK}$ flux, the perturbative contribution of the hypermultiplet in (2.98) is modified to

$$
\begin{aligned}
Z^{\mathrm{pert,U(1),hyp}}_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2} \times S^1}(\tilde{u}) = \prod_{n \in \mathbb{Z}} \prod_{k_1,k_2 \geq 0} & \left[\left(\tilde{u} + \frac{n}{r} + \frac{n\epsilon_1}{2} + \delta_1\epsilon_1 \left(k_1 + \frac{1}{2} \right) + \delta_2\epsilon_2 \left(k_2 + \frac{1}{2} \right) \right) \right. \\
& \left. \times \left(-\tilde{u} - \frac{n}{r} - \frac{n\epsilon_1}{2} + \delta_1\epsilon_1 \left(k_1 + \frac{1}{2} \right) + \delta_2\epsilon_2 \left(k_2 + \frac{1}{2} \right) \right) \right]^{-\delta_1\delta_2/2} \\
= \prod_{k_1,k_2 \geq 0} & \left[2i \sin \left[\pi\tilde{r} \left(\tilde{u} + \delta_1\epsilon_1 \left(k_1 + \frac{1}{2} \right) + \delta_2\epsilon_2 \left(k_2 + \frac{1}{2} \right) \right) \right] \right. \\
& \left. \times 2i \sin \left[\pi\tilde{r} \left(-\tilde{u} + \delta_1\epsilon_1 \left(k_1 + \frac{1}{2} \right) + \delta_2\epsilon_2 \left(k_2 + \frac{1}{2} \right) \right) \right] \right]^{-\delta_1\delta_2/2},
\end{aligned}
$$

(2.104)

where we defined

$$
\tilde{r} = \frac{r}{1 + \frac{r\epsilon_1}{2}} .
$$

(2.105)

This is the same function we obtained in (2.98) above, but now evaluated at the arguments

$$
x' = e^{2\pi i \tilde{r} \tilde{u}}, \qquad \mathfrak{q}'_i = e^{2\pi i \tilde{r} \epsilon_i} .
$$

(2.106)

Now to construct the full $S^3_b \times S^2$ partition function, we must include the contribution from the four fixed points, that is (writing this now as a function of x and the $\mathfrak{q}_i$),

$$
Z^{\mathrm{pert,U(1),hyp}}_{S^3_b \times S^2}(x)_{\mathfrak{m}_1,\mathfrak{m}_2} = \prod_\ell Z^{\mathrm{pert,U(1),hyp}}_{\mathbb{R}^2_{\mathfrak{q}^{(\ell)}_1} \times \mathbb{R}^2_{\mathfrak{q}^{(\ell)}_2} \times S^1}(x^{(\ell)}) ,
$$

(2.107)

where the parameters at the four fixed points can be read off from (2.70) and (2.106), and we find, for vanishing fluxes, and renaming $\epsilon_2 \to \epsilon$ as in (2.93),

$$
x^{(\ell)} = \begin{cases} e^{2\pi i \frac{r\tilde{u}}{1+\frac{1}{2}r\epsilon_b}} = e^{2\pi i b Q r\tilde{u}} \equiv x, & \ell = nn \text{ or } ns \\ e^{2\pi i \frac{r\tilde{u}}{1-\frac{1}{2}r\epsilon_b}} = e^{2\pi i b^{-1} Q r\tilde{u}} \equiv \bar{x}, & \ell = sn \text{ or } ss \end{cases}
$$

$$
\mathfrak{q}^{(\ell)}_1 = \begin{cases} e^{2\pi i \frac{r\epsilon_b}{1+\frac{1}{2}r\epsilon_b}} = e^{-2\pi i b^2} \equiv \mathfrak{q}, & \ell = nn \text{ or } ns \\ e^{-2\pi i \frac{r\epsilon_b}{1-\frac{1}{2}r\epsilon_b}} = e^{-2\pi i b^{-2}} \equiv \bar{\mathfrak{q}}, & \ell = sn \text{ or } ss \end{cases}
$$

(2.108)

$$
\mathfrak{q}^{(\ell)}_2 = \begin{cases} e^{2\pi i \frac{r\epsilon}{1+\frac{1}{2}r\epsilon_b}} = e^{2\pi i b Q r\epsilon}, & \ell = nn \\ e^{-2\pi i \frac{r\epsilon}{1+\frac{1}{2}r\epsilon_b}} = e^{-2\pi i b Q r\epsilon}, & \ell = ns \\ e^{2\pi i \frac{r\epsilon}{1-\frac{1}{2}r\epsilon_b}} = e^{2\pi i b^{-1} Q r\epsilon}, & \ell = sn \\ e^{-2\pi i \frac{r\epsilon}{1-\frac{1}{2}r\epsilon_b}} = e^{-2\pi i b^{-1} Q r\epsilon}, & \ell = ss \end{cases}
$$

where we introduced the parameters

$$
\mathfrak{q} = e^{-2\pi i b^2}, \quad \bar{\mathfrak{q}} = e^{-2\pi i b^{-2}}, \quad x = e^{2\pi i b Q r\tilde{u}}, \quad \bar{x} = e^{2\pi i b^{-1} Q r\tilde{u}},
$$

(2.109)

and $Q = \frac{1}{2}(b + b^{-1})$ as usual. We can then introduce fluxes, as above, by shifting $u^{(\ell)} \to u^{(\ell)} + \frac{1}{2}(\mathfrak{m}_1 \epsilon^{(\ell)}_1 + \mathfrak{m}_2 \epsilon^{(\ell)}_2)$. We will describe the dependence on fluxes in more detail below.

Perturbative contribution and holomorphic blocks. We now evaluate this perturbative contribution for the hypermultiplet explicitly using two methods. First, we may directly uplift the perturbative contribution on $S^2_{\epsilon_1} \times S^2_{\epsilon_2}$, as in (2.103), but now including the $U(1)_{KK}$ flux. We find

$$
Z^{\text{pert,U(1),hyp}}_{S^3_b \times S^2}(\tilde{u})_{\mathfrak{m}_1,\mathfrak{m}_2} = \prod_{n \in \mathbb{Z}} Z^{\text{pert,U(1),hyp}}_{S^2_{\epsilon_b} \times S^2_\epsilon}\left(\tilde{u} + \frac{n}{r}\right)_{\mathfrak{m}_1+n,\mathfrak{m}_2} = \prod_{n \in \mathbb{Z}} \left[\tilde{u} + \frac{n}{r}; \epsilon_b, \epsilon\right]^{-1}_{\mathfrak{m}_1+n,\mathfrak{m}_2}
$$

$$
\xrightarrow[\epsilon \to 0]{} \prod_{n \in \mathbb{Z}} \prod_{\substack{k_1 = \\ -\frac{1}{2}(|\mathfrak{m}_1+n|-1)}}^{\frac{1}{2}(|\mathfrak{m}_1+n|-1)} \left(\tilde{u} + \frac{n}{r} + k_1 \epsilon_b\right)^{-\operatorname{sgn}(\mathfrak{m}_1+n)\mathfrak{m}_2}, \qquad (2.110)
$$

where we used (2.86) to take the $\epsilon \to 0$ limit. Defining $\ell = |\mathfrak{m}_1 + n|$, this may be further rewritten as

$$
Z^{\text{pert,U(1),hyp}}_{S^3_b \times S^2}(\tilde{u})_{\mathfrak{m}_1,\mathfrak{m}_2} = \prod_{\ell=1}^{\infty} \prod_{k_1=-\frac{1}{2}(\ell-1)}^{\frac{1}{2}(\ell-1)} \left(\frac{\tilde{u} + \frac{-\ell-\mathfrak{m}_1}{r} + k_1\epsilon_b}{\tilde{u} + \frac{\ell-\mathfrak{m}_1}{r} + k_1\epsilon_b}\right)^{-\mathfrak{m}_2}. \qquad (2.111)
$$

Finally, after substituting the definition of ϵ_b and defining $j = \ell - 1 + 2k_1, k = \ell - 1 - 2k_1$, which take values in $\mathbb{Z}_{\geq 0}$, and comparing to (2.40), we see that

$$
Z^{\text{pert,U(1),hyp}}_{S^3_b \times S^2}(u)_{\mathfrak{m}_1,\mathfrak{m}_2} = s_b\left(-iQ(r\tilde{u} - \mathfrak{m}_1)\right)^{\mathfrak{m}_2}. \qquad (2.112)
$$

We may take another approach to this computation which will give another useful perspective on this partition function. Let us consider the contribution from the fixed points $\ell = nn$ and $\ell = ns$. We also take $|\mathfrak{q}| = |e^{-2\pi i b^2}| < 1$ for concreteness. Then one finds the infinite products over k_2 in (2.104) reduce to a finite product, and we can take a finite $\epsilon \to 0$ limit,

$$
Z^{\text{pert,U(1),hyp}}_{\mathbb{R}^2_{\epsilon_b} \times \mathbb{R}^2_{\epsilon} \times S^1} Z^{\text{pert,U(1),hyp}}_{\mathbb{R}^2_{\epsilon_b} \times \mathbb{R}^2_{-\epsilon} \times S^1} = \prod_{k_1 \geq 0} \prod_{\substack{k_2 = \\ -\frac{1}{2}(|\mathfrak{m}_2|-1)}}^{\frac{1}{2}(|\mathfrak{m}_2|-1)} \left(\frac{\sin\left[\pi\tilde{r}(\tilde{u} + \epsilon_b(k_1 + \frac{\mathfrak{m}_1}{2} + \frac{1}{2}) + \epsilon k_2)\right]}{\sin\left[\pi\tilde{r}(-\tilde{u} + \epsilon_b(k_1 - \frac{\mathfrak{m}_1}{2} + \frac{1}{2}) + \epsilon k_2)\right]}\right)^{-\operatorname{sgn}(\mathfrak{m}_2)/2}
$$

$$
\xrightarrow[\epsilon \to 0]{} \prod_{k_1 \geq 0} \left(\frac{\sin\left[\pi\tilde{r}\left(\tilde{u} + \epsilon_b(k_1 + \frac{\mathfrak{m}_1}{2} + \frac{1}{2})\right)\right]}{\sin\left[\pi\tilde{r}\left(-\tilde{u} + \epsilon_b(k_1 - \frac{\mathfrak{m}_1}{2} + \frac{1}{2})\right)\right]}\right)^{-\mathfrak{m}_2/2}
$$

$$
\propto \left(\frac{((-\mathfrak{q}^{\frac{1}{2}})^{1+\mathfrak{m}_1}x; \mathfrak{q})}{((-\mathfrak{q}^{\frac{1}{2}})^{1-\mathfrak{m}_1}x^{-1}; \mathfrak{q})}\right)^{-\mathfrak{m}_2/2}, \qquad (2.113)
$$

where $\mathfrak{q}$ and x are as in (2.109), and we introduced the $\mathfrak{q}$-Pochhammer symbol

$$
(x; \mathfrak{q}) = \prod_{k \geq 0}(1 - x\mathfrak{q}^k). \qquad (2.114)
$$

A similar argument for the contribution from the fixed points $\ell = sn$ and $\ell = ss$ gives

$$
Z^{\text{pert,U(1),hyp}}_{\mathbb{R}^2_{\epsilon_b} \times \mathbb{R}^2_{\epsilon} \times S^1} Z^{\text{pert,U(1),hyp}}_{\mathbb{R}^2_{\epsilon_b} \times \mathbb{R}^2_{-\epsilon} \times S^1} \to \left(\frac{((-\bar{\mathfrak{q}}^{\frac{1}{2}})^{-1-\mathfrak{m}_1}\bar{x}^{-1}; \bar{\mathfrak{q}}^{-1})}{((-\bar{\mathfrak{q}}^{\frac{1}{2}})^{-1+\mathfrak{m}_1}\bar{x}; \bar{\mathfrak{q}}^{-1})}\right)^{-\mathfrak{m}_2/2}. \qquad (2.115)
$$

Then, using the relation [13, 20]

$$s_b(irQ\tilde{u}) = e^{\frac{\pi i}{2}r^2Q^2\tilde{u}^2 - \frac{\pi i}{24}(b^2+b^{-2})}\frac{\left(-\mathfrak{q}^{1/2}x;\mathfrak{q}\right)}{\left(-\bar{\mathfrak{q}}^{-1/2}\bar{x};\bar{\mathfrak{q}}^{-1}\right)} = s_b(-irQ\tilde{u})^{-1}\,, \tag{2.116}$$

where the parameters are as in (2.109), we again arrive at

$$Z^{\text{pert,U(1),hyp}}_{S^3_b \times S^2} = s_b(-iQ(r\tilde{u} - \mathfrak{m}_1))^{\mathfrak{m}_2}\,. \tag{2.117}$$

This method demonstrates that the partition function of the hypermultiplet naturally factorizes into a contribution from the fixed points at the north and south poles of the S^2 base of S^3_b.

Note that the dependence on the flux $\mathfrak{m}_1$ can be absorbed into a shift of u, as expected since the S^3_b factor cannot support a topologically nontrivial flux. Thus we will henceforth set $\mathfrak{m}_1 = 0$, and rename the flux on S^2 by $\mathfrak{m}_2 \to \mathfrak{m}$. Then we see that this agrees with the result (2.46) derived above by KK reduction, where in the present case we are implicitly taking R-charge one for the hypermultiplet. We may also define, as in section 2.3, the rescaled parameters

$$u = -iQr\tilde{u}\,, \qquad \nu = -iQr\tilde{\nu}\,. \tag{2.118}$$

We will sometimes work in terms of these parameters below.

It is straightforward to extend the argument above to a hypermultiplet in a general representation, giving

$$Z^{\text{pert,hyp}}_{S^3_b \times S^2}(u,\nu)_{\mathfrak{m},\mathfrak{n}} = \prod_i \prod_{\rho \in R_i} s_b\left(\rho(u) + \nu_i\right)^{\rho(\mathfrak{m})+\mathfrak{n}_i}\,. \tag{2.119}$$

Similarly, for the vector multiplet we find

$$\begin{aligned}Z^{\text{pert,vec}}_{S^3_b \times S^2}(u)_{\mathfrak{m}} &= \prod_{\alpha>0} s_b\left(-iQ(r\alpha(\tilde{u}) - 1)\right)^{-\alpha(\mathfrak{m})-1} s_b\left(-iQ(r\alpha(\tilde{u}) + 1)\right)^{-\alpha(\mathfrak{m})+1} \\ &= \prod_{\alpha \in Ad(G)'} s_b\left(\alpha(u) - iQ\right)^{1-\alpha(\mathfrak{m})}\,.\end{aligned} \tag{2.120}$$

We note this perturbative piece precisely agrees with what we found by the naïve KK reduction of section 2.3. This confirms our claim that the previous computation only reproduced the perturbative contribution to the $S^3_b \times \Sigma_\mathfrak{g}$ partition function, and missed the instanton corrections. The latter can be seen in the present approach, and we will discuss them below.

The factorization of the perturbative contribution noted above also holds for the general hypermultiplet and vector multiplet in (2.119) and (2.120). This is closely related to the factorization of the S^3_b partition function of 3d $\mathcal{N} = 2$ theories into *holomorphic blocks* [20]. It is therefore natural to expect that a similar approach as the one here applies to the spaces $L(p,q)_b \times S^2$ for arbitrary (squashed) lens space $L(p,q)$, as such lens spaces can be constructed by gluing two such holomorphic blocks. Indeed, we have seen this already in the case of $S^2 \times S^2 \times S^1$, and will return to this example in section 5 below. We

also note that 3d holomorphic blocks have also appeared in 5d partition functions in other contexts — see, e.g., [56, 57] — and it would be interesting to understand the relation to their appearance here. As we will see in the next subsection, this factorization continues to hold at the non-perturbative level.

Instanton contribution and the Nekrasov-Shatashvili limit. The argument above extends to the 1-loop determinant of a general hypermultiplet or vector multiplet in a general instanton background. Specifically, we may similarly uplift the contribution in the background of point-like instantons on $\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}$, which are now configurations supported on loops wrapping the S^1 factor. Thus, we may write

$$Z_{\mathbb{R}^2_{\mathfrak{q}_1} \times \mathbb{R}^2_{\mathfrak{q}_2} \times S^1}(x,y,z) = Z^{\text{pert}}_{\mathbb{R}^2_{\mathfrak{q}_1} \times \mathbb{R}^2_{\mathfrak{q}_2} \times S^1}(x,y,z) \sum_{k=0}^{\infty} z^k Z^{(k)}_{\mathbb{R}^2_{\mathfrak{q}_1} \times \mathbb{R}^2_{\mathfrak{q}_2} \times S^1}(x,y)\,, \qquad (2.121)$$

where $Z^{\text{pert}}_{\mathbb{R}^2_{\mathfrak{q}_1} \times \mathbb{R}^2_{\mathfrak{q}_2} \times S^1}$ is the perturbative contribution from vector and hypermultiplets, and $Z^{(k)}_{\mathbb{R}^2_{\mathfrak{q}_1} \times \mathbb{R}^2_{\mathfrak{q}_2} \times S^1}$ is the 1-loop determinant in the k-instanton background. We have also defined the classical contribution to the instanton action, which is now integrated over S^1,

$$z = e^{2\pi i r \tilde{\gamma}}\,, \qquad \tilde{\gamma} = -\frac{2\pi i}{g_5{}^2}\,. \qquad (2.122)$$

This 5d instanton partition function was originally defined in [18], and plays an important role in many of the 5d partition functions mentioned at the beginning of this section. We see the same object controls the $S^3_b \times S^2$ partition function.

To form the integrand of the $S^3_b \times S^2$ partition function, we take the product of the contributions from the four fixed points,

$$\lim_{\epsilon \to 0} \prod_{\ell} Z_{\mathbb{R}^2_{\mathfrak{q}_1^{(\ell)}} \times \mathbb{R}^2_{\mathfrak{q}_2^{(\ell)}} \times S^1}(x^{(\ell)}, y^{(\ell)}, z^{(\ell)})\,, \qquad (2.123)$$

where the arguments of the instanton partition functions at the four fixed points are $x^{(\ell)}$ and $\mathfrak{q}_i^{(\ell)}$, defined as in (2.108), and $y^{(\ell)}$ and $z^{(\ell)}$, defined analagously. Explicitly,

$$y^{(\ell)} = \begin{cases} e^{2\pi i b Q r \tilde{\nu}} \equiv y\,, & \ell = nn \text{ or } ns \\ e^{2\pi i b^{-1} Q r \tilde{\nu}} \equiv \bar{y}\,, & \ell = sn \text{ or } ss \end{cases} \qquad z^{(\ell)} = \begin{cases} e^{2\pi i b Q r \tilde{\gamma}} \equiv z\,, & \ell = nn \text{ or } ns \\ e^{2\pi i b^{-1} Q r \tilde{\gamma}} \equiv \bar{z}\,, & \ell = sn \text{ or } ss \end{cases}\,. \qquad (2.124)$$

The expression (2.123) involves taking the limit of the instanton partition function where one of the equivariant parameters is sent to zero. This limit has been widely studied, beginning with [19], and is known as the *Nekrasov-Shatashvili limit* of the instanton partition function. As shown in [19], the leading behavior of the instanton partition function in this limit can be expanded as (considering the 4d case first)

$$Z_{\mathbb{R}^2_{\epsilon_1} \times \mathbb{R}^2_{\epsilon_2}}(\tilde{u}) \xrightarrow[\epsilon_2 \to 0]{} \exp\left\{ 2\pi i \left(\frac{1}{\epsilon_2} \mathcal{W}_{NS}(\tilde{u}, \epsilon_1) - \frac{1}{2} \Omega_{NS}(\tilde{u}, \epsilon_1) + \mathcal{O}(\epsilon_2) \right) \right\}\,, \qquad (2.125)$$

which implicitly defines functions, $\mathcal{W}_{NS}(\tilde{u}, \epsilon_1)$ and $\Omega_{NS}(\tilde{u}, \epsilon_2)$, which depend on the theory under consideration.[12] We may similarly define, for the 5d instanton partition function, adapted to the notation above (and setting $r = 1$ from now on),

$$
Z_{\mathbb{R}^2_{\mathfrak{q}_1 = e^{2\pi i \epsilon_1}} \times \mathbb{R}^2_{\mathfrak{q}_2 = e^{2\pi i \epsilon_2}} \times S^1} \left(x = e^{2\pi i \tilde{u}}, y = e^{2\pi i \tilde{\nu}}, z = e^{2\pi i \tilde{\gamma}} \right)
$$
$$
\xrightarrow[\epsilon_2 \to 0]{} \exp \left\{ 2\pi i \left(\frac{1}{\epsilon_2} \mathcal{W}_{NS}^{(5d)}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; \epsilon_1) - \Omega_{NS}^{(5d)}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; \epsilon_1) + O(\epsilon_2) \right) \right\}.
\tag{2.126}
$$

The notation we have chosen anticipates the role these functions will play as the twisted superpotential and effective dilaton below.

Let us consider again the expression (2.123) in light of this expansion. If we again consider the product of the $\ell = nn$ and $\ell = ns$ terms, we find (suppressing the dependence on $\tilde{\nu}$ and $\tilde{\gamma}$ from the notation)

$$
Z_{\mathbb{R}^2_{\mathfrak{q}=e^{-2\pi i b^2}} \times \mathbb{R}^2_{e^{2\pi i Q b \epsilon}} \times S^1} \left(e^{2\pi i b Q (\tilde{u} + \frac{1}{2}\mathfrak{m}\epsilon)} \right) Z_{\mathbb{R}^2_{\mathfrak{q}=e^{-2\pi i b^2}} \times \mathbb{R}^2_{e^{-2\pi i Q b \epsilon}} \times S^1} \left(e^{2\pi i b Q (\tilde{u} + \frac{1}{2}\mathfrak{m}\epsilon)} \right)
$$
$$
\xrightarrow[\epsilon \to 0]{} \exp \left\{ 2\pi i \left(\frac{1}{Qb\epsilon} \mathcal{W}_{NS}^{(5d)} \left(Qb \left(\tilde{u} + \frac{1}{2}\mathfrak{m}\epsilon \right), -b^2 \right) - \frac{1}{2} \Omega_{NS}^{(5d)} \left(Qb \left(\tilde{u} + \frac{1}{2}\mathfrak{m}\epsilon \right), -b^2 \right) \right. \right.
$$
$$
\left. \left. - \frac{1}{Qb\epsilon} \mathcal{W}_{NS}^{(5d)} \left(Qb \left(\tilde{u} - \frac{1}{2}\mathfrak{m}\epsilon \right), -b^2 \right) - \frac{1}{2} \Omega_{NS}^{(5d)} \left(Qb \left(\tilde{u} - \frac{1}{2}\mathfrak{m}\epsilon \right), -b^2 \right) + O(\epsilon) \right) \right\}
$$
$$
\longrightarrow \exp \left\{ 2\pi i \left(\mathfrak{m}\partial_{\tilde{u}} \frac{1}{Qb} \mathcal{W}_{NS}^{(5d)}(Qb\tilde{u}, -b^2) - \Omega_{NS}^{(5d)}(Qb\tilde{u}, -b^2) \right) \right\}.
\tag{2.127}
$$

Similarly, the other two fixed points contribute

$$
Z_{\mathbb{R}^2_{\mathfrak{q}=e^{-2\pi i b^{-2}}} \times \mathbb{R}^2_{e^{2\pi i Q b^{-1}\epsilon}} \times S^1} \left(e^{2\pi i b^{-1} Q (\tilde{u} + \frac{1}{2}\mathfrak{m}\epsilon)} \right) Z_{\mathbb{R}^2_{\mathfrak{q}=e^{-2\pi i b^{-2}}} \times \mathbb{R}^2_{e^{-2\pi i Q b^{-1}\epsilon}} \times S^1} \left(e^{2\pi i b^{-1} Q (\tilde{u} + \frac{1}{2}\mathfrak{m}\epsilon)} \right)
$$
$$
\xrightarrow[\epsilon \to 0]{} \exp \left\{ 2\pi i \left(\mathfrak{m}\partial_{\tilde{u}} \frac{1}{Qb^{-1}} \mathcal{W}_{NS}^{(5d)}(Qb^{-1}\tilde{u}, -b^{-2}) - \Omega_{NS}^{(5d)}(Qb^{-1}\tilde{u}, -b^{-2}) \right) \right\}.
\tag{2.128}
$$

Putting this together, we see these objects behave in precisely the same way as we expect the twisted superpotential and effective dilaton to behave in the context of the $S^2 \times S_b^3$ partition function. Thus it is natural to define

$$
\mathcal{W}_{S_b^3 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = \frac{1}{Qb} \mathcal{W}_{NS}^{(5d)}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; -b^2) + \frac{1}{Qb^{-1}} \mathcal{W}_{NS}^{(5d)}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; -b^{-2}),
\tag{2.129}
$$

$$
\Omega_{S_b^3 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = \Omega_{NS}^{(5d)}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; -b^2) + \Omega_{NS}^{(5d)}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; -b^{-2}).
\tag{2.130}
$$

Thus, we see the factorization of the perturbative contribution continues to hold at the non-perturbative level. We can now write, e.g.,

$$
\mathcal{W}_{S_b^3 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\text{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) + \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\text{inst}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}).
\tag{2.131}
$$

Then we have seen the perturbative piece is given by (2.54), and can in principle be further factorized into $\mathcal{W}_{NS}^{(5d),\text{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; -b^2)$ and $\mathcal{W}_{NS}^{(5d),\text{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; -b^{-2})$, although we do not write those expressions here. The instanton contribution is thus determined implicitly by studying the Nekrasov-Shatashvili limit of the instanton partition function.

---

[12]The notation here should not be confused with that of the fixed points at the north and south poles of the S^2 factors, which we denote with lower case letters.

Integration contour and Bethe sum. Having finally written an expression for the integrand of the $S_b^3 \times S^2$ partition function, we can now write the partition function itself as an integral over a suitable contour and sum over gauge fluxes on S^2, i.e.,

$$Z_{S_b^3 \times S^2}(\tilde{\nu}, \tilde{\gamma})_{\mathfrak{n}} = \frac{1}{|W_G|} \sum_{\mathfrak{m}_a \in \Lambda_G} \oint_{\mathcal{C}_{JK}} d\tilde{u} \, \Pi_a(\tilde{u}, \tilde{\nu}, \tilde{\gamma})^{\mathfrak{m}_a} \Pi_i(\tilde{u}, \tilde{\nu}, \tilde{\gamma})^{\mathfrak{n}_i} e^{-2\pi i \Omega_{S_b^3 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})} \,, \quad (2.132)$$

where

$$\Pi_a(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = e^{2\pi i \partial_{\tilde{u}_a} \mathcal{W}_{S_b^3 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})} \,, \qquad \Pi_i(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = e^{2\pi i \partial_{\tilde{\nu}_i} \mathcal{W}_{S_b^3 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})} \,. \quad (2.133)$$

We have so far not been careful to specify the precise contour of integration in the above integral. However, it is natural to conjecture that this contour is such that, by a similar manipulation as described above and in [11], we may evaluate this contour integral and perform the resulting geometric series over $\mathfrak{m}_a$, and obtain a "Bethe sum" formula,

$$Z_{S_b^3 \times S^2}(\tilde{\nu}, \tilde{\gamma})_{\mathfrak{n}} = \sum_{\hat{\tilde{u}} \in \mathcal{S}_{\mathrm{BE}}} \Pi_i(\hat{\tilde{u}}, \tilde{\nu}, \tilde{\gamma})^{\mathfrak{n}_i} \mathcal{H}(\hat{\tilde{u}}, \tilde{\nu}, \tilde{\gamma})^{-1} \,, \quad (2.134)$$

where $\mathcal{S}_{\mathrm{BE}}$ is defined as in (2.23), and we have defined

$$\mathcal{H}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = e^{2\pi i \Omega_{S_b^3 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})} \det_{a,b} \frac{\partial^2 \mathcal{W}_{S_b^3 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})}{\partial \tilde{u}_a \partial \tilde{u}_b} \,. \quad (2.135)$$

Finally, although the above formula has only been motivated for genus zero, it is very natural to conjecture the generalization to arbitrary genus, $\mathfrak{g}$, using the general A-twist formalism described above, namely,

$$Z_{S_b^3 \times \Sigma_\mathfrak{g}}(\tilde{\nu}, \tilde{\gamma})_{\mathfrak{n}} = \sum_{\hat{\tilde{u}} \in \mathcal{S}_{\mathrm{BE}}} \Pi_i(\hat{\tilde{u}}, \tilde{\nu}, \tilde{\gamma})^{\mathfrak{n}_i} \mathcal{H}(\hat{\tilde{u}}, \tilde{\nu}, \tilde{\gamma})^{\mathfrak{g}-1} \,. \quad (2.136)$$

In fact, this form is essentially determined by topological invariance on $\Sigma_\mathfrak{g}$, which means this observable must be computed by a 2d TQFT. This represents our result for the full, non-perturbative partition function on $S_b^3 \times \Sigma_\mathfrak{g}$.

Summary. We have computed the $S_b^3 \times \Sigma_\mathfrak{g}$ partition function by two methods. First, we performed a naïve reduction to 2d, leading to the perturbative result for the twisted superpotential in (2.54). However, this missed the contribution of instantons, and we then recovered the full non-perturbative result by a reduction to 4d, where we conjectured the answer is expressed in terms of the Nekrasov-Shatashvili limit of the instanton partition function, as in (2.129). From the twisted superpotential (and effective dilaton), we may then construct the full $S_b^3 \times \Sigma_\mathfrak{g}$ partition function as in the general discussion of section 2.2, leading to the formulae (2.56) for the perturbative partition function, and (2.136) for the non-perturbative result.

