

The Glueball Superpotential

Mina Aganagic¹, Ken Intriligator², Cumrun Vafa^{1,3}
and Nicholas P. Warner⁴

¹ Jefferson Physical Laboratory, Harvard University
Cambridge, MA 02138, USA

² Department of Physics, University of California
San Diego, La Jolla, CA 92093-0354, USA

³ California Institute of Technology 452-48
Pasadena, CA 91125, USA

⁴ Department of Physics and Astronomy, USC
Los Angeles, CA 90089-0484, USA

Abstract

We compute glueball superpotentials for four-dimensional, $\mathcal{N} = 1$ supersymmetric gauge theories, with arbitrary gauge groups and massive matter representations. This is done by perturbatively integrating out massive charged fields. The Feynman diagram computations simplify, and are related to the corresponding matrix model. This leads to a natural notion of “projection to planar diagrams” for arbitrary gauge groups and representations. We discuss a general ambiguity in the glueball superpotential $W(S)$ for terms, S^n , whose order, n is greater than the dual Coxeter number. This ambiguity can be resolved for all classical gauge groups (A, B, C, D), via a natural embedding in an infinite rank supergroup. We use this to resolve some recently raised puzzles. For exceptional groups, we compute the superpotential terms for low

powers of the glueball field and propose an all-order completion for some examples including $\mathcal{N} = 1^*$ for all simply-laced groups. We also comment on compactification of these theories to lower dimensions.

1 Introduction

New insights have recently been obtained into the non-perturbative dynamics of supersymmetric gauge theories in four and higher dimensions. It has been found that perturbative computation of the glueball superpotential in four-dimensional supersymmetric gauge theories admitting a large N description is related to matrix model amplitudes. Moreover, the extremization of the superpotential provides exact, non-perturbative results for the gauge theory [1]. This connection between the superpotential generated by integrating out massive fields and matrix model amplitudes has been explicitly demonstrated in [2]. In particular, using supergraph techniques, which mirror the corresponding topological superstring computations [3], [4], one can see a dramatic simplification of the Feynman amplitudes contributing to the glueball superpotential.

In this paper we follow the strategy of [2] to compute the exact glueball superpotential for general gauge group and massive matter content. We find that the dramatic simplifications occur for *any* group and *any* massive representation. Moreover this leads to an interesting interpretation of the “planar projection” for arbitrary groups and representations.

To simplify the analysis, as in [2], we use constant abelian gluino backgrounds. For such backgrounds one can compute the glueball superpotential $W(S)$ for terms S^n with $n \leq \text{rank}(G)$ ¹. A general aspect of the glueball superpotential $W(S)$, which we discuss in detail, is that it is ambiguous at terms of order S^n with $n \geq h$, the dual Coxeter number. These higher order terms depend on the details of how the theory is defined in the ultraviolet, the “UV completion of the theory.” These ambiguities have to do with instantons: Specifically, $\langle S^h \rangle$ is classically zero, but is non-zero in the quantum theory because of instantons. On the other hand, terms of order S^n , with $n < h$, generally do not depend on the UV details of the theory; these unambiguous terms correspond to fractional instanton effects.

We propose a natural completion for these F-terms for general classical gauge groups following the last reference in [1] which leads to an all or-

¹One can in principle develop this further to compute terms S^k with $r < k < h$ by including some non-abelian configurations.

der superpotential for S . This prescription is first motivated via a general expression derived here for the contribution of any Feynman graph to the superpotential. This expression is valid for any gauge theory with arbitrary matter content, and it leads to an F-term completion via analytic continuation in N for the classical groups. However, there is a more physical, refined and indeed predictive realization of this F-term completion in terms of supergroups.

The group $G(N) = U(N)$, or $SO(N)$, or $Sp(N)$ is completed into the supergroup $G(N+k|k)$, with large k . This is very natural from the viewpoint of branes/anti-brane systems [5]: We add a large number k of brane anti-brane pairs. For example, we embed $U(N)$ into $U(N+k|k)$, and in the latter theory the glueball superpotential is unambiguous up to order S^{N+k} . The supertrace structure ensures that the coefficients in the $U(N+k|k)$ glueball superpotential are independent of k , and so it is natural to compute the $U(N)$ glueball superpotential, including terms S^n with n arbitrarily large, by going to the $U(N+k|k)$ theory with $k \rightarrow \infty$. The $G(N+k|k)$ supergroup thus leads to a natural ‘‘F-term completion’’ of $G(N)$, in the sense that the $G(N)$ theory can be viewed as a Higgs branch of $G(N+k|k)$ theory, which is F-term complete as $k \rightarrow \infty$. For large k , this F-completion of the glueball superpotential is computable, as in [2], perturbatively to arbitrarily high order, it is exact and this higher power completion is consistent with a generalized Konishi anomaly [6].

In some instances the $G(N+k|k)$ F-completion differs from some of the standard UV completions of $G(N)$. As we discuss, this generally happens when there are residual instanton effects associated with the Higgsing of $G(N+k|k)$ to $G(N)$. These effects arise when the quotient $G(N+k|k)/G(N)$ has the appropriate topology to give rise to instantons. We will analyze precisely when our $G(N+k|k)$ F-completions differ from the standard $G(N)$ UV completions, and how this explains apparent discrepancies between the matrix model results and standard gauge theory results. For example, standard $U(1)$ gauge theory does not have a glueball superpotential, but its F-completion into $U(1+k|k)$ does, with the difference coming from instantons in $U(2|1)/U(1)$. This issue of the F-completion could explain the apparent discrepancy observed in [7] between the matrix model/perturbative glueball superpotential results (equivalently, the generalized Konishi anomaly results [8], [9]) and standard supersymmetric gauge theory results.

Gauge theories based upon supergroups are non-unitary, and so this might seem to be a rather unphysical way to resolve F-term ambiguities. However one should note that F-terms of broken gauge systems will involve ghost-like chiral fields as shown in [10]. Moreover, gauge theories for non-

unitary supergroups can have the same F-terms as unitary quiver theories, as demonstrated in the last reference in [1]. Thus, one cannot distinguish unitary and non-unitary gauge theories based solely upon F-terms. It is therefore natural to extend the class of theories whose F-terms are to be studied so as to include both unitary and non-unitary gauge systems.

It is also important to find F-completions of the glueball superpotential for non-classical groups. We point out at least two ways this may be done: If the theory has a branch with classical groups emerging as unbroken groups then the foregoing prescription leads to an answer. In some other instances, like $\mathcal{N} = 1^*$ theories, the amplitudes depend upon group theoretic factors in a universal way and one may “analytically continue” in the choice of the group to define an F-completion. In particular, we use our general expression for the diagrammatic contributions to obtain the superpotential to three loops for the $\mathcal{N} = 1^*$ theories for any simply-laced gauge group. This leads us to conjecture an analytic continuation prescription in which NS^ℓ is replaced by $S^\ell \sum (p_a)^\ell$, where p_a are the extended Dynkin labels of the extending root of the underlying algebra. This conjecture is further supported by considering compactifications to two and three dimensions.

We discuss the relationship between $\mathcal{N} = 1$ supersymmetric gauge theories in four dimensions and a particular class of $\mathcal{N} = 2$ supersymmetric sigma models in two dimensions. The latter is obtained via the moduli space of flat connections on the compactifying torus. One can then see a direct relationship between the instanton corrections to the chiral rings of both theories. This approach provides some insights as to how the ambiguities could be related to “gravitational descendants” in the two-dimensional theory. On a more straightforward level, in $\mathcal{N} = 1^*$ theories the mirror of this sigma model is naturally related to the integrable structures used in [11], [12], [13]. In particular, this enables us to recompute the diagrammatic expansion of the glueball superpotential by making a duality transformation of the elliptic Calogero-Moser superpotentials of [12]. We show that this agrees with the direct diagrammatic computation of the glueball superpotential.

The organization of this paper is as follows: In section 2 we discuss some general aspects of the glueball superpotential. In section 3 we compute the corrections to the superpotential for arbitrary groups and representation up to the glueball field to the power of the rank of the gauge group. One can make sense of the notion of the “projection to planar diagrams” for arbitrary groups provided the number of loops is less than the rank. In section 4 we propose the natural F-completion of our theories. We argue that the matrix model results should be understood as referring to this particular F-completion. In section 5 we discuss some special cases, where our

F-completion differs from more standard UV definitions of some gauge theories, at the non-perturbative level. The difference comes from instantons in the partially broken group $G(N + k|k)/G(N)$. In section 6 we discuss some further glueball superpotential examples and compute the glueball superpotential for the $\mathcal{N} = 1^*$ theory for arbitrary simply-laced gauge groups, to arbitrarily high order in terms of a proposed F-completion. In section 7 we consider compactifications to 2 and 3 dimensions and the meaning of the superpotential computation in these cases. In section 8 we make some final remarks. In appendix A we derive a group theory result that we need for the Feynman diagram computations.

Note added in revised version, Nov. 2003:

In the original version of this paper, we speculated that the residual instantons effects discussed in this paper, associated with the F-completion, should resolve the apparent discrepancy of [7] between the matrix model and standard gauge theory. We now know that this speculation was incorrect. We still claim that the matrix model refers to the $G(N + k|k)$ F-completion, and that the results thus obtained could, in principle, differ from standard gauge theory by residual instanton effects. But the matrix model side of the computation must be done appropriately, which requires glueball fields for $U(1)$ and $Sp(0)$ factors. Following the first version of the present paper, [14] appeared, which gave a particular treatment of the $Sp(0)$ factors introduced here; this treatment was later explained from the string theory perspective, and extended to a general prescription for all low-rank classical groups in [15].

Upon redoing the matrix model computation according to this new prescription, the results agree perfectly with standard gauge theory [14], [15]. This is consistent with the claim made here that the matrix model refers to the $G(N + k|k)$ completion, because further investigation of the theories of sect. 5.4 reveals that the residual instanton type effects, which in principle could have spoiled the agreement with standard gauge theory, can – and here do – exhibit remarkable cancellations [15].

2 The Glueball Superpotential

The glueball superfield of a four-dimensional, $\mathcal{N} = 1$ supersymmetric gauge theory is defined by

$$S = \epsilon^{\alpha\beta} g_{AB} \mathcal{W}_\alpha^A \mathcal{W}_\beta^B. \quad (2.1)$$

where A labels the Lie algebra elements, g_{AB} is the corresponding group invariant inner product and \mathcal{W}_α^A is the gluino field. The central idea in the proposal of [1] in gaining a perturbative window into non-perturbative dynamics of $\mathcal{N} = 1$ supersymmetric gauge theories has been to compute the glueball superpotential $W(S)$, perturbatively, by integrating out massive fields. One then treats S as a good order parameter in the IR physics and extremizes $W(S)$:

$$\frac{dW}{dS} = 0.$$

This yields, through the values of the superpotential W at the extrema, non-perturbative information about the gauge theory. This can also be extended to an exact computation of the coupling constants τ_{ij} for abelian factors.

For a pure gauge theory with dual Coxeter number h , the leading piece of the superpotential is given by the Veneziano-Yankielowicz superpotential [16]:

$$W_{VY}(S) = h S ((\log(S/\Lambda^3) - 1) + \tau S), \quad (2.2)$$

where τ is the gauge coupling constant at scale Λ . Upon extremization this yields

$$S = \Lambda^3 e^{-\tau/h}, \quad W = -h \Lambda^3 e^{-\tau/h}. \quad (2.3)$$

Note that this is consistent with the breaking of the non-anomalous \mathbf{Z}_{2h} subgroup of the $U(1)_R$ symmetry. That is, the \mathbf{Z}_{2h} symmetry acts on S via:

$$S \rightarrow S e^{-2\pi i/2h},$$

and is broken to \mathbf{Z}_2 by the vacuum expectation value of the glueball field.

In general, upon integrating out massive charged fields one finds corrections to the glueball superpotential. In particular, as argued in [2], an ℓ -loop diagram involving charged matter fields can contribute a term

$$\delta W(S) = c_\ell S^\ell.$$

For high powers of S there is an ambiguity in the definition of $W(S)$. Classically, S is a bilinear fermionic fields, and so if we raise S to a large enough power it will vanish. In particular $S^k = 0$ for $k > \dim(G)$. In a quantum theory it is natural to define powers of S by point splitting or smearing, and so we could instead consider a smeared glueball field

$$S_\rho(x) = \int d^4x' \rho_x(x') S(x'),$$

where $\rho_x(x')$ is a positive smearing function centered at x with $\int \rho_x(x') d^4x' = 1$. In the limit ρ becomes a δ -function the smeared glueball field goes back to

being the ordinary glueball field $S_\rho(x) \rightarrow S(x)$. There is *a priori* no reason for $S_\rho^k(x)$ to vanish for any k . Now consider an $\mathcal{N} = 1$ theory and change the UV action by turning on, by hand, a superpotential term

$$\delta(\text{action}) = \int d^4x d^2\theta \sum_{k \geq d} a_k S_\rho^k. \quad (2.4)$$

If $d = \dim(G) + 1$ then this term disappears in the limit $\rho_x(x') \rightarrow \delta(x - x')$. In fact one can say something stronger: As was discussed in [6] classically one expects

$$S^k = 0$$

for $k \geq h$. This equality is a statement in the chiral ring, that is, one actually has the *classical* relation:

$$S^k = \overline{D}\mathcal{O}_k \quad \text{for } k \geq h, \quad (2.5)$$

for some operators \mathcal{O}_k . Moreover, the difference $S(x) - S(y)$ is also trivial in the chiral ring. Equation (2.5) was established for $U(N)$ in [6] and for $SO(N)$ and $Sp(N)$ in [17]. In section 7 we will give a further argument in support of this classical ring relation for any group.

Note that if we have a chirally trivial operator then adding it to the superpotential does not change the action. This implies that if we include deformations with monomials S^k with $k \geq h$ (*i.e.* set $d = h$ in 2.4) in the superpotential, even with the smearing turned on, the action does not change classically. However, the quantum theory *will change* through such deformations because the classical chiral ring relation receives quantum corrections and S^k is no longer zero for $k \geq h$. This is apparent from the VY superpotential, which leads to $S \neq 0$, and so all the additional higher power terms in S will be relevant for the IR physics. Two $\mathcal{N} = 1$ theories can agree classically, but differ quantum mechanically by superpotential terms involving S^k with $k \geq h$. Note that for arbitrary addition of S^k with $k \geq h$ there may not be any corresponding UV complete theory. On the other hand, there could be several UV complete theories which agree classically, but differ quantum mechanically by such additions to the superpotential. Thus, in principle, specifying a classical description of an $\mathcal{N} = 1$ theory is not enough to determine all the F-terms unambiguously. This means that there is an inherent ambiguity in what one means by the quantum $\mathcal{N} = 1$ supersymmetric gauge theory.

The ambiguities set in at instanton number one, because $S^h \sim e^{-\tau}$. This means that if two quantum theories are classically the same then they will have same superpotential, W , for the fractional instantons. In particular, computations of the value of the superpotential, $W(q)$, where $q = e^{-\tau/h}$,

lead to an unambiguous answer for all q^k for $k < h$, but ambiguities can begin to show up at order q^h . That is, if W_1 and W_2 are the superpotentials of two theories with the same classical form, then:

$$W_1 - W_2 = \sum_{s \geq 0} c_s q^{s+h}$$

for some constants, c_s . The existence of such ambiguities has been noted recently in checking the matrix model proposal for computation of $\mathcal{N} = 1$ F-terms in [18], [7]. The obvious question is how to remove the quantum ambiguities?

Sometimes extra symmetry restricts the ambiguities: For example, for pure Yang-Mills theory, insisting on a non-anomalous \mathbf{Z}_{2h} symmetry allows only the addition of the terms

$$\Delta W = \sum_n a_n S^{1+nh}$$

to the superpotential. Note that if we add ΔW to the VY superpotential the vacuum structure does not qualitatively change. Or in fact, as was argued in [16], [19], if we impose the structure of the anomalous $U(1)$ R-symmetry in the superpotential (*i.e.* that the chiral $U(1)$ rotation by α leads to $\delta W = i \alpha h S$)² we can rule out any additional higher powers to the Veneziano-Yankielowicz potential. This is, however, a rare situation with a high degree of symmetry. For more general theories with less symmetries this is not possible.

One approach to fix the ambiguities would be to start with a conformal fixed point in the UV and flow down to the IR by the addition of some relevant operators. However, even in such cases there is room for ambiguity to develop in definitions of the relevant operators due to operator mixing. Such a possibility was already pointed out in [18] in the context of mass deformations of $\mathcal{N} = 4$ theories.

Thus one should not look for a unique IR answer, as it would depend on how the UV completion is achieved. For theories such as $\mathcal{N} = 1$, $U(N)$ supersymmetric gauge theories with one adjoint matter multiplet and arbitrary superpotential, the string theory embedding naturally provides a UV completion giving unambiguous higher order terms [20], [21] thus leading to

²Note that this result only uses the Adler-Bell-Jakiw anomaly and *does not* assume confinement. However to obtain chiral symmetry breaking, one has to make the non-trivial assumption that the glueball field S is a good order parameter for the IR physics. Note that there are theories for which there is no confinement but S is still a good order parameter, such as $U(N)$ with an adjoint field broken to $U(1)^N$.

the matrix model proposal in [1]. This was further extended to give an unambiguous proposal for arbitrary classical groups admitting large N description in [1]. The idea, motivated from string theory, basically reduces to computing the glueball superpotential W for the classical group admitting large N description by taking the large N limit to be exact for the computation of the glueball superpotential³. In other words, to compute S^k for any fixed k , consider a sequence of theories indexed by N and take N large enough and compute analytic expressions for $a_k(N)S^k$ and then substitute a finite value of N for $a_k(N)$ at the end of the computation. This gives an unambiguous completion of all F-terms, and this, in turn gets related to planar diagrams of the associated matrix model. Note that any other completion would lead to differences of order $q^h = O(e^{-N})$. In other words the computation of W at the extremum for any two possible UV completions would lead to the same exact result to all orders in the $(1/N)$ expansion.

The fact that classically $S^{h+n} = 0$ for any non-negative n , implies that perturbatively it makes sense only to compute the glueball superpotentials for powers up to h , that is, S^k with $k < h$. Beyond this, the perturbative computation is ambiguous. In other words we can compute, in principle, only a truncation of W unambiguously. We will call this the *reduced* superpotential, W_R . The higher powers of the glueball field lead to what one means quantum mechanically by the corresponding theory. We denote the corresponding piece of the superpotential consisting of terms with S^{n+h} with $n \geq 0$, by W_A . Thus any W_A can in principle arise in a quantum theory and should be viewed as part of the quantum definition of the theory. Put another way, we have the decomposition

$$W(S) = W_R(S) + W_A(S)$$

where W_R is unambiguously computable by integrating out matter fields and W_A is part of the definition of the quantum theory. Thus a conservative generalization of the proposal of [1] reduces to the statement that non-perturbative fractional instanton effects can be computed unambiguously from a perturbative definition of the theory. The rest can also be computed if we know the precise choice of the non-perturbative F-completion of the theory. For classical groups, even those which do not have a large N , 't Hooft description, we can follow the approach of [1] in defining a quantum completion by considering the large N regularization of the superpotential computation and substituting finite N in the analytic computations at the end. We will discuss the meaning of such a prescription in section 4.

