Perturbative analysis of gauged matrix models

Robbert Dijkgraaf,[1] Sergei Gukov,[2] Vladimir A. Kazakov,[3] and Cumrun Vafa[2]
[1]*Institute for Theoretical Physics & Korteweg–de Vries Institute for Mathematics, University of Amsterdam, 1018 XE Amsterdam, The Netherlands*
[2]*Jefferson Physical Laboratory, Harvard University, Cambridge, Massachusetts 02138, USA*
[3]*Laboratoire de Physique Théorique de l'Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris CEDEX, France*
(Received 5 May 2003; published 18 August 2003)

We analyze perturbative aspects of gauged matrix models, including those where classically the gauge symmetry is partially broken. Ghost fields play a crucial role in the Feynman rules for these vacua. We use this formalism to elucidate the fact that nonperturbative aspects of $\mathcal{N}=1$ gauge theories can be computed systematically using perturbative techniques of matrix models, even if we do not possess an exact solution for the matrix model. As examples we show how the Seiberg-Witten solution for $\mathcal{N}=2$ gauge theory, the Montonen-Olive modular invariance for $\mathcal{N}=1^*$, and the superpotential for the Leigh-Strassler deformation of $\mathcal{N}=4$ can be systematically computed in perturbation theory of the matrix model or gauge theory (even though in some of these cases an exact answer can also be obtained by summing up planar diagrams of matrix models).

DOI: 10.1103/PhysRevD.68.045007 PACS number(s): 11.30.Pb, 11.25.Yb

I. INTRODUCTION

In this paper we study perturbative aspects of matrix models as applied to nonperturbative dynamics of $\mathcal{N}=1$ supersymmetric gauge theories in four dimensions (admitting a large N description) [1–3]. The connection between the matrix model and the supersymmetric gauge theory proceeds by identifying the superpotential of the gauge theory with the potential of the matrix model. It was shown in [1–3], building on previous work [4–7], that the planar diagrams of the matrix model effectively compute the exact glueball superpotential for the associated supersymmetric gauge theory and thus yield, upon extremization, exact results for the gauge theory. There has been some further work in this direction [8–14].

In some cases the planar diagrams of matrix models can be summed up exactly. This then gives rise to a dual geometry at the planar limit, from which one can read off nontrivial holomorphic information about the associated supersymmetric gauge theory. In this respect it is interesting to note that up to now all the cases where the supersymmetric gauge theory can be solved using strong or weak coupling dualities fall in the class of exactly soluble matrix models. In all these cases the solution takes the form of a dual geometry. However, in most cases (i.e., for a generic matter content and interactions) the exact solution of the corresponding matrix model is not available, even in the planar limit.

But, even if the planar diagrams cannot be exactly summed, we still can resort to perturbative techniques of the matrix model. This yields, as noted in [3], a systematic instanton expansion in the gauge theory. Thus, for a large class of supersymmetric gauge theories for which we had no dual descriptions, we can now nevertheless compute in a systematic way instanton corrections to interesting holomorphic quantities. Thus, in a sense, we are going beyond duality, and we may hope that this will ultimately give us a new perspective about the meaning of duality in gauge theory and string theory.

Perturbative techniques for matrix models are not completely trivial. This is because we are dealing with a *gauged* matrix model, and it is crucial to take this gauging into account properly. For vacua where the gauge symmetry is not broken, this can be easily taken into account by dividing by the volume of the gauge group, which simply leads to an overall factor. However, for vacua where the gauge group is partially broken, not only do we have to divide by the volume of the unbroken gauge group, we also have to deal with naive flat directions of the matrix fields, which are pure gauge degrees of freedom. To address this, we can implement the standard method of Faddeev-Popov ghosts, now applied to the broken part of the gauge group. The main aim of this paper is to develop this further and apply it to a number of interesting examples. This will include examples where we know the exact solutions as well as some where we do not know how to sum up the planar diagrams. Since our emphasis in this paper is the applicability of perturbative techniques we illustrate the power of the perturbation theory, even for some of the examples where we do know how to sum up planar diagrams. We will consider in particular $\mathcal{N}=1^*$ and Leigh-Strassler deformation of the $\mathcal{N}=4$ super-Yang-Mills, as well as $\mathcal{N}=2$ Seiberg-Witten geometry.

As a byproduct of the results of this paper, which might be interesting to the matrix model specialists, we demonstrate how the matrix models with several eigenvalue supports in the large N limit can be studied by means of the planar diagram technique and established well-defined Feynman rules for it. (This subject is also discussed in [15].) Another novelty which is not well explored in the matrix model literature is the possibility of filling not only the minima but also the maxima of the matrix potential (the "unstable" cuts), by virtue of the analytical continuation in the filling parameters. We demonstrate this with the example of the one matrix model with the cubic potential where we fill by eigenvalues both the minimum and the maximum. One can show that this model is equivalent to a particular case of the models of random paths studied in [16], where the solution can be written in terms of elliptic functions.