While the latter method in principle gives the complete answer, in practice the instanton contributions may be difficult or impossible to compute analytically. For the remainder of this paper, we will therefore consider various simplifying limits in which their contribution is suppressed, or otherwise under control. Specifically, it is well-known that their

contribution is subleading in N when we take a large N limit, and in the next section we consider this limit in several examples, utilizing only the perturbative contribution to the partition function. In the following section, we will specialize to theories which we expect to have a 6d UV completion. For such theories we will argue that in the limit of large gauge coupling, corresponding to large radius of the emergent S^1 direction, the instantons contribute a simple factor to the partition function, which can be interpreted as computing the Casimir energy in this limit. Finally, in the special case of the maximal 5d $\mathcal{N} = 2$ SYM theory, we will see there exists a limit with enhanced supersymmetry, where the instanton contribution is very simple.

3 Large N limit and holography

In this section, we study the large N limit of the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function for a large class of 5d quiver gauge theories. We distinguish two classes of quivers; those leading to an $N^{5/2}$ scaling of the free energy and those leading to an N^3 scaling. The latter class of theories (which are expected to have UV completions as 6d theories on a circle) are discussed in detail at finite N in section 4. For the former class, we find that the matrix model determining the partition function exhibits an interesting structure at large N, becoming closely related to the matrix model determining the S^5 partition function [32, 33, 58–60]. Precisely, we will establish the large N relations

$$
\begin{aligned}
F_{S_b^3 \times \Sigma_{\mathfrak{g}}} &= -6\pi(\mathfrak{g} - 1) \left(\frac{\sum_{\alpha \in G} c_\alpha - \sum_{\rho \in R} c_\rho \hat{\mathfrak{n}}_\rho \tilde{\nu}_\rho}{\sum_{\rho \in R} c_\rho \left(1 - \tilde{\nu}_\rho^2\right)} \right) \mathcal{W}_{S_b^3 \times \mathbb{R}^2} , \\
\mathcal{W}_{S_b^3 \times \mathbb{R}^2} &= \frac{4Q^2}{27\pi} \left(\frac{\sum_{\rho \in R} c_\rho (1 - \tilde{\nu}_\rho^2)}{\sum_{\rho \in R} c_\rho} \right)^{3/2} F_{S^5} ,
\end{aligned}
\tag{3.1}
$$

where the sum $\sum_\alpha$ is over the vector multiplets and $\sum_\rho$ is over various hypermultiplets and F is the free energy of the theory on the corresponding manifold.[13] Thus, although the matrix models computing the partition functions on S^5 and on $S_b^3 \times \Sigma_{\mathfrak{g}}$ are distinct at finite N, they are closely related at large N.

As we discuss in detail below, the universal twist $\hat{\mathfrak{n}}_\rho = \tilde{\nu}_\rho = 0$ is special [21]. This corresponds to a topological twist purely along $U(1)_R \subset SU(2)_R$. Combining the expressions above it follows that[14]

$$
\mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\text{univ}} = \frac{4Q^2}{27\pi} F_{S^5} , \qquad F_{S_b^3 \times \Sigma_{\mathfrak{g}}}^{\text{univ}} = -\frac{8}{9}(\mathfrak{g} - 1)Q^2 F_{S^5} ,
\tag{3.2}
$$

for any 5d $\mathcal{N} = 1$ theory with a universal twist on $\Sigma_{\mathfrak{g}}$ at large N. The second relation above for $Q = 1$ is in agreement with the supergravity prediction of [21], valid for the compactification of any 5d $\mathcal{N} = 1$ theory with a universal twist. We will discuss holography in section 3.4.

---

[13] See sections 3.1 and 3.3 for notation and details.

[14] As shown in section 3.3, for this class of quivers $\sum_\alpha c_\alpha = \sum_\rho c_\rho$ and the universal relations follow.

Useful formulas. Before we proceed to the computation, let us collect the relevant results of section 2 for ease of reference. For a general theory with gauge group G and hypermultiplets in gauge representations R_I, and collecting the terms from (2.56), (2.57), the perturbative partition function is given by

$$Z_{S_b^3 \times \Sigma_{\mathfrak{g}}}^{\text{pert}} = \sum_{\hat{u} \in \mathcal{S}_{\text{BE}}} H^{\mathfrak{g}-1} \prod_{\alpha \in Ad(G)'} s_b \left(\alpha(\hat{u}) - iQ\right)^{1-\mathfrak{g}} \prod_I \prod_{\rho \in R_I} s_b \left(\rho(\hat{u}) + \nu_I\right)^{(\mathfrak{g}-1)\hat{\mathfrak{n}}_I}, \quad (3.3)$$

where H denotes the Hessian contribution to the handle-gluing operator in the second line of (2.57) and we have written the flavor flux as $\mathfrak{n}_I = \hat{\mathfrak{n}}_I(\mathfrak{g} - 1)$, with $\hat{\mathfrak{n}}_I$ integer-quantized. The Bethe equations are given by

$$\mathcal{S}_{\text{BE}} = \left\{ \hat{u} \ \Big| \ \Pi_a(\hat{u}) \equiv \exp\left(2\pi i \frac{\partial \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\text{pert}}}{\partial u_a}(\hat{u})\right) = 1, \quad a = 1, \ldots, r_G \right\} / W_G, \quad (3.4)$$

with the twisted superpotential (2.54), given by

$$\mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\text{pert}}(\tilde{u}) = \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\text{classical}}(\tilde{u}, \tilde{\nu}) + \sum_I \sum_{\rho \in R_I} g_b\big(\rho(\tilde{u}) + \tilde{\nu}_I\big) - \sum_\alpha g_b\big(\alpha(\tilde{u}) + 1\big), \quad (3.5)$$

where the classical contribution is given in (2.53), and for the theories considered in this section the Hessian contribution, H, given in (3.3) is subleading in the large N limit. Since instanton corrections are suppressed at large N, and to avoid clutter, in the remainder of this section we shall always omit the label "pert."

To study the large N limit, we will need the following asymptotic behaviors:

$$g_b(\tilde{u} + \tilde{\nu}) \rightarrow \pm\left(-\frac{1}{12}Q^2\tilde{u}^3 + \frac{1}{4}Q^2\tilde{\nu}\tilde{u}^2 - \frac{1}{4}\left[\frac{1+4Q^2}{6} - Q^2(1-\tilde{\nu}^2)\right]\tilde{u}\right), \quad (3.6)$$

$$\ell_b(\tilde{u} + \tilde{\nu}) \rightarrow \pm\frac{Q^2}{2}\left(\frac{1}{2}\tilde{u}^2 + \tilde{\nu}\tilde{u}\right), \qquad \text{for } \text{Im}(\tilde{u}) \rightarrow \pm\infty, \quad (3.7)$$

where we have assumed that $\tilde{\nu} \in \mathbb{R}$ and $|\tilde{\nu}| \leq 1$, and took the principal branch. We note also the expansion

$$s_b(x) \rightarrow e^{\pm \frac{i\pi}{2}\left(x^2 + \frac{1}{3}Q^2 - \frac{1}{6}\right)}, \qquad \text{for } \text{Re}(x) \rightarrow \pm\infty. \quad (3.8)$$

3.1 Seiberg theory and its orbifolds

The Seiberg theory [5] (see also [31]) consists of a single 5d $\mathcal{N} = 1$ vector multiplet in the adjoint of the gauge group $\text{Sp}(N) \simeq \text{USp}(2N)$,[15] $N_f < 8$ hypermultiplets in the fundamental representation of the gauge group, and one hypermultiplet in the antisymmetric representation (see figure 1). The global symmetry of the theory is $\text{SO}(5) \times \text{SU}(2)_R \times$

---

[15] Since the terminology regarding symplectic groups differs in the literature, let us clarify the one followed here. The group of real $2N \times 2N$ matrices U such that $U^{\mathsf{T}}\Omega U = \Omega$, with $\Omega^2 = -1$, is denoted $\text{Sp}(2N, \mathbb{R})$. The group of such complex matrices is denoted $\text{Sp}(2N, \mathbb{C})$. If one adds to the later the condition that U is unitary, one obtains the group $\text{USp}(2N) \equiv \text{Sp}(2N, \mathbb{C}) \cap \text{U}(2N)$, also denoted $\text{Sp}(N)$. It is useful to know that $\text{SU}(N) \times \text{U}(1) \subset \text{USp}(2N) \subset \text{SU}(2N)$.
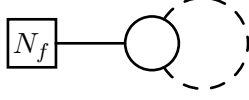
Figure 1. The 5d Seiberg theory. The node represents a $\mathrm{USp}(2N)$ gauge group, the solid line N_f hypermultiplets in the fundamental representation and the dashed line an antisymmetric hypermultiplet.

$\mathrm{SU}(2)_M \times \mathrm{SO}(2N_f) \times \mathrm{U}(1)_{\mathrm{top}}$.[16] The antisymmetric hypermultiplet transforms as a doublet of $\mathrm{SU}(2)_M$ and as a singlet of $\mathrm{SO}(2N_f)$. We turn general flavor fluxes along the Cartan of $\mathrm{SU}(2)_M$ and $\mathrm{SO}(2N_f)$, denoted $\hat{\mathfrak{n}}_M$ and $\hat{\mathfrak{n}}_I$, $I = 1, \ldots, N_f$, respectively.

The fundamental representation of $\mathrm{USp}(2N)$ has weights $\pm e_i$, where e_i are unit vectors of $\mathbb{R}^N$. The antisymmetric representation has weights $\pm(e_i - e_j)$ and $\pm(e_i + e_j)$ with $i < j$. The adjoint has the same weights as the antisymmetric and also $\pm 2e_i$. Then, the twisted superpotential (3.5) reads

$$
\begin{aligned}
\mathcal{W}_{S_b^3 \times \mathbb{R}^2} = &-\sum_{\pm} \sum_{i<j} \left[g_b\big(1 \pm (\tilde{u}_i - \tilde{u}_j)\big) + g_b\big(1 \pm (\tilde{u}_i + \tilde{u}_j)\big) \right] - \sum_i g_b\big(1 \pm 2\tilde{u}_i\big) \\
&+ \sum_{\pm} \sum_{i<j} \left[g_b\big(\tilde{\nu}_{AS} \pm (\tilde{u}_i - \tilde{u}_j)\big) + g_b\big(\tilde{\nu}_{AS} \pm (\tilde{u}_i + \tilde{u}_j)\big) \right] \\
&+ \sum_{I=1}^{N_f} \sum_{\pm} \sum_i g_b\big(\tilde{\nu}_I \pm \tilde{u}_i\big) ,
\end{aligned}
\tag{3.9}
$$

where we have omitted the classical piece $\mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathrm{classical}}$ since one can check that it is subleading in N. To analyze the large N behavior of this quantity we follow the approach introduced in [61], replacing

$$
\tilde{u}_i \to iN^\alpha x , \qquad \sum_i \to N \int dx \, \rho(x) ,
\tag{3.10}
$$

where $\alpha > 0$, x is a continuous variable of order $\mathcal{O}(N^0)$, and $\rho(x)$ is the eigenvalue density, normalized as $\int dx \, \rho(x) = 1$. Using the large $\tilde{u}$ expansion (3.6) and then going to the continuum variables, we have[17]

$$
\begin{aligned}
\mathcal{W}_{S_b^3 \times \mathbb{R}^2} \approx &\frac{(8 - N_f)}{6} Q^2 \, N^{1+3\alpha} \int dx \, \rho(x) \, |x|^3 \\
&- \frac{1}{2} Q^2 \left(1 - \tilde{\nu}_{AS}^2\right) N^{2+\alpha} \int_{y<x} dx \, dy \, \rho(x)\rho(y) \left(|x+y| + |x-y|\right) .
\end{aligned}
\tag{3.11}
$$

For this function to have a nontrivial saddle at large N requires both terms to be of the same order in N and hence $\alpha = \frac{1}{2}$, which we set in what follows. We note that, to this order in N, only the mixing parameter for the antisymmetric hypermultiplet, $\tilde{\nu}_{AS}$, is visible while the parameters for the fundamental fields, $\tilde{\nu}_I$, are not. We also note there has been

---

[16]The $\mathrm{U}(1)_{\mathrm{top}}$ is a topological symmetry which together with the $\mathrm{SO}(2N_f)$ is expected to combine into an enhanced E_{N_f+1} at the conformal fixed point [5].

[17]Due to the structure of roots and weights for $\mathrm{USp}(2N)$ and $\mathrm{SU}(2N)$, only the real part of $g_b(z)$ contributes from the expansion (3.6) and $\mathcal{W}_{S_b^3 \times \mathbb{R}^2}$ is real. The same holds for $\ell_b(z)$ appearing in the free energy.

a cancellation of the nonlocal cubic terms $|x \pm y|^3$ among the vector multiplet and the antisymmetric hypermultiplet. The same cancellation occurs for the free energy of this theory on S^5 [62]. In fact, by a simple rescaling of the coordinates

$$x \to \frac{2}{3}\sqrt{1 - \tilde{\nu}_{AS}^2}\, x\,, \tag{3.12}$$

and a corresponding inverse rescaling of ρ to preserve its normalization, we have[18]

$$\mathcal{W}_{S_b^3 \times \mathbb{R}^2} \approx \frac{4Q^2}{27\pi}\left(1 - \tilde{\nu}_{AS}^2\right)^{3/2}\left[\frac{\pi(8 - N_f)}{3}\, N^{5/2} \int dx\, \rho(x)\, |x|^3 \right.$$
$$\left. - \frac{9\pi}{8}\, N^{5/2} \int dx\, dy\, \rho(x)\rho(y)\left(|x + y| + |x - y|\right)\right]. \tag{3.13}$$

We note the quantity inside the brackets is precisely the free energy functional of this theory on S^5 (see eq. (3.4) in [62]). Thus, we find the simple relation

$$\mathcal{W}_{S_b^3 \times \mathbb{R}^2} = \frac{4Q^2}{27\pi}\left(1 - \tilde{\nu}_{AS}^2\right)^{3/2} F_{S^5}\,. \tag{3.14}$$

Since this is a functional relation in the eigenvalue density $\rho(x)$, extremization of the Bethe potential is equivalent to the extremization of F_{S^5}. As shown in [62] the saddle configuration is given by

$$\rho(x) = \frac{2|x|}{x_*^2}\,, \qquad x_* = \frac{3}{\sqrt{2(8 - N_f)}}\,, \tag{3.15}$$

with $x \in [0, x_*]$ and at the extremum

$$F_{S^5} = -\frac{9\sqrt{2}\pi N^{5/2}}{5\sqrt{8 - N_f}}\,. \tag{3.16}$$

Now let us evaluate the free energy on $S_b^3 \times \Sigma_\mathfrak{g}$ in this Bethe vacuum. Taking the logarithm of (3.3) for this theory it is easy to see that the Hessian is subleading in this vacuum and only the flux operators contribute, which using the expansion (3.7) and vacuum (3.15) becomes

$$F_{S_b^3 \times \Sigma_\mathfrak{g}} = -\frac{8}{9}(\mathfrak{g} - 1)Q^2(1 - \hat{\mathfrak{n}}_M \tilde{\nu}_{AS})\sqrt{1 - \tilde{\nu}_{AS}^2}\, F_{S^5}\,. \tag{3.17}$$

As we will show in section 3.3 the relations (3.14) and (3.17) are not particular to the Seiberg theory but can be generalized to a large class of quiver gauge theories. This is rather nontrivial as it requires a number of special cancellations in the Bethe potential. We note that relations analogous to these hold for 3d $\mathcal{N} = 2$ theories on $S^1 \times \Sigma_\mathfrak{g}$; in that case the relations are among the topological twisted index, corresponding Bethe potential, and the free energy on the round S^3 [63, 64].

Before discussing the general case we prove these relations for orbifolds of the Seiberg theory.

---

[18]We have used the symmetry of the integrand to replace $\int_{y<x} dxdy \to \frac{1}{2}\int dxdy$, where now both integration variables x, y are integrated over their full domain.
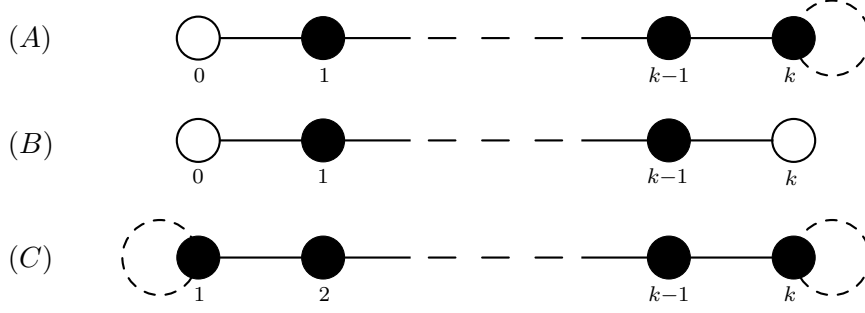
Figure 2. The 5d quiver gauge theories of discussed in [65]. Black nodes represent $SU(2N)$ gauge groups and white ones $USp(2N)$ gauge groups. Solid lines denote bifundamental hypermultiplets and dashed lines hypermultiplets in the antisymmetric representation. In addition, at any given node, a, one may have N_f^a number of fundamental hypermultiplets, which we have not depicted.

Orbifolds. These theories are obtained as $\mathbb{Z}_n$ orbifolds of the Seiberg theory discussed above and consist of linear quivers with $SU(2N)$ nodes, in addition to $USp(2N)$ nodes [65]. There are three classes of quivers: class (A) for $n = 2k + 1$ odd and; $(B), (C)$ when $n = 2k$ is even, shown in figure 2. The Seiberg theory corresponds to class (A) with $k = 0$.

Analyzing each class separately one finds that the Bethe equations for each quiver can be obtained from the following Bethe potentials:[19]

$$\mathcal{W}^{(A)} = \mathcal{W}_{W_0}^\circ + \sum_{a=1}^{k} \mathcal{W}_{W_a}^\bullet + \sum_{a=0}^{k-1} \mathcal{W}_{BF_{(a,a+1)}} + \mathcal{W}_{AS_k} + \sum_{a=0}^{k} N_f^a \, \mathcal{W}_{F_a} \, , \qquad (3.18)$$

$$\mathcal{W}^{(B)} = \mathcal{W}_{W_0}^\circ + \sum_{a=1}^{k-1} \mathcal{W}_{W_a}^\bullet + \mathcal{W}_{W_k}^\circ + \sum_{a=0}^{k-1} \mathcal{W}_{BF_{(a,a+1)}} + \sum_{a=0}^{k} N_f^a \, \mathcal{W}_{F_a} \, , \qquad (3.19)$$

$$\mathcal{W}^{(C)} = \sum_{a=1}^{k} \mathcal{V}_{W_a}^\bullet + \sum_{a=1}^{k-1} \mathcal{W}_{BF_{(a,a+1)}} + \mathcal{W}_{AS_1} + \mathcal{W}_{AS_k} + \sum_{a=1}^{k} N_f^a \, \mathcal{W}_{F_a} \, , \qquad (3.20)$$

where each term is given by

$$\mathcal{W}_W^\circ = - \sum_{\pm, i, j < i} \left[g_b(1 \pm (\tilde{u}^i - \tilde{u}^j)) + g_b(1 \pm (\tilde{u}^i + \tilde{u}^j)) \right] - \sum_{\pm, i} g_b(1 \pm 2\tilde{u}^i) \, ,$$

$$\mathcal{W}_W^\bullet = -2 \sum_{\pm, i, j < i} \left[g_b(1 \pm (\tilde{u}^i - \tilde{u}^j)) + g_b(1 \pm (\tilde{u}^i + \tilde{u}^j)) \right] - \sum_{\pm, i} g_b(1 \pm 2\tilde{u}^i) \, ,$$

$$\mathcal{W}_{BF_{(a,b)}} = 2 \sum_{\pm, i, j < i} \left[g_b(\tilde{\nu}_{(a,b)} \pm (\tilde{u}^i - \tilde{u}^j)) + g_b(\tilde{\nu}_{(a,b)} \pm (\tilde{u}^i + \tilde{u}^j)) \right] + \sum_{\pm, i} g_b(\tilde{\nu}_{(a,b)} \pm 2\tilde{u}^i) \, ,$$

$$\mathcal{W}_{AS} = \sum_{\pm, i, j < i} \left[g_b(\tilde{\nu}_{AS} \pm (\tilde{u}^i - \tilde{u}^j)) + g_b(\tilde{\nu}_{AS} \pm (\tilde{u}^i + \tilde{u}^j)) \right] \, ,$$

$$\mathcal{W}_F = \sum_{\pm, i} g_b(\tilde{\nu}_F \pm \tilde{u}^i) \, . \qquad (3.21)$$

To write the above expressions, we start with fundamental weights in the $\mathbb{R}^{2N}$ basis subject to the constraint $\sum_{A=1}^{2N} \tilde{u}^A = 0$ for $SU(2N)$ group and $\tilde{u}^i = -\tilde{u}^{N+i}$ with $i = 1, \ldots, N$ for

---

[19]To avoid clutter we denote $\mathcal{W}_{S_b^3 \times \mathbb{R}^2} = \mathcal{W}$ in what follows.

USp($2N$) group, as explained in [62]. To extremize the Bethe potentials, we chose the same Ansatz for $\tilde{u}$'s corresponding to different nodes, i.e., $\tilde{u}_a^i \equiv \tilde{u}^i$. The final result in terms of $\tilde{u}^i$ then follows for all representations straightforwardly. The continuum limit now follows as before with $\tilde{u}^i \to iN^{1/2}x$ and using the large $\tilde{u}$ expansion (3.6). Let us work out the coefficients of different terms at various orders of N in $\mathcal{W}^{(A)}$ (ignoring overall signs and factors):

$$\mathcal{O}(N^{7/2}) \text{ with } |x \pm y|^3 : 1 + 2k - 2k - 1 + 0 = 0 \,,$$

$$\mathcal{O}(N^{5/2}) \text{ with } |x|^3 : 8 + 8k - 8k + 0 - \sum_a N_f^a = 8 - N_f \,,$$

$$\mathcal{O}(N^{5/2}) \text{ with } |x \pm y| : \left(\frac{1+4Q^2}{6}\right) + 2k\left(\frac{1+4Q^2}{6}\right) - 2\sum_{a=0}^{k-1}\left(\frac{1+4Q^2}{6} - Q^2(1-\tilde{\nu}_{(a,a+1)}^2)\right)$$

$$- \left(\frac{1+4Q^2}{6} - Q^2(1-\tilde{\nu}_{AS_k}^2)\right) + 0 = Q^2\left(\sum_{a=0}^{k-1}\left(2(1-\tilde{\nu}_{(a,a+1)}^2)\right) + (1-\tilde{\nu}_{AS_k}^2)\right), \quad (3.22)$$

where we defined $N_f = \sum_a N_f^a$. Note that the field content is such that the terms at $\mathcal{O}(N^{7/2})$ cancel leaving the expected terms at $\mathcal{O}(N^{5/2})$ to be dominant. This happens for classes (B) and (C) as well as can be easily checked. In fact, for all the three classes, the cubic term has the same coefficient as the Seiberg theory and only the coefficient of the nonlocal linear term changes depending on the precise matter field content. Repeating the same extremization procedure done for the Seiberg theory for each class, we obtain the following Bethe potentials:

$$\mathcal{W}^{(A)} = \frac{4Q^2}{27\pi}\left(\frac{\sum_{a=0}^{k-1}\left(2(1-\tilde{\nu}_{(a,a+1)}^2)\right) + (1-\tilde{\nu}_{AS_k}^2)}{2k+1}\right)^{3/2} F_{S^5}^{(A)}, \quad (3.23)$$

$$\mathcal{W}^{(B)} = \frac{4Q^2}{27\pi}\left(\frac{\sum_{a=0}^{k-1}\left(2(1-\tilde{\nu}_{(a,a+1)}^2)\right)}{2k}\right)^{3/2} F_{S^5}^{(B)}, \quad (3.24)$$

$$\mathcal{W}^{(C)} = \frac{4Q^2}{27\pi}\left(\frac{\sum_{a=1}^{k-1}\left(2(1-\tilde{\nu}_{(a,a+1)}^2)\right) + (1-\tilde{\nu}_{AS_1}^2) + (1-\tilde{\nu}_{AS_k}^2)}{2k}\right)^{3/2} F_{S^5}^{(C)}, \quad (3.25)$$

where $F_{S^5}^{(A,B,C)} = n^{3/2}F_{S^5}^{\text{Seiberg}}$ [62]. The free energy calculation is similar to what was done for the Seiberg theory. We turn on flavor fluxes for all the $\text{U}(1)_{(a,a+1)}$ and $\text{U}(1)_{AS}$ global symmetries of the quiver theories. This leads to the following free energies

$$F_{S_b^3 \times \Sigma_{\mathfrak{g}}}^{(A)} = -6\pi(\mathfrak{g}-1)\frac{2k+1 - \sum_{a=0}^{k-1} 2\hat{\mathfrak{n}}_{(a,a+1)}\tilde{\nu}_{(a,a+1)} - \hat{\mathfrak{n}}_{AS_k}\tilde{\nu}_{AS_k}}{\left(\sum_{a=0}^{k-1}\left(2(1-\tilde{\nu}_{(a,a+1)}^2)\right) + (1-\tilde{\nu}_{AS_k}^2)\right)}\mathcal{W}^{(A)}, \quad (3.26)$$

$$F_{S_b^3 \times \Sigma_{\mathfrak{g}}}^{(B)} = -6\pi(\mathfrak{g}-1)\frac{2k - \sum_{a=0}^{k-1} 2\hat{\mathfrak{n}}_{(a,a+1)}\tilde{\nu}_{(a,a+1)}}{\sum_{a=0}^{k-1}\left(2(1-\tilde{\nu}_{(a,a+1)}^2)\right)}\mathcal{W}^{(B)}, \quad (3.27)$$

$$F_{S_b^3 \times \Sigma_{\mathfrak{g}}}^{(C)} = -6\pi(\mathfrak{g}-1)\frac{2k - \sum_{a=1}^{k-1} 2\hat{\mathfrak{n}}_{(a,a+1)}\tilde{\nu}_{(a,a+1)} - \hat{\mathfrak{n}}_{AS_1}\tilde{\nu}_{AS_1} - \hat{\mathfrak{n}}_{AS_k}\tilde{\nu}_{AS_k}}{\left(\sum_{a=1}^{k-1}\left(2(1-\tilde{\nu}_{(a,a+1)}^2)\right) + (1-\tilde{\nu}_{AS_1}^2) + (1-\tilde{\nu}_{AS_k}^2)\right)}\mathcal{W}^{(C)}. \quad (3.28)$$

We note that for the universal twist, $\hat{\mathfrak{n}} = \tilde{\nu} = 0$, the formulas above reduce to $\mathcal{W} = \frac{4Q^2}{27\pi} F_{S^5}$ and $F_{S_b^3 \times \Sigma_\mathfrak{g}} = -6\pi(\mathfrak{g}-1)\mathcal{W}$ for all classes of orbifolds. We will see this observation generalizes to a large class of quivers with an $N^{5/2}$ scaling in section 3.3.

3.2 Extremization

We recall that when $Z_{S_b^3 \times \Sigma_\mathfrak{g}}$ is viewed as direct sum of 3d theories, the fugacity parameters $\tilde{\nu}$ are associated to the 3d R-charge, $\Delta = 1 - \tilde{\nu}$, of the 3d $\mathcal{N} = 2$ chiral fields obtained by reduction of the hypermultiplets on $\Sigma_\mathfrak{g}$. Since the theory is topological on $\Sigma_\mathfrak{g}$ we can shrink it to zero size and $F_{S_b^3 \times \Sigma_\mathfrak{g}} = F_{S_b^3}$ coincides with the free energy of the 3d effective theory on S_b^3. In view of F-maximization [66, 67] it is thus natural to propose that in order to obtain information about the 3d IR fixed point (assuming such fixed point exists), the fugacities must be set to those values extremizing the free energy:[20]

$$\frac{\partial F_{S_b^3 \times \Sigma_\mathfrak{g}}}{\partial \tilde{\nu}_\rho} = 0 \,. \tag{3.29}$$

Carrying this out for each class of quivers leads to a coupled set of quadratic equations with two sets of solutions:

$$\tilde{\nu}_I^{(A,B,C)} = \frac{\hat{\mathfrak{n}}_I^{(A,B,C)}}{4\hat{\mathfrak{n}}_{(A,B,C)}^2} \left(1 \pm \sqrt{1 + 8\hat{\mathfrak{n}}_{(A,B,C)}^2}\right), \tag{3.30}$$

where we defined

$$\hat{\mathfrak{n}}_{(A)}^2 = \frac{1}{2k+1} \left(\sum_{a=0}^{k-1} 2\hat{\mathfrak{n}}_{(a,a+1)}^2 + \hat{\mathfrak{n}}_{AS_k}^2 \right), \tag{3.31}$$

$$\hat{\mathfrak{n}}_{(B)}^2 = \frac{1}{2k} \left(\sum_{a=0}^{k-1} 2\hat{\mathfrak{n}}_{(a,a+1)}^2 \right), \tag{3.32}$$

$$\hat{\mathfrak{n}}_{(C)}^2 = \frac{1}{2k} \left(\sum_{a=1}^{k-1} 2\hat{\mathfrak{n}}_{(a,a+1)}^2 + \hat{\mathfrak{n}}_{AS_1}^2 + \hat{\mathfrak{n}}_{AS_k}^2 \right). \tag{3.33}$$

We emphasize the index I in (3.30) runs over all antisymmetric and bifundamental hypermultiplets in the corresponding quiver, and not fundamental ones.[21] Plugging the solutions back into $F_{S_b^3 \times \Sigma_\mathfrak{g}}^{(A,B,C)}$ gives the common formula

$$F_{S_b^3 \times \Sigma_\mathfrak{g}}^{(A,B,C)} = -\frac{8}{9}(\mathfrak{g}-1)Q^2 \left(\pm \frac{\sqrt{2}\left|\hat{\mathfrak{n}}_{(A,B,C)}^2 - 1\right|^{3/2}\left(\sqrt{1 + 8\hat{\mathfrak{n}}_{(A,B,C)}^2} \pm 1\right)}{\left|4\hat{\mathfrak{n}}_{(A,B,C)}^2 - 1 \mp \sqrt{1 + 8\hat{\mathfrak{n}}_{(A,B,C)}^2}\right|^{3/2}} \right) F_{S^5}^{(A,B,C)} \,. \tag{3.34}$$

Depending on the values of the flavor fluxes, one of the roots in (3.30) may be discarded by the requirement $|\tilde{\nu}_I| \leq 1$ (i.e., $0 \leq \Delta_I \leq 2$), which we have assumed.