One aim of this paper is to compute W_R using perturbative techniques.

³This structure was anticipated from string theory where on the large N gravitational dual, the glueball superpotential is exact at genus zero.

We will use the technique of [2] which considers a constant, abelian gluino backgrounds. For such backgrounds, S^k is non-zero for $k \leq r$ where r is the rank of the group, so one can compute all monomials in $W_R(S)$ up to S^r . In general the dual Coxeter number, h , is greater than the rank, r . For $SU(N)$ and $Sp(N)$ we have $h = r + 1$, and thus the abelian computation suffices to determine W_R completely. For other groups one has to extend the computation of [2] to certain non-abelian configurations of the background gauge fields in order to obtain the other monomials S^k for $r < k < h$. We will content ourselves in this paper with the abelian configurations, leaving the non-abelian configurations for future work. In the context of the classical groups we give a proposal of how to extend this superpotential computation not only to the full W_R but to all powers of S .

3 Computing the Fractional Instanton Part of the Glueball Superpotential

In this section we show, following [2], how to compute the glueball superpotential up to the power S^r , where r is the rank of the group, for arbitrary $\mathcal{N} = 1$ supersymmetric gauge theories in four dimensions⁴. Even though we state this in the context of $\mathcal{N} = 1$ theories in four dimensions, as noted in the last paper in [1] this can also be extended to $\mathcal{N} = 1$ theories coming from higher dimensions.

We consider turning on a constant, abelian gluino background. This leads to a particularly simple perturbation theory in which there are no path ordered exponentials. As noted in [2], the propagators of the charged fields, in the Schwinger formulation are given by

$$\int \prod_i ds_i \exp(-s_i [p_i^2 + \vec{W}_\alpha \cdot \vec{\lambda}_i \pi_i^\alpha + m_i]), \quad (3.1)$$

where i denotes the edge, p_i denotes the four-dimensional bosonic momentum, π_i denotes the fermionic momentum, and $\vec{\lambda}_i$ denotes the charge under the Cartan generators (*i.e.* the group theory weight) flowing along the i^{th} edge. Note that the bosonic and fermionic momenta, p_i , π_i , and the Cartan charges, λ_i , are all conserved, and in particular they are conserved by each vertex in a Feynman graph. We can therefore encode the independent

⁴There seems to be some confusion in the literature on the meaning of the computation in [2]: This computation is also non-perturbatively exact. In particular the computation of the path integral contribution to glueball superpotential reduces to perturbative configurations of charged fields, as is clear from the derivation of [2].

variables by passing to the corresponding loop quantities: p_a, π_a, λ_a , where $a = 1, \dots, \ell$ and ℓ is the number of loops of the Feynman graph.

Consider a particular graph, \mathcal{G}_ℓ , with ℓ loops. Introduce the incidence matrix, L_{ia} , where $L_{ia} = 0$ if the i -th edge does not belong to the loop a and it is ± 1 (taking into account their relative orientation) if the edge is part of the loop. One can then write:

$$p_i = \sum_a p_a L_{ia}, \quad \pi_i = \sum_a \pi_a L_{ia}, \quad \lambda_i = \sum_a \lambda_a L_{ia}.$$

Defining an $\ell \times \ell$ matrix

$$M_{ab}(s) = \sum_i s_i L_{ia} L_{ib}$$

the propagators can be written as

$$\int \prod_i ds_i e^{-s_i m_i} \exp(-[p_a M_{ab} p_b + \vec{W}_\alpha \cdot \vec{\lambda}_a M_{ab} \pi_b^\alpha]).$$

Integrating out the p_a and π_a leads to (up to factors of 2π):

$$\int \prod_i ds_i e^{-s_i m_i} \det(M)^{-2} \epsilon_{a_1 \dots a_\ell} \epsilon_{b_1 \dots b_\ell} (M_{1a_1} M_{2a_2} \dots M_{\ell a_\ell}) (M_{1b_1} M_{2b_2} \dots M_{\ell b_\ell}) \\ \times (\mathcal{W}_1^{\mu_1} \mathcal{W}_1^{\mu_2} \dots \mathcal{W}_1^{\mu_\ell}) (\mathcal{W}_2^{\nu_1} \dots \mathcal{W}_2^{\nu_\ell}) (\lambda_1^{\mu_1} \lambda_2^{\mu_2} \dots \lambda_\ell^{\mu_\ell}) (\lambda_1^{\nu_1} \dots \lambda_\ell^{\nu_\ell})$$

where the μ_i, ν_j are vector indices on the Cartan subalgebra. The ϵ 's and the combination of M 's generate a $\det(M)^2$ which cancels the $\det(M)^{-2}$ coming from the bosonic momentum integration. Thus, the s_i dependence of the amplitudes trivializes just as in the computation in [2]. The s_i integration can now be performed trivially leading to $1/m_i$ for each propagator. The complete amplitude is then obtained by combining this with the combinatorial factor, F_λ , coming from the interaction vertices and the symmetry factors of the graph, with fixed internal charges λ_a , which leads to the matrix model amplitude $F(\lambda)$. We are thus left with:

$$F(\lambda) (\mathcal{W}_1^{\mu_1} \mathcal{W}_1^{\mu_2} \dots \mathcal{W}_1^{\mu_\ell}) (\mathcal{W}_2^{\nu_1} \dots \mathcal{W}_2^{\nu_\ell}) (\lambda_1^{\mu_1} \lambda_2^{\mu_2} \dots \lambda_\ell^{\mu_\ell}) (\lambda_1^{\nu_1} \dots \lambda_\ell^{\nu_\ell}).$$

The factor $F(\lambda)$ is the corresponding matrix model amplitude with loop weights specified by λ . This needs to be summed over all the weights in the representations running around the loops, and over all the ways the representations that can run through the graph, \mathcal{G}_ℓ . Thus, associated to each such graph, \mathcal{G}_ℓ , there is a group theory factor

$$T^{\mu_1 \dots \mu_\ell \nu_1 \dots \nu_\ell} \equiv \frac{1}{(\ell!)^2} \sum_{\lambda_1, \dots, \lambda_\ell} F(\lambda) \lambda_1^{[\mu_1} \lambda_2^{\mu_2} \dots \lambda_\ell^{\mu_\ell]} \lambda_1^{[\nu_1} \lambda_2^{\nu_2} \dots \lambda_\ell^{\nu_\ell]}, \quad (3.2)$$

where the μ 's and ν 's are totally antisymmetrized among themselves. The tensor is symmetric under any exchange of members in a pair (μ_k, ν_k) , but the very important point is that this tensor is invariant under the Weyl group of the underlying gauge group. One can then show that for a simple group one has:

$$T^{\mu_1 \dots \mu_\ell \nu_1 \dots \nu_\ell} = C (\delta^{\mu_1 \nu_1} \delta^{\mu_2 \nu_2} \dots \delta^{\mu_\ell \nu_\ell} \pm \text{permutations}), \quad (3.3)$$

for some constant C ⁵. A proof of this statement is given in Appendix A. Contracting T with the \mathcal{W} 's and using the definition of the glueball field we get a contribution of $C \ell! S^\ell$ to the glueball superpotential. To find C , one can contract both (3.2) and (3.3) with $\delta_{\mu_1 \nu_1} \delta_{\mu_2 \nu_2} \dots \delta_{\mu_\ell \nu_\ell}$. The left-hand side of (3.3) gives $C r!/(r - \ell)!$ and from (3.2) this is

$$\frac{1}{\ell!} \sum_{\lambda_1, \dots, \lambda_\ell} F(\lambda) \det(\vec{\lambda}_a \cdot \vec{\lambda}_b),$$

where $\vec{\lambda}_a \cdot \vec{\lambda}_b$ denotes the $\ell \times \ell$ matrix of inner products of weights in each loop.

Putting all this together, we see that the graph \mathcal{G}_ℓ gives a contribution to the superpotential of:

$$(\Delta W)_{\mathcal{G}_\ell} = \sum_{\lambda_1, \dots, \lambda_\ell} F(\lambda) \det(\vec{\lambda}_a \cdot \vec{\lambda}_b) (r - \ell)! * \text{over} * r! S^\ell. \quad (3.4)$$

This must then be summed over all graphs in the usual loop expansion to obtain W_R .

As a special case, consider the results of [2] for $U(N)$ with adjoint matter fields. The determinant factor in (3.4) implies that the ℓ Cartan charges λ_a should be linearly independent. This means that the corresponding 't Hooft diagrams contributing to $F(\lambda)$ should be planar (having fewer 't Hooft index loops is equivalent to a linear relation between the λ_a). With no loss of generality we can thus restrict to planar diagrams with $\ell + 1$ distinct choices for the 't Hooft index loops. For each such choice the determinant in (3.4) is the same as the determinant for the Cartan matrix of $SU(\ell + 1)$, which is $\ell + 1$. In (3.4) this is to be summed over all the roots of the adjoint of $SU(N)$, and the number of such index loop choices is:

$$N(N - 1) \dots (N - \ell) = \frac{N!}{(N - \ell - 1)!},$$

⁵Throughout this discussion we are taking the metric on the Cartan subalgebra to be $g_{\mu\nu} = \delta_{\mu\nu}$.

Substituting these into (3.4) (with $r = N - 1$) we get:

$$\frac{N!}{(N - \ell - 1)!} (\ell + 1) \frac{(N - 1 - \ell)!}{(N - 1)!} S^\ell = N (\ell + 1) S^\ell$$

and there are also the combinatorial factors $F_{0,h}$ of the planar graphs with $h = l + 1$ holes.

In particular if we define $F_0(S) = \sum F_{0,h} S^h$ then this gives rise to the $W = NdF_0/dS$ as expected. Note, however, that this makes sense even if $\ell \geq N$. For $\ell \geq N$ computation of (3.4) is ambiguous: On the one hand the determinant in (3.4) vanishes for $\ell \geq N$; on the other hand $(r - \ell)! = (N - 1 - \ell)!$ if analytically continued is infinite. Thus naively defining the $U(N)$ theory to be part of a sequence of theories indexed by N gives a way to regularize this computation. This was the proposal of [1] motivated from some cases which was realized in string theory. Note that we can define a similar completion for all the classical groups in the same way. Namely, for a given theory for (A, B, C, D) groups, we can view it as a sequence of theories indexed by N . Taking the large N limit we can compute the superpotential to arbitrary high powers of S , and in the end substitute in the coefficient of the monomials the finite value of N in the analytic expressions. This will give a potential F-complete definition of the theory. In the next section we discuss one meaning for this prescription⁶.

Note that the computation above can easily be generalized when we have more than one simple group. One then wants to compute the superpotential $W(S_1, \dots, S_d)$ for the glueball fields S_i of the d simple group factors. The tensor, (3.2) is then only skew on each simple factor, and it decomposes into δ 's as in (3.3) on each such factor. One obtains:

$$\begin{aligned} (\Delta W(S_1, \dots, S_d))_{\mathcal{G}_\ell} &= \sum_{\ell_1 + \dots + \ell_d = \ell} \sum_{\lambda_1, \dots, \lambda_\ell} \sum_{\omega, \omega' \in \frac{S_\ell}{S_{\ell_1} \times \dots \times S_{\ell_d}}} F(\lambda^1, \dots, \lambda^d) \epsilon(\omega) \epsilon(\omega') \\ &\times \prod_{i=1}^d \det_{\ell_i \times \ell_i} (\vec{\lambda}_{\omega(a)}^i \cdot \vec{\lambda}_{\omega'(b)}^i) \frac{(r_i - \ell_i)!}{r_i!} S_i^{\ell_i} \end{aligned} \tag{3.5}$$

where one chooses a partition of ℓ loops into ℓ_i loops and $\epsilon(\omega) = \pm 1$ denote the corresponding permutation factors of the symmetric group $\frac{S_\ell}{S_{\ell_1} \times \dots \times S_{\ell_d}}$.

⁶An interesting test of this for $SO(N)$ can be done as follows: We have computed the correction to glueball superpotential only up to $S^{\lfloor N/2 \rfloor}$ terms, whereas we can, in principle, compute up to S^{N-2} term. Could the extra terms be obtained by embedding at a larger N and analytically continuing the expressions to the smaller N ? This passes the test for the examples studied in [7].

Note that this expression also applies when the choice of the vacuum breaks a gauge symmetry. As discussed in [10] the only new ingredient is that superpotential will include contributions of the charged ghost field as well. One can extend the foregoing computation to include the computation of the $U(1)$ gauge coupling constants as well, as was done for the $U(N)$ in [2], and here we leave it to the reader.

4 A Resolution of Ambiguities for Classical Groups

As noted above, the glueball superpotential is classically ambiguous for powers of the glueball field beyond the dual Coxeter number of the group. In [1] a prescription was given, using the 't Hooft double line notation, to obtain all powers of the glueball field. Let us first review how this is done.

Consider, for concreteness, a $U(N)$ gauge theory with adjoint fields. We can instead consider the same theory, with the same matter content, but with group $U(Nk)$. For low enough powers of S , the dependence of the glueball superpotential on k is simply a multiplicative factor of k . But, in the $U(Nk)$ theory, the glueball superpotential is unambiguous up to S^{Nk} . Thus, by taking k large enough, computing the coefficient of S^l , and dividing it by k , gives a prescription for completing the full perturbative series. This amounts to getting one factor of S for each 't Hooft index loop except one, and not using the classical relation to set $S^h = 0$. It is as if we treat different index loops as giving distinct S 's.

There is an alternative, and we believe more fundamental, way to understand this, which was pointed out in the last reference in [1]: Consider, instead of the $U(N)$ theory, an $\mathcal{N} = 1$ supersymmetric theory based on the non-unitary supergroup $U(N+k|k)$, with the same matter content and superpotential. (See also [22] for another discussion on supergroups in the $\mathcal{N}=1$ context.) This is very natural from the viewpoint of brane/anti-brane system [5], where we have added k such pairs. For this theory, the rank is $N+2k$, so the computation of the glueball superpotential is unambiguous up to S^{N+2k} . However the coefficients will not depend on k at all. This is clear because in each 't Hooft index loop the supertrace gives $(N+k) - k = N$. So there is no k -dependence. In particular the answer is the same as that of $k = 0$. This k -independence (not even a simple multiplicative factor) of the coefficients of the glueball superpotential suggests that we define the $k = 0$ theory in terms of this theory by taking the $k \rightarrow \infty$ limit, which yields unambiguous answer for the superpotential to all orders. This is physically

analogous to saying that the k brane/anti-brane pairs disappear.

We can also apply the supergroup F-completion to SO and Sp theories. For these groups using the embedding into $SO(kN)$ or $Sp(kN)$ will not work as nicely because the superpotential contains terms both linear in k ($g = 0$ contribution) and independent of k (\mathbb{RP}^2 contribution) [23], [24], [25]. But the idea of the supergroup works just as well as it does for $U(N)$: That is, we embed the theory in $SO(N + 2k|k)$ and $Sp(N + k|k)$ and then the k -independence works exactly as it does for $U(N)$. This way of resolving the ambiguity boils down to treating distinct 't Hooft loops as giving rise to distinct S 's, which is the same as computing it for large N , and then analytically continuing the coefficients of each power of the glueball field to finite N .

For $\mathcal{N} = 1$ supersymmetric $G(N) = U(N)$, $SO(N)$ or $Sp(N)$, with a single adjoint, the k -independence can be further justified: The non-unitary theory based on the supergroup $G(N + k|k)$ is equivalent to a unitary one [1]. For example, for $U(N + k|k)$, one can take block diagonal vevs for the adjoint field breaking to $U(N + k) \times U(k)$, with ghost fields which are bifundamental. However, because of the super-group nature of the gauge group, the ghost fields are ordinary rather than fermionic fields. This is the same as the matter content of a *unitary* $\mathcal{N} = 2$ supersymmetric quiver theory based on \widehat{A}_1 . Thus, as far as F -terms are concerned, this non-unitary system is equivalent to that of a unitary theory.⁷

Finally, all of the $G(N + k|k)$ theories have a Higgs branch, where they can be Higgsed down to the original $G(N)$ theory in the IR. Moving the vacuum to be out along this Higgs branch corresponds to moving away the k added brane-antibrane pairs. The IR theory is then the original $G(N)$ theory, along with an approximately decoupled $U(k)$ gauge theory with enhanced supersymmetry. Infinitely far along this Higgs branch, the $U(k)$ gauge dynamics completely decouples, and we are left with the original $G(N)$ theory – up to the possibility of the residual instanton effects that we will discuss and classify. Even at finite distances along the Higgs branch, we expect that no dynamically generated superpotential lifts the moduli space degeneracy, i.e. the superpotential is constant along the Higgs branch, and thus k inde-

⁷Higgsing ordinary gauge theories, e.g. $U(N) \rightarrow \prod U(N_i)$ by the vev of an adjoint field, are also examples where a unitary theory has F-terms which are indistinguishable from those of a non-unitary theory. The original, unitary theory has massive W-bosons in the bi-fundamental representations. This unitary theory is F-equivalent to the gauge fixed version, which is a non-unitary $\prod U(N_i)$ gauge theory, which instead has bi-fundamental ghost fields, viewed as matter fields with the wrong statistics. These ghosts couple to the $U(N_i)$ adjoints via superpotential terms [10]. Again we see that, considering only the F-terms, it is not easy to distinguish a unitary theory from a non-unitary one.

pendent. This is expected because of the enhanced supersymmetry of the low-energy $U(k)$ theory, and because the Higgs branch is the probe realization of the fact that the added brane-antibrane pairs can move around freely in the geometry. We thus view the $G(N+k|k)$ as a particularly natural F-completion of the theory $G(N)$.

In the remainder of this section, we will outline the F-completion of several examples. The reader need not get lost in all of the examples: the main idea is apparent in the $U(N)$ with adjoint example. In the next section, we will further analyze these examples, to determine in precisely which (rare) cases there are residual instanton effects when $G(N+k|k)$ is Higgsed to $G(N)$. In that context, the $Sp(N)$ with anti-symmetric tensor example will also be especially interesting.

4.1 $U(N)$ with an adjoint and general superpotential

Consider $U(N)$ with an adjoint ϕ and superpotential $W = TrW(\phi)$, with $W(\phi) = \sum_{p=1}^{n+1} g_p \phi^p / p$. When the superpotential is just a mass term, we can integrate out the adjoint to obtain pure $U(N)$ $\mathcal{N} = 1$ Yang-Mills in the IR; more generally, the IR gauge group is $\prod_{i=1}^n U(N_i)$ with $\sum_i N_i = N$.