The organization of this paper is as follows. In Sec. II we

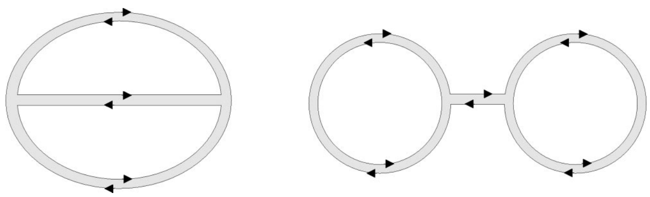

FIG. 1. The two planar two-loop diagrams, with combinatorial weight $\frac{1}{6}$ and $\frac{1}{2}$, that contribute to the order S^3 term in the free energy.

show how gauge fixing in the one matrix model with the cubic potential is done, when the classical vacuum partially breaks the gauge symmetry. We establish the planar diagrammatic rules for this model. We show the importance of ghosts for matrix models in this context and relate it to the ghosts of the supersymmetric gauge theory. We also demonstrate that the Feynman rules for the multicut solutions have a nice geometric interpretation in terms of domain walls on the closed string world sheet. In Sec. III we study various examples. In Appendix A we recall how the exact solution can be obtained in the case of the cubic superpotential as well as some connections with $c=1$ strings on the self-dual radius. In Appendix B we show how to setup the perturbation theory for massive vacua of $\mathcal{N}=1^*$ where the rank of the gauge group is reduced.

II. GAUGE FIXING IN FIELD THEORY AND MATRIX MODELS

A. The problem

To explain the setup and review the proposal of [1–3], let us start with a simple integral over a single $M\times M$ matrix Φ

$$Z=\frac{1}{\mathrm{vol}\,U(M)}\int d\Phi\,\exp\left(\frac{1}{g_s}\mathrm{tr}\,W(\Phi)\right), \qquad (2.1)$$

where $W(x)$ is a cubic polynomial with two critical points at $x=a_1$ and $x=a_2$

$$W'(x)=(x-a_1)(x-a_2). \qquad (2.2)$$

It was explained in [1] how to compute the genus zero free energy in this model if we put all the eigenvalues of the matrix Φ at one critical point, say at a_1. Shifting the matrix as $\Phi\rightarrow a_1\mathbf{1}+\Phi$ we obtain (up to a constant)

$$W=\mathrm{tr}\left(\frac{1}{2}\Delta\Phi^2+\frac{1}{3}\Phi^3\right) \qquad (2.3)$$

with

$$\Delta=a_1-a_2.$$

From this action we easily read off the Feynman rules: a propagator $1/\Delta$ for the Φ variable and a three-point vertex with weight 1. This gives, for example, the following two-loop contribution to the perturbative part of the genus zero free energy, with contributions $\frac{1}{6}$ and $\frac{1}{2}$ from the two planar diagrams of Fig. 1,

$$\mathcal{F}_0^{\mathrm{pert}}=\frac{2}{3}\,\frac{1}{\Delta^3}S^3+\cdots. \qquad (2.4)$$

Here $S=g_sM$ plays the role of the 't Hooft parameter.

According to [1], the planar limit of this matrix model can be used to obtain exact holomorphic quantities in the corresponding $\mathcal{N}=1$ gauge theory, which in this case is simply a $U(N)$ supersymmetric gauge theory with a single adjoint superfield and a tree-level superpotential $\mathrm{tr}\,W(\Phi)$ given by Eq. (2.2). For example, the effective superpotential is essentially given by the derivative of the $\mathcal{F}_0(S)$,

$$W_{\mathrm{eff}}(S)=NS\,\log(S/\Lambda^3)-2\pi i\tau_0S+N\frac{\partial\mathcal{F}_0^{\mathrm{pert}}(S)}{\partial S}, \qquad (2.5)$$

where the first term can be seen as coming from the contribution of the measure factor to the free energy $\mathcal{F}_0$ [17]. Here the variable S is identified with the chiral glueball field,

$$S=\frac{1}{32\pi^2}\mathrm{tr}\,\mathcal{W}_\alpha\mathcal{W}^\alpha.$$

From the effective superpotential W_{eff} one can read off nonperturbative information about the infrared dynamics and vacuum structure of $\mathcal{N}=1$ theory. Thus, critical points of W_{eff} generically correspond to massive vacua in the low-energy theory. On the other hand, the difference ΔW_{eff} between the value of the superpotential at two different critical points determines the tension of the BPS domain wall separating the two vacua.

In order to find the value of W_{eff} at each vacuum, one should extremize it with respect to S and then reexpress the result in terms of the (bare) gauge coupling τ_0. As a result, one typically finds an instanton expansion, in which the n-instanton terms are fixed by the perturbative contributions to $\mathcal{F}_0$ up to the n-loop order. For example, already the two-loop result (2.4) can be used to determine W_{eff} *exactly* up to two-instanton order.