---

[20]We are assuming there are no accidental flavor symmetries in the IR.

[21]Just as in the Seiberg theory, the fugacities for all fundamental fields are not visible at this order in N and thus are not fixed by extremization.

Note that for the universal twist $\hat{\mathfrak{n}} = \tilde{\nu} = 0$ this reduces to (3.2) for all classes. For non-universal twists the relation among free energies depends on the theory under consideration. Specifying (3.34) for the case of the Seiberg theory with a nonzero flux for its $SU(2)_M$ flavor symmetry matches the supergravity result recently found in [22] where the relevant supergravity solution was constructed.

3.3 General quivers

Consider a 5d quiver gauge theory with a number N_V° and $N_V^\bullet$ of symplectic and unitary gauge groups, respectively, and matter fields in the fundamental, bifundamental, antisymmetric, or adjoint representations of the gauge groups. The perturbative Bethe potential (3.5) receives contributions from all these fields, but the scaling with N of each contribution depends on the particular weights of the representations. We recall that for $USp(2N)$ gauge group, the fundamental representation has weights $\pm e_i$, where e_i are unit vectors of $\mathbb{R}^N$. The antisymmetric representation has weights $\pm(e_i - e_j)$ and $\pm(e_i + e_j)$ with $i < j$. The adjoint has the same weights as the antisymmetric and also $\pm 2e_i$. For $U(N)$ gauge group the fundamental has weights e_i, the antisymmetric $\pm(e^i + e^j)$ and the adjoint $\pm(e_i - e_j)$ with $i < j$. We refer to weights with two nonzero entries as "nonlocal weights" as in the continuum limit they lead to nonlocal terms in the Bethe potential. Since a given representation may have both kinds of weights we introduce the continuum notation $\rho(u) \to \boldsymbol{\rho}(x, y)$ for the nonlocal terms and $\rho(u) \to \boldsymbol{\rho}(x)$ for the local ones,[22] and similarly for the roots $\alpha(u)$. With this notation, using the expansion (3.6) and rearranging the various terms that appear, the Bethe potential reads

$$
\begin{aligned}
\mathcal{W} = &-\frac{1}{2}Q^2\tilde{\gamma}N^{1+2\alpha}\sum_{\boldsymbol{\alpha}} N_{\boldsymbol{\alpha}} \int dx \rho(x) x^2 \\
&+ \frac{1}{2}\sum_{\boldsymbol{\alpha}}\left(N^2 \int dx dy \rho(x)\rho(y)\left[\frac{Q^2}{3}N^{3\alpha}|\boldsymbol{\alpha}(x,y)|^3 - \frac{1+4Q^2}{6}N^\alpha|\boldsymbol{\alpha}(x,y)|\right]\right. \\
&\qquad\qquad \left.+ N \int dx \rho(x)\left[\frac{Q^2}{3}N^{3\alpha}|\boldsymbol{\alpha}(x)|^3 - \frac{1+4Q^2}{6}N^\alpha|\boldsymbol{\alpha}(x)|\right]\right) \\
&-\frac{1}{2}\sum_{\boldsymbol{\rho}}\left(N^2 \int dx dy \rho(x)\rho(y)\left[\frac{Q^2}{3}N^{3\alpha}|\boldsymbol{\rho}(x,y)|^3 - \left(\frac{1+4Q^2}{6} - Q^2(1-\tilde{\nu}_{\boldsymbol{\rho}}^2)\right)N^\alpha|\boldsymbol{\rho}(x,y)|\right]\right. \\
&\qquad\qquad \left.+ N \int dx \rho(x)\left[\frac{Q^2}{3}N^{3\alpha}|\boldsymbol{\rho}(x)|^3 - \left(\frac{1+4Q^2}{6} - Q^2(1-\tilde{\nu}_{\boldsymbol{\rho}}^2)\right)N^\alpha|\boldsymbol{\rho}(x)|\right]\right).
\end{aligned}
\tag{3.35}
$$

We see that the various terms have different scalings in N: $1+2\alpha, 2+3\alpha, 2+\alpha, 1+3\alpha, 1+\alpha$. In order to have a nontrivial saddle at large N both a quadratic and linear term in the eigenvalue density must appear at a given order in N. Although in principle there seem to be six cases, demanding $\alpha > 0$ there are in fact only two scalings that provide a nontrivial saddle point:

(I) $2 + \alpha = 1 + 3\alpha \Rightarrow \alpha = \frac{1}{2} \Rightarrow N^{5/2}$ Higher order to cancel: $N^{2+3\alpha=7/2}$

(II) $2 + \alpha = 1 + 2\alpha \Rightarrow \alpha = 1 \Rightarrow N^3$ Higher orders to cancel: $N^{2+3\alpha=5}$ and $N^{1+3\alpha=4}$

---

[22]We denote the weights by boldface in order not to be confused with the eigenvalue density ρ. We will switch to normal font in expressions we get after the integration is done.

In each case the cancellation of higher orders is required to avoid a trivial saddle point, which would otherwise render the large N method used here inapplicable. We consider each case in turn.

(I) Theories with $N^{5/2}$ scaling. We assume the generic quivers have $\mathrm{USp}(2N)$ and $\mathrm{SU}(2N)$ gauge groups,[23] with $N_V^\circ, N_V^\bullet$ number of vector multiplets corresponding to respective gauge groups and $N_{Ad}, N_{AS}, N_{BF}, N_f$ number of hypermultiplets in the adjoint, antisymmetric, bifundamental and fundamental representations of given gauge groups, respectively. Setting $\alpha = \frac{1}{2}$ in (3.35) the cancellation of nonlocal cubic terms at $\mathcal{O}(N^{7/2})$ requires

$$\left(\sum_\alpha c_\alpha - \sum_\rho c_\rho \right) \int dx dy \rho(x) \rho(y) \left(|x+y|^3 + |x-y|^3 \right) = 0$$

$$\Rightarrow \sum_\alpha c_\alpha - \sum_\rho c_\rho = 0 \,, \qquad (3.36)$$

where c_α and c_ρ are numerical constants for each vector and hypermultiplet that appear when collecting all the contributions to the integral shown above. For instance, $c_V^\circ = 1$, $c_V^\bullet = 2$ for each vector multiplet in the quiver and $c_{Ad} = 1$, $c_{AS} = 1$, $c_{BF} = 2$, $c_f = 0$ for each hypermultiplet. Thus, a constraint is imposed on the number of various fields present in the theory. We have already seen an example of such a relation for type (A) quiver theory in (3.22), where $\sum_\rho c_\rho = 1 + 2k = N_{AS} + 2N_{BF}$, which is also equal to $\sum_\alpha c_\alpha = N_V^\circ + 2N_V^\bullet$ for this quiver. Thus, the Bethe potential for these theories with an $N^{5/2}$ scaling is given by[24]

$$\mathcal{W} = N^{5/2} Q^2 \left[\frac{1}{6} \int dx \rho(x) \left(\sum_{\boldsymbol{\alpha}} |\boldsymbol{\alpha}(x)|^3 - \sum_{\boldsymbol{\rho}} |\boldsymbol{\rho}(x)|^3 \right) \right.$$

$$\left. - \frac{1}{2} \int dx dy \, \rho(x) \rho(y) \sum_{\boldsymbol{\rho}} \left(1 - \tilde{\nu}_\rho^2 \right) |\boldsymbol{\rho}(x,y)| \right]$$

$$= N^{5/2} Q^2 \left[\frac{1}{6} \left(\sum_\alpha c'_\alpha - \sum_\rho c'_\rho \right) \int dx \rho(x) |x|^3 \right.$$

$$\left. - \frac{1}{2} \sum_\rho c_\rho \left(1 - \tilde{\nu}_\rho^2 \right) \int dx dy \, \rho(x) \rho(y) \left(|x+y| + |x-y| \right) \right]. \qquad (3.37)$$

Here c'_α and c'_ρ are another set of numerical constants appearing when collecting all the contributions to the local $|x|^3$ term. These are in general different from the c's, for example, $c_V'^\circ = c_V'^\bullet = 8$, $c'_{Ad} = 8$, $c'_{AS} = 0$, $c'_{BF} = 8$, $c'_f = 1$, assuming the same set of gauge groups as before. After the rescaling $x \to \frac{2}{3} \sqrt{\sum_\rho c_\rho \left(1 - \tilde{\nu}_\rho^2 \right)} \, x$ (and corresponding inverse rescaling of ρ), we have

$$\mathcal{W} = \frac{4 Q^2}{27\pi} \left(\sum_\rho c_\rho \left(1 - \tilde{\nu}_\rho^2 \right) \right)^{3/2} N^{5/2} \left[\frac{\pi (8 N_v - N_f)}{3} \int_0^{x_*} dx \, \rho(x) |x|^3 \right.$$

$$\left. - \frac{9\pi}{8} \int_0^{x_*} dx dy \rho(x) \rho(y) \left(|x+y| + |x-y| \right) \right], \quad (3.38)$$

---

[23]It should be possible to study quivers with different ranks but we do not consider this here.
[24]We note that due to (3.36), the $-\frac{1}{6}(1 + 4Q^2)$ part in the terms of order $N^{2+\alpha}$ in (3.35) also cancels.

where we defined $N_v = N_V^\circ + N_V^\bullet - N_{Ad} - N_{BF}$. By comparison with the free energy functional on S^5 for the same theory, it follows that

$$\mathcal{W} = \frac{4Q^2}{27\pi} \left(\frac{\sum_\rho c_\rho \left(1 - \tilde{\nu}_\rho^2\right)}{\sum_\rho c_\rho} \right)^{3/2} F_{S^5}, \tag{3.39}$$

with $F_{S^5} = \left(\sum_\rho c_\rho \right)^{3/2} F_{S^5}^{Seiberg}$, which is extremized for

$$\rho(x) = \frac{2|x|}{x_*^2}, \qquad x_* = \frac{3}{\sqrt{2(8N_v - N_f)}}. \tag{3.40}$$

We now evaluate the on-shell free energy. Taking the logarithm of (3.3) one can see that the Hessian part of the handle-gluing operator does not contribute to leading order in N and one finds[25]

$$F_{S_b^3 \times \Sigma_{\mathfrak{g}}} = -6\pi(\mathfrak{g} - 1) \left(\frac{\sum_\alpha c_\alpha - \sum_\rho c_\rho \hat{\mathfrak{n}}_\rho \tilde{\nu}_\rho}{\sum_\rho c_\rho \left(1 - \tilde{\nu}_\rho^2\right)} \right) \mathcal{W}, \tag{3.41}$$

which using (3.39) can also be written as

$$F_{S_b^3 \times \Sigma_{\mathfrak{g}}} = -\frac{8}{9} Q^2 (\mathfrak{g} - 1) \left(\sum_\alpha c_\alpha - \sum_\rho c_\rho \hat{\mathfrak{n}}_\rho \tilde{\nu}_\rho \right) \left(\sum_\rho c_\rho (1 - \tilde{\nu}_\rho^2) \right)^{1/2} \left(\sum_\rho c_\rho \right)^{-3/2} F_{S^5}. \tag{3.42}$$

Finally, extremizing with respect to the fugacities gives

$$\tilde{\nu}_\rho^\pm = \frac{\hat{\mathfrak{n}}_\rho}{4\hat{\mathfrak{n}}^2} \left(1 \pm \sqrt{1 + 8\hat{\mathfrak{n}}^2} \right), \tag{3.43}$$

where $\hat{\mathfrak{n}}^2 \equiv \left(\sum_\rho c_\rho \right)^{-1} \left(\sum_\rho c_\rho \hat{\mathfrak{n}}_\rho^2 \right)$ and evaluating at the extremum,

$$F_{S_b^3 \times \Sigma_{\mathfrak{g}}} = -\frac{8}{9} (\mathfrak{g} - 1) Q^2 \left(\pm \frac{\sqrt{2} \left| \hat{\mathfrak{n}}^2 - 1 \right|^{3/2} \left(\sqrt{1 + 8\hat{\mathfrak{n}}^2} \pm 1 \right)}{\left| 4\hat{\mathfrak{n}}^2 - 1 \mp \sqrt{1 + 8\hat{\mathfrak{n}}^2} \right|^{3/2}} \right) F_{S^5}. \tag{3.44}$$

Since F_{S^5} is negative and F_{S^3} should be positive, we see that for $\mathfrak{g} > 1$, the solution with $+$ sign should be chosen above and for $\mathfrak{g} = 0$, the solution with $-$ sign. This sign then suggests the introduction of κ, the normalized curvature of $\Sigma_{\mathfrak{g}}$ defined below (2.60), by rescaling $\hat{\mathfrak{n}}_\rho \to \frac{\hat{\mathfrak{n}}_\rho}{\kappa}$, giving

$$F_{S_b^3 \times \Sigma_{\mathfrak{g}}} = \frac{8}{9} (\mathfrak{g} - 1) Q^2 \left(\frac{\sqrt{2} \left| \hat{\mathfrak{n}}^2 - \kappa^2 \right|^{3/2} \left(\sqrt{\kappa^2 + 8\hat{\mathfrak{n}}^2} - \kappa \right)}{\kappa \left| 4\hat{\mathfrak{n}}^2 - \kappa^2 + \kappa \sqrt{\kappa^2 + 8\hat{\mathfrak{n}}^2} \right|^{3/2}} \right) F_{S^5}. \tag{3.45}$$

One can also make sense of the above expression for $\kappa = 0$ (as in [22] for the Seiberg theory) by considering the limit $\frac{\mathfrak{g}-1}{\kappa} \to -4\pi$, $\hat{\mathfrak{n}}_\rho \to \frac{\hat{\mathfrak{n}}_\rho}{2\pi}$, resulting in the relation $F_{S_b^3 \times \Sigma_{\mathfrak{g}=1}} = -\frac{8}{9} Q^2 \sqrt{\hat{\mathfrak{n}}^2} F_{S^5}$.

---

[25]Although in principle one may worry that exponentially subleading/diverging terms in H of the form $\exp\left(\mp N^\alpha x\right)$ may contribute, to leading order in N these are of the form $H \approx \exp\left(\sum_i \partial_{u_i} \mathcal{W} \right)$ and thus $H \approx 1$ when evaluated in the Bethe vacua of the theory.
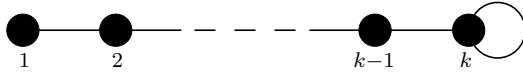
Figure 3. A simple class of 5d quiver gauge theories with free energy scaling as N^3. Black nodes represent $SU(2N)$ gauge groups, a line connecting two nodes denotes a bifundamental hypermultiplet, and the line ending on the same node denotes an adjoint hypermultiplet. No fundamental hypermultiplets are allowed in this case.

(II) Theories with N^3 scaling. Setting $\alpha = 1$ in (3.35) and requiring the cancellation of nonlocal cubic terms at $\mathcal{O}(N^5)$ gives the same constraint as before, i.e., (3.36). The cancellation of local cubic terms at $\mathcal{O}(N^4)$ leads to the additional condition

$$\sum_\alpha c'_\alpha - \sum_\rho c'_\rho = 0\,. \tag{3.46}$$

The constraints (3.36) and (3.46) together mean no cubic terms (either nonlocal and local) survive in the expansion of the function g_b. For pure $U(N)$ or $SU(N)$ theories, this constraint simply becomes $N_f = 0$. The maximal 5d $\mathcal{N} = 2$ $SU(N)$ SYM is such an example. Another known example is class $\mathcal{S}_k$; circular quivers consisting of k $SU(N)$ gauge groups and k bifundamental hypermultiplets. The maximal theory is obtained for $k = 1$ when the bifundamental multiplet turns into an adjoint one. We also note a linear quiver with k $SU(N)$ gauge groups and $k - 1$ bifundamentals, and an additional adjoint matter on one node (see figure 3) satisfies (3.36) and (3.46). An example with $USp(2N)$ group is the Seiberg theory with $N_f = 8$ or E-string theory. In addition, all the three classes of orbifold theories studied above with $N_f = 8$ also give rise to N^3 scaling of the free energy.

3.4 Holography

On general grounds, we expect the gravity dual of a 5d SCFT on $S^3 \times \Sigma_{\mathfrak{g}}$ with a partial topological twist on $\Sigma_{\mathfrak{g}}$ to be given by a supersymmetric solution interpolating between asymptotically locally AdS_6 at infinity (with an $S^3 \times \Sigma_{\mathfrak{g}}$ boundary) and an $AdS_4 \times \Sigma_{\mathfrak{g}}$ geometry for small values of the radial coordinate.[26] In the case of a universal topological twist ($\mathfrak{n} = 0$) it was argued in [21] that such a solution, originally found in [68, 69], is given by an extremal 2-brane solution in minimal 6d $F(4)$ gauged supergravity, with metric of the form

$$ds^2 = e^{2f(r)}(-dt^2 + dz_1^2 + dz_2^2 + dr^2) + e^{2g(r)}ds_{\Sigma_{\mathfrak{g}}}^2\,, \tag{3.47}$$

which is supported by a nontrivial magnetic flux for a $U(1) \subset SU(2)$ graviphoton on $\Sigma_{\mathfrak{g}}$ and a nontrivial scalar $\varphi(r)$.[27] The interpolating solution, which exists only for $\mathfrak{g} > 1$, preserves four real supercharges and can be found numerically. As the radial coordinate goes to $r \to \infty$ the metric is locally asymptotic to AdS_6 with unit radius and as $r \to 0$ it

---

[26]The holographic solutions described here correspond to the case of a round 3-sphere $S_{b=1}^3$.

[27]This supergravity theory [70] has 16 real supercharges and bosonic field content: the graviton $g_{\mu\nu}$, an $SU(2)$ gauge potential A_μ^I, an Abelian one-form potential A_μ, a massive two-index tensor gauge field $B_{\mu\nu}$, and a scalar field φ.

is of the form $AdS_4 \times \Sigma_{\mathfrak{g}}$, with sizes $e^{2f(r)} = \frac{1}{r^2}2^{3/2}3^{-3/2}$ and $e^{2g(r)} = 2^{1/2}3^{-3/2}$, and the free energy given by [21]

$$F_{S^3 \times \Sigma_{\mathfrak{g}}}^{\text{sugra}} = -\frac{8}{9}\left(\mathfrak{g} - 1\right)F_{S^5}, \qquad (3.48)$$

where F_{S^5} is the free energy of the 5d field theory on S^5. The crucial point, emphasized in [21], is that upon uplift to massive IIA on topologically S^4, the solution (3.47) describes the twisted compactification of *any* 5d $\mathcal{N} = 1$ theory with a gravity dual and thus (3.48) holds for any such compactification.[28] Indeed, this is corroborated by our field theory results; setting $Q = 1$ for the round S^3, $\kappa = -1$, and $\mathfrak{n}_\rho = 0$ in (3.45) we see this exactly matches (3.48). The existence of this holographic flow is thus strong evidence for the existence of a large class of 3d SCFTs arising from compactification of 5d $\mathcal{N} = 1$ theories, at least at large N, for the universal twist. We emphasize that the free energy of these 3d theories scales as $N^{5/2}$ rather than the more standard $N^{3/2}$ or $N^{5/3}$ scaling of 3d theories with gravity duals.

The general result (3.45) suggests that a larger class of such holographic flows should exist, whose endpoint is described by a discrete family of 3d SCFTs labeled by the integers $\mathfrak{n}_\rho$ for each 5d parent theory. The simplest example is the Seiberg theory with a nonzero flux for the Cartan of $SU(2)_M$. The explicit holographic RG flow in this case was constructed for $Q = 1$ in [22]. Indeed, specifying (3.45) for this case exactly matches the supergravity result (see eq. (1.1) in [22]). It would be interesting to construct the analogous holographic RG flows for the orbifold theories considered in detail in section 3.1, which now have a number of flavor symmetries visible at large N and whose free energy is given by (3.34). In general, one should be able to do this for arbitrary values of Q, which would be interesting.

These holographic checks give strong support for the existence a novel class of 3d SCFTs. The results of this paper provide a method for computing the exact partition function of such theories, in principle at finite N.

4   5d theories, 6d SCFTs, and the 4d index

As reviewed in the Introduction, some 5d gauge theories are believed to be low energy descriptions of certain 6d SCFTs compactified on a circle, S_β^1. The radius of the 6d circle is related to the 5d gauge coupling constant by

$$\beta = \frac{g_5^2}{2\pi\lambda}, \qquad (4.1)$$

where λ is a numerical factor which depends on the specific theory under consideration. The prototypical example of this phenomenon, which we will discuss in detail in section 4.1, is that of the maximally supersymmetric Yang-Mills theory in 5d with an ADE gauge group, which is expected to have a UV completion as the circle compactification of the corresponding $\mathcal{N} = (2,0)$ SCFT. There are also other examples of this phenomenon with $\mathcal{N} = (1,0)$ supersymmetry, including theories obtained by orbifolding the maximal theory, and the E-string, which uplifts to the 6d E-string theory.

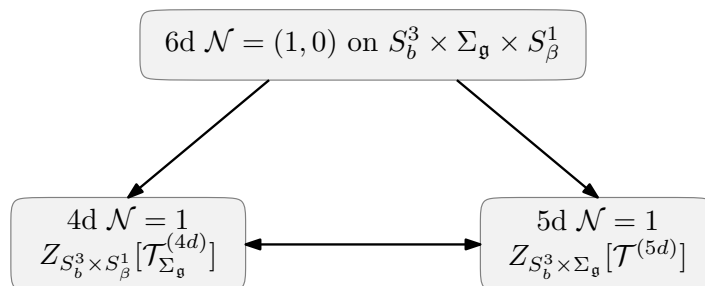[28]See also [71] and [72] for similar discussions in other dimensions.

Figure 4. The partition function of 4d $\mathcal{N} = 1$ class $\mathcal{S}$ theories on $S_b^3 \times S_\beta^1$ from the partition function of $\mathcal{N} = 1$ gauge theories on $S_b^3 \times \Sigma_\mathfrak{g}$. The precise relation is given in (4.3) with the mapping of parameters in (4.7).

When one computes protected observables in these 5d theories on a compact manifold, $\mathcal{M}_5$, one generally expects that these may be interpreted as observables in the "parent" 6d theory on $\mathcal{M}_5 \times S_\beta^1$. Indeed, this philosophy was applied in the case of $\mathcal{M}_5$ being the squashed five-sphere to study the superconformal index, or $S^5 \times S_\beta^1$ partition function [34, 35], as well as other examples; see [4] for a review. In the case of the $S_b^3 \times \Sigma_\mathfrak{g}$ partition function, we expect this to compute the partition function of the 6d SCFT on

$$S_b^3 \times \Sigma_\mathfrak{g} \times S_\beta^1 \,. \tag{4.2}$$

This leads to another perspective on this object, as follows. By compactifying the parent 6d $\mathcal{N} = (1,0)$ theory on the Riemann surface instead leads to a class of 4d offspring theories labeled by the compactification manifold, $\Sigma_\mathfrak{g}$,[29] and flavor fluxes, $\mathfrak{n}$, which we denote here by $\mathcal{T}_{\Sigma_\mathfrak{g},\mathfrak{n}}^{(4d)}$. Then our computation can be interpreted as giving the partition function of these 4d theories on $S_b^3 \times S_\beta^1$ (see figure 4), i.e.,

$$Z_{S_b^3 \times \Sigma_\mathfrak{g}}(\nu, \gamma)_\mathfrak{n} \left[\mathcal{T}^{(5d)} \right] = Z_{S_b^3 \times S_\beta^1}(p, q, \mu) \left[\mathcal{T}_{\Sigma_\mathfrak{g},\mathfrak{n}}^{(4d)} \right] \,. \tag{4.3}$$

The latter is closely related to the 4d supersymmetric index as [73]

$$Z_{S_b^3 \times S_\beta^1}(p, q, \mu) = e^{-\beta E_{\text{Casimir}}} \mathcal{I}_{4d}(p, q, \mu) \,, \tag{4.4}$$

where the 4d index may be defined as a trace in radial quantization [74, 75],

$$\mathcal{I}_{4d}(p, q, \mu) = \text{Tr}(-1)^F e^{-\beta'\{\mathcal{Q}, \mathcal{Q}^\dagger\}} p^{j_1+j_2-\frac{R}{2}} q^{j_1-j_2-\frac{R}{2}} \prod_i \mu_i^{F_i} \,, \tag{4.5}$$

where j_1 and j_2 are Cartan generators for the $\text{SO}(4) \cong \text{SU}(2)_1 \times \text{SU}(2)_2$ rotation symmetry, and F_i run over a basis of the flavor symmetries. The quantity E_{Casimir}, to which we will return to more detail in section 4.2.1, is the Casimir energy of the 4d theory quantized on S_b^3 [73].

---

[29]In the general case the Riemann surface may contain nontrivial punctures, which we do not consider in this work. These should appear as local defect operators in the TQFT on $\Sigma_\mathfrak{g}$.

The identification of parameters to those appearing naturally in the $S_b^3 \times S_\beta^1$ partition function is [45, 76]

$$p = e^{-2\pi b Q^{-1}\beta}, \qquad q = e^{-2\pi b^{-1} Q^{-1}\beta}, \qquad \mu_i = e^{2\pi Q^{-1}\beta(\nu_i + iQ)}, \qquad (4.6)$$

where ν_i are the effective real mass parameters on S_b^3. Using (4.1) we find these are related to the parameters of the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function by

$$p = e^{2\pi b \lambda^{-1}\gamma^{-1}}, \qquad q = e^{2\pi b^{-1}\lambda^{-1}\gamma^{-1}}, \qquad \mu = e^{2\pi i \lambda^{-1}\gamma^{-1}(\nu_i + iQ)}, \qquad (4.7)$$

where $\gamma = -\frac{2\pi Q}{g_5^2}$. To summarize, this chain of reasoning leads us to identify

$$Z_{S_b^3 \times \Sigma_{\mathfrak{g}}}(\nu, \gamma)_{\mathfrak{n}}\left[\mathcal{T}^{(5d)}\right] = Z_{S_b^3 \times S_\beta^1}(p, q, \mu)\left[\mathcal{T}_{\Sigma_{\mathfrak{g}}, \mathfrak{n}}^{(4d)}\right] = e^{-\beta E_{\mathrm{Casimir}}}\mathcal{I}_{4d}(p, q, \mu)\left[\mathcal{T}_{\Sigma_{\mathfrak{g}}, \mathfrak{n}}^{(4d)}\right]. \qquad (4.8)$$

In this section, we study this relation in more detail. Although in principle this is an exact result, valid for any choice of parameters appearing in (4.8), in practice we are only able to compute the l.h.s. in certain simplifying limits, where the contributions from instantons are under control. We first consider in detail the case of the maximal, $\mathcal{N} = 2$ SYM theory. Due to the extra supersymmetry, this admits a limit where the instanton contribution is very simple, and the partition function can be computed exactly. Then we study the partition function in the "Casimir limit," $\beta \sim g_5{}^2 \to \infty$. Although naïvely the instantons are important in this limit, we find their contribution can be explicitly characterized.

4.1 The 5d $\mathcal{N} = 2$ Yang-Mills theory

Let us start by considering the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function for the maximally supersymmetric 5d $\mathcal{N} = 2$ Yang-Mills theory with simply laced gauge group G.[30] We will often specialize to the case $G = \mathrm{SU}(N)$ for concreteness, but will keep the discussion general when possible. This theory contains an adjoint hypermultiplet, acted on by an $\mathrm{SU}(2)_F$ flavor symmetry,[31] for which we include a mass ν and flux $\mathfrak{n}$. Then the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function of this theory can be written as a sum over Bethe vacua as

$$Z_{S_b^3 \times \Sigma_{\mathfrak{g}}}(\nu, \gamma)_{\mathfrak{n}} = \sum_{\hat{u} \in \mathcal{S}_{\mathrm{BE}}} \mathcal{H}(\hat{u}, \nu, \gamma)^{\mathfrak{g}-1}\Pi_\nu(\hat{u}, \nu, \gamma)^{\mathfrak{n}}. \qquad (4.9)$$

These operators can be constructed out of the twisted superpotential

$$\mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathcal{N}=2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = \frac{1}{Qb}\mathcal{W}_{NS}^{\mathcal{N}=2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; -b^2) + \frac{1}{Qb^{-1}}\mathcal{W}_{NS}^{\mathcal{N}=2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; -b^{-2}), \qquad (4.10)$$

in the notation of section 2.4, and similarly for the effective dilaton. We may alternatively decompose this into a perturbative and instanton contribution,

$$\mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathcal{N}=2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathcal{N}=2,\mathrm{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) + \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathcal{N}=2,\mathrm{inst}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}), \qquad (4.11)$$

---

[30]As mentioned in section 2.1, this computation was also considered in [37–39], where they also observed a relation to the Schur limit of the 4d index. Below we clarify the precise relation of this observable to the 4d index of certain 4d $\mathcal{N} = 1$ SCFTs.