The F-completion of this theory is $U(N+k|k)$, with adjoint Φ and superpotential $W = StrW(\Phi)$. We gauge fix this theory to $U(N+k) \times U(k)$, choosing a gauge where the adjoint Φ takes the form

$$\Phi = \begin{pmatrix} \phi & 0 \\ 0 & \phi' \end{pmatrix}, \quad (4.1)$$

with ϕ a $U(N+k)$ adjoint and ϕ' a $U(k)$ adjoint. Associated with this gauge choice are some bi-fundamental ghost fields which, because of the statistics of the original supergroup, transform as bi-fundamentals of ordinary statistics. We thus obtain the the following *ordinary* \widehat{A}_1 quiver gauge theory:

$$\begin{array}{ccc} & U(N+k) & \times & U(k) \\ \phi & Ad & & \cdot \\ \phi' & \cdot & & Ad \\ Q_i & \square & & \overline{\square} \\ \tilde{Q}_i & \overline{\square} & & \square, \end{array} \quad (4.2)$$

with $i = 1, 2$. The superpotential is

$$W = TrW(\phi) - TrW(\phi') + \sqrt{2} \sum_i \left(Tr Q_i \phi \tilde{Q}_i - Tr \tilde{Q}_i \phi' Q_i \right), \quad (4.3)$$

with the relative sign between the ϕ and ϕ' terms coming from that of the Str of $U(N+k|k)$. The cubic interaction in (4.3) between the ghosts Q_i and \tilde{Q}_i and the adjoints ϕ and ϕ' arises in the standard Faddeev-Popov procedure, as discussed in [10]. The $\sqrt{2}$, which is henceforth not explicitly written, is the $\mathcal{N} = 2$ value for the coefficient and can be absorbed into the normalization of the fields, rescaling the mass and other superpotential couplings.

When $W = \frac{1}{2}m\Phi^2$, we can integrate out the adjoints to obtain

$$W = -\frac{1}{m}\text{Tr}(Q_1\tilde{Q}_1Q_2\tilde{Q}_2 - Q_1\tilde{Q}_2Q_2\tilde{Q}_1). \quad (4.4)$$

This is the system studied in [26], [27], corresponding to k D3's and N wrapped D5's on the resolved conifold. This theory has duality cascades [27], where k shifts by N units at a time. (As do the generalizations with general higher order superpotential [28], [29].) This can be used to relate the $k = 0$ and $k = \infty$ theories. So the k independence of the original $U(N+k|k)$ theory, which was a simple consequence of the relative sign in the Str , is related to Seiberg duality in the gauge fixed version 4.2. Further, for large k the theory is almost conformal. So resolving the ambiguities of the glueball superpotential by embedding the $k = 0$ theory in the $k = \infty$ theory is related to fixing the UV completion by starting from a perturbed conformal fixed point.

Independent of the cascade, reducing the k by one unit at a time also makes sense physically: It corresponds to going to a particular Higgs branch of the theory (equivalent to removing brane/anti-brane pairs with a net 3-brane charge of 1 unit). For example we can take $(Q_1)_c^{c'} = (\tilde{Q}_2)_c^{c'} = v_c\delta_c^{c'}$, with $\phi = \phi' = 0$. This is a solution of the D and F term equations (c is a $U(N+k)$ fundamental index and c' is a $U(k)$ fundamental index). For example, if $v_c = v$ for $c = 1, \dots, \ell$, and zero otherwise, we Higgs $U(N+k) \times U(k)$ to $U(N+k-\ell) \times U(k-\ell)$, along with an approximately decoupled $U(\ell)$, $\mathcal{N} = 4$ theory (which does not contribute to the F terms), corresponding to pulling away ℓ units of 3-brane charge.

The F terms of the $U(N+k|k)$ theory, or its gauge fixed version (4.2) are independent of the location on the foregoing Higgs branch, which is to say that they are independent of k . So it is reasonable to compute F terms for the original, $k = 0$, $U(N)$ theory in terms of the large k $U(N+k|k)$ theory. However, as will be discussed in the next section, this F completion of $U(N)$ into $U(N+k|k)$ can differ, in some rare instances, from the more standard UV completion. The potential differences arise at the last steps of Higgsing, for small k .

Now consider the dynamical scales of the foregoing theories. The original $U(N)$ theory has a scale Λ , with the instanton factor Λ^{2N} . We take the $U(N+k|k)$ theory to have this same scale Λ , so that is the scale of both $U(N+k)$ and $U(k)$ upon gauge fixing. The $U(N+k)$ theory has an adjoint and $N_f = 2k$ flavors, so its instanton factor is $\Lambda_1^{2(N+k)-2k} = \Lambda^{2N}$, which we equate with the instanton factor in the original $U(N)$. On the other hand, the $U(k)$ factor has $N_f = 2(N+k)$, so its instanton factor is $\Lambda^{2k-2(N+k)} = \Lambda^{-2N}$, which is inverse to the original $U(N)$ instanton factor. This relative sign is related to that of the beta functions, coming ultimately from the relative sign in the supertraces.

4.2 $SO(N)$ with a single adjoint ϕ and superpotential $TrW(\phi)$.

By considering an orientifold of the foregoing $U(N+k) \times U(k)$ theory, one sees that the natural F-completion of $SO(N)$ is $SO(N+2k) \times Sp(k)$. (We label Sp groups by their rank, so that $Sp(1) \cong SU(2)$.) This group is indeed the gauge-fixed version of the supergroup $SO(N+2k|k)$ ⁸. So our F-completion of $SO(N)$ is $SO(N+2k|k)$ with an adjoint Φ and superpotential $StrW(\Phi)$. We can gauge fix $SO(N+2k|k)$ to $SO(N+2k) \times Sp(k)$ by choosing the adjoint as 4.1. This gauge choice requires Faddeev Popov ghosts which are, as before, ordinary superfields. Because the $SO(N+2k|k)$ adjoint Φ satisfies a reality condition, so do the ghosts. This implies that, rather than two hypermultiplets, as in 4.2, the ghosts are now two half-hypermultiplets. So the gauge fixed version of our F-completion is the following *ordinary* gauge theory:

$$\begin{array}{ccc}
 SO(N+2k) & \times & Sp(k) \\
 \phi & Ad & \cdot \\
 \phi' & \cdot & Ad \\
 Q_i & \square & \square
 \end{array} \tag{4.5}$$

with $i = 1, 2$. The superpotential is

$$W = TrW(\phi_1) - TrW(\phi_2) + \sum_{i=1}^2 \left(TrQ_i \phi \tilde{Q}_i - Tr\tilde{Q}_i \phi' Q_i \right), \tag{4.6}$$

where we define $(\tilde{Q}_i)_{c'}^c = \delta^{cd} J_{c'd'} (Q_i)_d^{d'}$, which is the reality condition for half-hypermultiplets, with c, d SO fundamental indices and c', d' the Sp fundamental indices.

⁸We write $SO(N|M)$ for the supergroup with bosonic part $SO(N) \times Sp(M)$. The names $B(n, m)$ and $D(n, m)$ are sometimes used for what we call $SO(2n+1|m)$ and $SO(2n|m)$, respectively.

When $W = \frac{1}{2}m\Phi^2$, we can integrate out the adjoints to obtain

$$W = -\frac{1}{m}\text{Tr}(Q_1\tilde{Q}_1Q_2\tilde{Q}_2 - Q_1\tilde{Q}_2Q_2\tilde{Q}_1). \quad (4.7)$$

The $SO(N + 2k)$ theory has $N_f = 4k$ vectors and the $Sp(k)$ theory has $N_f = N + 2k$ flavors. There is a duality cascade where $SO(N + 2k)$ group gets strong and dualized [30] into $SO(2k - N + 4)$. Then the $Sp(k)$ gets strong and is dualized [31] into $Sp(k - N + 2)$. After these two steps, we get back the same theory with $k \rightarrow k - N + 2$. This cascade, and its string theory realization, was discussed in [32]. Exactly the same cascade occurs for theory (4.10) with adjoint and general tree-level superpotential; this is seen via a deformation of the dualities discussed in [33], [34].

In terms of orientifolding N wrapped D5's and k D3's on the conifold, the gauge theory would naturally be similar to 4.10, but with gauge group $SO(N + 2k) \times Sp(k - 1)$. (For example, this theory is superconformal for $N = 0$, whereas (4.10) is not.) In the large k limit, these two F-completions are, of course, completely equivalent. But we prefer 4.10, because it recovers the original $SO(N)$ theory for $k = 0$.

Independent of the cascades, the theory (4.10) has a Higgs branch, along which we can successively reduce $k \rightarrow k - 1$. E.g. we can take $(Q_1)_c' = (\tilde{Q}_2)_c' = v_c\delta_c'$; taking ℓ equal non-zero v_c , we get a similar theory, with $k \rightarrow k - \ell$, along with an approximately decoupled $U(\ell)$ $\mathcal{N} = 4$ theory. The F terms are independent of the location of the theory along this Higgs branch. So the large k $SO(N + 2k|k)$ theory indeed provides an F-completion of the original $SO(N)$ theory.

Again, we take the $SO(N + 2k|k)$ theory to have the same scale, Λ , as the original $SO(N)$ theory. For all k the $SO(N + 2k)$ gauge theory in (4.10) has instanton factor $\Lambda^{2(N+2k-2)-4k} = \Lambda^{2(N-2)}$, which is that of the original $SO(N)$ theory. On the other hand, the $Sp(k)$ group in (4.10) has instanton factor $\Lambda^{2(k+1)-(N+2k)} = \Lambda^{-(N-2)}$, which is the inverse square-root of the original $SO(N)$ instanton factor. This will play a small role in the discussion in the following section.

4.3 $Sp(N)$ with a single adjoint ϕ and superpotential $W(\phi)$.

We write the supergroup F-completion as $Sp(N + k|k)$, which denotes the same group as $SO(2k|N + k) = D(k, N + k)$ in the above notation. There is an $Sp(N + k|k)$ adjoint Φ and superpotential $StrW(\Phi)$. Upon gauge fixing,

we obtain the ordinary gauge theory

$$\begin{array}{ccc}
 & Sp(N+k) & \times & SO(2k) \\
 \phi & Ad & & \cdot \\
 \phi' & \cdot & & Ad \\
 Q_i & \square & & \square
 \end{array} \quad (4.8)$$

with $i = 1, 2$. The superpotential is

$$W = \text{Tr}W(\phi) - \text{Tr}W(\phi') + \sum_{i=1}^2 \left(\text{Tr}Q_i\phi\tilde{Q}_i - \text{Tr}\tilde{Q}_i\phi'Q_i \right). \quad (4.9)$$

As in the previous subsection, the fields Q_i , which arise as the ghosts in the gauge fixing, satisfy a half-hypermultiplet reality condition, $\tilde{Q}_i^c = J^{cd}\delta_{c'd'}Q_i^{d'}$, where c, d are Sp fundamental indices and c', d' are SO fundamental indices. When $W = \frac{1}{2}m\Phi^2$, we can integrate out the adjoints to obtain the quartic superpotential 4.7.

As mentioned in the previous section, in an orientifold the group would actually naturally be $Sp(N+k) \times SO(2k+2)$. Again, this would be equivalent to (4.8) in the large k limit, but we prefer (4.8) because it reduces to our original gauge theory for $k = 0$.

The foregoing theory, with any $W(\phi)$, undergoes a RG cascade. First, the $Sp(N+k)$ gets strong and is dualized to $Sp(k-N-2)$; next, the $SO(2k)$ gets strong and is dualized to $SO(2k-4N-4)$. After these two steps, we are back to a similar theory, with $k \rightarrow k-2N-2$. This cascade can be regarded as the continuation, to negative N , of that of the previous section.

Again, independent of the cascade, and for all $W(\phi)$, there is a Higgs branch where we can successively reduce $k \rightarrow k-1$. The F terms are independent of the location of the vacuum on this Higgs branch moduli space of vacua. So the large k $Sp(N+k|k)$ theory is indeed an F-completion of the original $Sp(k)$ theory.

We take the $Sp(N+k|k)$ theory to have the same scale, Λ , as the original $Sp(N)$ theory. The $Sp(N+k)$ theory has instanton factor $\Lambda_1^{2(N+k+1)-2k} = \Lambda^{2(N+1)}$, which is that of the original $Sp(N)$ theory. The $SO(2k)$ has $4(N+k)$ vectors, so its instanton factor is $\Lambda_2^{2(2k-2)-4(N+k)} = \Lambda^{-4(N+1)}$, which is the inverse square of the $Sp(N)$ instanton factor.

4.4 $SO(N)$ with a symmetric tensor ϕ and superpotential $W(\phi)$

The F-completion is again $SO(N + 2k|k)$, with a two-index tensor with $\Phi = \Phi^T$, as opposed to $\Phi = -\Phi^T$ for the adjoint (transpose is defined with the usual sign in the upper block and the opposite in the lower), and superpotential $W = StrW(\Phi)$. Gauge

$$\begin{array}{ccc}
 SO(N + 2k) & \times & Sp(k) \\
 \phi & \square & \cdot \\
 \phi' & \cdot & \square \\
 Q_i & \square & \square
 \end{array} \tag{4.10}$$

with $i = 1, 2$. The gauge fixing procedure yields the tree-level superpotential

$$W = TrW(\phi) - TrW(\phi'J) + \epsilon^{ij} \left(Tr\tilde{Q}_i\phi Q_j + TrQ_i\phi'\tilde{Q}_j \right), \tag{4.11}$$

with repeated indices summed and $\tilde{Q}_i \equiv Q_i^T J$, and $J^{[c'd']}$ is the $Sp(k)$ defining symplectic tensor.

When the superpotential is a mass term, we can integrate out ϕ and ϕ' to obtain

$$W = -\frac{1}{m} Tr(Q_1\tilde{Q}_1Q_2\tilde{Q}_2 - Q_1\tilde{Q}_2Q_2\tilde{Q}_1). \tag{4.12}$$

For the general theory with superpotential $W(\Phi) = Str\Phi^{p+1}$ there is a duality cascade, which can be obtained by a deformation (corresponding to the cubic interactions in 4.11) of that discussed in [35], [34]. Unlike the above cases, the groups in the cascade now depend on p . First the SO gets strong and is dualized to $SO(2k - N + 4p)$, then the Sp group gets strong and is dualized to $Sp(k - N + 2p)$; after these two steps we are back to a theory which is similar to the original theory, but with $k \rightarrow k - N + 2p$.

Again, independent of the cascade reduction of k , there is a Higgs branch, along which we can successively reduce $k \rightarrow k - 1$. For example, taking $\langle Q_1 \rangle = diag(v_1, v_1, v_2, v_2, \dots, v_l, v_l, 0, 0 \dots)$, with all other entries zero and with $\langle Q_2 \rangle = \langle \phi \rangle = \langle \phi' \rangle = 0$ solves the D and F term equations. Along this Higgs branch we reduce to a similar theory, with $k \rightarrow k - \ell$, together with an approximately decoupled $\mathcal{N} = 2 U(\ell)$ theory. The exact F terms are again expected to be independent of the location of the vacuum on this Higgs branch; showing that the large k $SO(N + 2k|k)$ theory is indeed a sensible F-completion of the $k = 0$ theory.

We take $SO(N + 2k|k)$ to have the same scale Λ as the original $SO(N)$ theory. For all k , the $SO(N + 2k)$ gauge group has instanton factor

$\Lambda^{3(N+2k-2)-(N+2k+2)-4k} = \Lambda^{2N-8}$, which is the instanton factor in the original $SO(N)$. The $Sp(k)$ has instanton factor $\Lambda^{3(k+1)-(k-1)-N-2k} = \Lambda^{4-N}$, which is the inverse square-root of the $SO(N)$ instanton factor.

4.5 $Sp(N)$ with an antisymmetric tensor ϕ and superpotential $W(\phi)$.

This can be regarded as the being the continuation of the previous example to negative N . The F-completion is $Sp(N + k|k)$, and gauge fixing yields the ordinary gauge theory

$$\begin{array}{ccc}
 & Sp(N + k) & \times & SO(2k) \\
 \phi & \square & & \cdot \\
 \phi' & \cdot & & \square \\
 Q_i & \square & & \square
 \end{array} \tag{4.13}$$

$i = 1, 2$, with superpotential

$$W = \text{Tr}W(\phi J) - \text{Tr}W(\phi') + \epsilon^{ij} \left(\text{Tr}\tilde{Q}_i \phi Q_j + \text{Tr}Q_i \phi' \tilde{Q}_j \right), \tag{4.14}$$

with $\tilde{Q}_i \equiv Q_i^T J$, and J^{cd} is the $Sp(N + k)$ symplectic tensor.

Again, this theory has a RG cascade where, after two steps, we return to the same theory with $k \rightarrow k - 2N - 2p$ for superpotential $W(\Phi) = \text{Str}\Phi^{p+1}$. Also, independent of the cascade, there is a Higgs branch where we can successively reduce $k \rightarrow k - 1$. An example direction in this Higgs branch of vacua is $\langle Q_1 \rangle = \text{diag}(v_1, v_1, v_2, v_2, \dots, v_\ell, v_\ell, 0, 0\dots)$, with all other entries zero and with $\langle Q_2 \rangle = 0$ and $\langle \phi \rangle = \langle \phi' \rangle = 0$. Along this direction we Higgs to a similar theory, with $k \rightarrow k - \ell$, along with an approximately decoupled $\mathcal{N} = 2$ $U(\ell)$ theory, having Coulomb branch moduli $(v_1, \dots v_\ell)$. The F terms are expected to be independent of the location of the vacuum along this Higgs branch.

The $Sp(N + k)$ has instanton factor $\Lambda^{3(N+k+1)-(N+k-1)-2k} = \Lambda^{2N+4}$, which is identified with the instanton factor for the original $Sp(N)$ theory with anti-symmetric tensor. The $SO(2k)$ has instanton factor $\Lambda^{3(2k-2)-(2k+2)-4(N+k)} = \Lambda^{-4N-8}$, which is the inverse square of the original $Sp(N)$ instanton factor.