It is important to stress here that the rank M of the gauge group in the matrix model is completely unrelated to the rank N of the gauge group in the corresponding $\mathcal{N}=1$ theory. In order to appreciate this point, note that M enters the effective superpotential (2.5) in a very complicated manner (via the S dependence), whereas the N dependence is very simple (linear). In particular, the value of N does not have to be large; the result (2.5) can be applied just as well to a $U(2)$ gauge theory. Henceforth, we will be very careful to distinguish between M and N.

Now let us proceed to a more general classical vacuum with M_1 eigenvalues at a_1 and M_2 eigenvalues at a_2

$$\Phi=\begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}.$$

So in the matrix model we break the gauge symmetry as

$$U(M)\rightarrow U(M_1)\times U(M_2).$$

Within the string theory realization this corresponds to a background with two clusters of D-branes of charge M_1 and M_2, respectively. Taking both M_1 and M_2 to be large, we obtain a so-called two-cut solution of the matrix model. To find the perturbative expansion of this solution it is too naive to simply expand the matrix Φ around this point. Indeed, if we shift

$$\Phi \to \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} + \Phi, \tag{2.6}$$

and decompose the matrix Φ in blocks

$$\Phi = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} \tag{2.7}$$

(where Φ_{ij} corresponds to an ij string, going from the ith D-brane to the jth D-brane) then the quadratic piece in the action takes the form

$$\frac{1}{2}\Delta\ \mathrm{tr}(\Phi_{11}^2 + \Phi_{21}\Phi_{12} - \Phi_{12}\Phi_{21} - \Phi_{22}^2) = \frac{1}{2}\Delta\ \mathrm{tr}(\Phi_{11}^2 - \Phi_{22}^2).$$

So, the kinetic terms for the "off-block diagonal" components Φ_{12} and Φ_{21} will vanish. This makes it problematic to keep track of the 12 and 21 degrees of freedom.

This vanishing of the kinetic term for the off-diagonal components is not surprising since they are zero modes. The original $U(M)$ gauge symmetry still acts on the matrix configurations and the broken gauge transformations will transform a vacuum with two clusters of eigenvalues into a gauge equivalent state. More precisely, we now have a nontrivial vacuum manifold parametrized by the coset

$$U(M)/U(M_1) \times U(M_2).$$

Since the action is $U(M)$ invariant, the matrix integral will not depend on the choice of point on this vacuum manifold. The corresponding $2M_1M_2$ zero modes are exactly the components Φ_{12} and Φ_{21}.

The correct way to treat the semiclassical expansion, keeping track of the M_1 and M_2 dependence, is by the method of Faddeev-Popov ghosts. We will see in a moment how this emerges both from the four-dimensional gauge theory and from the matrix model. But let us here remark that the role played by the ghosts is also suggested by going back to the topological string derivation of the matrix model as described in [1].

There one starts from a reduction to two dimensions of six-dimensional holomorphic Chern-Simons theory [18]. The six-dimensional open string field theory contains fields of various ghost numbers that correspond geometrically to differential forms of different degrees on the Calabi-Yau manifold. If we reduce the theory down to two dimensions, we find at the physical ghost level (among other fields) a gauged chiral scalar field $\Phi(z)$, whose zero mode is the variable Φ in the matrix integral.

But there is also a contribution of the ghosts in this two-dimensional world-volume theory. One finds in particular a scalar ghost $C(z)$ and a conjugate ghost $B(z)$, that is a (1,1) form on the world volume. Both are adjoint valued, with action

$$\frac{1}{g_s}\int d^2z\ \mathrm{tr}(B\bar{D}_A C + B[\Phi, C]).$$

Since both scalars Φ and C reduce to their constant zero modes, only the overall volume factor in the two-form B contributes in the path integral. So we get an additional ghost contribution to the matrix integral of the form

$$W_{\mathrm{ghost}} = \mathrm{tr}(B[\Phi, C]), \tag{2.8}$$

where B, C are now anticommuting $M \times M$ matrices. Let us now explain in more detail the origin of this term more directly in the four-dimensional $\mathcal{N}=1$ gauge theory and in the corresponding matrix model.

B. Gauge fixing in $\mathcal{N}=1$ supersymmetric gauge theory

Consider $\mathcal{N}=1$ gauge theory with a $U(N)$ vector multiplet and one chiral matter multiplet in the adjoint representation of the gauge group. In $\mathcal{N}=1$ superspace the field content of such theory is represented by a vector superfield V and an adjoint chiral scalar superfield Φ. Let $S_{\mathrm{inv}}(V, \Phi, \bar{\Phi})$ be the action of the superfields V and Φ, invariant under $U(N)$ gauge transformations

$$e^V \to e^{i\bar{\Lambda}} e^V e^{-i\Lambda}, \tag{2.9}$$

where Λ is a chiral gauge parameter.