[31]Specifically, we have $\mathrm{SU}(2)_R \times \mathrm{SU}(2)_F \subset \mathrm{USp}(4)$, where the latter is the full R-symmetry group for this theory, and $\mathrm{SU}(2)_R$ is the 5d $\mathcal{N} = 1$ R-symmetry.

and we may write an explicit expression for the first term, as in section 2.3:

$$\mathcal{W}^{\mathcal{N}=2,\text{pert}}_{S^3_b \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = \frac{1}{2}\tilde{\gamma}K^{ab}\tilde{u}_a\tilde{u}_b + \sum_{\alpha \in Ad(G)} g_b(\alpha(\tilde{u}) + \tilde{\nu}) - \sum_{\alpha \in Ad(G)'} g_b(\alpha(\tilde{u}) + 1), \quad (4.12)$$

where $\tilde{\gamma} = -\frac{2\pi i}{g_5^2}$. In general it is difficult to evaluate the non-perturbative contribution analytically. However, we will describe some simplifying limits below.

To see what the $S^3_b \times \Sigma_{\mathfrak{g}}$ partition function of this theory corresponds to in 4d, we recall that the 6d UV completion of this 5d theory is the $\mathcal{N} = (2,0)$ theory compactified on a circle of radius

$$\beta = \frac{g_5^2}{2\pi}. \qquad (4.13)$$

In other words, (4.1) holds with $\lambda = 1$ in this case. The compactification of the 6d SCFT on a Riemann surface with a topological twist gives rise, in general, to a class of 4d $\mathcal{N} = 1$ SCFTs, which were described as the A_{N-1} case in [25]. The specific theory depends on the choice of R-symmetry used to perform the twist. If we mix the $U(1)_R \subset SU(2)_R$ symmetry used to perform the topological twist on $\Sigma_{\mathfrak{g}}$ with the maximal torus of the $SU(2)_F$ flavor symmetry with a coefficient $\hat{\mathfrak{n}}$, this is equivalent to inserting a flux,[32]

$$\mathfrak{n} = \hat{\mathfrak{n}}(\mathfrak{g} - 1), \qquad (4.14)$$

on $\Sigma_{\mathfrak{g}}$. In the special case

$$\hat{\mathfrak{n}} = \pm 1, \qquad (4.15)$$

the twist preserves $\mathcal{N} = 2$ supersymmetry in 4d, and gives rise to the theories of class $\mathcal{S}$ [23]. Let us denote the theory for general choices of flux, $\mathfrak{n}$, by $\mathcal{T}^{(4d)}_{\Sigma_{\mathfrak{g}},\mathfrak{n}}$.

As a special case of the 4d-2d correspondence, we may relate the index of this 4d theory to the partition function of a suitable 2d TQFT [77, 78]. Then it is clear from the logic above that this TQFT is precisely the A-twist of the effective 5d $\mathcal{N} = 2$ theory compactified on S^3_b. In other words, we have

$$Z_{S^3_b \times S^1_\beta}\left[\mathcal{T}^{(4d)}_{\Sigma_{\mathfrak{g}},\mathfrak{n}}\right] = Z_{S^3_b \times \Sigma_{\mathfrak{g}}}(\nu)_{\mathfrak{n}}\left[\mathcal{T}^{(5d)}\right], \qquad (4.16)$$

where the r.h.s. may be interpreted as a TQFT living on $\Sigma_{\mathfrak{g}}$. In particular, in writing the r.h.s. as a sum over Bethe vacua, as in (4.9), we exhibit this TQFT structure explicitly.

While we expect the relation (4.16) to hold for general parameters, it is in general difficult to evaluate the $S^3_b \times \Sigma_{\mathfrak{g}}$ partition function explicitly due to the nontrivial instanton contributions. Thus we first consider a special limit, where their contribution simplifies significantly. We will see this limit turns out to be related to the "Schur limit" of the 4d $\mathcal{N} = 2$ superconformal index [28]. We then briefly comment on the general case.

---

[32]In [25] the parameter we call $\hat{\mathfrak{n}}$ was denoted by z.

The Schur limit. Let us first consider the perturbative contribution to the (ungauged) partition function in more detail. Using (4.12), and working in terms of the untilded variables, we have[33]

$$Z^{\mathcal{N}=2,\text{pert}}_{S^3_b \times \Sigma_{\mathfrak{g}}}(u,\nu,\gamma)_{\mathfrak{m},\mathfrak{n}} = e^{2\pi i \gamma K^{ab} u_a \mathfrak{m}_b} \prod_{\alpha \in Ad(G)} s_b\left(\alpha(u) + \nu\right)^{\alpha(\mathfrak{m})+\mathfrak{n}}$$
$$\times \prod_{\alpha \in Ad(G)'} s_b\left(\alpha(u) - iQ\right)^{-\alpha(\mathfrak{m})+1-\mathfrak{g}}. \tag{4.17}$$

Now let us consider the limit[34]

$$\nu = \frac{i}{2}(b - b^{-1}). \tag{4.18}$$

Then, using the following identities of the double sine function,

$$s_b(x) = s_b(-x)^{-1}, \tag{4.19}$$

$$s_b\left(x + \frac{i}{2}b^{\pm}\right) = \frac{1}{2\cosh(\pi b^{\pm} x)} s_b\left(x - \frac{i}{2}b^{\pm}\right), \tag{4.20}$$

we find the perturbative contribution simplifies to

$$Z^{\mathcal{N}=2,\text{pert}}_{S^3_b \times \Sigma_{\mathfrak{g}}}(u,\nu,\gamma)_{\mathfrak{m},\mathfrak{n}}$$
$$= e^{2\pi i \gamma K^{ab} u_a \mathfrak{m}_b} b^{-\mathfrak{n} r_G} \prod_{\alpha > 0} \left[2\sinh\left(\pi b \alpha(u)\right)\right]^{1-\mathfrak{g}-\mathfrak{n}} \left[2\sinh\left(\pi b^{-1} \alpha(u)\right)\right]^{1-\mathfrak{g}+\mathfrak{n}}. \tag{4.21}$$

where we simplified the Cartan contribution of the adjoint hypermultiplet using $s_b(\nu) = s_b(\frac{i}{2}(b - b^{-1})) = b^{-1}$. In particular, we see the dependence on the gauge flux $\mathfrak{m}$, and hence also the perturbative Bethe equations, are very simple.

Next, we consider the instanton contribution, which can be written as

$$Z^{\mathcal{N}=2,\text{inst}}_{S^3_b \times \Sigma_{\mathfrak{g}}}(u,\nu,\gamma)_{\mathfrak{m},\mathfrak{n}} = \prod_{\ell} Z^{\mathcal{N}=2,\text{inst}}_{\mathbb{R}^2_{\mathfrak{q}_1^{(\ell)}} \times \mathbb{R}^2_{\mathfrak{q}_2^{(\ell)}} \times S^1}\left(x^{(\ell)}, y^{(\ell)}, z^{(\ell)}\right). \tag{4.22}$$

Here the parameter, $y^{(\ell)}$, for the flavor symmetry is given by, using (2.108),

$$y^{(\ell)} = \begin{cases} e^{2\pi b \nu} \mathfrak{q}_2^{(\ell)\mathfrak{n}/2} = -\mathfrak{q}_1^{(\ell)1/2} \mathfrak{q}_2^{(\ell)\mathfrak{n}/2} & \ell = nn \text{ or } ns \\ e^{2\pi b^{-1}\nu} \mathfrak{q}_2^{(\ell)\mathfrak{n}/2} = -\mathfrak{q}_1^{(\ell)-1/2} \mathfrak{q}_2^{(\ell)\mathfrak{n}/2} & \ell = sn \text{ or } ss \end{cases} \tag{4.23}$$

Let us now specialize to the case $G = \text{SU}(N)$. Then we note the following simplification of the instanton contribution to the Nekrasov partition function of the $\mathcal{N} = 2$ theory:

$$Z^{\mathcal{N}=2,\text{inst}}_{\mathbb{R}^2_{\mathfrak{q}_1} \times \mathbb{R}^2_{\mathfrak{q}_2} \times S^1}\left(x, y = -(\mathfrak{q}_1\mathfrak{q}_2)^{\pm 1/2}, z\right) = \frac{1}{\eta(z)^{N-1}}, \tag{4.24}$$

$$Z^{\mathcal{N}=2,\text{inst}}_{\mathbb{R}^2_{\mathfrak{q}_1} \times \mathbb{R}^2_{\mathfrak{q}_2} \times S^1}\left(x, y = -(\mathfrak{q}_1\mathfrak{q}_2^{-1})^{\pm 1/2}, z\right) = 1. \tag{4.25}$$

---

[33]Here we recall the product with the prime includes only the non-zero roots, while the unprimed product includes all roots.

[34]In [79] this limit of the S^3_b partition function was shown to correspond to the dimensional reduction of the Schur limit of the supersymmetric index.

This is derived in appendix B, and was also noted in the context of the S^5 [34] and $\mathbb{CP}^2 \times S^1$ [35], partition functions. Then we see that if we set $\mathfrak{n} = 1$ we have

$$
\begin{aligned}
Z_{S_b^3 \times S^2}^{\mathcal{N}=2,\text{inst}} & \left(u, \nu = \frac{i}{2}(b - b^{-1}), \gamma \right)_{\mathfrak{m},\mathfrak{n}=1} \\
& = \prod_{\ell=nn,ns} Z_{\mathbb{R}^2_{\mathfrak{q}_1^{(\ell)}} \times \mathbb{R}^2_{\mathfrak{q}_2^{(\ell)}} \times S^1}^{\mathcal{N}=2,\text{inst}} \left(x^{(\ell)}, y^{(\ell)} = -\mathfrak{q}_1^{(\ell)1/2} \mathfrak{q}_2^{(\ell)1/2}, z^{(\ell)} \right) \\
& \quad \times \prod_{\ell=sn,ss} Z_{\mathbb{R}^2_{\mathfrak{q}_1^{(\ell)}} \times \mathbb{R}^2_{\mathfrak{q}_2^{(\ell)}} \times S^1}^{\mathcal{N}=2,\text{inst}} \left(x^{(\ell)}, y^{(\ell)} = -\mathfrak{q}_1^{(\ell)-1/2} \mathfrak{q}_2^{(\ell)1/2}, z^{(\ell)} \right) \\
& = \frac{1}{\eta(z)^{2(N-1)}} \,,
\end{aligned}
\tag{4.26}
$$

where we recall from (2.124) that

$$
z^{(nn,ns)} = z = e^{-2\pi b \gamma} \,, \qquad z^{(sn,ss)} = \bar{z} = e^{-2\pi b^{-1} \gamma} \,.
\tag{4.27}
$$

Similarly, for $\mathfrak{n} = -1$, we have

$$
Z_{S_b^3 \times S^2}^{\mathcal{N}=2,\text{inst}} \left(u, \nu = \frac{i}{2}(b - b^{-1}), \gamma \right)_{\mathfrak{m},\mathfrak{n}=-1} = \frac{1}{\eta(\bar{z})^{2(N-1)}} \,.
\tag{4.28}
$$

On the other hand, we expect the instanton contribution to the integrand of the $S_b^3 \times S^2$ partition function to be equal to

$$
e^{-2\pi i \Omega_{\text{inst}}} \left(\Pi_\nu^{\text{inst}} \right)^{\mathfrak{n}} \left(\Pi_a^{\text{inst}} \right)^{\mathfrak{m}_a} \,.
\tag{4.29}
$$

Equating these expressions for $\mathfrak{n} = \pm 1$, we deduce that (up to a sign)

$$
e^{2\pi i \Omega_{\text{inst}}} = \eta(z)^{N-1} \eta(\bar{z})^{N-1} \,, \qquad \Pi_\nu^{\text{inst}} = \frac{\eta(\bar{z})^{N-1}}{\eta(z)^{N-1}} \,, \qquad \Pi_a^{\text{inst}} = 1 \,.
\tag{4.30}
$$

The last relation also implies that there is no instanton contribution to the Hessian determinant appearing in the handle-gluing operator.

It will be convenient to rewrite the η functions appearing in the handle-gluing and flux operators to relate them to the index parameters in (4.7),

$$
p = e^{2\pi \gamma^{-1} b} \,, \qquad q = e^{2\pi \gamma^{-1} b^{-1}} \,.
\tag{4.31}
$$

Namely, using the modular properties of the η functions, we have

$$
\begin{aligned}
\eta(z) &= \eta(e^{-2\pi b \gamma}) = (b\gamma)^{-1/2} \eta(e^{2\pi \gamma^{-1} b^{-1}}) = (b\gamma)^{-1/2} \eta(q) \,, \\
\eta(\bar{z}) &= \eta(e^{-2\pi b^{-1} \gamma}) = (b^{-1}\gamma)^{-1/2} \eta(e^{2\pi \gamma^{-1} b}) = (b^{-1}\gamma)^{-1/2} \eta(p) \,.
\end{aligned}
\tag{4.32}
$$

We may now write the full, non-perturbatively complete operators for the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function in the limit (4.18). We have

$$
\Pi_a = e^{2\pi i \gamma K^{ab} u_b} \,, \qquad \Pi_\nu = \frac{\eta(p)^{N-1}}{\eta(q)^{N-1}} \prod_{\alpha>0} \frac{2 \sinh\left(\pi b^{-1} \alpha(u)\right)}{2 \sinh\left(\pi b \alpha(u)\right)} \,,
\tag{4.33}
$$

$$
\mathcal{H} = \eta(p)^{N-1} \eta(q)^{N-1} \prod_{\alpha>0} \left[2 \sinh\left(\pi b \alpha(u)\right) 2 \sinh\left(\pi b^{-1} \alpha(u)\right) \right]^{-1} \,.
\tag{4.34}
$$

Note that the Hessian determinant contributes a factor of γ^{N-1}, which precisely cancels against the factors of γ from (4.32), and similarly the factor of $b^{-(N-1)}$ from the Cartan component of the adjoint hypermultiplet precisely cancels the factors of b from (4.32).

It is now trivial to write the solutions to the Bethe equations, $\Pi_a = 1$, which are given by

$$\hat{u}_a = \gamma^{-1} K_{ab}^{-1} n^b \equiv \gamma^{-1} \lambda_a \,, \qquad n^b \in \mathbb{Z} \;\Leftrightarrow\; \lambda \in \Lambda_{cr} \,, \tag{4.35}$$

where $\lambda_a \equiv K_{ab}^{-1} n^b$ runs over the coroot lattice of G as we vary over $n^b \in \mathbb{Z}$. Without loss of generality, we may restrict to coroots in the interior of the fundamental Weyl chamber,[35] or equivalently, those which can be written as $\delta + \lambda$, for λ a dominant coroot, and δ the Weyl vector of G. Then the set of Bethe vacua is

$$\mathcal{S}_{\mathrm{BE}} = \left\{ \; \hat{u}_a^{(\lambda)} \equiv \gamma^{-1}(\lambda_a + \delta_a) \;\; \Big| \;\; \lambda \in \Lambda_{cr}^+ \right\} \,. \tag{4.36}$$

It remains to compute the partition function by evaluating the flux and handle-gluing operators at the above Bethe vacua. First we observe

$$\prod_{\alpha > 0} 2\sinh\left(\pi b \alpha(\hat{u}^{(\lambda)})\right) = \prod_{\alpha > 0} (p^{\alpha(\lambda+\delta)/2} - p^{-\alpha(\lambda+\delta)/2}) = V(p) \dim_p R_\lambda \,, \tag{4.37}$$

and similarly

$$\prod_{\alpha > 0} 2\sinh\left(\pi b^{-1} \alpha(\hat{u}^{(\lambda)})\right) = V(q) \dim_q R_\lambda \,, \tag{4.38}$$

where R_λ is the representation with highest weight λ, where we have identified the coroot and weight lattices using the Killing metric, and $\dim_p R$ is the "quantum dimension" of a representation R of G. Here we have made the identification to the parameters of the 4d index in (4.31). Finally, $V(p)$ is the polynomial (here we specialize again to SU(N), but the generalization is straightforward)

$$V(p) \equiv \det_{1 \le a,b \le N} p^{\left(\frac{N+1}{2} - a\right)(N-b)} = (-1)^{N(N-1)} p^{-\frac{1}{12}N(N^2-1)} \prod_{i=1}^{N-1} (1-p^i)^{N-i} \,. \tag{4.39}$$

Putting this together, we see that the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function for the $\mathcal{N} = 2$ SU(N) theory, in the limit (4.18), is given by

$$\begin{aligned}
Z_{S_b^3 \times \Sigma_{\mathfrak{g}}}^{\mathcal{N}=2}\left(\nu = \frac{i}{2}(b - b^{-1})\right)_{\mathfrak{n}} &= \sum_{\hat{u} \in \mathcal{S}_{\mathrm{BE}}} \Pi_\nu(\hat{u})^{\mathfrak{n}} \mathcal{H}(\hat{u})^{\mathfrak{g}-1} \\
&= \sum_{\lambda \in \Lambda_{cr}^+} \left(\frac{V(p) \dim_p(R_\lambda)}{\eta(p)^{N-1}}\right)^{-\mathfrak{n}+1-\mathfrak{g}} \left(\frac{V(q) \dim_q(R_\lambda)}{\eta(q)^{N-1}}\right)^{\mathfrak{n}+1-\mathfrak{g}} \,.
\end{aligned} \tag{4.40}$$

---

[35]Those on the boundary of the Weyl chamber will lead to Bethe vacua with enhanced Weyl symmetry, which we are instructed to discard.

4d interpretation. As described above, we expect this observable to compute the 4d superconformal index of the theory $\mathcal{T}_{\Sigma_{\mathfrak{g},\mathfrak{n}}}^{(4d)}$. Specifically, using the map (4.7), we can see that the limit (4.18) corresponds to setting the fugacity μ corresponding to this flavor symmetry in 4d as

$$\mu = p\,. \tag{4.41}$$

In the case of $\mathcal{N} = 2$ class $\mathcal{S}$ theories, μ is usually referred to as t, and is the fugacity for the $\mathrm{U}(1)_r$ symmetry in the 4d $\mathcal{N} = 2$ algebra. This limit of the index was first studied for the $\mathcal{N} = 2$ class $\mathcal{S}$ theories in [28], where it was referred to as the "Schur limit" of the index. It was later generalized to the 4d $\mathcal{N} = 1$ theories corresponding to a more general choice of flux, $\mathfrak{n}$, in [29], where it was referred to as the "mixed Schur index." The general result they found for the 4d index can be written in our notation as

$$\mathcal{I}_{4d}(p,q)[\mathcal{T}_{\Sigma_{\mathfrak{g},\mathfrak{n}}}^{(4d)}] = \left(\frac{p^{\frac{1}{12}N(N^2-1)}V(p)}{(p;p)^{N-1}}\right)^{-\ell_1}\left(\frac{q^{\frac{1}{12}N(N^2-1)}V(q)}{(q;q)^{N-1}}\right)^{-\ell_2}\sum_{\lambda\in\Lambda_w^+}\dim_p(R_\lambda)^{-\ell_1}\dim_q(R_\lambda)^{-\ell_2}, \tag{4.42}$$

where the sum is over representations, R_λ, of $\mathrm{SU}(N)$, labeled by weights λ, p and q are the parameters of the index, and

$$\ell_1 = \mathfrak{g} - 1 + \mathfrak{n}\,, \qquad \ell_2 = \mathfrak{g} - 1 - \mathfrak{n}\,. \tag{4.43}$$

Comparing to (4.40), we see these agree for general choices of genus $\mathfrak{g}$ and flux $\mathfrak{n}$, up to an overall factor of[36]

$$p^{-\ell_1\left(\frac{N(N^2-1)}{12}+\frac{N-1}{24}\right)}q^{-\ell_2\left(\frac{N(N^2-1)}{12}+\frac{N-1}{24}\right)}\,. \tag{4.44}$$

Using (4.6), we may rewrite the powers of p and q appearing here as

$$p^{-\ell_1\left(\frac{N(N^2-1)}{12}+\frac{N-1}{24}\right)}q^{-\ell_2\left(\frac{N(N^2-1)}{12}+\frac{N-1}{24}\right)} = e^{-\beta\hat{E}_{\mathrm{Casimir}}}\,;$$

$$\hat{E}_{\mathrm{Casimir}} = 2\pi Q^{-1}\left(\frac{N(N^2-1)}{12}+\frac{N-1}{24}\right)(\ell_1 b + \ell_2 b^{-1})\,. \tag{4.45}$$

As we show in appendix A this prefactor precisely matches the Casimir energy of the 4d theory (see (A.20)) and thus we precisely recover the expected relation (4.4) between the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function and the 4d index of these 4d $\mathcal{N} = 1$ theories. This serves as a strong consistency check of our calculation. We discuss the Casimir energy for general parameters and more general theories in section 4.2.1.

General parameters. Above we considered the partition function in a very special limit of parameters, where it drastically simplified, leading to an explicit evaluation formula. A natural and (as far as the authors are aware) open question is to compute the 4d index of class $\mathcal{S}$ theories for more general choices of the parameters, and in particular, find the dual TQFT that describes it.

---

[36]The terms proportional to $\frac{N-1}{24}$ in the exponents arise due to the relation between the Dedekind eta function and the q-Pochhammer symbol, i.e., $\eta(q) = q^{\frac{1}{24}}(q;q)$.

In principle the results derived above imply that the answer to this question is controlled by the Nekrasov-Shatashvili limit of the 5d instanton partition function of the maximal $\mathcal{N} = 2$ SYM theory. That is, the TQFT in question is governed by the twisted superpotential

$$\mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathcal{N}=2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathcal{N}=2,\text{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) + \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathcal{N}=2,\text{inst}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}). \tag{4.46}$$

Then the equations determining the supersymmetric vacua of the effective 2d theory, and hence the states of the 2d TQFT, are given by solving the Bethe equations,

$$e^{2\pi i \partial_{\tilde{u}_a} \mathcal{W}_{S_b^3 \times \mathbb{R}^2}^{\mathcal{N}=2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})} = 1, \qquad a = 1, \dots, r_G. \tag{4.47}$$

By the gauge-Bethe correspondence, it is known [19] that the vacuum equations associated to the NS-limit of the instanton partition function are equivalent to the Bethe equations of a certain integrable system. In the case of the 5d $\mathcal{N} = 2$ theory, this system is a quantization of the *relativistic Calogero-Moser system*, also known as the *Ruijsenaars-Schneider* (RS) model [80]. The relation between the 4d index and this integrable system was already pointed out in [81], where they observed that the indices of class $\mathcal{S}$ theories are naturally eigenfunctions of certain difference operators associated to the RS model. We expect this relation can be naturally understood in the above framework. Moreover, it would be interesting to explore whether the perspective above allows one to practically compute the 4d index of non-Lagrangian class $\mathcal{S}$ theories, which is a long-standing open problem.[37] We hope to return to this in future work.

4.2   5d $\mathcal{N} = 1$ theories

We now consider the case of $\mathcal{N} = 1$ theories with a 6d $\mathcal{N} = (1,0)$ UV completion. We consider two known examples in detail. The first example is a set of 5d circular quiver gauge theories whose UV completion is a $\mathbb{Z}_k$ orbifold of the 6d A_{N-1} theory, and which give rise, upon compactification of the 6d theory, to the 4d theories of class $\mathcal{S}_k$ [84]. The second example is the Seiberg theory in the special case $N_f = 8$, whose UV completion is the so-called E-string theory [5, 85]. The twisted compactifications of these two classes of theories were considered in [26] and [27], respectively. Then, as for the maximal theory above, we expect that the $S_b^3 \times \Sigma_\mathfrak{g}$ partition function of these 5d theories computes the 4d index of the corresponding 4d theories.

We also expect the $S_b^3 \times \Sigma_\mathfrak{g}$ partition function of these 5d theories, and therefore the supersymmetric indices of the corresponding 4d theories, to be related to an appropriate integrable system, through the gauge-Bethe correspondence for the Nekrasov-Shatashvili limit of the 5d instanton partition function [19]. Indeed, a connection of the corresponding 4d models to integrable systems was observed in these two examples, namely, to a generalization of the RS model discussed in [84, 86–88] for the class $\mathcal{S}_k$ case, and to the *van Diejen model* in the E-string case [89]. It would be interesting to use the computation

---

[37]There has been some partial progress via various indirect methods, such as constructing Lagrangians with reduced supersymmetry which flow to these non-Lagrangian class $\mathcal{S}$ fixed points; see, e.g., [82, 83].

above to shed more light on the connection between these gauge theories and integrable systems. As in the $\mathcal{N} = 2$ case discussed above, this requires a detailed understanding of the instanton corrections to the partition function.

In the present work, we will limit ourselves to a simplifying limit where we may evaluate the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function analytically. Specifically, we will consider the partition function in the strong coupling limit, $g_5 \to \infty$. Although naïvely in this limit we expect the contributions from instantons to be large, we will see that in fact their contribution in this limit can be simply characterized, and so we may compute the leading behavior of the partition function analytically.

We first study the general behavior of 5d theories with 6d completions in the strong coupling limit in section 4.2.1. We then describe the instanton corrections arising in this limit, and in the process conjecture a general condition for 5d theories to admit such 6d completions in section 4.2.2. Finally, we compute the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function this limit for the $\mathcal{N} = 2$ theory, as well as two examples mentioned above, and compare to the corresponding 4d computations, in section 4.2.3.

4.2.1 Casimir energy

We begin by showing that for generic theories with a 6d UV completion, the strong-coupling limit $g_5 \to \infty$ of the partition function exhibits the behavior

$$\lim_{g_5 \to \infty} Z_{S_b^3 \times \Sigma_{\mathfrak{g}}} \approx e^{-\beta \hat{E}_{\text{Casimir}}} \,, \tag{4.48}$$

where β is the radius of the 6d circle, related to g_5 via (4.1). The quantity $\hat{E}_{\text{Casimir}}$ appearing here has the following interpretation. In theories with a 6d UV completion, the limit $g_5 \to \infty$ corresponds to the $\beta \to \infty$ limit of the 6d partition function on $S_b^3 \times \Sigma_{\mathfrak{g}} \times S_\beta^1$. Since in this limit the partition function is dominated by the vacuum state, $\hat{E}_{\text{Casimir}}$ denotes the vacuum energy of the 6d theory quantized on $S_b^3 \times \Sigma_{\mathfrak{g}}$. Equivalently, by shrinking $\Sigma_{\mathfrak{g}}$ to zero size, it also corresponds to the Casimir energy of the 4d theory thus obtained, quantized on S_b^3.

On the other hand, this Casimir energy can be independently computed from the anomaly polynomial, I_8, of the parent 6d theory (see appendix A for details) by performing the topological twist, integrating over the Riemann surface, and applying the method of equivariant integration of [90], i.e.,

$$E_{\text{Casimir}} = \int_\varepsilon \int_{\Sigma_{\mathfrak{g}}} I_8^{\text{twisted}} \,. \tag{4.49}$$

For consistency, the two quantities obtained from (4.48) and (4.49) should coincide: $\hat{E}_{\text{Casimir}} = E_{\text{Casimir}}$. We have already seen an example of this in (4.45) for the Schur limit.

In the case of $\mathcal{N} = 1$, however, to make the computation feasible we limit ourselves to the perturbative partition function in (4.48). Thus, we expect the Casimir energy extracted from there to differ from the exact result (4.49). Nonetheless, the difference can be easily characterized. Namely, decomposing $I_8 = I_{8,\text{free}} + I_{8,\text{int}}$, where $I_{8,\text{free}}$ is the contribution

from all the free multiplets in the tensor branch of the theory, and $I_{8,\text{int}}$ is the remaining part, each term gives a corresponding contribution to (4.49).

Then, we will find that the perturbative result matches exactly the latter and thus, for consistency, the instanton contribution must be given by the former:

$$\hat{E}_{\text{Casimir}}^{\text{pert}} = \int_{\varepsilon} \int_{\Sigma_{\mathfrak{g}}} I_{8,\text{int}}^{\text{twisted}}, \qquad \hat{E}_{\text{Casimir}}^{\text{inst}} = \int_{\varepsilon} \int_{\Sigma_{\mathfrak{g}}} I_{8,\text{free}}^{\text{twisted}}. \qquad (4.50)$$

The analogous observation for the case of S^5 was already made in [90], suggesting a deeper understanding of this fact. We now proceed with the explicit computation.

Computation of the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function in the $g_5 \to \infty$ limit. To study this limit, it will be useful to go to the Bethe sum formulation of the partition function. Let us consider a general 5d theory, with gauge group G and hypermultiplets in a representation $R = \bigoplus_i R_i$ of G.[38] Then recall the Bethe equations are given by

$$\Pi_a = e^{2\pi i \gamma K^{ab} u_b} \prod_{\rho \in R_i} s_b(\rho(u) + \nu_i)^{\rho_a} \prod_{\alpha \in Ad(G)'} s_b(\alpha(u) - iQ)^{-\alpha_a} \Pi_a^{\text{inst}}(u, \nu, \gamma) = 1. \qquad (4.51)$$

where Π_a^{inst} is the contribution from instantons, and we recall

$$\gamma = -\frac{2\pi Q}{g_5^2}, \qquad (4.52)$$

Then the limit (4.48) corresponds to taking $\gamma \to 0$, so we consider the solutions to (4.51) in this limit. More precisely, we expect the partition function to be divergent in this limit, so we are interested in the solutions which contribute to the leading divergence.

Let us first ignore the contribution of the instantons. Then we see that, for any solutions at finite u, the exponent in the first factor of (4.51) becomes negligible as $\gamma \to 0$. Then we expect these solutions to have a finite contribution to the partition function as $\gamma \to 0$. Since we are interested in extracting the leading divergence in the limit (4.48), we instead look for solutions at large u, scaling with γ^{-1}. These can arise due to competition between this exponential factor and the double sine functions in (4.51).