4.6 $U(N) \mathcal{N} = 1^*$

The F-completion of the $U(N) \mathcal{N} = 1^*$ theory is a $U(N + k|k)$ with three adjoints Φ_I , $I = 1, 2, 3$, and superpotential

$$W = S\text{tr}[\sum_{IJK} (\Phi_I[\Phi_J, \Phi_K])\epsilon^{IJK} + \frac{1}{2}m \sum_I \Phi_I^2]. \tag{4.15}$$

We write the Φ_I as:

$$\Phi_i = \begin{pmatrix} \phi_i & R_i \\ \tilde{R}_i & \phi'_i \end{pmatrix}, \quad i = 1, 2, \quad \Phi_3 = \begin{pmatrix} \phi_3 & 0 \\ 0 & \phi'_3 \end{pmatrix}, \tag{4.16}$$

where the zeros in Φ_3 are our gauge choice in breaking $U(N + k|k)$ to $U(N + k) \times U(k)$. The bi-fundamentals R_i and \tilde{R}_i are fermionic. Also, as before, the Faddeev-Popov ghosts are ordinary matter bi-fundamentals Q_i and \tilde{Q}_i ; so the gauge fixed gauge group and matter content is the non-unitary theory (because of the R_i and \tilde{R}_i):

	$U(N + k)$	\times	$U(k)$	
ϕ_I	Ad		\cdot	
ϕ'_I	\cdot		Ad	
Q_i	\square		$\overline{\square}$	(4.17)
\tilde{Q}_i	$\overline{\square}$		$\square,$	
R_i	\square		$\overline{\square}$	
\tilde{R}_i	$\overline{\square}$		$\square,$	

with $I = 1, 2, 3$ and $i = 1, 2$. The superpotential (4.15) yields

$$W = \sum_{IJK} \epsilon^{IJK} [\text{Tr}(\phi_I[\phi_J, \phi_K]) - \text{Tr}(\phi'_I[\phi'_J, \phi'_K])] + \sum_I \frac{m}{2} (\text{Tr}\phi_I^2 - \text{Tr}\phi'^2_I) + \sum_i m R_i \tilde{R}_i + \text{Tr}\phi_3(Q_i \tilde{Q}_i + R_1 \tilde{R}_2 - R_2 \tilde{R}_1) - \text{Tr}\phi'_3(\tilde{Q}_i Q_i + \tilde{R}_1 R_2 - \tilde{R}_2 R_1). \tag{4.18}$$

The fermionic fields R_i and \tilde{R}_i are massive, as are the adjoints ϕ_I and ϕ'_I . The massless spectrum consists of just the bi-fundamentals Q_i and \tilde{Q}_i with a quartic superpotential. This is the same massless spectrum as for the F-completion of $\mathcal{N} = 1$ super-Yang-Mills, as expected. We need to keep the effects of the massive fields in (4.17) to see the difference between $\mathcal{N} = 1^*$ and $\mathcal{N} = 1$.

Once again, the full $U(N + k|k)$ F-completion of $\mathcal{N} = 1^*$ has a Higgs branch, along which we effectively reduce k one unit at a time. Again, the superpotential is expected to be independent of the location of the vacuum along this Higgs branch, implying that the different Higgs branches, with different values of k , all have the same superpotential.

4.7 Classical groups with *arbitrary massive representations*.

We will assume that there are no baryon operators in the superpotential; if there is one can presumably treat it using ideas similar to [36]. We will use analytic continuation as a regularization scheme: Compute the coefficient of S^k for fixed k as a function of N for $k < N$. This computation is unambiguous; one can then analytically continue the coefficient to $k > N$. For representations obtained from tensor products of fundamental representations (*e.g.*, not the spinorial representations of the SO groups) this is equivalent to considering multi-line Feynman graphs and putting only up to two gluino fields per index loop and treating each index loop independently of the other. Again, a justification/meaning of this regularization scheme is to embed the theory in the corresponding supergroup, $G(N + k|k)$ and take $k \rightarrow \infty$ to remove the ambiguities. This prescription is equivalent to defining the $G(N)$ theory as a Higgs branch of the arbitrarily large rank group $G(N + k|k)$.

4.8 Prescription for Non-Classical Groups (E, F, G) .

For non-classical groups (E, F, G) there is no obvious canonical choice for the resolution of the ambiguities in the glueball superpotential. In some cases, with certain matter content, there are phases where the gauge group breaks to subgroups involving only classical groups, in which case we can write the theory in terms of the broken factors and use the above prescription to remove the ambiguities. For example, this is true for the G_2 example studied in [37],[38], where there is a phase where G_2 breaks to $SU(2)$. In some other cases “analytic continuation” in groups suggests an answer as we will see in the next section where we discuss the $\mathcal{N} = 1^*$ theory for arbitrary simply-laced groups.

5 Comparison with Other UV Completions

As seen in the previous section, $G(N + k|k)$ is a sensible F-completion of $G(N)$ because the $G(N + k|k)$ theory always has a Higgs branch⁹, where we can successively reduce $k \rightarrow k - 1$. In this way, we can Higgs from $G(N + k|k)$ back down to $G(N)$, and the exact superpotential is expected to be independent of the location of the vacuum on this Higgs branch. However,

⁹If necessary one can always add a very massive adjoint to achieve this, without affecting the IR dynamics of the theory.

as we discuss in this section, the F-terms obtained in this way for the $G(N)$ theory can differ from that of the standard UV completion of $G(N)$. In terms of our Higgsing from $G(N+k|k)$ to $G(N)$, this is seen as residual instanton effects, associated with the broken part of the group, $G(N+k|k)/G(N)$.

Suppose that a certain theory has gauge group \mathcal{G} in the extreme UV, but the group is Higgsed at some high scale to a subgroup \mathcal{H} in the IR. Usually, at least when the Higgsing scale is taken to the extreme UV, the IR dynamics can be understood purely in terms of that of the IR theory \mathcal{H} , without having to include additional interactions, which require knowing about the original UV completion \mathcal{G} . We refer to this as “decoupling”. For example, for $SU(N_c)$ SQCD with $N_f < N_c - 1$, the superpotential [39] for vacua far out along the classical moduli space (as well as exactly) simply comes from gaugino condensation in the low energy theory with $\mathcal{H} = SU(N_c - N_f)$. One known exception to this decoupling intuition are the Wess-Zumino terms, which can be necessary to add to the IR theory to account for any discrepancies between the symmetries or anomalies of the UV and IR theories. For F-terms, there is only one known exception to the decoupling intuition: instantons in the broken part, \mathcal{G}/\mathcal{H} , of the UV group \mathcal{G} . For example, this is how the superpotential arises for SQCD with $N_f = N_c - 1$, where the IR group \mathcal{H} is trivial. Another example is $\mathcal{N} = 2$ SYM, where \mathcal{H} is the Cartan subalgebra, and instantons in \mathcal{G}/\mathcal{H} are essential to understanding the IR dynamics of the Coulomb branch [40].

Instantons in partially broken UV groups can appear to violate decoupling intuition, since the particular UV completion can affect the IR results. For example, $\mathcal{N} = 1$, $\mathcal{H} = SU(2)$, super-Yang-Mills has a UV completion, based on $\mathcal{G} = SU(2) \times SU(2)$ which is broken to \mathcal{H} at an arbitrarily high scale in the UV, where instanton effects in the broken UV group \mathcal{G}/\mathcal{H} are of the same size as the leading non-perturbative effects in the IR $SU(2)$, and can cancel the IR \mathcal{H} gaugino condensation superpotential [41], [30]. (In this case, the instantons in the broken \mathcal{G}/\mathcal{H} actually compete with the fractional instantons in \mathcal{H} . This is unusual, since generally \mathcal{G}/\mathcal{H} instantons are similar to whole \mathcal{H} instantons; here it’s fractional because of how the IR group \mathcal{H} is diagonally embedded in the UV group \mathcal{G} .)

When can such instantons in the partially broken group \mathcal{G}/\mathcal{H} potentially contribute? A standard lore about instantons, which is borne out in the exact results for supersymmetric gauge theories, is that they never contribute to the effective action unless the gauge group is sufficiently Higgsed¹⁰. The Higgsing is needed to regulate the divergence in the integral over instanton

¹⁰Technically, instantons don’t exist in this case – one has to consider constrained instantons.

size¹¹. What “sufficiently Higgsed” means is that \mathcal{G} instantons can contribute only if the Higgsed part \mathcal{G}/\mathcal{H} has $\pi_3(\mathcal{G}/\mathcal{H}) \neq 0$. See [42] for a nice discussion.

Applying these ideas to our F-completion, we have $\mathcal{G} = G(N + k|k)$ and $\mathcal{H} = G(N)$. Our F-completion $G(N + k|k)$ of $G(N)$ can include residual instanton F-term contributions, coming from $G(N + k|k)$, which are not present in the standard UV completion of the $G(N)$ theory. Such residual instanton effects can arise when $\pi_3(G(N + k|k)/G(N)) \neq 0$. As we will discuss, this is a necessary, but not sufficient, condition for such effects. In a nutshell, the $\pi_3(\mathcal{G}/\mathcal{H}) \neq 0$ condition here implies that \mathcal{G} has to have some $SU(2)$ factor, which is either completely broken, or broken to $U(1)$, in \mathcal{H} . The additional condition is on the number of massless flavors in this broken $SU(2)$. The usual condition for an $SU(2)$ instanton contribution to the superpotential, based on a zero-mode analysis, is that the $SU(2)$ must have precisely $N_f = 1$ flavor. For our examples, because of a tree-level superpotential, the relevant number of $SU(2)$ flavors for an instanton contribution is instead $N_f = 2$ or $N_f = 3$ massless flavors. These are the only two possibilities; for example, there is no superpotential contribution for $N_f = 4$ massless flavors.

To give an example of the kind of effect which these residual instanton contributions from $G(N + k|k)$ have, consider the glueball superpotential for pure supersymmetric Yang-Mills:

$$W = hS \left(\log \left(\frac{S}{\Lambda^3} \right) - 1 \right). \quad (5.1)$$

The standard gauge theory (“sgt”) glueball superpotential for the classical gauge groups is given by (5.1) with

$$\begin{aligned} U(N) : & \quad h_{sgt} = N - \delta_{N,1} \\ Sp(N) : & \quad h_{sgt} = N + 1 - \delta_{N,0} \\ SO(N) : & \quad h_{sgt} = N - 2 + \delta_{N,1} + 2\delta_{N,0}, \end{aligned} \quad (5.2)$$

where the $\delta_{N,*}$ are some exceptions, corresponding, for example, to the fact that standard $U(1)$ gauge theory does not have any non-perturbative effects which could lead to photino condensation and an associated non-zero superpotential. There are similar exceptions for the trivial groups $Sp(0)$, $SO(1)$ and $SO(0)$.

¹¹Alternatively, this divergence can be regulated by considering correlation functions at separated points, in which case no Higgsing is needed. This is how instantons contribute to the glueball h -point function $\langle S^h \rangle$ in supersymmetric Yang-Mills, where there is no Higgsing of the gauge group. We will only consider superpotential terms, in which case the sufficient Higgsing is required for an instanton contribution.

On the other hand, our $G(N + k|k)$ F-completion can not have any such exceptions, for any N values, since we can avoid the low N exceptions by making k sufficiently large. The N dependence of the $G(N + k|k)$ F-completion must be completely smooth! For example, the F-completion for pure Yang-Mills gives the glueball superpotential (5.1) with

$$\begin{aligned} U(N) : & \quad h = N \\ Sp(N) : & \quad h = N + 1 \\ SO(N) : & \quad h = N - 2, \end{aligned} \tag{5.3}$$

for all $N \geq 0$, with no exceptions. In particular, we find non-trivial glueball superpotentials for $U(1)$, $Sp(0)$, $SO(1)$, and $SO(0)$. As we will discuss, the difference between (5.2) and (5.3) can be understood as residual instanton effects, with the effect of the residual instantons being precisely that needed to smooth out the exceptions. For example, the glueball superpotential for $U(1)$ can be seen in its $U(1 + k|k)$ F-completion, which arises as a residual instanton effect when we Higgs $U(2|1)$ to $U(1)$. Another UV completion of $U(1)$ gauge theory that leads to the same non-zero superpotential of 5.1, with $h = 1$, is non-commutativity. For this, the instanton leading to the non-zero W can be seen directly in the non-commutative $U(1)$ gauge theory.

We now discuss various examples. The main idea is illustrated in the $U(N)$ case, so the reader can feel free to skip the later subsections. One highlight, though, is the case of $Sp(N)$ with an anti-symmetric tensor, discussed in sect. 5.4. For that theory, the residual instanton effects of our F-completion play an especially prominent role.

5.1 The $U(N + k|k)$ F-completion of $U(N)$ with an adjoint

The $U(N + k|k)$ F-completion was discussed in section 4.1. We first consider pure $\mathcal{N} = 1$ $U(N)$ Yang-Mills, which we get from the theory with an added adjoint by taking the superpotential to be a large mass term. The low energy theory is then the \widehat{A}_1 quiver $U(N + k) \times U(k)$ theory, with no adjoints, and with a quartic low-energy superpotential 4.4, which we write as

$$W = -\frac{1}{m} \text{Tr}(M_{11}M_{22} - M_{12}M_{21}), \tag{5.4}$$

with $(M_{ij})_a^b = (Q_i)_{a'}^{b'} (\widetilde{Q}_j)_{b'}^a$.

When we Higgs $k \rightarrow k - 1$, there can only be a residual instanton effect if $\pi_3(U(N + k)/U(N + k - 1)) \neq 0$ or $\pi_3(U(k)/U(k - 1)) \neq 0$. This only happens when $N + k = 2$ or when $k = 2$, i.e. when a $U(2)$ factor is Higgsed to $U(1)$. Even for these cases, there is generally no residual instanton effect: if the

$U(2)$ has $N_f > 3$ massless flavors, the classical moduli space is unmodified and there is a supersymmetric vacuum of the theory with superpotential (5.4) with zero contribution to the superpotential. There is only the possibility of a residual instanton effect when $N_f \leq 3$.

Noting that the $U(N+k)$ has $N_f = 2k$ and the $U(k)$ has $N_f = 2(N+k)$, we see that for $k = 2$ one has $N_f = 2N + 4$ massless flavors, which is too many for a residual instanton contribution for any $N \geq 0$. So the only possibility for a residual instanton contribution is in the $U(N+k)$ factor, which requires $N+k = 2$ and $0 < 2k < 4$, *i.e.* $N = k = 1$. To summarize, there is only the possibility of a residual instanton effect for $U(N)$ and that is when $N = 1$, where the residual instanton can only contribute at the very last stage in the $k \rightarrow k-1$ Higgsing of $U(N+k|k)$, namely $U(2|1) \rightarrow U(1)$.

So we consider the $U(2) \times U(1)$ theory with $N_f = 2$ bi-fundamentals and tree level superpotential 5.4. Since $U(2)$ has $N_f = 2$ flavors there is an instanton effect, which leads to a quantum modified moduli space constraint [43], $\det M - B\tilde{B} = \Lambda_L^4$, where Λ_L is the scale of the low energy $U(2)$ gauge theory, in which the adjoint has been integrated out. The matching relation to the scale Λ of the high energy $U(2)$ theory which includes the adjoint of mass m is $\Lambda_L^4 = m^2\Lambda^2$. We write the superpotential as

$$W = S \log \left(\frac{\det M - B\tilde{B}}{m^2\Lambda^2} \right) - \frac{1}{m} \det M, \quad (5.5)$$

where the glueball field S term is to enforce the quantum moduli space constraint, and the other term is 4.4.

We now integrate out the fields M and B by their equations of motion

$$SM^{-1} = \frac{1}{m} M^{-1} \det M, \quad SB = S\tilde{B} = 0. \quad (5.6)$$

There are two branches of vacua. One is an isolated vacuum at $\langle M \rangle = 0$, $\langle S \rangle = 0$, and $\langle B\tilde{B} \rangle = -\Lambda^4$. This is the wrong solution for our purposes, because it does not have the Higgs branch connecting to our original $U(1)$ theory in the IR. The relevant solution, which does have the expected Higgs branch, is

$$\langle M_{11}M_{22} - M_{12}M_{21} \rangle = m\langle S \rangle, \quad \text{and} \quad \langle B \rangle = \langle \tilde{B} \rangle = 0. \quad (5.7)$$

The $\langle M_{ij} \rangle$ of (5.7) is a three complex dimensional moduli space of vacua, which is the deformed conifold geometry, with $m\langle S \rangle$ giving the deformation. This is the expected moduli space of the added D3 brane, probing the geometry of the string theory realization of this theory. Everywhere on this

moduli space, even infinitely far from the origin, where $U(2) \times U(1)$ is Higgsed to our original $U(1)$ theory (along with an approximately decoupled, $\mathcal{N} = 4$, $U(1)$ theory) in the extreme UV, there is the constant superpotential for the glueball superfield

$$W = S \left(\log \left(\frac{S}{\Lambda_{U(1)}^3} \right) - 1 \right), \tag{5.8}$$

where $\Lambda_{U(1)}^3 = m\Lambda^2$ is naturally regarded as the scale of the low energy $U(1)$ pure Maxwell theory, after integrating out the massive adjoint. Thus our F-completion of $U(1)$ has the superpotential 5.1, with $h = 1$. If we integrate out S , we obtain $\langle S \rangle = \Lambda_{U(1)}^3$ and the constant superpotential along the above Higgs branch

$$W_{low} = -\Lambda_{U(1)}^3, \tag{5.9}$$

coming from the instanton in the broken $U(2)$ of the extreme UV completion.

Let us compare this with $U(N + k|k)$ F-completion of $U(N)$ for $N > 1$. Consider, in particular, the gauge fixed $U(N + k) \times U(k)$ theory (4.2) for $N > 1$ and $k = 1$. The superpotential is

$$W = S \left[\log \left(S^{N-1} \det_{ij} M_{ij} * over * * m^{N+1} \Lambda^{2N} \right) - (N - 1) \right] - \frac{1}{m} \det M, \tag{5.10}$$

where $i = 1, 2$. The first term in (5.10) is dynamically generated by the $U(N + 1)$ theory with $N_f = 2$ flavors and the second is the tree-level term 4.4. The M_{ij} equation of motion is $SM^{-1} = \frac{1}{m}M^{-1} \det M$, as in 5.6. The solution is as in 5.7: $\langle \det M \rangle = mS$, giving the expected deformed conifold moduli space of vacua, along which the gauge group can be Higgsed to the original $U(N)$ theory (along with some decoupled additional fields) at an arbitrarily high scale in the UV. Along this entire moduli space, the superpotential has the constant value:

$$W = S \left[\log \left(\frac{S^N}{m^N \Lambda^{2N}} \right) - N \right], \tag{5.11}$$

which, is precisely the expected gaugino condensation superpotential for the $U(N)$ theory obtained in the IR, out along the Higgs branch. The IR $U(N)$ theory has scale $\Lambda_{U(N)}$ given by $\Lambda_{U(N)}^{3N} = m^N \Lambda^{2N}$, upon integrating out the massive adjoint. No residual instanton contribution is present or needed for $N > 1$, since (5.11) can be seen directly in terms of the IR $U(N)$ theory for $N > 1$.

The upshot is that the residual instanton term only arises for $U(N)$ with $N = 1$, and it is precisely such that it eliminates the $\delta_{N,1}$ in 5.2, yielding the

superpotential (5.1) with h given by 5.3, which is smooth for all $N \geq 0$. We note, as a special case, that there is no residual instanton superpotential for $N = 0$, since then the UV completion is $U(k) \times U(k)$ and, for the relevant case of Higgsing $k = 2$ to $k = 1$, each $U(2)$ has $N_f = 4$ massless flavors, which is too many to lead to a residual superpotential. This agrees with the fact that (5.2) is already smooth for $N = 0$.