Our goal will be to study (partial) gauge fixing in the functional integral

$$Z = \int \mathcal{D}V \mathcal{D}\Phi \mathcal{D}\bar{\Phi}\, e^{S_{\mathrm{inv}}(V, \Phi, \bar{\Phi})} \tag{2.10}$$

by imposing a gauge fixing constraint on the adjoint scalar Φ. Implementing the standard Faddeev-Popov procedure, one finds (a) that (partial) fixing of the $U(N)$ gauge symmetry leads to new anticommuting chiral ghost superfields B and C, and (b) that the ghost action can be written as an F term of the form (2.8).

The first statement does not depend on the particular way of gauge fixing. It is simply related to the fact that the gauge parameter Λ is a chiral scalar and, therefore, the gauge-fixing function $F = F(V, \Phi, \bar{\Phi})$ should also be a chiral superfield. Namely, the gauge constraint should be of the form [19]

$$F = f, \quad \bar{F} = \bar{f}, \tag{2.11}$$

where $f = f(x, \theta)$ is some chiral function. As we review below, this implies that the ghost superfields are also chiral.

On the other hand, the second statement above relies on the assumption that the gauge-fixing function F does not depend on the vector superfield V. Since, as we just explained, F has to be chiral we conclude that $F = F(\Phi)$. In particular, a convenient choice of gauge is given by a linear function $F(\Phi)$. Then, it follows from the gauge transformation of Φ, that under $U(N)$ gauge symmetry F transforms as

$$\delta F=[\Phi,\Lambda].$$

Now, in order to apply the usual Faddeev-Popov method to the gauge condition (2.11), we introduce the functional determinant

$$\Delta_F=\int \mathcal{D}\Lambda\mathcal{D}\bar{\Lambda}\,\delta(F-f)\,\delta(\bar{F}-\bar{f}).$$

Inserting 1 into the path integral (2.10) in the form $\Delta_F\Delta_F^{-1}$, we obtain

$$Z=\int \mathcal{D}V\mathcal{D}\Phi\mathcal{D}\bar{\Phi}\Delta_F^{-1}\delta(F-f)\,\delta(\bar{F}-\bar{f})e^{S_{\rm inv}(V,\Phi,\bar{\Phi})}.$$

Introducing the chiral ghost fields B,C and expressing the Faddeev-Popov determinant Δ_F^{-1} in terms of the ghost action

$$\begin{aligned}\Delta_F^{-1}=&\int \mathcal{D}B\mathcal{D}\bar{B}\mathcal{D}C\mathcal{D}\bar{C}\exp\left[\mathrm{tr}\int d^4x d^2\theta B\left(\frac{\delta F}{\delta\Lambda}C\right.\right.\\ &\left.\left.+\frac{\delta F}{\delta\bar{\Lambda}}\bar{C}\right)+\mathrm{tr}\int d^4x d^2\overline{\theta B}\left(\frac{\delta\bar{F}}{\delta\Lambda}C+\frac{\delta\bar{F}}{\delta\bar{\Lambda}}\bar{C}\right)\right]\\ =&\int \mathcal{D}B\mathcal{D}\bar{B}\mathcal{D}C\mathcal{D}\bar{C}\exp\left[\mathrm{tr}\int d^4x d^2\theta B[\Phi,C]+\text{c.c.}\right]\\ =&\int \mathcal{D}B\mathcal{D}\bar{B}\mathcal{D}C\mathcal{D}\bar{C}e^{S_{\rm ghost}},\end{aligned}$$

leads to the path integral

$$Z=\int \mathcal{D}V\mathcal{D}\Phi\mathcal{D}\bar{\Phi}\mathcal{D}B\mathcal{D}\bar{B}\mathcal{D}C\mathcal{D}\bar{C}e^{S_{\rm inv}+S_{\rm GF}+S_{\rm ghost}}, \tag{2.12}$$

where $S_{\rm GF}$ is the gauge-fixing action and $S_{\rm ghost}$ is given by

$$S_{\rm ghost}=\int d^4x d^2\theta\,\mathrm{tr}(B[\Phi,C])+\text{c.c.}$$

This is the tree-level contribution to the superpotential that we were after. Specifically, it shows that for a (partial) gauge fixing via imposing constraints on the adjoint chiral superfield Φ, the ghost action can indeed be written as the F term. Moreover, the form of this term is exactly the same as the form of the ghost term (2.8) in the matrix model action, which is in line with the general statement that potential in matrix model should be identified with classical superpotential in $\mathcal{N}=1$ gauge theory [1].