Since we are working at large u, it will be useful to recall the expansion of the double sine function, as in (3.8),

$$s_b(x) \to e^{\pm \frac{\pi i}{2}\left(x^2 + \frac{1}{3}Q^2 - \frac{1}{6}\right)} \qquad \text{for } \text{Re}(x) \to \pm\infty, \qquad (4.53)$$

where $Q = \frac{1}{2}(b + b^{-1})$. At this point we must determine the sign of $\rho(u)$ for each weight ρ. Without loss of generality we may take u to lie in the fundamental Weyl chamber. For certain representations, such as the adjoint, this fixes the sign of $\rho(u)$ for all weights ρ. In this case we will call a weight "positive," and write $\rho > 0$, if $\rho(u) > 0$ for u in the chosen fundamental Weyl chamber, generalizing the notation for weights of the adjoint representation. For general representations, we may need to say more about the

---

[38]Here we do not include a bare 5d Chern-Simons term, as one does not generally appear in 5d theories with 6d UV completions.

region where u lies before making this split into positive and negative weights, but we will not consider such representations here.[39] We also restrict to representations that are self conjugate, and then we may group the weights into pairs, $(\rho, -\rho)$, where $\rho > 0$. Then let us define, for a general such representation R of G,

$$\mathcal{C}_R^{abc} = \frac{1}{2}\sum_{\rho > 0} \rho^a \rho^b \rho^c, \qquad \mathcal{C}_R^a = \frac{1}{2}\sum_{\rho > 0} \rho^a. \tag{4.54}$$

For example, for $R = Ad$, the adjoint representation, $\mathcal{C}_{Ad}^a = \delta^a$, where $\delta = \frac{1}{2}\sum_{\alpha > 0}\alpha$ is the Weyl vector of the group.

With this background, and using (4.53), we may approximate the l.h.s. of the Bethe equations by

$$\Pi_a \underset{u \to \infty}{\longrightarrow} \exp\left\{2\pi i\left(\gamma K^{ab}u_b + \sum_i\left(\mathcal{C}_{R_i}^{abc}u_b u_c + \mathcal{C}_{R_i}^a\left(\nu_i{}^2 + \frac{1}{3}Q^2 - \frac{1}{6}\right)\right)\right.\right.$$
$$\left.\left. - \mathcal{C}_{Ad}^{abc}u_b u_c - \mathcal{C}_{Ad}^a\left(-\frac{2}{3}Q^2 - \frac{1}{6}\right)\right)\right\}. \tag{4.55}$$

Generically, the term quadratic in u will dominate the first term at large u, and then the solutions will not scale with γ^{-1} as $\gamma \to 0$. However, suppose we consider a theory with field content such that[40]

$$\mathcal{C}_{Ad}^{abc} - \sum_i \mathcal{C}_{R_i}^{abc} = 0, \tag{4.56}$$

then (4.55) simplifies to (after using $\nu = -iQ\tilde{\nu}$)

$$\Pi_a \underset{u \to \infty}{\longrightarrow} \exp\left\{2\pi i\left(\gamma K^{ab}u_b + \sum_i \mathcal{C}_{R_i}^a\left(Q^2(1 - \tilde{\nu}_i^2) - \frac{1 + 4Q^2}{6}\right) + \mathcal{C}_{Ad}^a\left(\frac{1 + 4Q^2}{6}\right)\right)\right\}. \tag{4.57}$$

The solutions in this limit are then found by setting the exponent equal to $-2\pi i n^a$, $n^a \in \mathbb{Z}$, and we have

$$\hat{u}_a^{(n)} \approx -\gamma^{-1}K_{ab}\left(\sum_i \mathcal{C}_{R_i}^b\left(Q^2(1 - \tilde{\nu}_i^2) - \frac{1 + 4Q^2}{6}\right) + \mathcal{C}_{Ad}^b\left(\frac{1 + 4Q^2}{6}\right) + n^a\right), \tag{4.58}$$

and we indeed find the expected scaling with γ^{-1}. Here we must choose the n^a such that u lies in the interior of the fundamental Weyl chamber, as in the Schur limit discussed above.

So far we have ignored the contribution of instantons to the Bethe equations. As we will argue in section 4.2.2 below, we expect these to have a subleading contribution at large u, and so we assume now that we may ignore their contribution in the Bethe equations above.

---

[39]For example, for the fundamental representation of SU(N), specifying we are in the fundamental Weyl chamber, with $u_1 > \ldots > u_N$ does not specify the sign of all u_i; there are still $N-1$ regions in this chamber where the separation into positive and negative weights are distinct.

[40]We note the same relation can be derived from the conditions (3.36) and (3.46) discussed in section 3 as the condition to obtain N^3 scaling of the free energy.

It now remains to evaluate

$$Z_{S_b^3 \times \Sigma_{\mathfrak{g}}} = \sum_{\hat{u} \in \mathcal{S}_{\mathrm{BE}}} \mathcal{H}(\hat{u}, \nu, \gamma)^{\mathfrak{g}-1} \Pi_i(\hat{u}, \nu, \gamma)^{\mathfrak{n}_i} \tag{4.59}$$

at the Bethe vacua (4.58). Since we are interested in extracting the leading divergence as $\gamma \to 0$, we will look for the vacuum which has the dominant contribution in this limit. As before, we first consider the perturbative contribution

$$\mathcal{H}^{\mathrm{pert}}(\hat{u}, \nu, \gamma)^{\mathfrak{g}-1} \Pi_i^{\mathrm{pert}}(\hat{u}, \nu, \gamma)^{\mathfrak{n}_i} = (H^{\mathrm{pert}})^{\mathfrak{g}-1} \prod_i \prod_{\rho \in R_i} s_b(\rho(\hat{u}) + \nu_i)^{\mathfrak{n}_i} \prod_{\alpha \in Ad(G)'} s_b(\alpha(\hat{u}) - iQ)^{1-\mathfrak{g}}, \tag{4.60}$$

where

$$H^{\mathrm{pert}} = \det_{a,b} \frac{1}{2\pi i} \frac{\partial \log \Pi_a^{\mathrm{pert}}}{\partial u_b}. \tag{4.61}$$

Expanding these at large u using (4.53), we find (defining $\hat{\mathfrak{n}}_i = \frac{\mathfrak{n}_i}{\mathfrak{g}-1}$ as in the previous subsection)

$$\mathcal{H}^{\mathrm{pert}}(u, \nu)^{\mathfrak{g}-1} \Pi_i^{\mathrm{pert}}(u, \nu)^{\mathfrak{n}_i} \approx (\gamma^{r_G} K)^{\mathfrak{g}-1} \exp \left\{ 4\pi i(\mathfrak{g}-1) \left(-\sum_i iQ \mathcal{C}_{R_i}^a \hat{\mathfrak{n}}_i \tilde{\nu}_i + iQ \mathcal{C}_{Ad}^a \right) u_a \right\}, \tag{4.62}$$

where we have used

$$H^{\mathrm{pert}} \approx \det_{a,b} \gamma K^{ab} = K \gamma^{r_g}, \qquad K \equiv \det K^{ab}. \tag{4.63}$$

For $\gamma \to 0$, the contribution from the Hessian factor is subleading, and we may approximate the contribution from the Bethe vacuum, (4.58), as

$$\mathcal{H}^{\mathrm{pert}}(\hat{u}^{(n)}, \nu, \gamma)^{\mathfrak{g}-1} \Pi_i^{\mathrm{pert}}(\hat{u}^{(n)}, \nu, \gamma)^{\mathfrak{n}_i} \approx \exp \left\{ -4\pi i(\mathfrak{g}-1)\gamma^{-1} K_{ab} \right.$$
$$\left. \left(-\sum_i iQ \mathcal{C}_{R_i}^a \hat{\mathfrak{n}}_i \tilde{\nu}_i + iQ \mathcal{C}_{Ad}^a \right) \left(\sum_i \mathcal{C}_{R_i}^b \left(Q^2(1-\tilde{\nu}_i^2) - \frac{1+4Q^2}{6} \right) + \mathcal{C}_{Ad}^b \left(\frac{1+4Q^2}{6} \right) + n^b \right) \right\}. \tag{4.64}$$

The final step is to determine the vacuum, $\hat{u}^{(n)}$, which has the leading contribution to the Bethe sum. Recall that we assume u lies in a fundamental Weyl chamber, and so we may choose the n^b to take u arbitrarily large within this chamber. For concreteness, let us take the basis, e_a, of the Cartan to be dual to the fundamental weights, so that $c^a e_a$ spans the interior of the fundamental Weyl chamber as we take $c^a > 0$. Then we see that, in this basis we must impose

$$\mathrm{Im} \left[(\mathfrak{g}-1)\gamma^{-1} K_{ab} \left(-\sum_i iQ \mathcal{C}_{R_i}^b \hat{\mathfrak{n}}_i \tilde{\nu}_i + iQ \mathcal{C}_{Ad}^b \right) \right] > 0, \qquad a = 1, \ldots, r_G; \tag{4.65}$$

otherwise, by taking an appropriate n^a large we can arrange the r.h.s. of (4.64) to be arbitrarily large, and we do not find a well-defined $\gamma \to 0$ limit.

– 53 –

With this assumption, we can see the dominant contribution comes from taking the n^a as small as possible, while keeping u in the fundamental Weyl chamber. In the basis above, this is achieved by setting $n^a = 0$, and so we finally arrive at the leading behavior of the partition function,

$$Z^{\text{pert}}_{\Sigma_{\mathfrak{g}} \times S^3_b} \underset{\gamma \to 0}{\approx} \exp \left\{ 4\pi(\mathfrak{g}-1)\gamma^{-1} K_{ab} \left(-\sum_i Q\mathcal{C}^a_{R_i} \hat{\mathfrak{n}}_i \tilde{\nu}_i + Q\mathcal{C}^a_{Ad} \right) \right.$$
$$\left. \times \left(\sum_i \mathcal{C}^b_{R_i} \left(Q^2(1-\tilde{\nu}_i^2) - \frac{1+4Q^2}{6} \right) + \mathcal{C}^b_{Ad} \left(\frac{1+4Q^2}{6} \right) \right) \right\}. \quad (4.66)$$

Given the relation (4.52) and identification (4.1) we see this has the expected behavior (4.48). Specifically, we find

$$\hat{E}^{\text{pert}}_{\text{Casimir}} = 4\pi\lambda(\mathfrak{g}-1) K_{ab} \left(-\sum_i \mathcal{C}^a_{R_i} \hat{\mathfrak{n}}_i \tilde{\nu}_i + \mathcal{C}^a_{Ad} \right)$$
$$\times \left(\sum_i \mathcal{C}^b_{R_i} \left(Q^2(1-\tilde{\nu}_i^2) - \frac{1+4Q^2}{6} \right) + \mathcal{C}^b_{Ad} \left(\frac{1+4Q^2}{6} \right) \right), \quad (4.67)$$

where the label "pert" denotes that this has been computed using the perturbative approximation to the $S^3_b \times \Sigma_{\mathfrak{g}}$ partition function. We turn to consider the instanton contributions next.

4.2.2 Instantons, the 5d prepotential, and a 6d uplift condition

So far we have ignored the contribution of instantons in the above analysis, and only used the perturbative approximation to the partition function. We argue that these have a subleading contribution to the twisted superpotential in the regions of large u. Therefore we expect they will not modify the analysis of the Bethe solutions obtained using the perturbative twisted superpotential above.

To motivate this claim, note that taking large u is equivalent to exploring the asymptotic region of the Coulomb branch, where we expect the theory to be weakly coupled, and dominated by the perturbative calculation. In fact, we claim that the large u behavior of the twisted superpotential, which controls the flux operator Π_a appearing in the Bethe equations, is related to the effective 5d prepotential, $\mathcal{F}_{\text{eff}}$, which is known to be perturbatively exact. Namely, using (4.53) we may write

$$g_b(\tilde{u}) \underset{\text{Re}(\tilde{u}) \to \pm\infty}{\longrightarrow} \mp \frac{1}{12} Q^2 \tilde{u}^3 + \mathcal{O}(\tilde{u}). \quad (4.68)$$

This means the behavior of the perturbative twisted superpotential of the theory at large $\tilde{u}$ is given by[41]

$$\mathcal{W}^{\text{pert}}_{S^3_b \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = -\frac{1}{2}Q^2\tilde{\gamma}\tilde{u}^2 + \sum_i \sum_{\rho \in R_i} g_b(\rho(\tilde{u})+\tilde{\nu}_i) - \sum_{\alpha \in Ad(G)'} g_b(\alpha(\tilde{u})+1)$$
$$\underset{|\tilde{u}| \to \infty}{\longrightarrow} Q^2 \left(\frac{\pi i}{g_5^2}\tilde{u}^2 - \frac{1}{12}\sum_i \sum_{\rho \in R_i} |\rho(\tilde{u})+\tilde{\nu}_i|^3 + \frac{1}{12}\sum_{\alpha \in Ad(G)} |\alpha(\tilde{u})|^3 \right) + \mathcal{O}(|\tilde{u}|).$$
$$(4.69)$$

---

[41]Here we assume $\text{Im}(\tilde{u}) \ll \text{Re}(\tilde{u})$, so that we may approximate $\pm\tilde{u}$ in (4.68) as $|\tilde{u}|$.

But the quantity in parentheses is precisely the effective prepotential of the 5d theory [5], as in (2.7). A similar relation was noted for the integrand of the S^5 partition function in [62]. Given that the prepotential of a 5d theory does not receive non-perturbative corrections, it is then natural to make a similar assumption for the large $|\tilde{u}|$ behavior of the twisted superpotential, and hence for the Bethe equations.

In light of the relation above, the condition (4.56) can be equivalently phrased by saying that, in the notation of (2.7),

$$c_{abc} = 0 \,, \tag{4.70}$$

i.e., the effective CS term vanishes for every direction in the Coulomb branch of the 5d theory. As we saw above, when this condition is satisfied, we find the partition function has a leading divergence with $\log Z \sim \gamma^{-1}$ as $\gamma \to 0$, which is the expected Casimir behavior of a theory with an emergent circle of radius $\beta \sim \gamma^{-1}$. Thus we may conjecture that the condition (4.56) (or equivalently (4.70)) is the relevant condition for the 5d theory to admit a 6d UV completion. We will see in examples below that this is indeed satisfied in several examples where a 6d UV completion is expected to exist. It would be interesting to explore this conjecture further.

In addition to the Bethe equations, we did not include instanton contributions when substituting the Bethe solutions into the sum over vacua in (4.59). In fact, we will see below that these do have a nontrivial contribution, but they can be explicitly characterized in terms of free fields of the 6d theory as already mentioned above (4.50).

4.2.3 Examples

Let us now consider some examples of the Casimir energy computation in 5d theories that are believed to admit a 6d UV completion.

Maximal theory. We begin with maximal 5d $\mathcal{N} = 2$ super Yang-Mills with gauge group G, consisting of an $\mathcal{N} = 1$ vector and an adjoint hypermultiplet, with mass $\tilde{\nu}$. In this case (4.66) gives

$$Z_{\Sigma_{\mathfrak{g}} \times S_b^3} \underset{\gamma \to 0}{\approx} \exp\left\{ \frac{\pi}{3}(\mathfrak{g} - 1)\gamma^{-1} h_G d_G (1 - \hat{\mathfrak{n}}\tilde{\nu})(1 - \tilde{\nu}^2)Q^3 \right\} \,, \tag{4.71}$$

where we used

$$K_{ab} \mathcal{C}^a_{Ad} \mathcal{C}^b_{Ad} = K_{ab} \delta^a \delta^b = \frac{1}{12} h_G d_G \,, \tag{4.72}$$

where d_G and h_G are the dimension and dual Coxeter number of G, respectively. For example, for $G = \mathrm{SU}(N)$, we have

$$K_{ab} \delta^a \delta^b = \frac{1}{12} N(N^2 - 1) \,. \tag{4.73}$$

Using (4.52) and (2.41) we thus have

$$\log Z_{\Sigma_{\mathfrak{g}} \times S_b^3} \approx -\frac{g_5^2}{2\pi} \frac{\pi}{3}(\mathfrak{g} - 1)N(N^2 - 1)(\tilde{\nu}^2 - 1)(\hat{\mathfrak{n}}\tilde{\nu} - 1)Q^2 \,. \tag{4.74}$$

Now, with the identification $\beta = \frac{g_5^2}{2\pi}$ (i.e., $\lambda = 1$) we see that the partition function has the expected behavior (4.48), with

$$\hat{E}_{\text{Casimir}}^{\text{pert}} = \frac{\pi}{3}(\mathfrak{g} - 1)N(N^2 - 1)(\tilde{\nu}^2 - 1)(\hat{\mathfrak{n}}\tilde{\nu} - 1)Q^2 \,. \tag{4.75}$$

This may be compared to the exact Casimir energy, computed from the anomaly polynomial of the 6d theory in appendix A.1. As discussed there, separating the contribution from free fields to the anomaly polynomial and the remaining part one can write $E_{\text{Casimir}} = E^{\text{pert}} + E^{\text{non-pert}}$, with

$$E^{\text{pert}} \equiv \int_\varepsilon \int_{\Sigma_\mathfrak{g}} I_{8,\text{int}}^{\text{twisted}} = \frac{\pi}{3}(\mathfrak{g} - 1)N(N^2 - 1)\left(\epsilon^2 - 1\right)(\hat{\mathfrak{n}}\epsilon - 1)Q^2 \,, \tag{4.76}$$

$$E^{\text{non-pert}} \equiv \int_\varepsilon \int_{\Sigma_\mathfrak{g}} I_{8,\text{free}}^{\text{twisted}} = \frac{\pi}{6}(\mathfrak{g} - 1)(N - 1)\left(1 + \hat{\mathfrak{n}}\epsilon + 2\hat{\mathfrak{n}}\epsilon\left(\epsilon^2 - 1\right)Q^2\right) \,, \tag{4.77}$$

where ϵ is the parameter controlling the mixing of R-symmetry with the flavor symmetry acting on the (adjoint) hypermultiplet; see (A.3). We see that identifying $\tilde{\nu} = \varepsilon$, the perturbative localization calculation (4.75) coincides with (4.76). Thus, for consistency, the instanton contribution to the partition function should match the remaining free field contribution (4.77), as claimed at the beginning of section 4.2.1. Next, we provide more examples of the same phenomenon.

Class $\mathcal{S}_k$ Theories. These theories have k SU(N) vector multiplets and k bifundamental hypermultiplets, forming a circular quiver [84]. Let us assume that all the u_a's have the same solution. In this case each bifundamental contribution to the BAE is essentially that of an adjoint in which case we simply obtain k times the contribution of maximal theory in the Casimir limit,

$$\log Z_{\Sigma_\mathfrak{g} \times S_b^3} \approx -\frac{g_5^2}{2\pi k}\frac{\pi k^2}{3}(\mathfrak{g} - 1)N(N^2 - 1)(1 - \tilde{\nu}^2)(1 - \hat{\mathfrak{n}}\tilde{\nu})Q^2 \,. \tag{4.78}$$

Again, identifying $\tilde{\nu} = \varepsilon$ and the 6d radius with $\lambda = k$, i.e.,

$$\beta = \frac{g_5^2}{2\pi k} \,, \tag{4.79}$$

we see that the perturbative Casimir energy reads

$$\hat{E}_{\text{Casimir}}^{\text{pert}} = \frac{\pi k^2}{3}(\mathfrak{g} - 1)N(N^2 - 1)(1 - \varepsilon^2)(1 - \hat{\mathfrak{n}}\varepsilon)Q^2 \,, \tag{4.80}$$

which precisely matches (A.30). Thus, we see that the missing instanton part (A.31) is again identified with the contribution of free fields to the anomaly polynomial (A.32).

E-string. Next, we consider the 5d E-string theory, a 5d $\mathcal{N}=1$ USp($2N$) theory with $N_f = 8$ fundamental hypermultiplets and an antisymmetric hyper. First we will need to compute the quantities $\mathcal{C}_R^{abc}$ and $\mathcal{C}_R^a$ for the various representations. We find, working with $u = (u_a)$, $a = 1, \ldots, N$, in the fundamental representation

$$
\mathcal{C}_R^{abc} u_a u_b u_c = \begin{cases} \frac{1}{2} \sum_{a=1}^{N} (u_a)^3, & R = F \\ \frac{1}{2} \sum_{a<b} \left((u_a + u_b)^3 + (u_a - u_b)^3 \right), & R = AS \\ \frac{1}{2} \sum_{a<b} \left((u_a + u_b)^3 + (u_a - u_b)^3 \right) + \frac{1}{2} \sum_{a=1}^{N} (2u_a)^3, & R = Ad \end{cases} \tag{4.81}
$$

$$
\mathcal{C}_R^a u_a = \begin{cases} \frac{1}{2} \sum_{a=1}^{N} u_a, & R = F \\ \frac{1}{2} \sum_{a<b} \left((u_a + u_b) + (u_a - u_b) \right) = \sum_{a=1}^{N} (N-a) u_a, & R = AS \\ \sum_{a=1}^{N} (N - a + 1) u_a, & R = Ad \end{cases} \tag{4.82}
$$

Then note that

$$
\mathcal{C}_{Ad}^{abc} = \mathcal{C}_{AS}^{abc} + 8\, \mathcal{C}_F^{abc} \,, \tag{4.83}
$$

and thus (4.56) is satisfied precisely for $N_f = 8$ fundamental hypers. In this case, plugging these into (4.66), we have

$$
Z_{\Sigma_{\mathfrak{g}} \times S_b^3} \underset{\gamma \to 0}{\approx} \exp \left\{ 4\pi(\mathfrak{g}-1) \gamma^{-1} K_{ab} \left(-\sum_{i=1}^{8} \mathcal{C}_F^a \hat{\mathfrak{n}}_i \tilde{\nu}_i - \mathcal{C}_{AS}^a \hat{\mathfrak{n}} \tilde{\nu} + \mathcal{C}_{Ad}^a \right) Q \right.
$$
$$
\left. \times \left(\sum_{i=1}^{8} \mathcal{C}_F^b \left(Q^2 (1 - \tilde{\nu}_i^2) - \frac{1 + 4Q^2}{6} \right) + \mathcal{C}_{AS}^b \left(Q^2 (1 - \tilde{\nu}^2) - \frac{1 + 4Q^2}{6} \right) + \mathcal{C}_{Ad}^b \left(\frac{1 + 4Q^2}{6} \right) \right) \right\} . \tag{4.84}
$$

We may simplify this using

$$
\mathcal{C}_{Ad}^a = \mathcal{C}_{AS}^a + 2\mathcal{C}_F^a \,, \tag{4.85}
$$

which gives

$$
Z_{\Sigma_{\mathfrak{g}} \times S_b^3} \underset{\gamma \to 0}{\approx} \exp \left\{ 4\pi(\mathfrak{g} - 1) \gamma^{-1} K_{ab} \left(\sum_{i=1}^{8} \mathcal{C}_F^a \left(\frac{1}{4} - \hat{\mathfrak{n}}_i \tilde{\nu}_i \right) + \mathcal{C}_{AS}^a (1 - \hat{\mathfrak{n}} \tilde{\nu}) \right) Q \right.
$$
$$
\left. \times \left(\sum_{i=1}^{8} \mathcal{C}_F^b \left(Q^2 (1 - \tilde{\nu}_i^2) - \frac{1}{8} (1 + 4Q^2) \right) + \mathcal{C}_{AS}^b Q^2 (1 - \tilde{\nu}^2) \right) \right\} . \tag{4.86}
$$

We may further expand this, using

$$
K_{ab} \mathcal{C}_F^a \mathcal{C}_F^b = \frac{1}{8} N \,, \quad K_{ab} \mathcal{C}_{AS}^a \mathcal{C}_{AS}^b = \frac{1}{12} N(N-1)(2N-1) \,, \quad K_{ab} \mathcal{C}_F^a \mathcal{C}_{AS}^b = \frac{1}{8} N(N-1) \,, \tag{4.87}
$$

which yields

$$
\begin{aligned}
Z_{\Sigma_{\mathfrak{g}} \times S_b^3} \underset{\gamma \to 0}{\approx} \exp \Bigg\{ & 4\pi(\mathfrak{g}-1)\gamma^{-1}Q \left(\frac{1}{8}N \sum_{i,j=1}^{8} \left(\frac{1}{4} - \hat{\mathfrak{n}}_i \tilde{\nu}_i \right) \left(Q^2 \left(\frac{1}{2} - \tilde{\nu}_j^2 \right) - \frac{1}{8} \right) \right. \\
& + \frac{1}{8}N(N-1) \sum_{i=1}^{8} \left(\left(\frac{1}{4} - \hat{\mathfrak{n}}_i \tilde{\nu}_i \right) Q^2 \left(1 - \tilde{\nu}^2\right) + \left(1 - \hat{\mathfrak{n}}\tilde{\nu}\right) \left(Q^2 \left(\frac{1}{2} - \tilde{\nu}_j^2 \right) - \frac{1}{8} \right) \right) \quad \text{(4.88)} \\
& \left. + \frac{1}{12}N(N-1)(2N-1)\left(1 - \hat{\mathfrak{n}}\tilde{\nu}\right)Q^2 \left(1 - \tilde{\nu}^2\right) \right) \Bigg\} \, .
\end{aligned}
$$

The result of the anomaly polynomial calculation for the case of two flavor fluxes is given in (A.37) and (A.38). To compare with that calculation, we set $\tilde{\nu} = \varepsilon$, $\tilde{\nu}_1 = \varepsilon_E$, $\hat{\mathfrak{n}}_1 = \hat{\mathfrak{n}}_E$ with all other $\hat{\mathfrak{n}}_{i=2,\dots,8} = \tilde{\nu}_{i=2,\dots,8} = 0$ in (4.88) to get

$$
\begin{aligned}
\hat{E}_{\text{Casimir}}^{\text{pert}} = \pi(\mathfrak{g}-1)\lambda \Bigg[& \frac{1}{2}N(N-1)Q^2 \left(1 - \varepsilon^2\right) \left(\frac{2}{3}(2N-1)\left(1 - \hat{\mathfrak{n}}\varepsilon\right) + \left(2 - \hat{\mathfrak{n}}_E \varepsilon_E\right) \right) \\
& + \frac{1}{2}N \left(Q^2 \left(4 - \varepsilon_E^2\right) - 1\right) \left((N-1)\left(1 - \hat{\mathfrak{n}}\varepsilon\right) + \left(2 - \hat{\mathfrak{n}}_E \varepsilon_E\right)\right) \Bigg] , \quad \text{(4.89)}
\end{aligned}
$$

which matches (A.37) provided $\lambda = 2$. This means the 6d radius is identified, as in [90], to be

$$
\beta = \frac{g_5^2}{4\pi} \, . \tag{4.90}
$$

Thus, once again, the instanton part (A.38) missing from the above calculation is identified with the contribution of free fields to the anomaly polynomial (A.39).

Summary. We have argued that computing the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function for theories with a UV completion as 6d $\mathcal{N} = (1,0)$ theories is equivalent to computing the $S_b^3 \times S^1$ partition function (or index, up to the Casimir energy), of the corresponding 4d $\mathcal{N} = 1$ theories obtained by compactification of the 6d theory on $\Sigma_{\mathfrak{g}}$. We have checked this explicitly for the case of 5d maximal SYM and the corresponding 4d $\mathcal{N} = 1$ and $\mathcal{N} = 2$ class $\mathcal{S}$ theories of [23–25] in the mixed Schur limit, finding perfect agreement with [29]. We then considered examples of 5d $\mathcal{N} = 1$ theories. In this case we do not have an explicit handle on instanton contributions and we limited ourselves to computing the partition function asymptotically on the Coulomb branch, where instanton contributions are suppressed and the partition function is expected to be dominated by the Casimir energy, which we extracted. We find an exact match with an independent calculation of the Casimir energy from anomaly polynomials, up to the expected instanton corrections which, for consistency, should be identified with the contribution of free fields to the corresponding 6d anomaly polynomials, as noted in the case of S^5 in [90].

5 Twisted partition function on $S^2 \times S^1 \times \Sigma_\mathfrak{g}$ and AdS$_6$ black holes

Although the main focus of this paper has so far been the partition function on $S^3_b \times \Sigma_\mathfrak{g}$, we observed in section 2.4 that there is a natural generalization of the above computation to $\mathcal{M}_3 \times \Sigma_\mathfrak{g}$ for $\mathcal{M}_3$ a general lens space, $L(p,q)$, using the factorization of the twisted superpotential into "holomorphic block"-like contributions, as in (2.129). In this section we briefly consider the simplest example, that of the space $\mathcal{M}_3 = S^2 \times S^1$, corresponding to the topological index in 3d [8].

After writing the result for the partition function on $S^2 \times S^1 \times \Sigma_\mathfrak{g}$ according to the above framework in section 5.1, we conjecture a natural generalization for the partition function on $\Sigma_{\mathfrak{g}_1} \times \Sigma_{\mathfrak{g}_2} \times S^1$, which we claim is valid perturbatively, and so also in the large N limit. Based on this conjecture, we reproduce in section 5.2 the microscopic entropy of an infinite class of black holes in AdS$_6$.