Now consider the $U(N)$ theory with adjoint ϕ and higher order superpotential $W = \text{Tr}W(\phi)$, e.g. $W(\phi) = \frac{g}{3}\phi^3 + \frac{m}{2}\phi^2$. Since $W'(z) = gz(z - m/g)$, the gauge group is Higgsed in the various vacua as $U(N_1 + N_2) \rightarrow U(N_1) \times U(N_2)$. Our F-completion of this is $U(N_1 + N_2 + k_1 + k_2|k_1 + k_2) \rightarrow U(N_1 + k_1|k_1) \times U(N_2 + k_2|k_2)$. Gauge fixing, the relevant IR gauge theory is a $U(N_1 + k_1) \times U(N_2 + k_2) \times U(k_1) \times U(k_2)$ quiver gauge theory, with two bi-fundamentals and two anti-bifundamentals connecting every pair of gauge groups. These bi-fundamentals come from the various ghosts, with statistics, and masses given by

Groups connected	statistics	mass	
$U(N_1 + k_1)$ and $U(k_1)$	ordinary	massless	
$U(N_2 + k_2)$ and $U(k_2)$	ordinary	massless	
$U(N_1 + k_1)$ and $U(k_2)$	ordinary	$\Delta = m/g$	(5.12)
$U(N_2 + k_2)$ and $U(k_1)$	ordinary	$\Delta = m/g$	
$U(N_1 + k_1)$ and $U(N_2 + k_2)$	ghost	$\Delta = m/g$	
$U(k_1)$ and $U(k_2)$	ghost	$\Delta = m/g$	

The non-zero masses are seen by the same sort of analysis as in [10].

We now ask where there can be residual instanton effects in Higgsing between $k_1, k_2 = \infty$ and $k_1, k_2 = 0$. Exactly as above, this can happen when some $U(2)$ with $N_f = 2$ massless flavors is Higgsed to $U(1)$. The $U(k_1)$ and $U(k_2)$ groups have too many massless flavors, and the only possibility is when $N_1 = k_1 = 1$ or $N_2 = k_2 = 1$. So we consider $U(N_2 + k_2)$, with $N_2 = k_2 = 1$; this gauge group is $U(2)$ with $N_f = 2k_2 = 2$ massless flavors, which leads to a superpotential exactly as in 5.5. The only difference from (5.5) and (5.8) is that the dynamical scale Λ^2 there should be regarded as the low energy scale, after we've integrated out the massive bi-fundamentals in 5.12. Since $U(N_2 + k_2)$, for $N_2 = k_2 = 1$, has $2k_1$ ordinary matter fields with mass Δ and $2(N_1 + k_1)$ ghost matter fields with mass Δ , the correct replacement is $m\Lambda^2 \rightarrow m\Lambda^{2N} \Delta^{(2k_1 - (2N_1 + 2k_1))} = m\Delta^{-2N_1} \Lambda^{2N}$. The ordinary and ghost matter fields contribute with opposite sign in the exponent because of their opposite sign contributions to the one-loop beta function.

The upshot is that when we Higgs $U(N_1 + N_2) \rightarrow U(N_1) \times U(N_2)$ we get

a glueball superpotential

$$W = N_1 S_1 \log S_1 + N_2 S_2 \log S_2 + \dots, \quad (5.13)$$

where the $N_i S_i \log S_i$ is present even for $N_i = 1$; in that case, it arises as described above, from a residual instanton term when our large k F-completion is Higgsed down to $k = 0$. For example, if we consider $U(2) \rightarrow U(1) \times U(1)$, the glueball superpotential contains terms $S_1 \log S_1 + S_2 \log S_2$.

These residual instanton terms, which we just found for the particular example of $U(N) \rightarrow U(N-1) \times U(1)$, are actually also present in the standard gauge theory description of this breaking, for all $N > 1$. We described the residual instanton contributions above as arising from the F-completion of the $U(1)$ factor into $U(2|1)$. But, in the standard gauge theory description of this breaking pattern, these residual instanton contributions are also present, coming from the broken part of the UV group, *i.e.* $U(N)/U(N-1) \times U(1)$. This is why the “ $U(1)$ instanton” factor found above was $m\Delta^{-2N_1}\Lambda^{2N}$, which indeed has the correct Λ exponent to be a $U(N)$ instanton. As an example, in $U(2) \rightarrow U(1) \times U(1)$, the residual instantons seen in our F-completion of the $U(1)$ factors are actually just the ordinary instantons in the original $U(2)$, *i.e.* they are residual instantons associated with $\pi_3(U(2)/U(1) \times U(1)) \neq 0$. The effect of these instantons in $U(2) \rightarrow U(1) \times U(1)$ was, for example, already included in the analysis of [44], where the result was verified to agree with the standard $U(2)$ gauge theory result of [40].

5.2 $Sp(N)$ with an adjoint and superpotential $W(\phi)$

The $Sp(N+k|k)$ F-completion was discussed in sect. 4.3. We first consider the theory where the adjoint field is very massive and can be integrated out, to describe the case of pure $\mathcal{N} = 1$ $Sp(N)$ Yang-Mills. As discussed above, we will see that the F-completion includes a residual instanton contribution, which occurs only for $N = 0$ and is precisely that needed to eliminate the non-smooth $\delta_{N,0}$ term in the standard UV completion of gauge theory 5.2. The F-completion is smooth for all N , and in particular leads to a non-zero superpotential, given by (5.1) with $h = 1$, for the F-completion of $Sp(0)$.

Integrating out the massive adjoints, the F-completion gauge theory is 4.8, without the adjoints, and with a quartic tree-level superpotential as in 4.7. We consider when there can be residual instanton contributions in a Higgsing $k \rightarrow k-1$, looking for non-trivial $\pi_3(\mathcal{G}/\mathcal{H})$, with $\mathcal{G}/\mathcal{H} = Sp(N+k)/Sp(N+k-1)$ or $SO(2k)/SO(2k-2)$. The only non-trivial cases are $Sp(1)/Sp(0)$, which occurs for $N = 0$ and $k = 1$, or $SO(4)/SO(2)$, corresponding to $k = 2$. Since the $SO(2k)$ generally has $4(N+k)$ massless

flavors, the $k = 2$ case of $SO(4)/SO(2)$ always has far too many massless flavors to lead to a residual instanton superpotential. So the only possible case for a residual instanton contribution is $N = 0$, Higgsing $Sp(1) \times SO(2)$ to nothing.

So consider the $Sp(1) \times SO(2)$ theory. We write the bi-fundamentals as $(Q_i)_c^{c'}$, with $i = 1, 2$; c is the $Sp(1)$ fundamental index and $c' = 1, 2$ the $SO(2)$ fundamental index. The $Sp(1) \cong SU(2)$ has $N_f = 2$, so the theory is described by exactly the same superpotential as in 5.5. We organize the gauge invariants in terms of mesons $M^{c'd'} = (Q_1)_c^{c'}(Q_2)_d^{d'}\epsilon^{cd}$ and baryons $B = \det_{cc'}((Q_1)_c^{c'})$, $\tilde{B} = \det_{cc'}((Q_2)_c^{c'})$. As described after (5.6) there are two solution branches and, again, the vacuum with $\langle M^{c'd'} \rangle = 0$ and $\langle B\tilde{B} \rangle = 0$ is not the right one for our purposes, because it not connected to the moduli space where the gauge group can be Higgsed to $Sp(0)$ at an arbitrarily high scale. As before, the correct branch is $\langle \det M \rangle = mS$ and $\langle B \rangle = \langle \tilde{B} \rangle = 0$. As in 5.8, this leads to the usual gaugino condensation glueball superpotential 5.1, with coefficient $h = 1$ for $Sp(N = 0)$. Again, this precisely corresponds to eliminating the non-smooth behavior of (5.2) in favor of the smooth result $h = N + 1$, for all $N \geq 0$.

Now consider the $Sp(N+k|k)$ F-completion with an adjoint and a quartic superpotential $W = StrW(\Phi)$, with $W(\Phi) = \frac{g}{4}\Phi^4 + \frac{m}{2}\Phi^2$. Now in the general vacuum we have $Sp(N_1 + N_2 + k_1 + k_2|k_1 + k_2) \rightarrow Sp(N_1 + k_1|k_1) \times U(N_2 + k_2|k_2)$, where $2N_1$ is the number of zero eigenvalues of ϕ_1 and ϕ_2 , and $2N_2$ is the number with eigenvalue corresponding to the vacuum at $a = -m/g$. Our F-completion of the IR group is $Sp(N_1 + k_1) \times SO(2k_1) \times U(N_2 + k_2) \times U(k_2)$. Taking into account the numbers of massless and massive bi-fundamental flavors, as in 5.12, we find that there are residual instanton effects only in the cases of $N_1 = 0$, where the instanton comes when we Higgs $k_1 = 1$ to $k_1 = 0$, and $N_2 = 1$, where the instanton comes when we Higgs $k_2 = 1$ to $k_2 = 0$.

So the residual instanton effects, associated with our F-completion, are only present for $Sp(N) \rightarrow Sp(N-1) \times U(1)$ and for $Sp(N) \rightarrow U(N)$. For the case $Sp(N) \rightarrow Sp(N-1) \times U(1)$, the residual instanton effects look like instantons in the $U(1)$ factor, associated with the F-completion of $U(1)$. But, as in the $U(N)$ case, these same instanton effects are already present in the standard gauge theory UV completion of $Sp(N)$; they are ordinary instanton effects in the original $Sp(N)$. There they again arise as residual instanton effects, associated with the case $Sp(N) \rightarrow Sp(N-1) \times U(1)$. Likewise, for $Sp(N) \rightarrow U(N)$, one might suspect there to be residual instanton effects associated with our F-completion, because it's really $Sp(N) \rightarrow Sp(0) \times U(N)$, and the F-completion of $Sp(0)$ to $Sp(1) \times SO(2)$ has a residual instanton

effect, as described above. But, again, this same residual instanton contribution is already present in the standard UV completion of $Sp(N)$ gauge theory. From that perspective it is a residual instanton effect associated with $\pi_3(Sp(N)/U(N)) = Z_2$.

So our F-completion of $Sp(N)$ with an adjoint differs from the standard gauge theory UV completion for one, and only one, case: $Sp(0)$. We thus expect agreement between the matrix model results, which corresponds to the large k F-completion, and standard gauge theory results for $Sp(N)$ gauge theory with an adjoint, for all Higgs breaking patterns, and for any $N > 0$.

5.3 $SO(N)$ with an adjoint and superpotential $W(\phi)$

The F-completion to $SO(N + 2k|k)$ was described in sect. 4.2. We consider first the case where the adjoint is massive and can be integrated out. So the gauge theory is the ordinary $SO(N + 2k) \times Sp(k)$ gauge theory 4.10, with the adjoints integrated out and the resulting quartic superpotential (4.7) for the bi-fundamentals. As before, we can immediately determine when there can be an residual instanton contribution to the superpotential in the Higgsing $k \rightarrow k - 1$: the only possibilities for a non-trivial $\pi_3(\mathcal{G}/\mathcal{H})$ are $SO(4) \rightarrow SO(2)$ and $Sp(1)$ breaking completely. The $SO(N + 2k)$ has $N_f = 4k$ vectors and can never lead to a superpotential. So the only possibility for a residual instanton effect is in the $Sp(k)$ factor, when $k = 1 \rightarrow k = 0$. The $Sp(1)$ factor has $N_f = N + 2$, and can only lead to a superpotential when $N_f < 4$, *i.e.* $N = 0$ and $N = 1$. Again, these are precisely the cases where the standard gauge theory results are non-smooth, and we'll see that the F-completion has precisely the correct residual instanton contributions needed to smooth out 5.2.

Consider first the F-completion of $SO(0)$ to $SO(2) \times Sp(1)$. Higgsing the F-completion back down to $SO(0)$ leads to a residual contribution, precisely as in the discussion of the F-completion of $Sp(0)$ in the previous section. There is, however, a small difference from our previous examples: this time the instanton is not in the $G(N + k)$ factor, namely $SO(2)$, but rather the other factor in $G(N + k|k)$, *i.e.* the instanton in Higgsing $Sp(1) \rightarrow Sp(0)$. As discussed at the end of sect. 4.2, the $Sp(k)$ instanton factor is actually the inverse square root of the $SO(N)$ instanton factor. In particular, the $Sp(1)$ instanton leads to a superpotential contribution $W = -\Lambda^3$, but the $SO(0)$ instanton factor is naturally $\Lambda^{3(0-2)}$. We thus write the residual instanton term in the standard glueball superpotential form 5.1, with $h = -2$. The minus sign comes essentially from that of the Str , since the instanton is in the $Sp(k)$ factor of $SO(N + 2k|k)$, and the 2 is because the $Sp(k)$ instanton

is the inverse square root of the $SO(N)$ instanton. The result $h = -2$ is precisely the correct result smoothing out $h(SO(N)) = N - 2$, to apply even for $N = 0$.

The F-completion of $SO(1)$ involves a slight variant of the above discussion, since now the F-completion is $SO(3) \times Sp(1)$, and the $Sp(1)$ has $N_f = 3$, rather than $N_f = 2$, flavors. Instead of the quantum-modified moduli space constraint, this theory is described by the superpotential [43]

$$W = -\frac{1}{m^2 \Lambda} (M^{ab} B_a \tilde{B}_b - \det_{ab} M^{ab}) + W_{tree}, \quad (5.14)$$

where Λ is the scale of the theory including the adjoint of mass m . We organize the gauge invariants as $M^{ab} = (Q_1)_{a'}^a (Q_2)_{b'}^b \epsilon^{a'b'}$ and $B_a = \epsilon_{abc} (Q_1)_{a'}^b (Q_1)_{b'}^c \epsilon^{a'b'}$, and $\tilde{B}_a = \epsilon_{abc} (Q_2)_{a'}^b (Q_2)_{b'}^c \epsilon^{a'b'}$, with a the $SO(3)$ index and a' the $Sp(1)$ fundamental index. We can write the tree-level quartic superpotential, obtained by integrating out the massive adjoints, as

$$W_{tree} = -\frac{1}{m} (B_a \tilde{B}^a - M_{ab} M^{ab}). \quad (5.15)$$

Integrating out the fields M and B we have a moduli space of vacua, along which we can Higgs to $SO(1)$, with the constant low-energy superpotential

$$W = m \Lambda^2 = \Lambda_L^3, \quad (5.16)$$

where Λ_L is the scale of the low energy $SO(N)$ Yang-Mills theory. In terms of the $Sp(1)$ theory, the superpotential (5.16) has the quantum numbers of a two-instanton contribution. But, because the $Sp(k)$ is the inverse square-root of the $SO(N)$ instanton factor, the superpotential (5.16) looks like the inverse of an $SO(N)$ instanton. In particular, the $SO(N)$ instanton factor is $\Lambda^{3(N-2)} = \Lambda^{-3}$ for $N = 1$. So the residual instanton term (5.16) can be described via the usual glueball superpotential 5.1, but with $h = -1$. This is precisely the value associated with the smooth result 5.3, continued to $N = 1$.

Continuing on to $SO(N)$ with an adjoint and more general superpotential, the discussion is very similar to that of the $Sp(N)$ theory discussed in the previous subsection. For example, with a superpotential including a quartic term, we break $SO(N + 2M) \rightarrow SO(N) \times U(M)$ in the general vacuum. Our F-completion of this is $SO(N + 2M + 2k_1 + 2k_2 | k_1 + k_2) \rightarrow SO(N + 2k_1 | k_1) \times U(M + k_2 | k_2)$. The residual instantons only occur for $k_1 = 1$, with $N = 0$ or $N = 1$ as in the above discussion, or for $M = 1$, with $k_2 = 1$, *i.e.* $SO(2M) \rightarrow SO(0) \times U(M)$, $SO(2M + 1) \rightarrow SO(1) \times U(M)$, and $SO(N + 2) \rightarrow SO(N) \times U(1)$. In all of these cases, these residual instanton

contributions from our F-completion are already present in the standard UV completion of $SO(N + 2M) \rightarrow SO(N) \times U(M)$. For example, the standard UV completion gives a residual instanton associated with $\pi_3(G/H)$ when we break $\mathcal{G} = SO(2M)$ to $\mathcal{H} = U(M)$.

5.4 $Sp(N)$ with an antisymmetric tensor ϕ and superpotential $W(\phi)$

The F-completion was discussed in sect. 4.5. When the superpotential is a mass term for the field ϕ , the low energy theory is the same as with $Sp(N)$ with a massive adjoint, so we already know what happens in that case: there is a residual instanton term only for $N = 0$, making $Sp(0)$ contribute as expected based on smoothly applying $h = N + 1$ for all $Sp(N)$, including $N = 0$.

On the other hand, for higher order in ϕ tree-level superpotentials, $W(\phi) = \sum_{p=1}^{n+1} \frac{1}{p} g_p \text{Tr}(\phi J)^p$, with $n > 1$, our F-completion contains additional residual instanton contributions, which are not already present in the standard UV completion of these gauge theories. This sets these examples apart from those of the previous sub-sections. This could, in principle, lead to a discrepancy with standard gauge theory, as was found in [7].

To motivate the additional residual instanton contributions, consider the simplest interesting case, a cubic superpotential: $W = \text{Tr}(\frac{g}{3}(\phi J)^3 + \frac{m}{2}(\phi J)^2)$. In the general vacuum, we break $Sp(N_1 + N_2) \rightarrow Sp(N_1) \times Sp(N_2)$. Our F completion of this is $Sp(N_1 + N_2 + k_1 + k_2|k_1 + k_2) \rightarrow Sp(N_1 + k_1|k_1) \times Sp(N_2 + k_2|k_2)$. The IR theory can be gauge fixed to a quiver gauge theory with gauge group $Sp(N_1 + k_1) \times SO(2k_1) \times Sp(N_2 + k_2) \times SO(2k_2)$, with two bi-fundamentals connecting every pair of gauge groups, with mass and statistics given by

Groups connected	statistics	mass	
$Sp(N_1 + k_1)$ and $SO(2k_1)$	ordinary	massless	
$Sp(N_2 + k_2)$ and $SO(2k_2)$	ordinary	massless	
$Sp(N_1 + k_1)$ and $SO(2k_2)$	ordinary	$\Delta = m/g$	(5.17)
$Sp(N_2 + k_2)$ and $SO(2k_1)$	ordinary	$\Delta = m/g$	
$Sp(N_1 + k_1)$ and $Sp(N_2 + k_2)$	ghost	$\Delta = m/g$	
$SO(2k_1)$ and $SO(2k_2)$	ghost	$\Delta = m/g$.	

The instanton factors for $Sp(N_1 + k_1)$ and $Sp(N_2 + k_2)$ are

$$\begin{aligned} \Lambda_1^{3(N_1+1)} &= g^{N_1-1} \Lambda^{2N+4} \Delta^{N_1-1-2N_2}, \\ \Lambda_2^{3(N_2+1)} &= g^{N_2-1} \Lambda^{2N+4} (-\Delta)^{N_2-1-2N_1}. \end{aligned} \tag{5.18}$$

Taking $N_1 = N$ and $N_2 = 0$, we should expect additional, residual instanton contributions to W , coming from the $Sp(k_2|k_2)$ completion of the $Sp(0)$ factor. Since the residual instanton contributions occur in the last step of Higgsing $k \rightarrow k - 1$, it suffices to consider the F-completion to just $Sp(N+1|1) \rightarrow Sp(N) \times Sp(1|1)$, *i.e.* $Sp(N+1) \times SO(2) \rightarrow Sp(N) \times Sp(1) \times SO(2)$. The matter content is as in 4.13, with the interesting dynamics, which leads to the residual instanton contributions to the superpotential, that associated with the $Sp(N+1)$ factor. The $Sp(N+1)$ theory in (4.13) has an antisymmetric tensor ϕ and $N_f = 2$ flavors (four fundamentals). That theory has a quantum moduli space, with one of the classical constraints modified by a $Sp(N+1)$ instanton, and N classical constraints unmodified; see [45], [46] for a general discussion, and explicit analysis of some low N examples. To this we can add the tree level superpotential in (4.14) and explicitly compute the full low-energy superpotential, including all residual instanton corrections to the original $Sp(N)$ theory.