C. Gauge fixing in matrix models

The ghost term (2.8) can also be derived directly in the matrix model by gauge fixing the $U(M)$ gauge symmetry that acts by conjugation on Φ

$$\Phi\rightarrow U\cdot\Phi\cdot U^{-1}.$$

A convenient gauge choice is putting Φ to diagonal form. This gives the condition

$$\Phi_{ij}=0,\quad i\neq i.$$

Implementing this gauge fixing through the BRST formalism introduces exactly the above ghost fields; see [20,21] for more discussion of ghost fields and gauge fixing in matrix models.

Decomposing the ghosts also in the block form (2.7), we see that after the shift (2.6) the kinetic term of the ghosts is given by

$$\Delta\ \mathrm{tr}(B_{21}C_{12})-\Delta\ \mathrm{tr}(B_{12}C_{21}).$$

So, in the case of the ghosts it is the 11 and 22 blocks that are not propagating and the 12 and 21 block that are "physical."

We conclude that in the reduction to the matrix integral the 11 and 22 strings represent physical matter fields and that the 12 and 21 strings represent ghost degrees of freedom. This makes sense physically, since, as we already explained, in this two-cut classical vacuum with reduced gauge symmetry $U(M_1)\times U(M_2)$ the matrix elements in the 11 and 22 blocks cannot be obtained by gauge transformations and thus they are classically not pure gauge, whereas the 12 and 21 blocks are pure gauge. In perturbation theory we therefore are left with only the ghosts in the 12 and 21 blocks.

Before we turn to the Feynman rules that all this implies, let us point out that this interpretation is consistent with the multicut solution of the large M limit of the matrix integral. Here we first reduce the matrix integral to eigenvalues

$$Z=\int\prod_I d\lambda_I\prod_{I<J}(\lambda_I-\lambda_J)^2\exp\frac{1}{g_s}\sum_I W(\lambda_I). \tag{2.13}$$

In the case of a two-cut solution we can split the eigenvalues λ_I in two subsets. The first subset of M_1 eigenvalues $\lambda_I^{(1)}$ are located around the first critical point a_1, the second subset of M_2 eigenvalues $\lambda_J^{(2)}$ are located around the second critical point a_2. In a semiclassical expansion these two critical points and the corresponding eigenvalues can be thought to be well separated. We can regard the two sets $\{\lambda_I^{(1)}\}$ and $\{\lambda_J^{(2)}\}$ as eigenvalues of two matrices, a $M_1\times M_1$ matrix Φ_{11} and a $M_2\times M_2$ matrix Φ_{22} with matching potentials W. In the saddle-point approximation after the shift (2.6) this gives the action

$$W_{\rm tree}=\mathrm{tr}\left(\frac{1}{2}\Delta\Phi_{11}^2+\frac{1}{3}\Phi_{11}^3\right)+\mathrm{tr}\left(-\frac{1}{2}\Delta\Phi_{22}^2+\frac{1}{3}\Phi_{22}^3\right). \tag{2.14}$$

From the eigenvalue representation of the matrix integral it is clear that the only way these matrices Φ_{11} and Φ_{22} interact is through the Jacobian factor

$$\prod_{I,J}(\lambda_I^{(1)}-\lambda_J^{(2)})^2.$$

(This is clearly true for arbitrary W.) This term can be exponentiated directly in the action (see also [15]) giving the effective action

$$2\,\mathrm{tr}\log(\Phi_{11}\otimes\mathbf{1}-\mathbf{1}\otimes\Phi_{22}).$$

To bring out clearly the M_1 and M_2 dependence, this part of the Vandermonde determinant can also be exponentiated by using the two pairs of ghosts (B_{21},C_{12}) and (B_{12},C_{21}). [We have two pairs because of the square of the Vandermonde in Eq. (2.13).] In order to reproduce the right determinant the action of these ghosts should be

$$W_{\mathrm{ghost}}=\mathrm{tr}(B_{21}\Phi_{11}C_{12}+C_{21}\Phi_{11}B_{12}) +\mathrm{tr}(B_{12}\Phi_{22}C_{21}+C_{12}\Phi_{22}B_{21}). \tag{2.15}$$

But this is exactly the action (2.8) restricted to the propagating fields: the 11 and 22 blocks of Φ and the 12 and 21 blocks of B,C.

From the two contributions to the action (2.14) and (2.15) we can read off the Feynman rules. We have propagators (we suppress the obvious matrix indices)

$$\langle\Phi_{11}\Phi_{11}\rangle=\frac{1}{\Delta},$$

$$\langle\Phi_{22}\Phi_{22}\rangle=-\frac{1}{\Delta},$$

$$\langle B_{12}C_{21}\rangle=\langle B_{21}C_{12}\rangle=\frac{1}{\Delta},$$

and all three-point vertices have weight **1**.