5.1 Partition function on $S^2 \times S^1 \times \Sigma_\mathfrak{g}$

Let us first consider the case where $\mathcal{M}_3$ corresponds to the "refined" topological index, $S^2_{\mathfrak{q}_1} \times S^1$, with a fugacity $\mathfrak{q}_1$ for the angular momentum.[42] Then we may build the integrand of the full $S^2_{\mathfrak{q}_1} \times S^2 \times S^1$ partition function by including the contribution of the 5d Nekrasov partition function at the four fixed points, as in (2.123), namely,

$$\lim_{\epsilon \to 0} \prod_\ell Z_{\mathbb{R}^2_{\mathfrak{q}_1^{(\ell)}} \times \mathbb{R}^2_{\mathfrak{q}_2^{(\ell)}} \times S^1}(x^{(\ell)}, y^{(\ell)}, z^{(\ell)}) \,, \tag{5.1}$$

where now the parameters can be identified by a straightforward modification of the argument leading to (2.108), giving

$$\begin{aligned}
x^{(\ell)} &= e^{2\pi i \tilde{u}} \mathfrak{q}_1^{(\ell)\mathfrak{m}_1/2} \mathfrak{q}_2^{(\ell)\mathfrak{m}_2/2} \,, \\
\mathfrak{q}_1^{(\ell)} &= \begin{cases} \mathfrak{q}_1 \,, & \ell = nn \text{ or } ns \\ \mathfrak{q}_1^{-1} \,, & \ell = sn \text{ or } ss \end{cases} \\
\mathfrak{q}_2^{(\ell)} &= \begin{cases} e^{2\pi i \epsilon} \,, & \ell = nn \text{ or } sn \\ e^{-2\pi i \epsilon} \,, & \ell = ns \text{ or } ss \end{cases}
\end{aligned} \tag{5.2}$$

Then, using (2.126) and arguing as before, we find that the twisted superpotential controlling the partition function on $S^2_{\mathfrak{q}_1} \times S^1 \times \Sigma_\mathfrak{g}$ is given by

$$\begin{aligned}
\mathcal{W}_{S^2_{\mathfrak{q}_1} \times S^1 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1} = \mathcal{W}^{(5d)}_{NS}&\left(\tilde{u} + \frac{1}{2}\mathfrak{m}_1\epsilon_1, \tilde{\nu} + \frac{1}{2}\mathfrak{n}_1\epsilon_1, \tilde{\gamma}; \epsilon_1\right) \\
&+ \mathcal{W}^{(5d)}_{NS}\left(\tilde{u} - \frac{1}{2}\mathfrak{m}_1\epsilon_1, \tilde{\nu} - \frac{1}{2}\mathfrak{n}_1\epsilon_1, \tilde{\gamma}; -\epsilon_1\right) \,,
\end{aligned} \tag{5.3}$$

and similarly for the effective dilaton, $\Omega_{S^2_{\mathfrak{q}_1} \times S^1 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1}$. An important difference from the S^3_b case is that the dependence on the fluxes, $\mathfrak{m}_1, \mathfrak{n}_1$, on $S^2_{\mathfrak{q}_1}$ cannot be absorbed

---

[42]The subscript "1" denotes this S^2 sits as the first factor in the manifold, to avoid confusion below.

into a shift of $\tilde{u}$ and $\tilde{\nu}$. As a result, we obtain a direct sum of 2d theories indexed by the gauge flux, $\mathfrak{m}_1 \in \Lambda_G$. Then we may write the full partition function as[43]

$$Z_{S^2_{\mathfrak{q}_1} \times S^1 \times \Sigma_{\mathfrak{g}}}(\tilde{\nu}, \tilde{\gamma})_{\mathfrak{n}_1, \mathfrak{n}_2} = \sum_{\mathfrak{m}_1 \in \Lambda_G} \sum_{\hat{\tilde{u}} \in \mathcal{S}_{\mathrm{BE}\mathfrak{m}_1, \mathfrak{n}_1}} \Pi_i(\hat{\tilde{u}}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1}^{\mathfrak{n}_{2,i}} \mathcal{H}(\hat{\tilde{u}}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1}^{\mathfrak{g}-1}, \qquad (5.4)$$

where we must sum over the 2d theories labeled by the gauge fluxes, $\mathfrak{m}_1$ (as well as the flavor fluxes, $\mathfrak{n}_1$), whose vacua are given by

$$\mathcal{S}_{\mathrm{BE}\mathfrak{m}_1, \mathfrak{n}_1} = \{\hat{\tilde{u}} \mid \Pi_a(\hat{\tilde{u}}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1} = 1\}/W_G, \qquad (5.5)$$

and, e.g.,

$$\Pi_a(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1} = e^{2\pi i \partial_{\tilde{u}_a} \mathcal{W}_{S^2_{\mathfrak{q}_1} \times S^1 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1}}, \qquad (5.6)$$

and similarly for the other operators.

Let us now focus on the unrefined limit, $\mathfrak{q}_1 \to 1$, corresponding to the ordinary A-twist background on S^2. We first discuss the perturbative contribution to the partition function in this limit. For a single hypermultiplet, we have, by taking the $\epsilon_{1,2} \to 0$ limit of (2.101) using (2.86),

$$Z_{S^2 \times S^2 \times S^1}^{\mathrm{pert}, \mathrm{U}(1), \mathrm{hyp}}(\tilde{u})_{\mathfrak{m}_1, \mathfrak{m}_2} = \lim_{\epsilon_{1,2} \to 0} (x; \mathfrak{q}_1, \mathfrak{q}_2)_{\mathfrak{m}_1, \mathfrak{m}_2}^{-1} = e^{\mathfrak{m}_1 \mathfrak{m}_2 s(\tilde{u})}, \qquad (5.7)$$

where we introduced a function

$$s(\tilde{u}) \equiv \pi i \tilde{u} + \mathrm{Li}_1(e^{2\pi i \tilde{u}}) + \frac{i\pi}{2} \qquad \Rightarrow \qquad e^{s(\tilde{u})} = \frac{i e^{\pi i \tilde{u}}}{1 - e^{2\pi i \tilde{u}}}. \qquad (5.8)$$

The perturbative contribution for a general hyper and vector multiplet can be obtained similarly. Equivalently, these can be obtained as in section 2.3 by reduction on $\Sigma_{\mathfrak{g}}$ to a 3d theory on $S^2 \times S^1$. In either case, we find the perturbative approximation to the integrand of the $S^2 \times S^1 \times \Sigma_{\mathfrak{g}}$ partition function is given by

$$Z_{S^2 \times S^1 \times \Sigma_{\mathfrak{g}}}^{\mathrm{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1; \mathfrak{m}_2, \mathfrak{n}_2} = z^{\mathrm{tr}(\mathfrak{m}_1 \mathfrak{m}_2)} \prod_{\alpha \in Ad(G)'} \left(e^{-s(\alpha(\tilde{u}))} \right)^{[\alpha(\mathfrak{m}_1) + 1 - \mathfrak{g}][\alpha(\mathfrak{m}_2) + 1]}$$
$$\times \prod_I \prod_{\rho \in R_I} \left(e^{s(\rho(\tilde{u}) + \tilde{\nu})} \right)^{[\rho(\mathfrak{m}_1) + \mathfrak{n}_1][\rho(\mathfrak{m}_2) + \mathfrak{n}_2]}, \qquad (5.9)$$

where the index I runs over the hypermultiplets in the theory, in gauge representations R_I. Then the perturbative partition function is given by the integral formula

$$Z_{S^2 \times S^1 \times \Sigma_{\mathfrak{g}}}^{\mathrm{pert}}(\tilde{\nu}, \tilde{\gamma})_{\mathfrak{n}_1; \mathfrak{n}_2} = \sum_{\mathfrak{m}_1, \mathfrak{m}_2 \in \Lambda_G} \frac{1}{|W_G|} \oint_{\mathcal{C}_{JK}} d\tilde{u} \, Z_{S^2 \times \Sigma_{\mathfrak{g}} \times S^1}^{\mathrm{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1; \mathfrak{m}_2, \mathfrak{n}_2}. \qquad (5.10)$$

---

[43]Here we denote the fluxes on $\Sigma_{\mathfrak{g}}$ with a subscript "2" to distinguish them from those on $S^2_{\mathfrak{q}_1}$.

Equivalently, we may write this as a Bethe sum associated to the 2d theory obtained by compactification on $S^2 \times S^1$ as in (5.4), where now, to perturbative accuracy,

$$
\Pi_a^{\text{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1} = z^{\mathfrak{m}_1{}^a} \prod_{\alpha \in Ad(G)'} \left(e^{-s(\alpha(\tilde{u}))} \right)^{[\alpha(\mathfrak{m}_1)]\alpha^a} \prod_I \prod_{\rho \in R_I} \left(e^{s(\rho(\tilde{u})+\tilde{\nu})} \right)^{[\rho(\mathfrak{m}_1)+\mathfrak{n}_1]\rho^a} ,
$$

$$
\Pi_I^{\text{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1} = \prod_{\rho \in R_I} \left(e^{s(\rho(\tilde{u})+\tilde{\nu})} \right)^{[\rho(\mathfrak{m}_1)+\mathfrak{n}_1]} , \tag{5.11}
$$

$$
\mathcal{H}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = \prod_{\alpha \in Ad(G)'} \left(e^{-s(\alpha(\tilde{u}))} \right)^{-\alpha(\mathfrak{m}_2)} H^{\text{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1} ,
$$

where $H^{\text{pert}}_{\mathfrak{m}_1, \mathfrak{n}_1} = \det_{a,b} \frac{1}{2\pi i} \frac{\partial}{\partial \tilde{u}_b} \log \Pi_a^{\text{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1}$ is the Hessian factor.

5d prepotential and the Bethe equations. We can gain another perspective on this calculation, and characterize the instanton contributions, by observing the leading behavior of the Nekrasov-Shatashvili limit of the twisted superpotential in the $\epsilon \to 0$ limit [19],

$$
\mathcal{W}_{NS}^{(5d)}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}; \epsilon) \underset{\epsilon \to 0}{\cong} \frac{1}{\epsilon} \mathcal{F}_{\mathbb{R}^4 \times S^1}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) + \cdots , \tag{5.12}
$$

where $\mathcal{F}_{\mathbb{R}^4 \times S^1}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})$ is the effective prepotential of the 5d theory compactified on a circle [91]. Then we have, from (5.3), that the flux dependence of the twisted superpotential on $S^2 \times S^1 \times \mathbb{R}^2$ is governed by the 5d prepotential, via

$$
\begin{aligned}
\mathcal{W}_{S^2 \times S^1 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1} &= \lim_{\epsilon_1 \to 0} \mathcal{W}_{S^2_{\mathfrak{q}_1} \times S^1 \times \mathbb{R}^2}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1} \\
&= \mathfrak{m}_{1,a} \frac{\partial \mathcal{F}_{\mathbb{R}^4 \times S^1}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})}{\partial \tilde{u}_a} + \mathfrak{n}_{1,I} \frac{\partial \mathcal{F}_{\mathbb{R}^4 \times S^1}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})}{\partial \tilde{\nu}_I} + \cdots
\end{aligned} \tag{5.13}
$$

We can see this dependence on the fluxes explicitly at the perturbative level. For example, for a hypermultiplet, we have

$$
Z_{S^2 \times S^2 \times S^1}^{\text{pert,U(1),hyp}}(\tilde{u})_{\mathfrak{m}_1, \mathfrak{m}_2} = e^{2\pi i \mathfrak{m}_1 \mathfrak{m}_2 \partial_{\tilde{u}}^2 w(\tilde{u})}, \quad w(\tilde{u}) = \frac{1}{(2\pi i)^3} \text{Li}_3(e^{2\pi i \tilde{u}}) + \frac{1}{24} \tilde{u}(\tilde{u}+1)(2\tilde{u}+1) , \tag{5.14}
$$

where $w(\tilde{u})$ is the perturbative contribution to the effective prepotential of a 5d $\mathcal{N} = 1$ hypermultiplet. Then we indeed observe that the dependence on fluxes for the full perturbative contribution is governed by the perturbative prepotential

$$
\mathcal{F}_{\mathbb{R}^4 \times S^1}^{\text{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) = \tilde{\gamma}\tilde{u}^2 + \sum_I \sum_{\rho \in R_I} w(\rho(\tilde{u}) + \tilde{\nu}_I) - \sum_{\alpha \in Ad(G)'} w(\alpha(\tilde{u})) . \tag{5.15}
$$

Namely, one can check that the gauge flux operator in (5.11) is given by

$$
\Pi_a^{\text{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1, \mathfrak{n}_1} = \exp \left\{ 2\pi i \left(\mathfrak{m}_{1,b} \partial_{\tilde{u}_a} \partial_{\tilde{u}_b} \mathcal{F}_{\mathbb{R}^4 \times S^1}^{\text{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) + \mathfrak{n}_{1,I} \partial_{\tilde{\nu}_I} \partial_{\tilde{u}_b} \mathcal{F}_{\mathbb{R}^4 \times S^1}^{\text{pert}}(\tilde{u}, \tilde{\nu}, \tilde{\gamma}) \right) \right\} , \tag{5.16}
$$

and similarly for the flavor flux operator.

We expect that the instanton corrections to the $S^2 \times S^1 \times \Sigma_\mathfrak{g}$ partition function are given by replacing the perturbative expression for the prepotential above by the full, non-perturbative result, as discussed, e.g., in [91]. However, we will not require these instanton corrections when we consider the large N limit below, and leave a detailed discussion of them to future work.

Before moving on to discuss the large N limit, let us conjecture a generalization to the perturbative result above where S^2 is replaced by $\Sigma_{\mathfrak{g}_1}$ (and correspondingly we rename $\Sigma_\mathfrak{g} \to \Sigma_{\mathfrak{g}_2}$). This can be motivated as in section 2.3, by reducing the theory on $\Sigma_{\mathfrak{g}_2}$ and considering the $\Sigma_{\mathfrak{g}_1} \times S^1$ partition function of the resulting theory. This leads us to propose

$$Z^{\text{pert}}_{\Sigma_{\mathfrak{g}_1} \times \Sigma_{\mathfrak{g}_2} \times S^1}(\tilde{u}, \tilde{\nu}, \tilde{\gamma})_{\mathfrak{m}_1,\mathfrak{n}_1;\mathfrak{m}_2,\mathfrak{n}_2} = z^{\text{tr}(\mathfrak{m}_1 \mathfrak{m}_2)} \prod_{\alpha \in Ad(G)'} \left(e^{-s(\alpha(\tilde{u}))}\right)^{[\alpha(\mathfrak{m}_1)+1-\mathfrak{g}_1][\alpha(\mathfrak{m}_2)+1-\mathfrak{g}_2]}$$
$$\times \prod_I \prod_{\rho \in R_I} \left(e^{s(\rho(\tilde{u})+\tilde{\nu})}\right)^{[\rho(\mathfrak{m}_1)+\mathfrak{n}_1][\rho(\mathfrak{m}_2)+\mathfrak{n}_2]} . \quad (5.17)$$

Since the reduction to 3d gave us the correct perturbative contribution to the 1-loop determinant in the genus zero case, it is natural to conjecture the same holds here. However, we do not have a direct localization derivation of this claim, nor do we make any claims about the non-perturbative contribution.

5.2 A large N conjecture and AdS$_6$ black holes

In this final section, we propose a conjecture on the large N behavior of the partition function (5.17), motivated by holography and the entropy of black holes in AdS$_6$.

AdS$_6$ black holes. On general grounds, we expect the gravity dual of a 5d $\mathcal{N} = 1$ SCFT on $M_4 \times S^1$ with a topological twist on M_4 to be given by a supersymmetric solution interpolating between asymptotically locally AdS$_6$ (with an $M_4 \times S^1$ boundary) and an AdS$_2 \times M_4$ geometry for small values of the radial coordinate, i.e., an extremal black hole in locally AdS$_6$. Such a background was considered in 6d $F(4)$ minimal gauged supergravity in [69], with metric of the form

$$ds^2_{\text{BH}} = e^{2f(r)}(-dt^2 + dr^2) + e^{2g(r)}ds^2(M_4), \quad (5.18)$$

and a nonzero flux for the graviphoton through 2-cycles in M_4, which is Kähler. Since the graviphoton is the only nontrivial 1-form potential, and is dual to the R-symmetry of the field theory, it was argued in [21] that such a background should describe the IR behavior of generic 5d $\mathcal{N} = 1$ theories with AdS$_6$ gravity duals, with a universal topological twist on M_4. Taking $M_4 = \Sigma_{\mathfrak{g}_1} \times \Sigma_{\mathfrak{g}_2}$, a product of two negatively curved Riemann surfaces, i.e., $\mathfrak{g}_1 > 1$, $\mathfrak{g}_2 > 1$, the Bekenstein-Hawking entropy of the black hole reported in [69] can be written as[44]

$$S_{\text{BH}} = -\frac{4}{9}(\mathfrak{g}_1 - 1)(\mathfrak{g}_2 - 1) F_{S^5}, \quad (5.19)$$

where F_{S^5} is the free energy on S^5 of the UV 5d SCFT.

---

[44]However, see the comment at the end of this section, which was added in v2 of this paper.

A large N conjecture. In light of the recent success in accounting for the microscopic entropy of AdS_4 black holes from 3d twisted partition functions, initiated in [92] and followed up in [63, 64, 72, 93–96], it is natural to ask whether the entropy of AdS_6 black holes is similarly captured by the 5d partition function, i.e., $S_{BH} = \log Z_{M_4 \times S^1}$, which can be addressed with the results presented here for the case $M_4 = \Sigma_{\mathfrak{g}_1} \times \Sigma_{\mathfrak{g}_2}$.

The basic observation we make is the following. Recall that in section 3, we argued the dominant vacuum contributing to the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function was determined by extremizing the Bethe potential, $\mathcal{W}_{S_b^3 \times \mathbb{R}^2}$. In the present case (and specializing first to $\mathfrak{g}_1 = 0$), we can also attempt to extremize $\mathcal{W}_{S^2 \times S^1 \times \mathbb{R}^2}$, however, in this case we have seen the twisted superpotential and vacuum equations depend on the flux, $\mathfrak{m}_1$, through S^2. Then we must also extremize over this choice of flux. Given the form of (5.16), which implies this flux dependence is controlled by the prepotential, we will conjecture that at large N the dominant vacuum is found by extremizing the prepotential, $\mathcal{F}_{\mathbb{R}^4 \times S^1}$, itself. As we discuss below, this leads to the expected behavior of the partition function at large N. To fully justify this would require a detailed analysis of the matrix model (5.17) and the corresponding vacua arising from performing the double sum over magnetic fluxes $\mathfrak{m}_1, \mathfrak{m}_2$ and the contour integral over the Coulomb branch. This is an interesting problem, which lies beyond the scope of this paper.

Let us evaluate the prepotential in the large N limit, where it is dominated by its perturbative contribution. Then we simply need the expansion

$$w(iz + \tilde{\nu}) \approx -\frac{1}{8} (1 + 2\tilde{\nu}) z|z| + i \left(-\frac{1}{12}|z|^3 + \frac{1}{4} C_{\tilde{\nu}} |z| \right) \qquad \text{for } \text{Re}|z| \gg 1 \,, \qquad (5.20)$$

where we defined $C_{\tilde{\nu}} \equiv \frac{1}{6} + \tilde{\nu}(1 + \tilde{\nu})$.

Let us consider the Seiberg theory (the extension to the orbifold theories is straightforward). The manipulations are very similar to those of section 3 and thus here we are more concise. Adding the contributions to the prepotential (5.15) from the vector, antisymmetric and fundamental hypers (with corresponding real masses, which we denote by $\tilde{\nu}_I = -\Delta_I$), and using the Ansatz $\tilde{u}_i \to iN^\alpha x$ for the eigenvalue distribution we see that $\alpha = \frac{1}{2}$ is required to have a nontrivial extremum, hence an $N^{5/2}$ scaling of the prepotential, and the classical term in (5.15) is subleading. After a convenient rescaling of the coordinate, $x \to \frac{2}{3}\sqrt{\Delta_{AS}(1 - \Delta_{AS})}x$, and corresponding inverse rescaling of the density ρ, we find

$$\mathcal{F}_{\mathbb{R}^4 \times S^1}^{\text{pert}} = \frac{4i}{27\pi} N^{5/2} \left(\Delta_{AS}(1 - \Delta_{AS}) \right)^{3/2} \left[\frac{\pi(8 - N_f)}{3} \int dx \, \rho(x) \, |x|^3 \right.$$
$$\left. - \frac{9\pi}{8} \int dx \, dy \, \rho(x)\rho(y) \left(|x + y| + |x - y| \right) \right], \quad (5.21)$$

where all integrals run from 0 to x^*. Note that to this order in N this depends only on the real mass Δ_{AS} for the antisymmetric field, but not on the one for the fundamental fields and that the quantity inside the brackets is precisely the free energy functional on S^5 for the theory, with saddle point configuration (3.15). An analogous calculation can be done for the (A, B, C) orbifolds of the Seiberg theory, with the same conclusion. Given the similarity of the discussion in section 3 for $S_b^3 \times \Sigma_{\mathfrak{g}}$ this leads us to conjecture, as mentioned

above, that at large N the partition function is dominated by the eigenvalue distribution that extremizes the 5d prepotential.

Next, to determine the value of the partition function on the eigenvalue distribution above, one should determine the dominant contribution in different gauge flux sectors, $\mathfrak{m}_1, \mathfrak{m}_2$. Let us now observe that the zero gauge sector, $\mathfrak{m}_1 = \mathfrak{m}_2 = 0$, reproduces the r.h.s. of (5.19). Indeed, setting $\mathfrak{n}_1 = \mathfrak{n}_2 = 0$ in (5.17) for the universal twist, performing the rescaling of the coordinate mentioned above (5.21), and evaluating the zero gauge flux partition function on the configuration (3.15), we obtain[45]

$$
\begin{aligned}
\log Z^{\mathfrak{m}=0}_{\Sigma_{\mathfrak{g}_1} \times \Sigma_{\mathfrak{g}_2} \times S^1}\Big|_{\partial\mathcal{F}=0} &= \frac{\pi}{3}(\mathfrak{g}_1 - 1)(\mathfrak{g}_2 - 1)\sqrt{\Delta_{AS}(1 - \Delta_{AS})} \\
&\quad \times N^{5/2} \int dxdy\, \rho(x)\rho(y)\,(|x+y| + |x-y|) \\
&= -\frac{8}{9}(\mathfrak{g}_1 - 1)(\mathfrak{g}_2 - 1)\sqrt{\Delta_{AS}(1 - \Delta_{AS})}\, F_{S^5}\,.
\end{aligned}
\tag{5.22}
$$

Then, extremizing this expression with respect to the real mass parameter (or fugacity, Δ_{AS}), as in [92], sets $\Delta_{AS} = \frac{1}{2}$, with value at the extremum

$$
\log Z^{\mathfrak{m}=0}_{\Sigma_{\mathfrak{g}_1} \times \Sigma_{\mathfrak{g}_2} \times S^1}\Big|_{\partial\mathcal{F}=0} = -\frac{4}{9}(\mathfrak{g}_1 - 1)(\mathfrak{g}_2 - 1)\, F_{S^5}\,,
\tag{5.23}
$$

which coincides with the r.h.s. of (5.19).

Comment added in version 2. In light of the new paper [97], noting that the black hole reported in [69] needs to be corrected,[46] and, correspondingly, its entropy (5.19), we discuss how the results above relate to this observation. To summarize, the main points emphasized in v1 (and which are unchanged in v2) are as follows:

a) We have proposed that the eigenvalue configuration dominating the large N behavior of the partition function is obtained by extremizing the 5d prepotential, and showed (after appropriate rescalings) that this coincides with the configuration extremizing the free energy of the same theory on S^5.

b) Evaluating the partition function on this extremum, and keeping only the zero gauge flux sector, reproduces the entropy of the AdS_6 black hole described in [69].

Due to the exact match of (5.23) with (5.19), this suggests that the zero gauge flux sector completely accounts for the entropy of the black hole described in [69], as noted in v1 of this paper on the arXiv. However, as mentioned above it was recently pointed out in [97] that the background of [69] may not be valid, as it assumed the vanishing of the two-form gauge field, $B_{\mu\nu} = 0$, which is inconsistent with the equations of motion. Instead, it was shown

---

[45] Here we used the expansion $s(z) \approx -\pi|z|$ for z real and large, $|z| \gg 1$. Additionally, as in the case of the $S^3_b \times \Sigma_{\mathfrak{g}}$ partition function, we expect the contribution from the Hessian determinant to be subleading at large N, which we assume.

[46] We thank Minwoo Suh for sharing this result with us prior to publication.

that by solving for $B_{\mu\nu}$ from the equations of motion, that the corrected supersymmetric solution has twice the entropy [97]:

$$S'_{\text{BH}} = -\frac{8}{9}(\mathfrak{g}_1 - 1)(\mathfrak{g}_2 - 1)\, F_{S^5}\,. \tag{5.24}$$

Indeed, we note that this new result is consistent with first reducing from 5d to 3d via the universal relation $F_{S^3 \times \Sigma_{\mathfrak{g}_1}} = -\frac{8}{9}(\mathfrak{g}_1 - 1)F_{S^5}$ of section 3 and then reducing from 3d to 1d via the universal relation $F_{\Sigma_{\mathfrak{g}_2} \times S^1} = -(\mathfrak{g}_2 - 1)F_{S^3}$ discussed in [72]. Then, composing these two relations leads to $F_{\Sigma_{\mathfrak{g}_1} \times \Sigma_{\mathfrak{g}_2} \times S^1} = \frac{8}{9}(\mathfrak{g}_1 - 1)(\mathfrak{g}_2 - 1)\, F_{S^5}$, consistent with (5.24) and the expected identification $S_{\text{BH}} = \log Z_{\Sigma_{\mathfrak{g}_1} \times \Sigma_{\mathfrak{g}_2} \times S^1}$.

Assuming this new solution to be correct, the zero gauge flux sector (5.23) then accounts for only half the entropy of the corrected solution. Thus, it seems that to account for the full entropy one must also include contributions from non-zero gauge flux sectors. An interesting proposal on how to do so is discussed in [30] and, indeed, following the procedure described there one obtains the remaining half of the entropy, matching (5.24) and clarifying an initial numerical mismatch. It would be interesting to have a first-principles understanding of this proposal purely in field theory.

6 Outlook

In this work, we have considered five dimensional $\mathcal{N} = 1$ gauge theories and computed their exact partition function on various manifolds with partial topological twists using localization. As discussed, this is a fruitful vantage point from which we can also access the physics of field theories of various dimensions, from 1d up to 6d, arising at either the IR or UV end of RG flow. There are a number of possible directions for future work.

As discussed in section 2, one can consider partition functions on $\mathcal{M}_3 \times \Sigma_{\mathfrak{g}}$ with more general $\mathcal{M}_3$. Our formalism already suggests the result for $\mathcal{M}_3$ a lens space, and we discussed the example of $S^2 \times S^1$ in section 5, but in principle one can consider arbitrary Seifert manifolds [13]. It would also be interesting to better understand the relation of our computation to partition functions on more general five-manifolds considered in the literature, such as S^5, $\mathbb{CP}^2 \times S^1$, $S^4 \times S^1$, and $Y^{p,q}$, as well as studying new geometries. For example, $Y^{p,q}$ is topologically an S^1 fibration over $S^2 \times S^2$, and may be related to our results by introducing a suitable "fibering operator," as in [11–13].

For a given $\mathcal{M}_3$, the partition function of the 3d theory obtained by reduction on $\Sigma_{\mathfrak{g}}$ is computed by an appropriate 2d TQFT, in what may be called a "3d-2d correspondence." We described these 3d theories as a direct sum of ordinary 3d theories in section 2.3, however the full, non-perturbative computation suggests they are something more exotic. The results of section 3, in particular (3.34), suggest that a large class of novel 3d $\mathcal{N} = 2$ SCFTs exist, arising as the IR fixed point of the orbifolded Seiberg theories in 5d, compactified on a Riemann surface. These are labeled by the discrete flavor fluxes $\mathfrak{n}$ and their free energy scales as $N^{5/2}$. It would be worth investigating whether some of these theories can be understood purely in terms of simpler building block three-dimensional theories, analogous to T_N theories for four-dimensional class $\mathcal{S}$.

The results of section 4 provide a new method for computing the superconformal indices of non-Lagrangian theories in 4d, by Lagrangian methods in 5d. Our results coincide with those which have been previously computed and in principle provide a powerful tool to compute superconformal indices of these theories, for which very few methods are currently available. In practice, the feasibility of this calculation will depend on whether the instanton contribution is explicitly computable. Conversely, in cases where the 4d theory is Lagrangian, the exact 4d index is computable, and so gives a prediction for the $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function, which we may use to help characterize the instanton contributions. We may similarly compute the generalized indices of these 4d theories [12, 98] by studying the $\mathcal{M}_3 \times \Sigma_{\mathfrak{g}}$ partition function for more general $\mathcal{M}_3$. Another obvious generalization is to include punctures on the Riemann surface, which correspond to insertions of local defect operators in the TQFT, and give access to more general 4d theories. Finally, this computation gives a new entry in the gauge-Bethe correspondence dictionary [19], and it would be very interesting to understand in more detail and generality the relation between 4d compactifications of 6d theories and integrable systems, as discussed in some examples in section 4.

The results of section 5 and their relation to entropy of black holes in AdS_6 deserve a better understanding. On the field theory side, one should rigorously extend the exact computation of the $S^2 \times \Sigma_{\mathfrak{g}} \times S^1$ partition function to the case $\Sigma_{\mathfrak{g}_1} \times \Sigma_{\mathfrak{g}_2} \times S^1$, including instanton corrections. It would also be interesting to have a first-principles understanding on the proposed conjecture that the dominant eigenvalue distribution at large N is given by that extremizing the 5d prepotential. On the supergravity side, it would be interesting to construct extremal black hole solutions with an $\text{AdS}_2 \times \Sigma_{\mathfrak{g}_1} \times \Sigma_{\mathfrak{g}_2}$ near-horizon geometry, for generic $\mathfrak{g}_{1,2}$ and flavor fluxes. In particular, nonzero flavor fluxes may allow for solutions with $\mathfrak{g}_2 = 0$, in which case we have provided an expression for the full non-perturbative partition function, determined implicitly by studying the Nekrasov-Shatashvili limit of the twisted superpotential. This could allow for precision tests of the ideas presented here, including subleading corrections to the black hole entropy.