As an example, consider the simplest case, where the original theory is $Sp(1)$ with an anti-symmetric. Since the anti-symmetric is a gauge singlet, this is just $Sp(1) \cong SU(2)$ super-Yang-Mills, with an additional decoupled singlet. This case was not explicitly discussed in [7], because the gauge theory appears to be too trivial, with the $SU(2)$ dynamics completely decoupled from the parameters in W_{tree} . But the matrix model results of [7] apply to this case as well, we just set $N = 2$ in their matrix model formulae, which yields non-trivial results. We interpret these as applying to the $Sp(1+k|k)$ F-completion 4.13, whose residual instanton contributions are already apparent for the case $k = 1$, *i.e.* $Sp(2) \times SO(2)$. The relevant $Sp(2)$ gauge theory has gauge invariants $\mathcal{O}_1 \equiv \frac{1}{2}\text{Tr}(\phi J)$, $\mathcal{O}_2 = \frac{1}{2}\text{Tr}(\phi J)^2 - \mathcal{O}_1^2$, $M_{[ij]} = Q_i Q_j$, and $N_{[ij]} = Q_i \phi Q_j$, with $i, j = 1 \dots 4$. The superpotential is (we modify the results of [45], [46] to include the trace \mathcal{O}_1 .)

$$W = \lambda_1(2N_{ij}N^{ij} + M_{ij}M^{ij}\mathcal{O}_2 - 2\Lambda^6) + \lambda_2(M_{ij}N^{ij} - \frac{1}{2}\mathcal{O}_1M_{ij}M^{ij}) + W_{tree}, \quad (5.19)$$

where W_{tree} are all of the terms in 4.14. Here λ_1 and λ_2 are Lagrange multipliers and we use ϵ^{ijkl} to raise indices, so $N_{ij}N^{ij} \equiv Pf(N)$. Writing $W' = gx(x + m/g)$, the classical eigenvalues of ϕ and the $SO(2)$ symmetric tensor ϕ' are 0 and $-m/g$. We are interested in the vacuum where both ϕ and ϕ' classically have one eigenvalue zero and one $-m/g$, breaking $Sp(2|1) \rightarrow Sp(1) \times Sp(1|1)$. We leave a full analysis of the vacuum and residual superpotential to future work. This is currently under investigation.

5.5 $U(N) \mathcal{N} = 1^*$

The $U(N + k|k)$ F-completion has the exact effective superpotential

$$W_{mm} = -\frac{Nm^3}{12}E_2(\tau) = Nm^3(q + 3q^2 + 4q^3 + 7q^4 + \dots), \quad (5.20)$$

where $\tau = \tau_0/N$ and $q = e^{2\pi i\tau}$, with the instanton factor q^N . On the other hand, the standard gauge theory result is

$$W_{gt} = -\frac{Nm^3}{12}E_2(\tau) + \frac{N^2m^3}{12}E_2(N\tau). \quad (5.21)$$

The added term contributes at instanton order and higher, so this discrepancy could perhaps be related to residual instanton terms, associated with our F-completion.

In particular, consider $U(1) \mathcal{N} = 1^*$, where the standard gauge theory result (5.21) gives zero non-perturbative superpotential, as might have been expected for this free theory. On the other hand, the result (5.20) is non-zero. We would like to interpret this difference as coming from residual instanton contributions in breaking $U(1 + k|k)$ to $U(1)$; in particular, it comes from the Higgsing of $U(2|1)$ to $U(1)$. Clearly, there will indeed be such residual instanton contributions in this case: we can integrate out the massive fields in our F-completion (4.17) of the $\mathcal{N} = 1^*$ theory to obtain a low energy theory which coincides with our F-completion of the $\mathcal{N} = 1$ $U(1)$ theory. And we already saw that our F-completion of the $\mathcal{N} = 1$ $U(1)$ theory has a residual instanton contribution Λ^3 , with coefficient 1. Here Λ is the scale of our $\mathcal{N} = 1^*$ F-completion (4.17) with the massive fields integrated out, so $\Lambda^3 = m^3q$ in this $U(1)$ case. We thus precisely recover the q term in 5.20, with the correct coefficient 1. It should be possible to also recover the higher instanton terms in 5.20; since the ordinary gauge theory result for $U(1)$ is $W = 0$, the full set of residual instantons should sum to $m^3E_2(\tau)/12$. Explicitly computing these, and the higher N generalizations, from our F-completion is more challenging. The challenge is in how to properly include the effects of the additional massive matter fields in 4.17, because some are fermionic; this is currently under investigation.

6 Perturbative expansion for $\mathcal{N} = 1^*$ models

Here we apply the formalism of section 3 to $\mathcal{N} = 1^*$ theory in the maximally confining phase. For simplicity we will focus on the simply-laced Lie

algebras¹². For the $\mathcal{N} = 1^*$ models the weights, $\vec{\lambda}_a$ are in fact the roots, $\vec{\alpha}$, of the Lie algebra. We will focus on the group dependence of the Feynman diagrams and use the known answer for the $SU(N)$ case for fixing the combinatoric prefactors of the diagrams.

At one loop the determinant in (3.4) is simply the length-squared of a root, that is, 2. The number of roots in a Lie algebra is hr , where h is the Coxeter number and r is the rank. Thus, summing the determinant over the roots yields $2hr$. The factorial $(r - \ell)!/r!$ gives a factor of $1/r$, for an overall contribution of $2h$.

The two-loop diagram is shown in Figure 1. Let $\vec{\alpha}$ and $\vec{\beta}$ denote the root vectors running around each loop. The group theory structure constants are then given by the structure constants in the root basis:

$$[E_\alpha, E_\beta] = N_{\alpha\beta^\gamma} E_\gamma.$$

In a simply-laced Lie Algebra, $N_{\alpha\beta^\gamma}$ is non-zero if and only if $\vec{\alpha} \cdot \vec{\beta} = -1$, and in standard normalization the E_α 's can be chosen so that the $N_{\alpha\beta^\gamma}$'s are ± 1 . Since there are two vertices in the two-loop graph, it follows that the group theory factors is $(N_{\alpha\beta^\gamma})^2$ which is either zero or one depending on whether $\vec{\alpha} \cdot \vec{\beta} = -1$. Thus we only have to sum the determinant:

$$\det \begin{pmatrix} \vec{\alpha} \cdot \vec{\alpha} & \vec{\alpha} \cdot \vec{\beta} \\ \vec{\beta} \cdot \vec{\alpha} & \vec{\beta} \cdot \vec{\beta} \end{pmatrix} = \det \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = 3$$

over all the roots with $\vec{\alpha} \cdot \vec{\beta} = -1$.

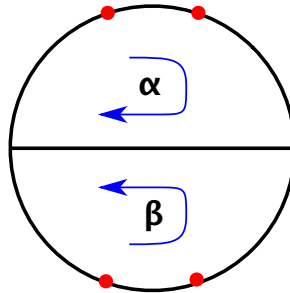


Figure 1: The two loop diagram showing the gauge theory charges, $\vec{\alpha}, \vec{\beta}$ in each loop. The dots denote the insertions of \mathcal{W}_α .

¹²Apart from having all the roots with $(\text{length})^2 = 2$, one should also remember that the Coxeter number and dual Coxeter number are equal for such Lie algebras.

To evaluate this sum, recall that in the root basis the quadratic Casimir operator on adjoint is:

$$C_2(X) = \sum_{I=1}^r [H_i, [H_i, X]] + \frac{1}{2} \sum_{\alpha} [E_{-\alpha}, [E_{\alpha}, X]] = hX.$$

Evaluating this on E_{β} , using the properties of the $N_{\alpha\beta\gamma}$ gives:

$$\vec{\beta} \cdot \vec{\beta} + \frac{1}{2} \sum_{\alpha: \vec{\alpha} \cdot \vec{\beta} = -1} 1 = h,$$

and hence, for a given root, $\vec{\alpha}$, there are precisely $2(h - 2)$ roots that have inner product -1 with it. Since the total number of roots is hr , the group theory factor for the two loop diagram is $3 \times 2h(h - 2)r$. Thus (3.4) at two loops gives:

$$(\Delta W)_{\mathcal{G}_2} = \frac{6h(h - 2)}{(r - 1)} S^2. \tag{6.1}$$

There are two diagrams at three loops, a “double-bar” graph, and a “peace-sign” graph. These are shown in figure 2. The “double-bar” graph involves three roots $\vec{\alpha}$, $\vec{\beta}$ and $\vec{\gamma}$ that satisfy $\vec{\alpha} \cdot \vec{\beta} = -1$ and $\vec{\beta} \cdot \vec{\gamma} = -1$. There is no constraint on $\vec{\alpha} \cdot \vec{\gamma}$ except that linear independence requires that this inner product not be ± 2 of -1 (if it is -1 then $\vec{\alpha} + \vec{\beta} + \vec{\gamma} = 0$). Thus $\vec{\alpha} \cdot \vec{\gamma}$ can be 0 or $+1$. We therefore have two types of root triples:

- (i) $(\vec{\alpha}, \vec{\beta}, \vec{\gamma})$ with $\vec{\alpha} \cdot \vec{\beta} = -1$, $\vec{\beta} \cdot \vec{\gamma} = -1$ and $\vec{\alpha} \cdot \vec{\gamma} = 0$
- (ii) $(\vec{\alpha}, \vec{\beta}, \vec{\gamma})$ with $\vec{\alpha} \cdot \vec{\beta} = -1$, $\vec{\beta} \cdot \vec{\gamma} = -1$ and $\vec{\alpha} \cdot \vec{\gamma} = +1$.

Suppose that $(\vec{\alpha}, \vec{\beta}, \vec{\gamma})$ is a type (ii) triple, then define $\vec{\gamma}' = -(\vec{\beta} + \vec{\gamma})$, then $(\vec{\alpha}, \vec{\beta}, \vec{\gamma}')$ is a type (i) triple. It thus follows that the numbers of each type of triple within a Lie algebra are equal. Also note that the determinant in (3.4) is equal to 4 for each type of triple. The “peace-sign” graph requires that $(\vec{\alpha}, \vec{\beta}, \vec{\gamma})$ be a triple of type (ii).

We therefore need to count all triples of type (i) in our Lie algebra, \mathcal{L} . We computed these explicitly, however we subsequently found a simpler way to encode the result. Observe that a triple of type (i) defines an A_3 (or $SU(4)$) subalgebra of \mathcal{L} , and so the problem amounts to counting all regularly embedded A_3 subalgebras. Our explicit computation shows that one can do this from the extended Dynkin diagram as follows: Take $\vec{\alpha}$ to be the extending node and then take $\vec{\beta}$ and $\vec{\gamma}$ to be connected nodes on the

extended Dynkin diagram. Delete these nodes from the diagram, *and* delete all nodes that are connected to the $\vec{\alpha}$, $\vec{\beta}$ and $\vec{\gamma}$ nodes. Call the residual Lie Algebra, K . It turns out the number of different A_3 subalgebras is then $|W_{\mathcal{L}}|/|W_{\mathcal{K}}|$, where $W_{\mathcal{X}}$ is the Weyl group of \mathcal{X} and $|W_{\mathcal{X}}|$ is its order. If there is more than one way to choose the $\vec{\alpha}$, $\vec{\beta}$ and $\vec{\gamma}$ nodes (starting with α as the extending root) then one sums over the corresponding $|W_{\mathcal{L}}|/|W_{\mathcal{K}}|$.

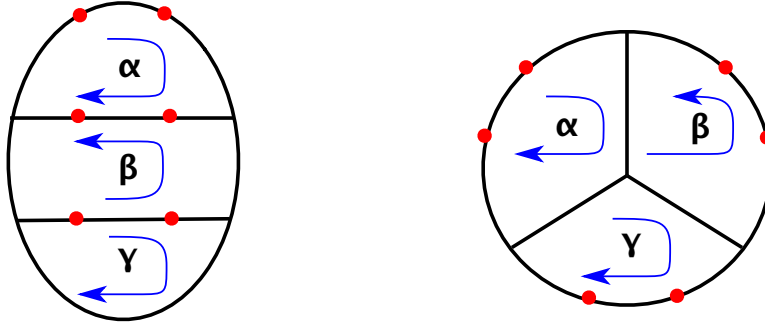


Figure 2: The two possible three loop graphs showing the gauge theory charges $\vec{\alpha}$, $\vec{\beta}$ and $\vec{\gamma}$.

We then arrive at the following numbers of triples of type (i):

$$\begin{aligned} A_{N-1} &: 2N(N-1)(N-2)(N-3), \\ D_N &: 8N(N-1)(N-2)(2N-5), \\ E_6 &: 2 \cdot 3^2 \cdot 6!, \quad E_7 : 2^2 \cdot 3 \cdot 7!, \quad E_8 : 3^2 \cdot 8!. \end{aligned} \tag{6.2}$$

These have to be multiplied by 4 because of the determinant, and by the factor of $\frac{(r-l)!}{r!}$ in 3.4. There is also a factor of 2 for the “double-bar” graph because triples of both types contribute. Finally there are symmetry factors for the graphs and the $\frac{1}{4!}$ from the Feynman diagram expansion. This gives a $\frac{3}{4}$ to the “double-bar” graph and a $\frac{1}{4}$ to the “peace-sign.” The foregoing group factors therefore get multiplied by $4 \times (2 \times \frac{3}{4} + \frac{1}{4}) \frac{1}{r(r-l)(r-2)} = \frac{7}{r(r-l)(r-2)}$. The end result is then a contribution to $W(s)$ of:

$$\begin{aligned} A_{N-1} &: 14N S^3, \quad D_N : 14 \times 4(2N-5) S^3, \\ E_6 &: 14 \times 54 S^3, \quad E_7 : 14 \times 144 S^3, \quad E_8 : 14 \times 540 S^3. \end{aligned} \tag{6.3}$$

While we have faithfully reproduced the result of [10] for $SU(N)$, it is important to realize that *graph by graph*, the results for other groups are obtained from the $SU(N)$ result by replacing N by integers that depend on the group and the loop order. These integers are shown in Table 1.

Gauge Group, G	$\ell = 1$	$\ell = 2$	$\ell = 3$
A_N	N	N	N
D_N	$2N - 2$	$4(N - 2)$	$4(2N - 5)$
E_6	12	24	54
E_7	18	48	144
E_8	30	120	540

Table 1: To obtain the contribution of an ℓ -loop graph to $W(S)$ with gauge group, G , the factors of N for the $SU(N)$ result are to be replaced by the integers from this table.

These integers are, in fact, well known for each gauge group. Consider the extended Cartan matrix, \hat{C}_{ij} , where $i, j = 0, \dots, r$, where 0 refers to the extending node. This matrix has a null vector, $\vec{p} \equiv (p_0, p_1, \dots, p_r)$, whose entries are integers, and are shown in Figure 3. The integers appearing in Table 1 are then precisely the integers:

$$N_\ell = \sum_{j=0}^r (p_j)^\ell. \quad (6.4)$$

Therefore, we find that the ℓ -loop result (for $\ell \leq 3$) is given, *graph by graph* by replacing:

$$NS^\ell \rightarrow \text{Tr}(\mathcal{S}^\ell), \quad (6.5)$$

where \mathcal{S} is a diagonal matrix whose entries are:

$$\mathcal{S} \equiv S \text{diag}(p_0, p_1, \dots, p_r). \quad (6.6)$$

We make the obvious conjecture that this is true to all loops for simply-laced groups. It is also natural to ask what this F-completion prescription means in terms of gauge theory. At least for the D series we can compare this notion of F-completion with that given before based on supergroups. We expect the amplitude to have a linear term in N and a constant piece, related to \mathbf{CP}^1 and \mathbf{RP}^2 contributions respectively. The above conjecture for all loop answer agrees with this structure. It would be interesting to check this explicitly.

7 Torus compactifications

7.1 Compactification on T^2 : A two-dimensional perspective

There is an interesting link between $\mathcal{N} = 1$ supersymmetric gauge theories in four dimensions and certain $\mathcal{N} = 2$ sigma models in two dimensions.

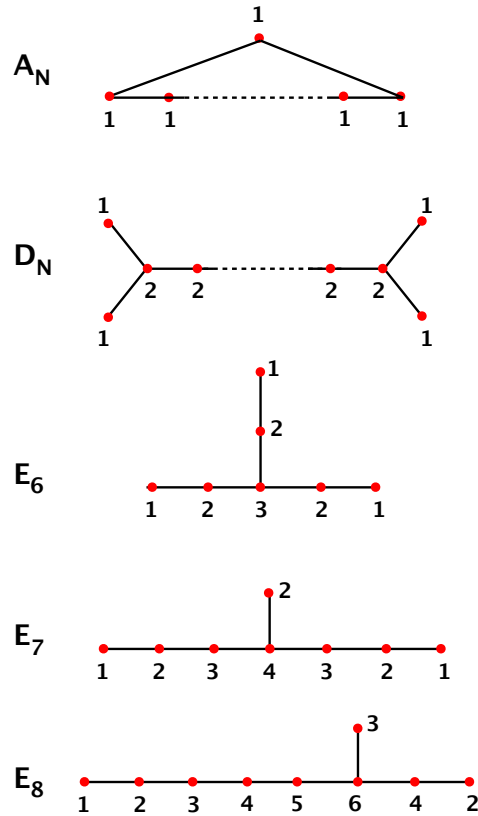


Figure 3: The extended Dynkin diagrams of the simply-laced lie algebras. The integers, p_a , on each node are the Dynkin indices that define the null vector of the extended Cartan matrix.

The idea, which has been noted before [47], [48], [50], [51], arises as follows: Consider compactification of the pure $\mathcal{N} = 1$ supersymmetric Yang-Mills theory on T^2 . We can now turn on a Wilson line on T^2 . The F-term data do not depend on the volume of T^2 . Thus we can relate the small volume description of F-terms, for which there would be an effective $\mathcal{N} = 2$ theory in two dimensions, to a large volume description, which is effectively a four-dimensional, $\mathcal{N} = 1$ supersymmetric F-term computation. In the small volume limit of the T^2 , we have a good description of the two-dimensional theory as a supersymmetric sigma model on the moduli space of flat connection of the corresponding group on T^2 . Moreover the Kähler class of the sigma model is identified with the coupling constant of the four-dimensional theory.