As a check of this perturbative prescription with the known properties of the two-cut solution we will compute in this case the two-loop contribution to the free energy $\mathcal{F}_0(S_1,S_2)$. From the explicit answer to the large M solution we know this term is given by [7]

$$\frac{1}{\Delta^3}\left(\frac{2}{3}S_1^3-5S_1^2S_2+5S_1S_2^2-\frac{2}{3}S_2^3\right). \tag{2.16}$$

The coefficients $\pm 2/3$ have already been computed. They come from the two diagrams in Fig. 1 in which only Φ_{11} and Φ_{22} (and no ghosts) propagate.

The coefficients ± 5 are given by the mixed diagrams in which also the ghosts B,C appear. Now there are three diagrams to consider, which are given in Fig. 2. Here the following factors contribute to the weight of the diagram: the symmetry factor of the (colored) graph, the extra minus signs of the ghost loops, the extra minus sign for the Φ_{22} propagator compared to the Φ_{11} propagator, and the fact that their are two flavors of ghosts (B and C) running through each ghost loop. With these considerations taken into account, the three diagrams give a total combinatorial weight to the $S_1^2S_2$ term of

$$\frac{1}{2}\cdot(-1)\cdot 2+1\cdot(-1)\cdot 2+\frac{1}{2}\cdot(-1)^2\cdot(-1)\cdot 4=-5.$$

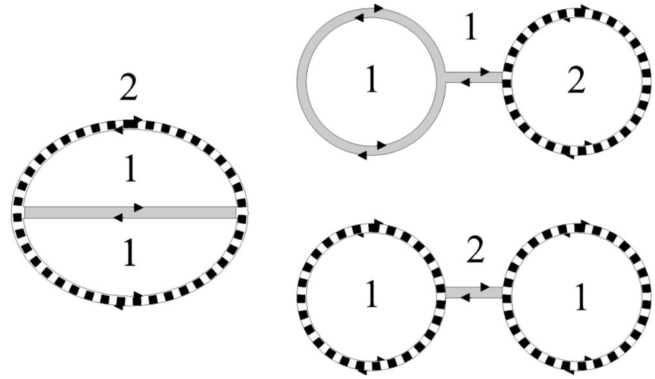

FIG. 2. The three planar two-loop diagrams, with combinatorial weight $\frac{1}{2}$, 1, and $\frac{1}{2}$, respectively, that contribute to the order $S_1^2S_2$ term in the free energy. The gray propagator indicates a bosonic Φ_{11} or Φ_{22} field; the dashed propagator indicates a B,C ghost of type 12 or 21. The labeling of the hole or index loop is also indicated.

This indeed reproduces the second and third term in Eq. (2.16).

D. Relation to $\hat{A}_2$ and $O(2)$ models on planar graphs

We will now argue that this two-cut model corresponds to the $\hat{A}_2$ "quiver" model[1] on planar graphs introduced and studied in [22]. Indeed, let us consider the Feynman rules of the previous section (we choose the dimensionful parameter $\Delta=1$): if we revert at the same time the sign of the propagator $\langle\Phi_{22}\Phi_{22}\rangle$ from $+1$ to -1 and the sign of the weight of each ghost loop from $+2$ to -2, it is the same as to revert the sign of S_2. The latter will lead to only positive coefficients in the formulas of the type (2.16) for the expansion for $\mathcal{F}_0(S_1,S_2)$ given in the next section. It is easy to check this statement inductively: if we add one $\langle\Phi_{22}\Phi_{22}\rangle$ to any diagram (like diagrams in Fig. 2) it adds up one extra loop weighted with the factor S_2, so their sign changes are compensated. The same about a ghost loop: its addition leads to a new loop with the S_2 factor, so their sign changes are again compensated.

Hence we can write down the equivalent matrix model with the potential

$$W=\mathrm{tr}\Big[\frac{1}{2}\Phi_1^2+\frac{1}{3}\Phi_1^3+\frac{1}{2}\Phi_2^2+\frac{1}{3}\Phi_2^3 +\frac{1}{2}\mathbf{C}^\dagger\mathbf{C}+\mathbf{C}^\dagger\mathbf{C}\Phi_1+\mathbf{C}\mathbf{C}^\dagger\Phi_2\Big],$$

where Φ_1 and Φ_2 are $M_1\times M_1$ and $M_2\times M_2$ matrices, respectively, and $\mathbf{C}=(C_1,C_2)$ is a vector of two $M_1\times M_2$ rectangular complex matrix *bosonic* ghosts. We recognize here actually the $\hat{A}_2$ "quiver" matrix model with a specific matrix potential.

[1] The corresponding Coxeter diagram consists of a circle with two nodes.