We plan to return to some of these questions in future work.

Acknowledgments

A Casimir energy on $S_b^3 \times \Sigma_{\mathfrak{g}} \times S_\beta^1$

Here we provide some general formulas for the Casimir energy for the 4d theories obtained by twisted compactification of 6d $\mathcal{N} = (1,0)$ theories. The expression is obtained by twisting the 6d anomaly polynomial, I_8, integrating over the Riemann surface to obtain the anomaly polynomial of the 4d theory, and then using the results of [90, 99], i.e.,

$$E_{\text{Casimir}} = \int_\varepsilon \int_{\Sigma_{\mathfrak{g}}} I_8^{\text{twisted}} , \tag{A.1}$$

where $\int_{\Sigma_{\mathfrak{g}}}$ is a regular integral over the Riemann surface, $\int_\varepsilon$ is an equivariant integral. We give an explicit expression below (namely (A.8)) for a rather general 6d $\mathcal{N} = (1,0)$ theories and consider in detail the examples of the 6d $\mathcal{N} = (2,0)$ theory, its orbifolds, and the E-string theory. As discussed in section 4.2.1 this quantity matches with the one extracted from the corresponding $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function.

We begin by assuming the untwisted 6d anomaly polynomial is of the form[47]

$$I_8 = \frac{1}{2} k^{AABB} C_2(A) C_2(B) + k^{AA} C_2(A) p_1(TM) + k_1 p_1(TM)^2 + k_2 p_2(TM) . \tag{A.2}$$

Here A, B run over all global symmetries of the theory (both R-symmetry and flavor symmetries), C_2 are their corresponding second Chern classes, and $k^{AABB} = k^{BBAA}$ the anomaly coefficients. $p_{1,2}(TM)$ are Pontryagin classes for the tangent bundle TM, and $k_{1,2}$ are the corresponding gravitational anomaly coefficients.

Performing a topological twist on $\Sigma_{\mathfrak{g}}$ by $U(1)_R \subset SU(2)_R$ and a generic $U(1)_F$ subgroup of the flavor symmetry amounts to the replacements

$$C_2(R) \rightarrow -\left(C_1(R) - \frac{\kappa}{2} t_{\mathfrak{g}} \right)^2 ,$$
$$C_2(F) \rightarrow -\left(\frac{\hat{\mathfrak{n}}_F}{2} t_{\mathfrak{g}} + \epsilon_F C_1(R) + C_1(F) \right)^2 + \sum_j C_2(G_j) , \tag{A.3}$$

where the sum over j is over all simple commutants of $U(1)_F$ in the full flavor group, and ϵ_F controls the amount of mixing of the R-symmetry with the flavor symmetry along the flow to the IR. We now implement the twist (A.3) in I_8 and integrate over the Riemann surface, normalized as $\int_{\Sigma_{\mathfrak{g}}} t_{\mathfrak{g}} = \frac{1}{2\pi} \text{Vol}(\Sigma_{\mathfrak{g}}) = \eta_\Sigma$, with $\eta_\Sigma = 2|\mathfrak{g} - 1|$ for $\mathfrak{g} \neq 1$ and $\eta_\Sigma = 1$ for $\mathfrak{g} = 1$, to obtain a six-form anomaly polynomial. Comparing this to the anomaly polynomial for a 4d theory,

$$I_6 = \frac{k_{RRR}}{6} C_1(R)^2 - \frac{k_R}{24} C_1(R) p_1(TM) + \cdots , \tag{A.4}$$

where ellipses denote contributions involving the flavor symmetries, we read off the 4d R-symmetry anomaly coefficients:

$$k_{RRR} = 12(\mathfrak{g} - 1) k^{AABB} \hat{\mathfrak{n}}_A \epsilon_A \epsilon_B^2 , \qquad k_R = 48(\mathfrak{g} - 1) k^{AA} \hat{\mathfrak{n}}_A \epsilon_A , \tag{A.5}$$

where we have defined $\epsilon_R = \hat{\mathfrak{n}}_R = 1$ to write the expressions in a compact form.

---

[47]This is a rather generic 6d anomaly polynomial which contains all the cases we consider below. If any of the flavor symmetries is Abelian, as will be the case for $\mathcal{S}_k$ theories, one may use $C_2(F) = -C_1(F)^2$.

We may now use these to compute the Casimir energy of the 4d theory on S_b^3. Recall this is generally given in terms of the R-symmetry anomaly coefficients by [90, 99, 100]

$$E_{\text{Casimir}} = \frac{1}{24} k_R \left(\omega_1 + \omega_2\right) + \frac{1}{48}(k_{RRR} - k_R) \frac{(\omega_1 + \omega_2)^3}{\omega_1 \omega_2}, \tag{A.6}$$

where $\omega_{1,2}$ are the squashing parameters of the S_b^3, which in our conventions are given by

$$\omega_1 = \frac{2\pi b}{Q}, \qquad \omega_2 = \frac{2\pi b^{-1}}{Q}; \qquad Q = \frac{1}{2}(b + b^{-1}). \tag{A.7}$$

Plugging the anomaly coefficients (A.5) into (A.6) then gives the desired expression

$$E_{\text{Casimir}} = 8\pi(\mathfrak{g} - 1) k^{AA} \hat{\mathfrak{n}}_A \epsilon_A + 16\pi(\mathfrak{g} - 1) \left(\frac{1}{4} k^{AABB} \hat{\mathfrak{n}}_A \epsilon_A \epsilon_B^2 - k^{AA} \hat{\mathfrak{n}}_A \epsilon_A\right) Q^2. \tag{A.8}$$

This represents the full Casimir energy for the theories at hand. It can be viewed as the Casimir energy of the 6d $\mathcal{N} = (1,0)$ theory, quantized on $S_b^3 \times \Sigma_{\mathfrak{g}}$ or as the Casimir energy of the 4d theory obtained by reduction on $\Sigma_{\mathfrak{g}}$, quantized on S_b^3.

As we discuss in the main text, this should coincide exactly with the expression derived from the non-perturbative partition function on $S_b^3 \times \Sigma_{\mathfrak{g}}$. However, since in practice we can often only evaluate the perturbative part explicitly, it is convenient to decompose this as

$$E_{\text{Casimir}} = E^{\text{pert}} + E^{\text{non-pert}}, \tag{A.9}$$

where E^{pert} is, by definition, the quantity extracted from the perturbative $S_b^3 \times \Sigma_{\mathfrak{g}}$ partition function and $E^{\text{non-pert}}$ is the remaining part. As we shall show below, in all the examples we study the remaining part is due to the contribution to the 6d anomaly polynomial from free multiplets. That is, decomposing $I_8 = I_{8,\text{int}} + I_{8,\text{free}}$ where $I_{8,\text{free}}$ is the contribution from all the free multiplets in the tensor branch of the theory and $I_{8,\text{int}}$ is the remaining part, we find

$$E^{\text{pert}} = \int_\varepsilon \int_{\Sigma_{\mathfrak{g}}} I_{8,\text{int}}^{\text{twisted}}, \qquad E^{\text{non-pert}} = \int_\varepsilon \int_{\Sigma_{\mathfrak{g}}} I_{8,\text{free}}^{\text{twisted}}. \tag{A.10}$$

The analogous observation for the case of S^5 was already made in [90].[48] The new examples we provide suggest there should be a deeper understanding of this fact, which we do not address here.

A.1 5d maximal theory and class $\mathcal{S}$

The simplest example is 5d $\mathcal{N} = 2$ SYM theory, consisting of an $\mathcal{N} = 1$ vector and a hypermultiplet in the adjoint representation of the gauge group G. This theory is believed [101] to correspond to the 6d $\mathcal{N} = (2,0)$ theory on a circle S_β^1, with radius

$$\beta = \frac{g_5^2}{2\pi}, \tag{A.11}$$

---

[48]To avoid possible confusions, we emphasize that the computations in [90] lead to the Casimir energy of the 6d theory on $S^1 \times S^5$ instead of the 4d Casimir energy computed here.

with g_5 the gauge coupling constant of the 5d gauge theory. The anomaly polynomial of the 6d $(2,0)$ theory is given by

$$I_8 = r_G A_8(1) + d_G h_G \frac{p_2(NM)}{24} \,, \tag{A.12}$$

where $A_8(1)$ is the anomaly polynomial of one free tensor multiplet,

$$A_8(1) = \frac{1}{48}\left(p_2(NM) - p_2(TM) + \frac{1}{4}\left(p_1(NM) - p_1(TM)\right)^2\right) \,, \tag{A.13}$$

and d_G, r_G, and h_G are the dimension, rank, and dual Coxeter number of the group G, respectively, and TM and NM refer to the tangent and normal $SO(5)_R$ bundles, respectively. To evaluate the full Casimir energy we first write (A.12) in $(1,0)$ language we take $SO(5)_R \supset SU(2)_R \times SU(2)_L$ and use the relations

$$p_1(NM) = -2(C_2(L) + C_2(R)) \,, \quad p_2(NM) = (C_2(L) - C_2(R))^2 \,. \tag{A.14}$$

Then, comparing the anomaly polynomial to (A.2) we read off the corresponding anomaly coefficients, and using these in (A.8) gives

$$\begin{aligned}
E_{\text{Casimir}} = {}& \frac{\pi}{6}(\mathfrak{g}-1)(N-1)(1+\hat{\mathfrak{n}}\epsilon) \\
&+ \frac{\pi}{3}(\mathfrak{g}-1)\left(\epsilon^2 - 1\right)(N-1)[N(N+1)(\hat{\mathfrak{n}}\epsilon - 1) + \hat{\mathfrak{n}}\epsilon]Q^2 \,,
\end{aligned} \tag{A.15}$$

where we have specified $G = SU(N)$ and used $d_G = N^2 - 1$, $r_G = N-1$, $h_G = N$. We wish to compare this to the quantity extracted from the 5d localization computation. With this in mind, let us split the quantity above into the two pieces:

$$E^{\text{pert}} = \frac{\pi}{3}(\mathfrak{g}-1)N(N^2-1)\left(\epsilon^2-1\right)(\hat{\mathfrak{n}}\epsilon - 1)Q^2 \,, \tag{A.16}$$

$$E^{\text{non-pert}} = \frac{\pi}{6}(\mathfrak{g}-1)(N-1)\left(1 + \hat{\mathfrak{n}}\epsilon + 2\hat{\mathfrak{n}}\epsilon\left(\epsilon^2 - 1\right)Q^2\right) \,. \tag{A.17}$$

As discussed in the main text one can easily check that the perturbative Casimir energy matches the contribution from the last term in (A.12), while the remaining non-perturbative part is entirely due to the free part,

$$E^{\text{non-pert}} = (N-1)\int_\varepsilon \int_{\Sigma_{\mathfrak{g}}} A_8^{\text{twisted}}(1) \,. \tag{A.18}$$

In general, we expect instantons in the 5d computation to contribute $r_G \int_\varepsilon \int_{\Sigma_{\mathfrak{g}}} A_8^{\text{twisted}}(1)$.

Mixed Schur limit. As we discuss in the main text, a special value for the fugacity is given by the mixed Schur limit:

$$\epsilon = \frac{b - b^{-1}}{b + b^{-1}} \,. \tag{A.19}$$

In this limit, the perturbative and non-perturbative contributions to the Casimir energy (A.16) and (A.17) simplify to

$$\begin{aligned}
E^{\text{pert}} &= \frac{\pi}{Q}\frac{N(N^2-1)}{6}\left((\mathfrak{g}-1-\mathfrak{n})b + (\mathfrak{g}-1+\mathfrak{n})b^{-1}\right) \,, \\
E^{\text{non-pert}} &= \frac{\pi}{Q}\frac{(N-1)}{12}\left((\mathfrak{g}-1-\mathfrak{n})b + (\mathfrak{g}-1+\mathfrak{n})b^{-1}\right) \,,
\end{aligned} \tag{A.20}$$

where we have used the definition $\mathfrak{n} = \hat{\mathfrak{n}}(\mathfrak{g} - 1)$. Note the values $\mathfrak{n} = \pm(\mathfrak{g} - 1)$, for which 4d supersymmetry is enhanced to $\mathcal{N} = 2$, are special. In this case, as we discuss below, the full Casimir energy is proportional to the central charge of the associated 2d chiral algebra, in the sense of [102].

Chiral algebra (or Schur) limit. We consider for concreteness $\mathfrak{n} = \mathfrak{g} - 1$. Recall for a generic 4d $\mathcal{N} = 2$ theory the Casimir energy is given by (see (4.36) in [90]):

$$E_{\text{Casimir}} = \frac{1}{2}(c_{4d} - 2a_{4d})\frac{\sigma(\sigma + \omega_1 + \omega_2)^2}{\omega_1 \omega_2} + (c_{4d} - a_{4d})\frac{\sigma(\sigma^2 - \omega_1^2 - \omega_2^2)}{\omega_1 \omega_2}, \qquad \text{(A.21)}$$

where $\sigma \equiv \gamma - \omega_1 - \omega_2$, with γ, σ fugacities for $\text{SU}(2)_R \times \text{U}(1)_r$, respectively, and we have set all flavor fugacities to zero. The "chiral algebra limit" corresponds to setting $\gamma = \omega_2$ and hence $\sigma = -\omega_1$, which gives the simple expression

$$E_{\text{Casimir}} = \frac{\omega_2}{2}c_{4d} = -\frac{\omega_2}{24}c_{2d}, \qquad \text{(A.22)}$$

where in the second equality we used the chiral algebra relation $c_{2d} \equiv -12c_{4d}$ [102]. Thus, in this limit the Casimir energy of the 4d $\mathcal{N} = 2$ theory is proportional to the central charge of its corresponding 2d chiral algebra.[49] To write this explicitly for the maximal theory on $\Sigma_{\mathfrak{g}}$ with no punctures, recall[50]

$$\begin{aligned}
c_{4d} &= \frac{1}{6}(\mathfrak{g} - 1)r_G(1 + 2h_G(1 + h_G)), \\
a_{4d} &= \frac{1}{24}(\mathfrak{g} - 1)r_G(5 + 8h_G(1 + h_G)),
\end{aligned} \qquad \text{(A.23)}$$

and thus

$$E^{\text{pert}} = \frac{\pi}{3Q}(\mathfrak{g} - 1)r_G h_G(1 + h_G)b^{-1}, \qquad E^{\text{non-pert}} = \frac{\pi}{6Q}(\mathfrak{g} - 1)r_G b^{-1}. \qquad \text{(A.24)}$$

For $G = \text{SU}(N)$ these are a special case of (A.20) with $\mathfrak{n} = \mathfrak{g} - 1$. For $\mathfrak{n} = 1 - \mathfrak{g}$ the roles of b and b^{-1} are exchanged.

A.2 Class $\mathcal{S}_k$ theories

These are a class of 4d $\mathcal{N} = 1$ theories which arise from the twisted compactification of $\mathbb{Z}_k$ orbifolds of the maximal 6d A_{N-1} theory [26]. For general values of k and N the theory has an $\text{SU}(k)_b \times \text{SU}(k)_c \times u(1)_s$ flavor symmetry. Before twisting the anomaly polynomial

---

[49]We note a similar relation was found in eq. (3.21) in [90] for the Casimir energy of the 6d theory on the squashed S^5.

[50]See, e.g., [25], where their parameter z is identified with $\hat{\mathfrak{n}}$ here.

of the 6d theory is given by[51]

$$
\begin{aligned}
I_8 = {} & \frac{k^2 N^3}{24} C_2(R)^2 - \frac{N(k^2-1)}{48} C_2(R) \left(4C_2(R) + p_1(TM)\right) \\
& + N \left(\frac{C_2(R)p_1(TM)}{48} - \frac{p_2(TM)}{48} + \frac{p_1(TM)^2}{192} \right) \\
& + (k^2-1) \left(\frac{1}{24} C_2(R)^2 + \frac{1}{48} C_2(R)p_1(TM) + \frac{7p_1(TM)^2 - 4p_2(TM)}{5760} \right) \\
& - \left(\frac{1}{24} C_2(R)^2 + \frac{1}{48} C_2(R)p_1(TM) + \frac{23}{5760} p_1(TM)^2 - \frac{29}{1440} p_2(TM) \right) \\
& - \frac{Nk^2}{48} p_1(TM)C_1(s)^2 + \frac{k^2 N(N^2-1)}{12} C_2(R)C_1(s)^2 + \frac{k^2 N^3}{24} C_1(s)^4 \\
& - k^2 I_{(1,0)\text{hyper}} + I_8(F_{b,c}) \,,
\end{aligned}
\tag{A.25}
$$

where $I_8(F_{b,c})$ denotes contributions involving at least one of the $\mathrm{SU}(k)_b$ or $\mathrm{SU}(k)_c$ symmetries. Since we do not turn background fields for these symmetries, we may ignore these terms in what follows. To identify the parts encoding the perturbative and non-perturbative terms in the Casimir energy it is convenient to recall the anomaly polynomial for a $(1,0)$ tensor multiplet (see, e.g., eq. (2.3) in [26]):

$$
I_{(1,0)\text{ tensor}} = \frac{1}{24} C_2(R)^2 + \frac{1}{48} C_2(R)p_1(TM) + \frac{23}{5760} p_1(TM)^2 - \frac{29}{1440} p_2(TM) \,, \tag{A.26}
$$

and for a free half-hypermultiplet charged under $\mathrm{U}(1)_s$ as[52]

$$
I_{(1,0)\text{hyper}} = \frac{1}{24} C_1(s)^4 - \frac{1}{48} C_1(s)^2 p_1(TM) + \frac{7p_1(TM)^2 - 4p_2(TM)}{5760} \,, \tag{A.27}
$$

and finally

$$
I_{(1,0)\text{vector}} = -\frac{1}{24} C_2(R)^2 - \frac{1}{48} C_2(R)p_1(TM) - \frac{7p_1(TM)^2 - 4p_2(TM)}{5760} \,. \tag{A.28}
$$

In terms of these combinations (A.25) takes the following form

$$
\begin{aligned}
I_8 = {} & (N-1) \left[I_{(1,0)\text{tensor}} + (k^2-1)I_{(1,0)\text{vector}} + k^2 I_{(1,0)\text{hyper}} \right] \\
& + k^2 N(N^2-1) \frac{\left(C_2(R) + C_1(s)^2\right)^2}{24} \,.
\end{aligned}
\tag{A.29}
$$

We note this has a very similar form to (A.12).[53] Indeed, from the 5d localization calculation we find the perturbative part of the Casimir energy arises entirely from the terms in the second line of (A.29), and is given by

$$
E^{\text{pert}} = \frac{\pi k^2}{3} (\mathfrak{g}-1) \left(\epsilon^2-1\right) [N(N^2-1)(\hat{\mathfrak{n}}\epsilon-1)]Q^2 \,, \tag{A.30}
$$

---

[51]Note that compared to eq. (2.1) in [26] we have a sign difference in the $C_2(R)C_1(s)^2$ term. Also, we have subtracted the contribution from k^2 hypermultiplets so that for $k=1$ this matches the expression used in (A.12) for the maximal theory.

[52]Here we have taken the anomaly polynomial for a free half-hypermultiplet in the doublet of $\mathrm{SU}(2)_L$ (see, e.g., (1.5) in [103]) and set $C_2(L) = -C_1(s)^2$.

[53]Indeed, in the special case $k=1$, $\mathrm{U}(1)_s$ is enhanced to $\mathrm{SU}(2)_L$ with $C_1(s)^2 = -C_2(L)$ and the contribution from the $(1,0)$ tensor combines with that of the hyper into that of a $(2,0)$ tensor to form the combination $A_8(1) = I_{(1,0)\text{tensor}} + I_{(1,0)\text{hyper}}$.

while the remaining part,

$$E^{\text{non-pert}} = \frac{\pi}{6}(\mathfrak{g} - 1)(N - 1)\left(2 - k^2 + k^2\hat{\mathfrak{n}}\epsilon + 2k^2\hat{\mathfrak{n}}\epsilon\left(\epsilon^2 - 1\right)Q^2\right), \tag{A.31}$$

arises from

$$E^{\text{non-pert}} = (N - 1)\int_\varepsilon\int_{\Sigma_\mathfrak{g}}\left(I^{\text{twisted}}_{(1,0)\text{tensor}} + (k^2 - 1)I^{\text{twisted}}_{(1,0)\text{vector}} + k^2 I^{\text{twisted}}_{(1,0)\text{hyper}}\right). \tag{A.32}$$

We thus see once again that the instanton correction to the Casimir energy is encoded in the 6d anomaly polynomial of free fields.

A.3   E-string theory

This is the 5d Seiberg theory in the special case $N_f = 8$. The twisted compactification of this theory on a Riemann surface was considered in [27]. The anomaly polynomial before twisting reads [103][54]

$$I_E = \frac{N^3}{6}p_2(NM) + \frac{N^2}{2}e(NM)A_4 + N\left(A_4^2 - \frac{p_2(NM)}{24}\right) + (N - 1)A_8(1),$$

$$I_{(1,0)\text{hyper}} = \frac{1}{24}C_2(L)^2 + \frac{1}{48}C_2(L)p_1(TM) + \frac{7p_1(TM)^2 - 4p_2(TM)}{5760},$$

$$A_4 = \frac{1}{4}\left(p_1(NM) + p_1(TM) + \frac{1}{15}C_2(E_8)\right). \tag{A.33}$$

To write the anomaly polynomial in the form (A.2) we use the relations (A.14) and $e(NM) = C_2(L) - C_2(R)$, which gives

$$\begin{aligned}
I_E = {}&\frac{N(4N^2 + 6N + 3)}{24}C_2^2(R) + \frac{(N - 1)(4N^2 - 2N + 1)}{24}C_2^2(L)\\
&- \frac{N(N^2 - 1)}{3}C_2(R)C_2(L) + \frac{(N - 1)(6N + 1)}{48}C_2(L)p_1(TM)\\
&- \frac{N(6N + 5)}{48}C_2(R)p_1(TM) + \frac{N(N - 1)}{120}C_2(L)C_2(E_8)_{\mathbf{248}}\\
&- \frac{N(N + 1)}{120}C_2(R)C_2(E_8)_{\mathbf{248}} + \frac{N}{240}p_1(TM)C_2(E_8)_{\mathbf{248}} + \frac{N}{7200}C_2^2(E_8)_{\mathbf{248}}\\
&+ (30N - 1)\frac{7p_1(TM)^2 - 4p_2(TM)}{5760} - I_{(1,0)\text{tensor}}.
\end{aligned} \tag{A.34}$$

This anomaly polynomial coincides with the one in [27], apart from the contribution $I_{(1,0)\text{tensor}}$ from a free tensor which we have included. We now perform the twisted reduction on the Riemann surface following this reference. In addition to the R-symmetry flux there are nine possible Abelian flavor fluxes; eight along the Cartan of E_8 and one for the Cartan of $\text{SU}(2)_L$. For simplicity we consider only two flavor fluxes, picking a generic

---

[54]Here we follow the conventions in [90] and we have included an additional contribution from a $(1,0)$ tensor in the $\mathcal{O}(1)$ part to complete $A_8(1)$. This is done so that for $k = 1$ this anomaly polynomial matches (A.12) for the maximal theory.

U(1)$_F \subset E_8$ and one along U(1)$_L \subset$ SU(2)$_L$. Twisting amounts to

$$C_2(R) \rightarrow - \left(C_1(R) - \frac{\kappa}{2} t_{\mathfrak{g}} \right)^2,$$

$$C_2(E_8)_{\mathbf{248}} \rightarrow -\xi' \left(\frac{\tilde{\mathfrak{n}}_E}{2} t_{\mathfrak{g}} + \tilde{\epsilon}_E C_1(R) + C_1(F) \right)^2 + d_{E_8} \sum_j C_2(G_j), \quad \text{(A.35)}$$

$$C_2(L) \rightarrow - \left(\frac{\hat{\mathfrak{n}}}{2} t_{\mathfrak{g}} + \epsilon C_1(R) + C_1(L) \right)^2,$$

where $\xi' \equiv 2\xi d_{E_8}$, with $d_{E_8} = 30$ the Dynkin index for the fundamental of E_8 and ξ is numerical constant that depends on the choice of U(1)$_F$ inside E_8. The sum over j is over all simple commutants of U(1)$_F$ in E_8. See [27] for details.

To write things in a compact form, let us write the anomaly polynomial in the form

$$\begin{aligned}
I &= \frac{k_{RRRR}}{2} C_2^2(R) + \frac{k_{LLLL}}{2} C_2^2(L) + k_{RRLL} C_2(R)C_2(L) \\
&\quad + k_{LL} C_2(L)p_1(TM) + k_{RR} C_2(R)p_1(TM) + \frac{k_{LLEE}}{\xi'} C_2(L)C_2(E_8)_{\mathbf{248}} \\
&\quad + \frac{k_{RREE}}{\xi'} C_2(R)C_2(E_8)_{\mathbf{248}} + \frac{k_{EE}}{\xi'} p_1(TM)C_2(E_8)_{\mathbf{248}} + \frac{k_{EEEE}}{2\xi'^2} C_2^2(E_8)_{\mathbf{248}} \\
&\quad + k_1 p_1(TM)^2 + k_2 p_2(TM).
\end{aligned} \quad \text{(A.36)}$$

The coefficients for the E-string theory are easily read off by comparing this expression to (A.34). We subtract the contribution of the free fields from the full expression (continuing the pattern from above) to get

$$\begin{aligned}
E^{\text{pert}} = \pi(\mathfrak{g}-1)\Bigg[&N(N-1)Q^2(1-\epsilon^2)\left(\frac{2}{3}(2N-1)(1-\hat{\mathfrak{n}}\epsilon) + (2-\hat{\mathfrak{n}}_E\epsilon_E)\right) \\
&+ N\left(Q^2(4-\epsilon_E^2)-1\right)\left((N-1)(1-\hat{\mathfrak{n}}\epsilon) + (2-\hat{\mathfrak{n}}_E\epsilon_E)\right) \Bigg], \quad \text{(A.37)}
\end{aligned}$$

where we defined $\tilde{\mathfrak{n}}_E = \frac{1}{\sqrt{2\xi}}\hat{\mathfrak{n}}_E$, $\tilde{\epsilon}_E = \frac{1}{\sqrt{2\xi}}\epsilon_E$. This exactly matches the perturbative result derived in the main text. The remaining piece,

$$E^{\text{non-pert}} = \frac{\pi}{6}(\mathfrak{g}-1)(N-1)\left(1 + \hat{\mathfrak{n}}\epsilon + 2\hat{\mathfrak{n}}\epsilon(\epsilon^2-1)Q^2\right), \quad \text{(A.38)}$$

is given by the free fields

$$E^{\text{non-pert}} = (N-1) \int_\varepsilon \int_{\Sigma_{\mathfrak{g}}} \left(I_{(1,0)\text{tensor}}^{\text{twisted}} + I_{(1,0)\text{hyper}}^{\text{twisted}} \right), \quad \text{(A.39)}$$

and must be generated entirely due to instanton corrections in the 5d partition function.

We have thus shown by explicit computation of the perturbative piece that the instanton corrections to the 4d Casimir energy arise from the free anomaly polynomial of the 6d theory in all the examples we have considered (for class $\mathcal{S}$, class $\mathcal{S}_k$, and the compactification of the E-string theory). It would be interesting to have a better understanding of whether this is always the case and, if so, the reason behind this.

B Instanton partition function for $\mathcal{N} = 2$ U(N) and SU(N) SYM

Here we consider the instanton partition function for the maximal SYM theory with gauge group U(N) or SU(N). This consists of an adjoint vector multiplet (as always) and an adjoint hypermultiplet, which we assign a mass $\tilde{\nu}$. Then, taking into account the extra factors needed to pass between the Dolbeault, self-dual, and Dirac complexes, we may write the total equivariant index for the fields of this theory as, first in the 4d case,

$$
\text{ind}_{4d\,\mathcal{N}=4}(\tilde{u}, \tilde{\nu}) = \left(\frac{1 + e^{i(\epsilon_1 + \epsilon_2)}}{2} - e^{i\frac{\epsilon_1 + \epsilon_2}{2}} \frac{e^{i\tilde{\nu}} + e^{-i\tilde{\nu}}}{2} \right) \left(\frac{\text{tr}_{\text{adj}}(e^{i\tilde{u}})}{(e^{i\epsilon_1} - 1)(e^{i\epsilon_2} - 1)} \right.
$$
$$
\left. - e^{-i\frac{\epsilon_1 + \epsilon_2}{2}} \left(\text{tr}_N(e^{i\tilde{u}}) \, \text{tr}_{\bar{k}}(e^{i\phi}) + c.c. \right) + (1 - e^{-i\epsilon_1})(1 - e^{-i\epsilon_2}) \, \text{tr}_{\text{adj}}(e^{i\phi}) \right),
$$
(B.1)

where ϵ_i are the equivariant parameters.