The moduli space of flat connection on T^2 has been obtained in [52], and

is given by the weighted projective space with weights given by the Dynkin numbers (p_0, \dots, p_r) of the corresponding affine Dynkin diagram, where $p_0 = 1$ corresponds to the extending root. Note that this is a space of complex dimension r as is expected. In this map, the chiral ring of the gauge theory gets identified with the chiral ring of the sigma model (*i.e.* the observables of the A-model topological strings). The statement that the Kähler class of the sigma model gets identified with the coupling constant means that the glueball chiral field S , which multiplies the bare gauge coupling in the action, gets identified with the chiral field corresponding to the Kähler class of the sigma model.

The sigma model can be realized in terms of a linear description involving a $U(1)$ gauge theory with $r + 1$ chiral fields with charges (p_0, \dots, p_r) . Let Σ denote the two-dimensional gauge field strength multiplet, corresponding to the Kähler class. In going from four dimensions to two dimensions, the glueball superfield S gets mapped to Σ ; there is a bare term $\int \tau \Sigma d^2\theta$ superpotential term. Integrating out the matter fields leads to a one loop induced superpotential for Σ :

$$W(\Sigma) = \left(\sum_a p_a \right) \Sigma \log(\Sigma - 1) - \tau \Sigma \quad (7.1)$$

This agrees with the VY superpotential where we use the fact that

$$\sum_{a=0}^r p_a = h$$

is the dual Coxeter number.

Thus here we also find that $\Sigma^h = e^{-\tau}$ holds as a ring relation. As $\tau \rightarrow \infty$ this leads to $\Sigma^h = 0$ as the classical ring relation. This relation may appear surprising at first sight, because the dimension of the manifold is r and one would have naively expected that $\Sigma^{r+1} = 0$ to be the ring relation. However, this is not quite true because the weighted projective space is not smooth. In particular there is a contribution to Σ from the twist fields at the singular loci, and they violate the naive expectation that $\Sigma^{r+1} = 0$. Similarly, it also appears that the Witten index of this supersymmetric theory is $r + 1$. However this is not quite correct: The weighted projective space with unequal weights has singularities. There are contributions to the index from the twisted sector and these contributions raise the naive result, $r + 1$, to the correct value, h .

For example, the weighted projective space for $SO(2n)$ has weights 1 and 2 with multiplicity 4 and $(n - 3)$ respectively. There is a twisted Z_2 sector, where the fixed point locus is a CP^{n-4} . In addition to the $n + 1$

cohomology classes of the underlying space (*i.e.* from the untwisted sector) we also get $n - 3$ classes from the twisted sector, from the cohomology of CP^{n-4} . Altogether this gives $2n - 2$ cohomology classes which is the dual Coxeter number of $SO(2n)$. This generalizes to arbitrary group. Consider a simple group. Let n_k denote the number of Dynkin indices of the affine Dynkin diagram with Dynkin index k . Suppose that all $k \neq 1$ are relatively prime. Then there is a Z_k twisted sector which contributes $(k - 1)$ twisted sectors, each with the cohomology of CP^{n_k-1} , *i.e.* n_k states. The untwisted sector contributes $r + 1 = \sum_k n_k$ to the cohomology. Thus the totality of contributions to the cohomology is given by

$$\sum_k n_k + (k - 1) n_k = \sum_k k n_k = h \quad (7.2)$$

as expected. When some of the k 's are not relatively prime the counting still works the same way but the reasoning is slightly different. For example, suppose we have n_2 nodes with label 2 and n_4 nodes with label 4. Then there are three twisted sectors related to Z_4 , given by 4th roots of unity. However the one given by the second root of unity is also part of the Z_2 sector. This implies that the two primitive fourth roots give rise to a CP^{n_4-1} fixed space, but the twisted sector associated to -1 gives a $CP^{n_2+n_4-1}$ fixed locus. Altogether this gives $2n_4 + (n_2 + n_4 - 1) = 3n_4 + n_2$ contribution to the cohomology, exactly as in 7.2. The extra contributions to the Witten index in this case are related to the puzzle raised in [53] that the Witten index of pure $N = 1$ Yang-Mills is naively equal to $r + 1$. This was resolved in the second paper in [53] where specific flat connections on T^3 were identified as the source of the discrepancy. Note that $T^3 = T^2 \times S^1$ and if we take T^2 to be small, we are discussing the sectors of the above sigma model on S^1 . The twisted sectors we have found are related to these extra contributions to the index.

It is also useful to consider using the mirror of this weighted projective space [50],[54]. We dualize the charged fields into chiral fields, Y_a , from which we obtain the superpotential:

$$W = \Sigma \left(\sum_a p_a Y_a - \tau \right) + \sum_a e^{-Y_a}.$$

Integrating out the Y_a gives the superpotential 7.1. If we integrated out Σ we get W in terms of the affine Toda superpotential, that is:

$$W = \sum_a e^{-Y_a}, \quad \text{where} \quad \sum_a p_a Y_a = \tau.$$

That is, we get

$$W = \sum_{a \neq 0} e^{-Y_a} + e^{-\tau} \prod_{a \neq 1} e^{p_a Y_a}.$$

This is also related to a three-dimensional description of F-terms, as was noted in [55] and generalized to many other examples in [56]. Namely we can also consider compactification from 4 down to 3 on a circle. In this case the Y_a have the interpretation of the complex field composed of gauge holonomies around the S^1 combined with the dual of the gauge field in three dimensions, which is a scalar.

So far we have discussed the reduction of pure Yang-Mills. If we have some massive charged fields we can integrate them out and get some correction to the foregoing superpotential. It is natural to expect that this can be done in two steps: First integrating out the charged fields, whose mass will vary with the choice of flat connection, and then dualizing the flat connections to chiral fields Y_a as above. This would give a theory involving the neutral fields interacting with the Y_a 's defined above. It would be interesting to work this out¹³. Integrating out all fields but the S we will get an effective superpotential as a function of S , $W(S)$. If we use our dictionary this means that this computes a deformation of the two-dimensional superpotential by the same function, that is, we get $W(\Sigma)$ in the two-dimensional theory. Note that here the monomials in the reduced superpotential $W_R(\Sigma)$ (*i.e.* $W(\Sigma)$ up to the Σ^h term) get mapped to observables of the chiral ring in the two-dimensional sigma model. The quantum ambiguity $W_A(\Sigma)$ can in principle be absorbed to a redefinition of the couplings. It is amusing to note that $W_A(\Sigma)$ would naturally correspond to “gravitational descendants” of the reduced phase space [58] where the notion of “gravitational descendent” is defined in the context of two-dimensional topological gravity [59].

7.2 Compactification to three dimensions and $\mathcal{N} = 1^*$ theories

For the $\mathcal{N} = 1^*$ theories with a simply-laced gauge group G , the exact superpotential was conjectured in [18], by means very different from ours.

In [12], [56] the theory is considered on $\mathbf{R}^3 \times \mathbf{S}^1$, where the light degrees of freedom of the theory can be written in terms of $r = \text{rank}(G)$ chiral superfields $Y_{1,\dots,r}$ which live on a torus of complex structure $\tau = \frac{8\pi^2}{g_{YM}^2} - i\theta$ [60]. It is argued that, in these variables, the superpotential is an elliptic

¹³As we were in the process of completing this work, an interesting paper appeared [13] which gives the proposal of how this superpotential should be modified for a single adjoint $U(N)$ theory, which we interpret as the effect on the superpotential after integrating out the massive charged fields. Their result suggests that in this case the massive charged modes are replaced by a specific quantum vev. See also the very recent work [57]

function on the torus, namely

$$W(\vec{Y}; \tau) = m^3 \sum_{\vec{\alpha} > 0} \mathcal{P}(\vec{\alpha} \cdot \vec{Y}), \quad (7.3)$$

where m is the mass used to deform the $\mathcal{N} = 4$ theory to $\mathcal{N} = 1^*$, \mathcal{P} is the Weierstrass function, and the sum is over all positive roots $\vec{\alpha}$. Note that natural variables, Y_i , correspond to the simple roots $\vec{\alpha}_i$, *i.e.*

$$Y_i = \vec{\alpha}_i \cdot \vec{Y}.$$

The Weierstrass function has an expansion

$$\mathcal{P}(Y) = \sum_{k=1}^{\infty} k \exp(-kY) + \sum_{k,n=1}^{\infty} k q^{kn} [\exp(-kY) + \exp(kY) - 2], \quad (7.4)$$

where $q = \exp(-\tau)$, and where we have subtracted a constant and rescaled to remove some factors of $4\pi^2$. Since the chiral field Y_j is the action of the three-dimensional instanton in the corresponding $U(1)$, the superpotential (7.3) is a sum over instantons. More precisely, $Y_j = \sigma_j + \tau\theta_j$ where θ_j is the holonomy of the flat connection on \mathbf{S}^1 which breaks the gauge group to $U(1)^r$, and σ_j are duals of the three-dimensional photons. The three-dimensional instantons come from the monopoles of the four-dimensional theory whose action is given by Y_j 's, and four-dimensional instantons whose action is τ .

Our purpose here is to show that the results for the superpotential, obtained by perturbative means in the previous sections, are compatible with the results of [12], [56]. We expect the agreement of terms up to S^h order, as it is these terms that are unambiguously computed by perturbation theory. In fact, for the gauge group $U(N)$, this was already checked in [1], [18], so here we will generalize this to arbitrary gauge groups.

First, note that in the limit of large mass, we should recover the affine-Toda superpotential. This is because in that limit the theory becomes pure $\mathcal{N} = 1$ SYM and moreover, the variables Y_j are *precisely* the same – the vector-scalar duality of the three-dimensional theory on a circle is the two-dimensional mirror symmetry which we used to relate the moduli space of flat connections on T^2 to the affine-Toda theory. Recalling the results of the previous section is natural to define Y_0 through the affine-Toda constraint

$$\tau = \sum_{a=0}^r p_a Y_a, \quad (7.5)$$

where $\sum_{i=1}^r p_i \alpha_i$ is the highest root of the Lie algebra, and $p_0 = 1$.

To show that the superpotentials coincide, we proceed as follows. We write the superpotential $W(\vec{Y}; \tau)$ as a superpotential $\mathcal{W}(Y)$ that depends upon the $Y_a = \{Y_i; Y_0\}$ by using the affine-Toda constraint to eliminate τ . This does not change the theory provided we introduce a Lagrange multiplier field S which imposes the affine-Toda constraint:

$$W(Y; S) = \mathcal{W}(Y) - S \left(\tau - \sum_{a=0}^r p_a Y_a \right). \tag{7.6}$$

The Lagrange multiplier field S must be identified with the glueball superfield, for the same reason as in the affine-Toda case. To get the glueball superpotential one simply integrates out the Y 's.

This can easily be done perturbatively, as there is a natural q -grading of roots. This is because, on the one hand, there is a grading corresponding to the charge lattice of instantons, generated by the instantons of smallest actions corresponding to Y_a 's and Y_0 themselves, and on the other hand, in the confining vacuum, to the leading order, these scale as $e^{-Y_a} \sim q^{1/h}$. Note that in keeping only the contributions to (7.3) coming from the monopoles of charges corresponding to the simple roots of the affine Lie algebra, one precisely recovers the affine-Toda superpotential to leading order.

The superpotential (7.3) has an expansion of the form:

$$\mathcal{W}(Y) = \sum_a y_a + \sum_{\ell=2}^{\infty} \sum_{a_1, \dots, a_\ell=0}^r B_{a_1, \dots, a_\ell}^{(\ell)} y_{a_1} \dots y_{a_\ell}, \tag{7.7}$$

where we define $y_a = e^{-Y_a}$ (For simplicity, we have set the mass, m , to one). A term of degree ℓ in the y_a corresponds to a term that scales $q^{\ell/h}$, and is thus related to the number, ℓ , of loops in the perturbative calculation. The one-loop term was computed in the previous section. At two loops, we should include monopoles whose charge corresponds to two times a simple (affine) root, and thus corresponds to a (non-affine) root that can be written as the sum of two simple affine roots. For simply-laced Lie algebras, this means that the affine simple roots must have an inner product of -1 . This is precisely what is encoded in the extended Dynkin diagram of the Lie algebra, and indeed, one has:

$$B_{ab}^{(2)} = 3\delta_{ab} - \frac{1}{2}\widehat{C}_{ab} \equiv 2\delta_{ab} + \frac{1}{2}\mathcal{I}_{ab}, \tag{7.8}$$

where \widehat{C} is the extended Cartan matrix, and \mathcal{I} is the incidence matrix of the extended diagram. So long as one has $\ell < h$ the expansion (7.7) is entirely characterized by this incidence matrix, and in terms of walks of a certain length on the extended Dynkin diagram. For example,

$$B_{abc}^{(3)} = \frac{1}{2} (\mathcal{I}_{(ab}\mathcal{I}_{bc)} - \delta_{(ab}\mathcal{I}_{bc)}), \quad \text{with no sum on } b.$$

To see how the result 6.5, (6.6) emerges here it is convenient to introduce more “auxilliary glueball fields.” That is, we introduce a dual variable, S_a , for each of the Y_a ’s, making sure, at each step, that the superpotentials contain equivalent data. Introduce $2(r + 1)$ new variables S_a and \hat{Y}_a , and write the superpotential as

$$W = \mathcal{W}(\hat{Y}) - S \left(\tau - \sum_{b=0}^r p_b Y_b \right) + \sum_{a=0}^r S_a (\hat{Y}_a - Y_a).$$

Integrating out the S_a sets $\hat{Y}_a = Y_a$, and the result reduces to 7.6. However, one could instead integrate out the \hat{Y}_a which gives:

$$W = W_D(S_a) - \sum_b Y_b (S_b - p_b S) - \tau S. \quad (7.9)$$

The $W_D(S_a)$ is the Legendre transform of $\mathcal{W}(Y_a)$:

$$W_D(S_a) = \{ \mathcal{W}(\hat{Y}_a) - \sum_b S_b \hat{Y}_b \}_{S_a = -\frac{\delta \mathcal{W}}{\delta \hat{Y}_a}},$$

and as such it contains exactly the same information as the original superpotential in 7.3. In particular, the Legendre transformation can be inverted to recover $\mathcal{W}(Y_a)$ from $W_D(S_a)$. Finally, note that the superpotential, W in (7.9) has Lagrange multipliers Y_a that impose the condition:

$$S_a = p_a S. \quad (7.10)$$

One can easily show that to third order one has:

$$\begin{aligned} W = & \sum_{a=0}^r S_a (\log(S_a) - 1) - \sum_{a,b=0}^r (2\delta_{ab} + \frac{1}{2}\mathcal{I}_{ab}) S_a S_b \\ & + \frac{1}{2} \sum_{a,b,c=0}^r (10\delta_{ab}\delta_{bc} + 9\delta_{ab}\mathcal{I}_{bc}) S_a S_b S_c + \mathcal{O}(S^4), \end{aligned} \quad (7.11)$$

where we have made use of the fact that the entries of the incidence matrix satisfy: $\mathcal{I}_{ab}\mathcal{I}_{ba} = \mathcal{I}_{ab}$ (with no sum on the indices). Using (7.10) and recalling that the p_a are a null vector of \hat{C}_{ab} , one finds:

$$\begin{aligned} W = & h S (\log(S) - 1) - 3 \left(\sum_{a=0}^r (p_a)^2 \right) S^2 \\ & + 14 \left(\sum_{a=0}^r (p_a)^3 \right) S^3 - 115 \left(\sum_{a=0}^r (p_a)^4 \right) S^4 + \mathcal{O}(S^5), \end{aligned} \quad (7.12)$$

where we have dropped a term $S \sum_a p_a \log p_a$ which can be absorbed to the definition of τ . Here the S^4 term has been obtained from the result for $U(N)$ and the observation that the fourth order result can be written entirely in terms of \mathcal{I}_{ab} and δ_{ab} , and so must be proportional to $\sum_{a=0}^r (p_a)^4$. It is also easy to convince oneself that this structure continues, namely the terms in $W_D(S)$ can be expressed solely in terms of \mathcal{I}_{ab} and δ_{ab} . For this, it is useful to use the eigenvectors of \widehat{C} , and the result is an expansion in powers of the entries of such eigenvectors. What is important is that (7.10) is precisely the null vector of \widehat{C} , and hence is an eigenvector of \mathcal{I} with eigenvalue 2. This gives further support for our conjecture at the end of section 6.

Note that it may be tempting to try to relate the S_a 's as glueball superfields of the abelian background (up to rescaling by p_a 's): $S_a = W_a^2$. Even though this should morally be correct, there are some subtleties to understand: the naive identification would lead one to expect the ℓ -loop contribution of the form $(\sum_a S_a)^\ell$, which is not the case.

8 Final Comments

In this paper we have seen that the computation of glueball superpotential can be carried out for all groups and representation for low powers of the glueball field explicitly and unambiguously. For higher powers we have discussed the existence of an ambiguity which relates to a UV complete definition of F-terms. For classical groups (and in some cases for non-classical groups) we have found a way to resolve the ambiguity by embedding the theory as a higgs branch of an arbitrarily large rank supergroup. Moreover we have discussed how this agrees or differs from more standard UV completions for some examples. It would be very interesting to find “all possible physically consistent F-term completions” and how they related to one another. This reminds one of the framing ambiguity of Chern-Simons theory, which is needed to complete the quantum definition of the theory [61]. It also reminds one of coupling topological matter to topological gravity in (1 + 1) dimensions: That is, one adds new physical observables, topological gravity, to extend and complete the topological free energy. This extension can then be computed from the “small phase space” of the topological matter via differential recurrence relations. It is thus tempting to conjecture that a given F-completion is equivalent to the choice of a “reduced phase space” inside this infinite dimensional phase space. It would be very interesting to characterize all such physically consistent choices and what relations this leads to for the computation of the coefficients of the glueball superpotential.

Another direction, currently under investigation, is to find alternative matrix models, which account for the possible residual instanton contributions to superpotentials. The occasional discrepancies associated with residual instantons effects can themselves be captured by a matrix model, within the original matrix models. One can perhaps regard this as “matrix model epicycles”.

Acknowledgements

We would like to thank A. Brandhuber, R. Dijkgraaf, J. Gomis, M. Grisaru, P. Kraus, H. Ooguri, C. Romelsberger, J. Schwarz and M. Shigemori for valuable discussions.

The research of MA and CV was supported in part by NSF grants PHY-9802709 and DMS-0074329. The work of KI was supported by DOE grant DOE-FG03-97ER40546. The work of NW was supported in part by funds provided by the DOE under grant number DE-FG03-84ER-40168. CV thanks the hospitality of the theory group at Caltech, where he is a Gordon Moore Distinguished Scholar. MA is also grateful to theory group at Caltech for hospitality during a part of this work.

Invariants defined on the whole Lie Algebra are equivalent to Weyl-invariant tensors defined on the Cartan subalgebra. The standard Casimir invariants reduce to *symmetric*, Weyl-invariant polynomials on the Cartan subalgebra, and the classification of such invariants is well-known.