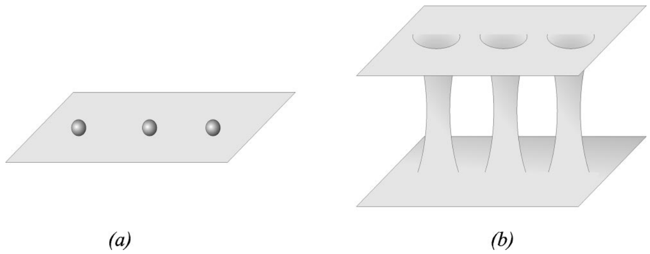

FIG. 3. (a) The distribution of eigenvalues at $g_s = 0$. (b) The dual geometry (spectral curve) at finite 't Hooft coupling.

In the symmetric case $S = S_1 = -S_2$ this model is equivalent (*only in the planar limit*, the difference due to the uncontractible ghost loops on graphs of a nontrivial topology) to the $O(2)$ model describing the statistics of self-avoiding (ghost) loops on planar Φ^3 type graphs, with the factor $+2$ for each loop [in the more general $O(n)$ model one has the weight n for each loop [16,23]]. This model is known to describe 2D quantum gravity coupled to the $c=1$ matter at the self-dual compactification radius. In Appendix A we review the full planar solution [7] of this model from the one matrix model setup. In the symmetric case the result is presented in terms of elliptic parametrization.

E. Multiple phases and domain walls on the world sheet

We would like to put the above construction into a bit more general perspective. As we already mentioned we are dealing with a toy model for a brane configuration where we have well-separated clusters of $M_1, M_2, \ldots$ D-branes. In our toy matrix model we can see clearly how such a multicenter geometry looks like from the open and closed string perspectives. This might be helpful for understanding gauge/gravity dualities for these kind of configurations in general.

In the matrix model at zero coupling ($g_s = 0$) such a vacuum state is simply given by the distribution of the eigenvalues in groups over the critical points of W in the complex eigenvalue plane as sketched in Fig. 3(a). The eigenvalue density is represented as a sum of delta functions.

With the use of the large N matrix model techniques we know that in the dual closed string picture this geometry gets modified at nonzero 't Hooft coupling [1]. The continuous eigenvalue density spreads out along branch cuts in the eigenvalue plane. In this way a nontrivial Calabi-Yau (CY) geometry emerges that is essentially given by a hyperelliptic curve obtained as a double cover of the eigenvalue plane as sketched in Fig. 3(b),

$$y^2 = W'(x)^2 + \text{deformations}. \tag{2.17}$$

Intuitively the following happens: if we insert a large number of eigenvalues M_i at the ith critical point of W this builds up a throat region in the dual geometry where the circumference of the neck is given by the 't Hooft coupling $g_s M_i$. This fact that the size of the geometry is proportional to the rank of the matrix, which is a measure of the total number of degrees of freedom, should be thought of as a version of the Bekenstein-Hawking geometric entropy, and it would be interesting to develop this interpretation further.

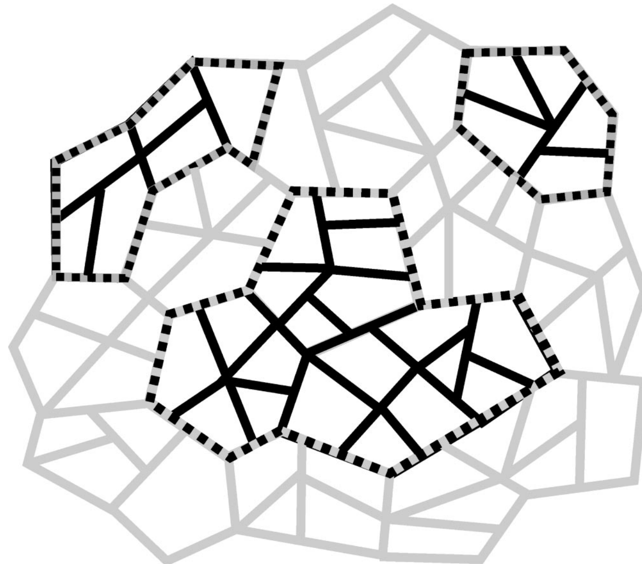

FIG. 4. In a two-cut solution the 11 strings and 22 strings (here indicated in gray and black) will build up world-sheet theories out of fishnet diagrams with interaction given by the (super)potential expanded around the relevant critical point. The 12 and 21 strings (here indicated by dashed lines) from self-avoiding loops, separating the two phases on the world sheet.

We have seen that in the open string picture the character of the *ij* strings, stretching from the ith to the jth D-brane, is very different depending on whether $j = i$ or $j \neq i$. The diagonal *ii* strings have interactions among themselves that are given by the expansion of the superpotential W around the ith critical point and can therefore be of arbitrary order. These interactions build up the fishnet double-line Feynman diagrams that in the large N limit will describe the closed string world sheet propagating in the local geometry around the ith D-brane, just as in the case of a single center geometry.