We are interested here in the 5d case, and for this we need to incorporate the KK modes, n, as in section 2.4. One additional subtlety is that, for the Dirac complex determining the hypermultiplets, the quantization of the KK momenta are shifted by $n \to n + \frac{1}{2}$ [34]. Then we find[55]

$$
\text{ind}_{5d\,\mathcal{N}=2}(\tilde{u}, \tilde{\nu}) = \sum_{n \in \mathbb{Z}} e^{\frac{in}{r}} \left(\frac{1 + e^{i(\epsilon_1 + \epsilon_2)}}{2} - e^{\frac{i}{2r}} e^{i\frac{\epsilon_1 + \epsilon_2}{2}} \frac{e^{i\tilde{\nu}} + e^{-i\tilde{\nu}}}{2} \right) \left(\frac{\text{tr}_{\text{adj}}(e^{i\tilde{u}})}{(e^{i\epsilon_1} - 1)(e^{i\epsilon_2} - 1)} \right.
$$
$$
\left. - e^{-i\frac{\epsilon_1 + \epsilon_2}{2}} \left(\text{tr}_N(e^{i\tilde{u}}) \, \text{tr}_{\bar{k}}(e^{i\phi}) + c.c. \right) + (1 - e^{-i\epsilon_1})(1 - e^{-i\epsilon_2}) \, \text{tr}_{\text{adj}}(e^{i\phi}) \right).
$$
(B.2)

where we write the equivariant parameters as $\mathfrak{q}_i = e^{2\pi i \epsilon_i}$, as in section 2.4.

Let us consider in more detail the instanton contribution, which comes from the second and third terms inside the second parantheses in (B.2). Using the rule (2.78) to pass from the equivariant index to the 1-loop determinant, and integrating over the equivariant parameters for the U(k) instanton symmetry, we may write the contribution to the partition function from the k instanton sector as

$$
Z_k = \frac{1}{k!} \left(\frac{\sin\left(\pi r(\tilde{\nu} + \frac{1}{2r} + \epsilon_-)\right) \sin\left(\pi r(\tilde{\nu} + \frac{1}{2r} - \epsilon_-)\right)}{\sin\left(\pi r(\tilde{\nu} + \frac{1}{2r} + \epsilon_+)\right) \sin\left(\pi r(\tilde{\nu} + \frac{1}{2r} - \epsilon_+)\right)} \frac{\sin\left(2\pi r \epsilon_+\right)}{\sin\left(\pi r \epsilon_1\right) \sin\left(\pi r \epsilon_2\right)} \right)^k
$$
$$
\times \int d^k\phi \prod_{i=1}^{N} \prod_{\alpha=1}^{k} \frac{\sin\left(\pi r(\tilde{u}_i - \phi_\alpha + \tilde{\nu} + \frac{1}{2r})\right) \sin\left(\pi r(\tilde{u}_i - \phi_\alpha - \tilde{\nu} - \frac{1}{2r})\right)}{\sin\left(\pi r(\tilde{u}_i - \phi_\alpha + \epsilon_+)\right) \sin\left(\pi r(\tilde{u}_i - \phi_\alpha - \epsilon_+)\right)}
$$
$$
\times \prod_{\alpha < \beta} \frac{\sin\left(\pi r(\phi_\alpha - \phi_\beta + \tilde{\nu} + \frac{1}{2r} + \epsilon_-)\right) \sin\left(\pi r(\phi_\alpha - \phi_\beta + \tilde{\nu} + \frac{1}{2r} - \epsilon_-)\right)}{\sin\left(\pi r(\phi_\alpha - \phi_\beta + \tilde{\nu} + \frac{1}{2r} + \epsilon_+)\right) \sin\left(\pi r(\phi_\alpha - \phi_\beta + \tilde{\nu} + \frac{1}{2r} - \epsilon_+)\right)}
$$
$$
\times \frac{\sin\left(\pi r(\phi_\alpha - \phi_\beta + 2\epsilon_+)\right) \sin\left(\pi r(\phi_\alpha - \phi_\beta)\right)}{\sin\left(\pi r(\phi_\alpha - \phi_\beta + \epsilon_1)\right) \sin\left(\pi r(\phi_\alpha - \phi_\beta + \epsilon_2)\right)},
$$
(B.3)

where we have defined $\epsilon_\pm = \frac{1}{2}(\epsilon_1 \pm \epsilon_2)$, and the prefactor is the contribution from the Cartan components of the adjoint representation of U(k). This is integrated over a suitable

---

[55]In this appendix we find it convenient to keep the factors of r, the radius of the S^1 fiber, explicit. To compare with the formulas in the main text one may set $r = 1$.

contour. As mentioned in section 2.4, the result can be expressed as a sum over N-colored Young diagrams, Y_i. Namely, we have [34, 52]

$$Z_k = \sum_{\vec{Y}, |\vec{Y}|=k} Z_{\vec{Y}}, \tag{B.4}$$

with

$$Z_{\vec{Y}} = \prod_{i,j=1}^{N} \prod_{s \in Y_i} \frac{\sin\left(\pi r(E_{ij} + \tilde{\nu} + \frac{1}{2r} - \epsilon_+)\right) \sin\left(\pi r(E_{ij} - \tilde{\nu} - \frac{1}{2r} - \epsilon_+)\right)}{\sin\left(\pi r E_{ij}\right) \sin\left(\pi r(E_{ij} - 2\epsilon_+)\right)}, \tag{B.5}$$

where we have defined

$$E_{ij} = \tilde{u}_i - \tilde{u}_j - \epsilon_1 h_i(s) + \epsilon_2(v_j(s) + 1), \tag{B.6}$$

where $h_i(s)$ is the horizontal distance from the box to the right side of the Young diagram, and $v_i(s)$ is the vertical distance to the bottom.

Let us note some simplifying limits of this instanton partition function. First, if we set

$$\tilde{\nu} = \pm\epsilon_+ - \frac{1}{2r} \quad \Leftrightarrow \quad y \equiv e^{2\pi i r \tilde{\nu}} = -(\mathfrak{q}_1 \mathfrak{q}_2)^{\pm 1/2}, \tag{B.7}$$

one can check from (B.5) that $Z_{\vec{Y}} = 1$ for all $\vec{Y}$, and so the full instanton partition function is given by

$$Z^{\text{inst}}_{\mathbb{R}^2_{\mathfrak{q}_1} \times \mathbb{R}^2_{\mathfrak{q}_2} \times S^1}\left(x, y = -(\mathfrak{q}_1 \mathfrak{q}_2)^{\pm 1/2}, z\right) \propto \sum_{k=0}^{\infty} z^k \sum_{\vec{Y}, |\vec{Y}|=k} 1. \tag{B.8}$$

This is simply the generating function for the N-colored Young diagrams, which can be expressed in terms of the q-Pochhammer symbol,

$$Z^{\text{inst}}_{\mathbb{R}^2_{\mathfrak{q}_1} \times \mathbb{R}^2_{\mathfrak{q}_2} \times S^1}\left(x, y = -(\mathfrak{q}_1 \mathfrak{q}_2)^{\pm 1/2}, z\right) \propto (z; z)^{-N}. \tag{B.9}$$

The precise coefficient depends on the detailed normalization of the instanton measure, and it is not known how to fix this from first principles. In [34, 35] it was found that the appropriate normalization is such that we obtain the Dedekind eta function, $\eta(z) = z^{1/24}(z; z)$, and we will make the same assumption here. Then

$$Z^{\text{inst}}_{\mathbb{R}^2_{\mathfrak{q}_1} \times \mathbb{R}^2_{\mathfrak{q}_2} \times S^1}\left(x, y = -(\mathfrak{q}_1 \mathfrak{q}_2)^{\pm 1/2}, z\right) = \eta(z)^{-N}. \tag{B.10}$$

Next, consider

$$\tilde{\nu} = -\frac{1}{2r} \pm \epsilon_- \quad \Leftrightarrow \quad y = -(\mathfrak{q}_1 \mathfrak{q}_2^{-1})^{\pm 1/2}. \tag{B.11}$$

Then we can see the prefactor in (B.3) vanishes, and so $Z_k = 0$ for $k > 0$, implying

$$Z^{\text{inst}}_{\mathbb{R}^2_{\mathfrak{q}_1} \times \mathbb{R}^2_{\mathfrak{q}_2} \times S^1}\left(x, y = -(\mathfrak{q}_1 \mathfrak{q}_2^{-1})^{\pm 1/2}, z\right) = 1. \tag{B.12}$$

Both of these identities remain true if we shift $\tilde{\nu} \to \tilde{\nu} + \frac{\ell}{r}$, $\ell \in \mathbb{Z}$, as this is a symmetry of the partition function.

In the SU(N) case, we expect the above identities to remain true, however we need to strip off the overall U(1) contribution from above, and find that $N \to N - 1$ in (B.10).

References

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

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

[3] J. Qiu and M. Zabzine, *Review of localization for 5d supersymmetric gauge theories*, *J. Phys.* **A 50** (2017) 443014 [arXiv:1608.02966] [INSPIRE].

[4] S. Kim and K. Lee, *Indices for 6 dimensional superconformal field theories*, *J. Phys.* **A 50** (2017) 443017 [arXiv:1608.02969] [INSPIRE].

[5] N. Seiberg, *Five-dimensional SUSY field theories, nontrivial fixed points and string dynamics*, *Phys. Lett.* **B 388** (1996) 753 [hep-th/9608111] [INSPIRE].

[6] H.-C. Kim, S.-S. Kim and K. Lee, *5-dim Superconformal Index with Enhanced En Global Symmetry*, *JHEP* **10** (2012) 142 [arXiv:1206.6781] [INSPIRE].

[7] N.A. Nekrasov and S.L. Shatashvili, *Bethe/Gauge correspondence on curved spaces*, *JHEP* **01** (2015) 100 [arXiv:1405.6046] [INSPIRE].

[8] F. Benini and A. Zaffaroni, *A topologically twisted index for three-dimensional supersymmetric theories*, *JHEP* **07** (2015) 127 [arXiv:1504.03698] [INSPIRE].

[9] C. Closset and H. Kim, *Comments on twisted indices in 3d supersymmetric gauge theories*, *JHEP* **08** (2016) 059 [arXiv:1605.06531] [INSPIRE].

[10] F. Benini and A. Zaffaroni, *Supersymmetric partition functions on Riemann surfaces*, *Proc. Symp. Pure Math.* **96** (2017) 13 [arXiv:1605.06120] [INSPIRE].

[11] C. Closset, H. Kim and B. Willett, *Supersymmetric partition functions and the three-dimensional A-twist*, *JHEP* **03** (2017) 074 [arXiv:1701.03171] [INSPIRE].

[12] C. Closset, H. Kim and B. Willett, $\mathcal{N} = 1$ *supersymmetric indices and the four-dimensional A-model*, *JHEP* **08** (2017) 090 [arXiv:1707.05774] [INSPIRE].

[13] C. Closset, H. Kim and B. Willett, *Seifert fibering operators in 3d* $\mathcal{N} = 2$ *theories*, arXiv:1807.02328 [INSPIRE].

[14] E. Witten, *Topological sigma models*, *Commun. Math. Phys.* **118** (1988) 411.

[15] N. Kurokawa, *Multiple sine functions and selberg zeta functions*, *Proc. Japan Acad.* **A 67** (1991) 61.

[16] N. Hama, K. Hosomichi and S. Lee, *SUSY Gauge Theories on Squashed Three-Spheres*, *JHEP* **05** (2011) 014 [arXiv:1102.4716] [INSPIRE].

[17] A. Bawane, G. Bonelli, M. Ronzani and A. Tanzini, $\mathcal{N} = 2$ *supersymmetric gauge theories on* $S^2 \times S^2$ *and Liouville Gravity*, *JHEP* **07** (2015) 054 [arXiv:1411.2762] [INSPIRE].

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

[19] N.A. Nekrasov and S.L. Shatashvili, *Quantization of Integrable Systems and Four Dimensional Gauge Theories*, in *Proceedings, 16th International Congress on Mathematical Physics (ICMP09)*, Prague, Czech Republic, August 3–8, 2009, pp. 265–289 (2009) [DOI:10.1142/9789814304634_0015] [arXiv:0908.4052] [INSPIRE].

[20] C. Beem, T. Dimofte and S. Pasquetti, *Holomorphic Blocks in Three Dimensions*, *JHEP* **12** (2014) 177 [arXiv:1211.1986] [INSPIRE].

[21] N. Bobev and P.M. Crichigno, *Universal RG Flows Across Dimensions and Holography*, *JHEP* **12** (2017) 065 [arXiv:1708.05052] [INSPIRE].

[22] I. Bah, A. Passias and P. Weck, *Holographic duals of five-dimensional SCFTs on a Riemann surface*, arXiv:1807.06031 [INSPIRE].

[23] D. Gaiotto, *N = 2 dualities*, *JHEP* **08** (2012) 034 [arXiv:0904.2715] [INSPIRE].

[24] F. Benini, Y. Tachikawa and B. Wecht, *Sicilian gauge theories and N = 1 dualities*, *JHEP* **01** (2010) 088 [arXiv:0909.1327] [INSPIRE].

[25] I. Bah, C. Beem, N. Bobev and B. Wecht, *Four-Dimensional SCFTs from M5-Branes*, *JHEP* **06** (2012) 005 [arXiv:1203.0303] [INSPIRE].

[26] I. Bah, A. Hanany, K. Maruyoshi, S.S. Razamat, Y. Tachikawa and G. Zafir, *4d N = 1 from 6d N = (1, 0) on a torus with fluxes*, *JHEP* **06** (2017) 022 [arXiv:1702.04740] [INSPIRE].

[27] H.-C. Kim, S.S. Razamat, C. Vafa and G. Zafir, *E-String Theory on Riemann Surfaces*, *Fortsch. Phys.* **66** (2018) 1700074 [arXiv:1709.02496] [INSPIRE].

[28] A. Gadde, L. Rastelli, S.S. Razamat and W. Yan, *The 4d Superconformal Index from q-deformed 2d Yang-Mills*, *Phys. Rev. Lett.* **106** (2011) 241602 [arXiv:1104.3850] [INSPIRE].

[29] C. Beem and A. Gadde, *The N = 1 superconformal index for class S fixed points*, *JHEP* **04** (2014) 036 [arXiv:1212.1467] [INSPIRE].

[30] S.M. Hosseini, I. Yaakov and A. Zaffaroni, *Topologically twisted indices in five dimensions and holography*, arXiv:1808.06626 [INSPIRE].

[31] K.A. Intriligator, D.R. Morrison and N. Seiberg, *Five-dimensional supersymmetric gauge theories and degenerations of Calabi-Yau spaces*, *Nucl. Phys.* **B 497** (1997) 56 [hep-th/9702198] [INSPIRE].

[32] J. Källén and M. Zabzine, *Twisted supersymmetric 5D Yang-Mills theory and contact geometry*, *JHEP* **05** (2012) 125 [arXiv:1202.1956] [INSPIRE].

[33] K. Hosomichi, R.-K. Seong and S. Terashima, *Supersymmetric Gauge Theories on the Five-Sphere*, *Nucl. Phys.* **B 865** (2012) 376 [arXiv:1203.0371] [INSPIRE].

[34] H.-C. Kim, J. Kim and S. Kim, *Instantons on the 5-sphere and M5-branes*, arXiv:1211.0144 [INSPIRE].

[35] H.-C. Kim, S. Kim, S.-S. Kim and K. Lee, *The general M5-brane superconformal index*, arXiv:1307.7660 [INSPIRE].

[36] J. Qiu and M. Zabzine, *5D Super Yang-Mills on $Y^{p,q}$ Sasaki-Einstein manifolds*, *Commun. Math. Phys.* **333** (2015) 861 [arXiv:1307.3149] [INSPIRE].

[37] T. Kawano and N. Matsumiya, *5D SYM on 3D Sphere and 2D YM*, *Phys. Lett.* **B 716** (2012) 450 [arXiv:1206.5966] [INSPIRE].

[38] Y. Fukuda, T. Kawano and N. Matsumiya, *5D SYM and 2D q-Deformed YM*, *Nucl. Phys.* **B 869** (2013) 493 [arXiv:1210.2855] [INSPIRE].

[39] T. Kawano and N. Matsumiya, *5D SYM on 3D Deformed Spheres*, *Nucl. Phys.* **B 898** (2015) 456 [arXiv:1505.06565] [INSPIRE].

[40] G. Festuccia and N. Seiberg, *Rigid Supersymmetric Theories in Curved Superspace*, *JHEP* **06** (2011) 114 [arXiv:1105.0689] [INSPIRE].

[41] Y. Pan, *Rigid Supersymmetry on 5-dimensional Riemannian Manifolds and Contact Geometry*, *JHEP* **05** (2014) 041 [arXiv:1308.1567] [INSPIRE].

[42] Y. Imamura and H. Matsuno, *Supersymmetric backgrounds from 5d N = 1 supergravity*, *JHEP* **07** (2014) 055 [arXiv:1404.0210] [INSPIRE].

[43] L.F. Alday, P. Benetti Genolini, M. Fluder, P. Richmond and J. Sparks, *Supersymmetric gauge theories on five-manifolds*, *JHEP* **08** (2015) 007 [arXiv:1503.09090] [INSPIRE].

[44] C. Closset, T.T. Dumitrescu, G. Festuccia and Z. Komargodski, *Supersymmetric Field Theories on Three-Manifolds*, *JHEP* **05** (2013) 017 [arXiv:1212.3388] [INSPIRE].

[45] Y. Imamura and D. Yokoyama, *N = 2 supersymmetric theories on squashed three-sphere*, *Phys. Rev.* **D 85** (2012) 025015 [arXiv:1109.4734] [INSPIRE].

[46] N.A. Nekrasov and S.L. Shatashvili, *Supersymmetric vacua and Bethe ansatz*, *Nucl. Phys. Proc. Suppl.* **192-193** (2009) 91 [arXiv:0901.4744] [INSPIRE].

[47] L.C. Jeffrey and F.C. Kirwan, *Localization for nonabelian group actions*, *Topology* **34** (1995) 291.

[48] F. Benini, R. Eager, K. Hori and Y. Tachikawa, *Elliptic Genera of 2d N = 2 Gauge Theories*, *Commun. Math. Phys.* **333** (2015) 1241 [arXiv:1308.4896] [INSPIRE].

[49] N. Drukker, T. Okuda and F. Passerini, *Exact results for vortex loop operators in 3d supersymmetric theories*, *JHEP* **07** (2014) 137 [arXiv:1211.3409] [INSPIRE].

[50] E. Witten, *Phases of N = 2 theories in two-dimensions*, *Nucl. Phys.* **B 403** (1993) 159 [hep-th/9301042] [INSPIRE].

[51] J. Gomis, T. Okuda and V. Pestun, *Exact Results for 't Hooft Loops in Gauge Theories on S^4*, *JHEP* **05** (2012) 141 [arXiv:1105.2568] [INSPIRE].

[52] N. Nekrasov and S. Shadchin, *ABCD of instantons*, *Commun. Math. Phys.* **252** (2004) 359 [hep-th/0404225] [INSPIRE].

[53] C.-M. Chang, M. Fluder, Y.-H. Lin and Y. Wang, *Spheres, Charges, Instantons and Bootstrap: A Five-Dimensional Odyssey*, *JHEP* **03** (2018) 123 [arXiv:1710.08418] [INSPIRE].

[54] C. Closset and M. Del Zotto, to appear.

[55] C. Closset, S. Cremonesi and D.S. Park, *The equivariant A-twist and gauged linear σ-models on the two-sphere*, *JHEP* **06** (2015) 076 [arXiv:1504.06308] [INSPIRE].

[56] F. Nieri, S. Pasquetti and F. Passerini, *3d and 5d Gauge Theory Partition Functions as q-deformed CFT Correlators*, *Lett. Math. Phys.* **105** (2015) 109 [arXiv:1303.2626] [INSPIRE].

[57] S. Pasquetti, *Holomorphic blocks and the 5d AGT correspondence*, *J. Phys.* **A 50** (2017) 443016 [arXiv:1608.02968] [INSPIRE].

[58] J. Källén, J. Qiu and M. Zabzine, *The perturbative partition function of supersymmetric* 5D *Yang-Mills theory with matter on the five-sphere*, *JHEP* **08** (2012) 157 [arXiv:1206.6008] [INSPIRE].

[59] H.-C. Kim and S. Kim, *M5-branes from gauge theories on the 5-sphere*, *JHEP* **05** (2013) 144 [arXiv:1206.6339] [INSPIRE].

[60] G. Lockhart and C. Vafa, *Superconformal Partition Functions and Non-perturbative Topological Strings*, *JHEP* **10** (2018) 051 [arXiv:1210.5909] [INSPIRE].

[61] C.P. Herzog, I.R. Klebanov, S.S. Pufu and T. Tesileanu, *Multi-Matrix Models and Tri-Sasaki Einstein Spaces*, *Phys. Rev.* **D 83** (2011) 046001 [arXiv:1011.5487] [INSPIRE].

[62] D.L. Jafferis and S.S. Pufu, *Exact results for five-dimensional superconformal field theories with gravity duals*, *JHEP* **05** (2014) 032 [arXiv:1207.4359] [INSPIRE].

[63] S.M. Hosseini and A. Zaffaroni, *Large N matrix models for* 3d $\mathcal{N} = 2$ *theories: twisted index, free energy and black holes*, *JHEP* **08** (2016) 064 [arXiv:1604.03122] [INSPIRE].

[64] S.M. Hosseini and N. Mekareeya, *Large N topologically twisted index: necklace quivers, dualities and Sasaki-Einstein spaces*, *JHEP* **08** (2016) 089 [arXiv:1604.03397] [INSPIRE].

[65] O. Bergman and D. Rodriguez-Gomez, 5d *quivers and their* AdS_6 *duals*, *JHEP* **07** (2012) 171 [arXiv:1206.3503] [INSPIRE].

[66] D.L. Jafferis, *The Exact Superconformal R-Symmetry Extremizes Z*, *JHEP* **05** (2012) 159 [arXiv:1012.3210] [INSPIRE].

[67] C. Closset, T.T. Dumitrescu, G. Festuccia, Z. Komargodski and N. Seiberg, *Contact Terms, Unitarity and F-Maximization in Three-Dimensional Superconformal Theories*, *JHEP* **10** (2012) 053 [arXiv:1205.4142] [INSPIRE].

[68] C. Núñez, I.Y. Park, M. Schvellinger and T.A. Tran, *Supergravity duals of gauge theories from F(4) gauged supergravity in six-dimensions*, *JHEP* **04** (2001) 025 [hep-th/0103080] [INSPIRE].

[69] M. Naka, *Various wrapped branes from gauged supergravities*, hep-th/0206141 [INSPIRE].

[70] L.J. Romans, *The F(4) Gauged Supergravity in Six-dimensions*, *Nucl. Phys.* **B 269** (1986) 691 [INSPIRE].

[71] F. Benini, N. Bobev and P.M. Crichigno, *Two-dimensional SCFTs from D3-branes*, *JHEP* **07** (2016) 020 [arXiv:1511.09462] [INSPIRE].

[72] F. Azzurli, N. Bobev, P.M. Crichigno, V.S. Min and A. Zaffaroni, *A universal counting of black hole microstates in* AdS_4, *JHEP* **02** (2018) 054 [arXiv:1707.04257] [INSPIRE].

[73] B. Assel, D. Cassani and D. Martelli, *Localization on Hopf surfaces*, *JHEP* **08** (2014) 123 [arXiv:1405.5144] [INSPIRE].

[74] J. Kinney, J.M. Maldacena, S. Minwalla and S. Raju, *An Index for 4 dimensional super conformal theories*, *Commun. Math. Phys.* **275** (2007) 209 [hep-th/0510251] [INSPIRE].

[75] C. Romelsberger, *Counting chiral primaries in* $N = 1$, $d = 4$ *superconformal field theories*, *Nucl. Phys.* **B 747** (2006) 329 [hep-th/0510060] [INSPIRE].

[76] O. Aharony, S.S. Razamat, N. Seiberg and B. Willett, 3d *dualities from 4d dualities*, *JHEP* **07** (2013) 149 [arXiv:1305.3924] [INSPIRE].

[77] A. Gadde, E. Pomoni, L. Rastelli and S.S. Razamat, *S-duality and 2d Topological QFT*, *JHEP* **03** (2010) 032 [arXiv:0910.2225] [INSPIRE].

[78] L. Rastelli and S.S. Razamat, *The Superconformal Index of Theories of Class S*, in *New Dualities of Supersymmetric Gauge Theories*, J. Teschner ed., Mathematical Physics Studies, pp. 261–305, Springer, Cham (2016) [DOI:10.1007/978-3-319-18769-3_9] [arXiv:1412.7131]

[79] S.S. Razamat and B. Willett, *Down the rabbit hole with theories of class S*, *JHEP* **10** (2014) 99 [arXiv:1403.6107] [INSPIRE].

[80] S.N. Ruijsenaars, *Elliptic integrable systems of calogero–moser type: a survey*, in *The proceedings of Workshop on Elliptic Integrable Systems*, pp. 201–221 (2004) [http://www.math.kobe-u.ac.jp/publications/rlm18/13.pdf].

[81] D. Gaiotto, L. Rastelli and S.S. Razamat, *Bootstrapping the superconformal index with surface defects*, *JHEP* **01** (2013) 022 [arXiv:1207.3577] [INSPIRE].

[82] A. Gadde, S.S. Razamat and B. Willett, *"Lagrangian" for a Non-Lagrangian Field Theory with $\mathcal{N} = 2$ Supersymmetry*, *Phys. Rev. Lett.* **115** (2015) 171604 [arXiv:1505.05834] [INSPIRE].

[83] P. Agarwal, K. Maruyoshi and J. Song, *A "Lagrangian" for the E_7 superconformal theory*, *JHEP* **05** (2018) 193 [arXiv:1802.05268] [INSPIRE].

[84] D. Gaiotto and S.S. Razamat, *$\mathcal{N} = 1$ theories of class $\mathcal{S}_k$*, *JHEP* **07** (2015) 073 [arXiv:1503.05159] [INSPIRE].

[85] E. Witten, *Small instantons in string theory*, *Nucl. Phys.* **B 460** (1996) 541 [hep-th/9511030] [INSPIRE].

[86] K. Maruyoshi and J. Yagi, *Surface defects as transfer matrices*, *PTEP* **2016** (2016) 113B01 [arXiv:1606.01041] [INSPIRE].

[87] Y. Ito and Y. Yoshida, *Superconformal index with surface defects for class $\mathcal{S}_k$*, arXiv:1606.01653 [INSPIRE].

[88] J. Yagi, *Surface defects and elliptic quantum groups*, *JHEP* **06** (2017) 013 [arXiv:1701.05562] [INSPIRE].

[89] B. Nazzal and S.S. Razamat, *Surface Defects in E-String Compactifications and the van Diejen Model*, *SIGMA* **14** (2018) 036 [arXiv:1801.00960] [INSPIRE].

[90] N. Bobev, M. Bullimore and H.-C. Kim, *Supersymmetric Casimir Energy and the Anomaly Polynomial*, *JHEP* **09** (2015) 142 [arXiv:1507.08553] [INSPIRE].

[91] N. Nekrasov, *Five dimensional gauge theories and relativistic integrable systems*, *Nucl. Phys.* **B 531** (1998) 323 [hep-th/9609219] [INSPIRE].

[92] F. Benini, K. Hristov and A. Zaffaroni, *Black hole microstates in AdS_4 from supersymmetric localization*, *JHEP* **05** (2016) 054 [arXiv:1511.04085] [INSPIRE].

[93] F. Benini, K. Hristov and A. Zaffaroni, *Exact microstate counting for dyonic black holes in AdS4*, *Phys. Lett.* **B 771** (2017) 462 [arXiv:1608.07294] [INSPIRE].

[94] A. Cabo-Bizet, V.I. Giraldo-Rivera and L.A. Pando Zayas, *Microstate counting of AdS_4 hyperbolic black hole entropy via the topologically twisted index*, *JHEP* **08** (2017) 023 [arXiv:1701.07893] [INSPIRE].

[95] S.M. Hosseini, K. Hristov and A. Passias, *Holographic microstate counting for AdS₄ black holes in massive IIA supergravity*, *JHEP* **10** (2017) 190 [`arXiv:1707.06884`] [INSPIRE].

[96] F. Benini, H. Khachatryan and P. Milan, *Black hole entropy in massive Type IIA*, *Class. Quant. Grav.* **35** (2018) 035004 [`arXiv:1707.06886`] [INSPIRE].

[97] M. Suh, *D4-branes wrapped on supersymmetric four-cycles*, `arXiv:1809.03517` [INSPIRE].

[98] T. Nishioka and I. Yaakov, *Generalized indices for $\mathcal{N} = 1$ theories in four-dimensions*, *JHEP* **12** (2014) 150 [`arXiv:1407.8520`] [INSPIRE].

[99] B. Assel, D. Cassani, L. Di Pietro, Z. Komargodski, J. Lorenzen and D. Martelli, *The Casimir Energy in Curved Space and its Supersymmetric Counterpart*, *JHEP* **07** (2015) 043 [`arXiv:1503.05537`] [INSPIRE].

[100] A. Arabi Ardehali, J.T. Liu and P. Szepietowski, *High-Temperature Expansion of Supersymmetric Partition Functions*, *JHEP* **07** (2015) 113 [`arXiv:1502.07737`] [INSPIRE].

[101] M.R. Douglas, *On $D = 5$ super Yang-Mills theory and $(2,0)$ theory*, *JHEP* **02** (2011) 011 [`arXiv:1012.2880`] [INSPIRE].

[102] C. Beem, M. Lemos, P. Liendo, W. Peelaers, L. Rastelli and B.C. van Rees, *Infinite Chiral Symmetry in Four Dimensions*, *Commun. Math. Phys.* **336** (2015) 1359 [`arXiv:1312.5344`] [INSPIRE].

[103] K. Ohmori, H. Shimizu and Y. Tachikawa, *Anomaly polynomial of E-string theories*, *JHEP* **08** (2014) 002 [`arXiv:1404.3887`] [INSPIRE].