In the computation of the superpotential we encountered a slightly different tensor on the Cartan subalgebra:

$$T_{\mu_1\mu_2\dots\mu_\ell}^{\nu_1\nu_2\dots\nu_\ell},$$

satisfying the conditions:

- (i) T is completely skew in $\mu_1\mu_2\dots\mu_\ell$,
- (ii) T is completely skew in $\nu_1\nu_2\dots\nu_\ell$,
- (iii) T is Weyl invariant.

Our purpose here is to show that such a tensor must have the form:

$$T_{\mu_1\mu_2\dots\mu_\ell}^{\nu_1\nu_2\dots\nu_\ell} = \text{const.} \delta_{[\mu_1}^{\nu_1} \delta_{\mu_2}^{\nu_2} \dots \delta_{\mu_\ell]}^{\nu_\ell],} \quad (8.1)$$

and hence, when contracted with W 's gives only an S^ℓ term.

We start by considering $SU(N)$, whose Weyl group is S_N , the permutation group on N elements. It is more convenient to go to $U(N)$ and introduce the standard, orthonormal weight basis, e_1, \dots, e_N . The cost of doing this is that there is a natural, Weyl-invariant vector, $V \equiv e_1 + e_2 + \dots + e_N$, which defines the overall $U(1)$ factor in $U(N)$.

In this instance we will show that:

$$T_{\mu_1 \mu_2 \dots \mu_\ell}^{\nu_1 \nu_2 \dots \nu_\ell} = a \delta_{[\mu_1}^{\nu_1} \delta_{\mu_2}^{\nu_2} \dots \delta_{\mu_\ell]}^{\nu_\ell]} + b \delta_{[\mu_1}^{\nu_1} \delta_{\mu_2}^{\nu_2} \dots \delta_{\mu_{\ell-1}}^{\nu_{\ell-1}} V^{\nu_\ell]} V_{\mu_\ell]}, \quad (8.2)$$

for some constants a and b . This establishes the $SU(N)$ result once (8.2) is projected onto the $SU(N)$ subalgebra.

The form of the argument is most easily seen by starting with $\ell = 1$, with a tensor T_μ^ν . Suppose that $T_2^1 \neq 0$ and define:

$$\tilde{T}_\mu^\nu \equiv T_\mu^\nu - T_2^1 V^\nu V_\mu.$$

By construction, $\tilde{T}_2^1 = 0$. Acting with the Weyl group shows that all the \tilde{T}_μ^ν must be zero for $\nu \neq \mu$. Consider \tilde{T}_μ^μ (no sum on indices): acting with the Weyl group implies that all of these elements must be equal. Thus $\tilde{T}_\mu^\nu = \delta_\mu^\nu$, and the result follows.

If $\ell > 1$, then the indices $\{\mu_1, \dots, \mu_\ell\}$ must be distinct, and so must the indices $\{\nu_1, \dots, \nu_\ell\}$. Moreover, *at most* one of the μ_k does not lie in the set $\{\nu_1, \dots, \nu_\ell\}$. Otherwise, if μ_p and μ_q are distinct from all the $\{\nu_1, \dots, \nu_\ell\}$, apply the permutation that interchanges μ_p and μ_q : Skew symmetry in the μ 's and Weyl invariance then imply

$$T_{\mu_1 \mu_2 \dots \mu_\ell}^{\nu_1 \nu_2 \dots \nu_\ell} = -T_{\mu_1 \mu_2 \dots \mu_\ell}^{\nu_1 \nu_2 \dots \nu_\ell},$$

and hence this component is zero.

Let ρ_1, \dots, ρ_ℓ and $\sigma_1, \dots, \sigma_\ell$ be one set of values of the indices for which $T_{\rho_1 \rho_2 \dots \rho_\ell}^{\sigma_1 \sigma_2 \dots \sigma_\ell} \neq 0$ and yet $\{\rho_1, \dots, \rho_\ell\}$ is not a just a permutation of $\{\sigma_1, \dots, \sigma_\ell\}$. Define:

$$\tilde{T}_{\mu_1 \mu_2 \dots \mu_\ell}^{\nu_1 \nu_2 \dots \nu_\ell} = T_{\mu_1 \mu_2 \dots \mu_\ell}^{\nu_1 \nu_2 \dots \nu_\ell} - b \delta_{[\mu_1}^{\nu_1} \delta_{\mu_2}^{\nu_2} \dots \delta_{\mu_{\ell-1}}^{\nu_{\ell-1}} V^{\nu_\ell]} V_{\mu_\ell]},$$

and since only one of the ρ 's does not lie in the set of σ 's, it is possible to choose b so that $\tilde{T}_{\rho_1 \rho_2 \dots \rho_\ell}^{\sigma_1 \sigma_2 \dots \sigma_\ell} = 0$. By acting with the permutation group one sees that the only non-zero elements of $\tilde{T}_{\mu_1 \mu_2 \dots \mu_\ell}^{\nu_1 \nu_2 \dots \nu_\ell}$ are those in which $\{\mu_1, \dots, \mu_\ell\}$ are a permutation of the indices $\{\nu_1, \dots, \nu_\ell\}$, and moreover all these tensor components must be equal. The result follows.

Once one has the result for $SU(N)$ it is relatively trivial for the other groups: One simply goes to a big enough $U(M)$ subgroup. For G_2, E_7

and E_8 it is trivial since they have $SU(3)$, $SU(8)$ and $SU(9)$ as maximal subgroups. The result then follows from the result for $SU(N)$. Every other Lie algebra of rank r has a maximal subalgebra of $U(r)$. From this we can deduce that T must have the form 8.2, however for a semi-simple Lie algebra there are no fixed vectors, V , and so the coefficient, b , must be zero.

References

- [1] R. Dijkgraaf and C. Vafa, “Matrix models, topological strings, and supersymmetric gauge theories,” Nucl. Phys. B **644**, 3 (2002), [arXiv:hep-th/0206255](#); “On geometry and matrix models,” Nucl. Phys. B **644**, 21 (2002), [arXiv:hep-th/0207106](#); “A perturbative window into non-perturbative physics,” [arXiv:hep-th/0208048](#); “ $N = 1$ supersymmetry, deconstruction, and bosonic gauge theories,” [arXiv:hep-th/0302011](#).
- [2] R. Dijkgraaf, M. T. Grisaru, C. S. Lam, C. Vafa and D. Zanon, “Perturbative computation of glueball superpotentials,” [arXiv:hep-th/0211017](#).
- [3] M. Bershadsky, S. Cecotti, H. Ooguri and C. Vafa, “Kodaira-Spencer theory of gravity and exact results for quantum string amplitudes,” Commun. Math. Phys. **165**, 311 (1994) [[arXiv:hep-th/9309140](#)].
- [4] H. Ooguri and C. Vafa, “The C-deformation of gluino and non-planar diagrams,” [arXiv:hep-th/0302109](#). “Gravity induced C-deformation,” [arXiv:hep-th/0303063](#).
- [5] C. Vafa, “Brane/anti-brane systems and $U(N|M)$ supergroup,” [arXiv:hep-th/0101218](#).
- [6] F. Cachazo, M. R. Douglas, N. Seiberg and E. Witten, “Chiral rings and anomalies in supersymmetric gauge theory,” JHEP **0212**, 071 (2002), [arXiv:hep-th/0211170](#).
- [7] P. Kraus and M. Shigemori, “On the matter of the Dijkgraaf-Vafa conjecture,” [arXiv:hep-th/0303104](#).
- [8] L. F. Alday and M. Cirafici, “Effective superpotentials via Konishi anomaly,” [arXiv:hep-th/0304119](#).
- [9] P. Kraus, A. V. Ryzhov and M. Shigemori, “Loop equations, matrix models, and $N = 1$ supersymmetric gauge theories,” [arXiv:hep-th/0304138](#).
- [10] R. Dijkgraaf, S. Gukov, V. A. Kazakov and C. Vafa, “Perturbative analysis of gauged matrix models,” [arXiv:hep-th/0210238](#).
- [11] E. J. Martinec and N. P. Warner, “Integrable systems and supersymmetric gauge theory,” Nucl. Phys. B **459**, 97 (1996) [[arXiv:hep-th/9509161](#)].

- R. Donagi and E. Witten, “Supersymmetric Yang-Mills Theory And Integrable Systems,” Nucl. Phys. B **460**, 299 (1996) [arXiv:hep-th/9510101].
E. J. Martinec, “Integrable Structures in Supersymmetric Gauge and String Theory,” Phys. Lett. B **367**, 91 (1996) [arXiv:hep-th/9510204].
- [12] N. Dorey, “An elliptic superpotential for softly broken $N = 4$ supersymmetric Yang-Mills theory,” JHEP **9907**, 021 (1999) [arXiv:hep-th/9906011].
- [13] R. Boels, J. de Boer, R. Duivenvoorden and J. Wijnhout, “Nonperturbative superpotentials and compactification to three dimensions,” arXiv:hep-th/0304061.
- [14] F. Cachazo, “Notes on supersymmetric $Sp(N)$ theories with an antisymmetric tensor,” arXiv:hep-th/0307063.
- [15] K. Intriligator, P. Kraus, A. V. Ryzhov, M. Shigemori and C. Vafa, “On Low Rank Classical Groups in String Theory, Gauge Theory and Matrix Models,” arXiv:hep-th/0311181.
- [16] G. Veneziano and S. Yankielowicz, “An Effective Lagrangian For The Pure $N=1$ Supersymmetric Yang-Mills Theory,” Phys. Lett. B **113**, 231 (1982).
- [17] E. Witten, “Chiral ring of $Sp(N)$ and $SO(N)$ supersymmetric gauge theory in four dimensions,” arXiv:hep-th/0302194.
- [18] N. Dorey, T. J. Hollowood, S. Prem Kumar and A. Sinkovics, “Exact superpotentials from matrix models,” JHEP **0211**, 039 (2002) [arXiv:hep-th/0209089].
- [19] T. R. Taylor, G. Veneziano and S. Yankielowicz, “Supersymmetric QCD And Its Massless Limit: An Effective Lagrangian Analysis,” Nucl. Phys. B **218**, 493 (1983).
- [20] C. Vafa, “Superstrings and topological strings at large N ,” J. Math. Phys. **42**, 2798 (2001) [arXiv:hep-th/0008142].
- [21] F. Cachazo, K. A. Intriligator and C. Vafa, “A large N duality via a geometric transition,” Nucl. Phys. B **603**, 3 (2001) [arXiv:hep-th/0103067].
- [22] H. Kawai, T. Kuroki and T. Morita, “Dijkgraaf-Vafa theory as large- N reduction,” arXiv:hep-th/0303210.
- [23] H. Ita, H. Nieder and Y. Oz, “Perturbative computation of glueball superpotentials for $SO(N)$ and $USp(N)$,” JHEP **0301**, 018 (2003) [arXiv:hep-th/0211261].

- [24] S. K. Ashok, R. Corrado, N. Halmagyi, K. D. Kennaway and C. Romelsberger, “Unoriented strings, loop equations, and $N = 1$ superpotentials from matrix models,” arXiv:hep-th/0211291.
- [25] A. Klemm, K. Landsteiner, C. I. Lazaroiu and I. Runkel, “Constructing gauge theory geometries from matrix models,” arXiv:hep-th/0303032.
- [26] I. R. Klebanov and E. Witten, “Superconformal field theory on three-branes at a Calabi-Yau singularity,” Nucl. Phys. B **536**, 199 (1998) [arXiv:hep-th/9807080]. , “AdS/CFT correspondence and symmetry breaking,” Nucl. Phys. B **556**, 89 (1999) [arXiv:hep-th/9905104].
- [27] I. R. Klebanov and M. J. Strassler, “Supergravity and a confining gauge theory: Duality cascades and chiSB-resolution of naked singularities,” JHEP **0008**, 052 (2000) [arXiv:hep-th/0007191].
- [28] S. Elitzur, A. Giveon, D. Kutasov, E. Rabinovici and A. Schwimmer, “Brane dynamics and $N = 1$ supersymmetric gauge theory,” Nucl. Phys. B **505**, 202 (1997) [arXiv:hep-th/9704104].
- [29] F. Cachazo, B. Fiol, K. A. Intriligator, S. Katz and C. Vafa, “A geometric unification of dualities,” Nucl. Phys. B **628**, 3 (2002) [arXiv:hep-th/0110028].
- [30] K. A. Intriligator and N. Seiberg, “Duality, monopoles, dyons, confinement and oblique confinement in supersymmetric $SO(N(c))$ gauge theories,” Nucl. Phys. B **444**, 125 (1995) [arXiv:hep-th/9503179].
- [31] K. A. Intriligator and P. Pouliot, “Exact superpotentials, quantum vacua and duality in supersymmetric $SP(N(c))$ gauge theories,” Phys. Lett. B **353**, 471 (1995) [arXiv:hep-th/9505006].
- [32] S. G. Naculich, H. J. Schnitzer and N. Wyllard, “A cascading $N = 1$ $Sp(2N+2M) \times Sp(2N)$ gauge theory,” Nucl. Phys. B **638**, 41 (2002) [arXiv:hep-th/0204023].
- [33] R. G. Leigh and M. J. Strassler, “Duality of $Sp(2N(c))$ and $SO(N(c))$ supersymmetric gauge theories with adjoint matter,” Phys. Lett. B **356**, 492 (1995) [arXiv:hep-th/9505088].
- [34] K. A. Intriligator, R. G. Leigh and M. J. Strassler, “New examples of duality in chiral and nonchiral supersymmetric gauge theories,” Nucl. Phys. B **456**, 567 (1995) [arXiv:hep-th/9506148].
- [35] K. A. Intriligator, “New RG fixed points and duality in supersymmetric $SP(N(c))$ and $SO(N(c))$ gauge theories,” Nucl. Phys. B **448**, 187 (1995) [arXiv:hep-th/9505051].

- [36] I. Bena, H. Murayama, R. Roiban and R. Tatar, “Matrix Model Description of Baryonic Deformations,” [arXiv:hep-th/0303115].
- [37] S. B. Giddings and J. M. Pierre, “Some exact results in supersymmetric theories based on exceptional groups,” *Phys. Rev. D* **52**, 6065 (1995) [arXiv:hep-th/9506196].
- [38] A. Brandhuber, H. Ita, H. Nieder, Y. Oz and C. Romelsberger, “Chiral rings, superpotentials and the vacuum structure of $N = 1$ supersymmetric gauge theories,” arXiv:hep-th/0303001.
- [39] I. Affleck, M. Dine and N. Seiberg, “Dynamical Supersymmetry Breaking In Supersymmetric QCD,” *Nucl. Phys. B* **241**, 493 (1984).
- [40] N. Seiberg and E. Witten, “Electric - magnetic duality, monopole condensation, and confinement in $N=2$ supersymmetric Yang-Mills theory,” *Nucl. Phys. B* **426**, 19 (1994) [Erratum-ibid. *B* **430**, 485 (1994)] [arXiv:hep-th/9407087].
- [41] K. A. Intriligator, R. G. Leigh and N. Seiberg, “Exact superpotentials in four-dimensions,” *Phys. Rev. D* **50**, 1092 (1994) [arXiv:hep-th/9403198]. HEP-TH 9403198;
- [42] C. Csaki and H. Murayama, “Instantons in partially broken gauge groups,” *Nucl. Phys. B* **532**, 498 (1998) [arXiv:hep-th/9804061].
- [43] N. Seiberg, “Exact results on the space of vacua of four-dimensional SUSY gauge theories,” *Phys. Rev. D* **49**, 6857 (1994) [arXiv:hep-th/9402044].
- [44] F. Cachazo and C. Vafa, “ $N = 1$ and $N = 2$ geometry from fluxes,” arXiv:hep-th/0206017.
- [45] P. L. Cho and P. Kraus, “Symplectic SUSY gauge theories with antisymmetric matter,” *Phys. Rev. D* **54**, 7640 (1996) [arXiv:hep-th/9607200].
- [46] C. Csaki, W. Skiba and M. Schmaltz, “Exact results and duality for $Sp(2N)$ SUSY gauge theories with an antisymmetric tensor,” *Nucl. Phys. B* **487**, 128 (1997) [arXiv:hep-th/9607210].
- [47] M. Bershadsky, A. Johansen, V. Sadov and C. Vafa, “Topological reduction of 4-d SYM to 2-d sigma models,” *Nucl. Phys. B* **448**, 166 (1995) [arXiv:hep-th/9501096].
- [48] J. A. Harvey, G. W. Moore and A. Strominger, “Reducing S duality to T duality,” *Phys. Rev. D* **52**, 7161 (1995) [arXiv:hep-th/9501022].

- [49] J. A. Harvey, G. W. Moore and A. Strominger, “Reducing S duality to T duality,” *Phys. Rev. D* **52**, 7161 (1995) [arXiv:hep-th/9501022].
- [50] K. Hori and C. Vafa, “Mirror symmetry,” arXiv:hep-th/0002222.
- [51] A. Losev, N. Nekrasov and S. L. Shatashvili, “Freckled instantons in two and four dimensions,” *Class. Quant. Grav.* **17**, 1181 (2000) [arXiv:hep-th/9911099].
- [52] E. Looijenga, “Root Systems And Elliptic Curves,” *Invent. Math.* **38**, 17 (1977) “Invariant Theory For Generalized Root Systems,” *Invent. Math.* **61**, 1 (1980).
- [53] E. Witten, “Constraints On Supersymmetry Breaking,” *Nucl. Phys. B* **202**, 253 (1982). “Toroidal compactification without vector structure,” *JHEP* **9802**, 006 (1998) [arXiv:hep-th/9712028].
- [54] M. Aganagic and C. Vafa, “Perturbative derivation of mirror symmetry,” arXiv:hep-th/0209138.
- [55] S. Katz and C. Vafa, “Geometric engineering of $N = 1$ quantum field theories,” *Nucl. Phys. B* **497**, 196 (1997) [arXiv:hep-th/9611090].
- [56] N. Dorey, T. J. Hollowood and S. Prem Kumar, “An exact elliptic superpotential for $N = 1^*$ deformations of finite $N = 2$ gauge theories,” *Nucl. Phys. B* **624**, 95 (2002) [arXiv:hep-th/0108221].
- [57] Mohsen Alishahiha, Amir E. Mosaffa, “On Effective Superpotentials and Compactification to Three Dimensions, hep-th/0304247.
- [58] A. Losev, “Descendants Constructed From Matter Field In Topological Landau-Ginzburg Theories Coupled To Topological Gravity,” *Theor. Math. Phys.* **95**, 595 (1993) [*Teor. Mat. Fiz.* **95**, 307 (1993)] [arXiv:hep-th/9211090].
- [59] E. Witten, “On The Structure Of The Topological Phase Of Two-Dimensional Gravity,” *Nucl. Phys. B* **340**, 281 (1990).
- [60] N. Seiberg and E. Witten, “Gauge dynamics and compactification to three dimensions,” arXiv:hep-th/9607163.
- [61] E. Witten, “Quantum Field Theory And The Jones Polynomial,” *Commun. Math. Phys.* **121**, 351 (1989).