The interactions of the off-diagonal *ij* strings with $j \neq i$ do *not* depend on the potential $W(\Phi)$. They are given entirely by the cubic interaction (2.15) that is dictated by gauge invariance. Note that the action is quadratic in these $i \neq j$ strings—ghost number is conserved—and therefore the ghost loops will form well-defined demarcation lines on the closed string world-sheet separating the "phase" where the string is propagating in the background of the ith D-brane from the phase where the string propagates in the background of the jth D-brane, as sketched in Fig. 4. Because the absence of interactions among the *ij* strings these loops are self-avoiding.

In this way we observe that the multicut solutions of the matrix model translated into a closed string picture naturally describe a system of dynamical domain walls on the world sheet. These domain walls connect different conformal field theories as was analyzed in [24]. In the open string channel the domain wall corresponds to an *ij* string stretching from one throat to another. This picture of different phases of the world sheet of a single closed topological string is a further application of the ideas in [17].

III. EXAMPLES

In this section we illustrate how matrix perturbation theory can be used to obtain nonperturbative instanton effects in various supersymmetric gauge theories. We start with some familiar examples, which include $\mathcal{N}=2$ Seiberg-Witten theory and $\mathcal{N}=1^*$ theory, where the exact answer is known to all orders. Despite the existence of the exact solution in these models, we will not need it here. Instead, our goal is to reproduce it by computing simple planar diagrams in the corresponding matrix model.

Of course, the real power of the perturbative technique is in those models where exact solution is not available. It is easy to come up with simple examples of such models. A particular example that we discuss in this section is a massive deformation of the Leigh-Strassler theory, which in turn is an (exactly marginal) deformation of the $\mathcal{N}=4$ super-Yang-Mills [25]. The case that we consider corresponds to a simple 3-matrix model with cubic interactions, solution to which is not known even in the planar limit. Nevertheless, one can systematically obtain instanton corrections to the effective superpotential from matrix perturbation theory. Similar perturbative analysis can be applied essentially to any $\mathcal{N}=1$ theory that admits a large N limit.

A. Seiberg-Witten solution from multicut matrix models

The fact that one can obtain the Seiberg-Witten solution from a perturbative analysis of the gauge theory, which in turn gets reduced to planar computations of a matrix model has already been noted in [3] as an interpretation of the string inspired derivation of Seiberg-Witten geometry in [26]. Our aim in this section is to show that even if the exact solution of matrix model were not available we could have nevertheless obtained a systematic instanton expansion for quantities of interest. So in this section we are tying one hand behind our back.

The basic idea of [26] is to deform $\mathcal{N}=2$ theory to $\mathcal{N}=1$ by a polynomial tree-level superpotential $W(x)$, which freezes the eigenvalues of the adjoint field Φ to a particular point on the Coulomb branch. For example, in the case of $U(2)$ gauge theory one deforms by a cubic superpotential of the form (2.2),

$$W'_{\rm tree}(x)=\epsilon(x-a)(x+a).$$

Here we explicitly introduced the deformation parameter ϵ, such that $\epsilon=0$ corresponds to the undeformed $\mathcal{N}=2$ theory. Choosing the configuration where one eigenvalue of Φ is at $+a$ and the other is at $-a$ determines a point on the Coulomb branch of the original $\mathcal{N}=2$ theory, and breaks the gauge group to an Abelian subgroup,

$$U(2)\rightarrow U(1)\times U(1).$$

This leads us precisely to the situation discussed in the previous section, where we studied vacua of $\mathcal{N}=1$ field theories with (partial) gauge symmetry breaking. Therefore, one should be able to compute all holomorphic quantities from the genus zero free energy $\mathcal{F}_0(S_1,S_2)$ of the corresponding two-cut matrix model. Evaluating the two-loop Feynman diagrams in the previous section we found the leading perturbative behavior of the genus zero free energy in the two-cut matrix model with a cubic interaction

$$\mathcal{F}_0^{\rm pert}(S_1,S_2)=\frac{1}{\Delta^3}\left(\frac{2}{3}S_1^3-5S_1^2S_2+5S_1S_2^2-\frac{2}{3}S_2^3\right)+\cdots.$$

One can go further and systematically compute higher-order corrections. In this way one finds a series expansion

$$\begin{aligned}\mathcal{F}_0(S_1,S_2)=&-\frac{1}{2}\sum_{i=1,2}S_i^2\log\left(\frac{S_i}{\Delta^3}\right)+(S_1+S_2)^2\log\left(\frac{\Lambda}{D}\right)\\&+\frac{1}{\Delta^3}\left(\frac{2}{3}S_1^3-5S_1^2S_2+5S_1S_2^2-\frac{2}{3}S_2^3\right)\\&+\frac{1}{\Delta^6}\left(\frac{8}{3}S_1^4-\frac{91}{3}S_1^3S_2+59S_1^2S_2^2\right.\\&\left.-\frac{91}{3}S_1S_2^3+\frac{8}{3}S_2^4\right)+\frac{1}{\Delta^9}\left(\frac{56}{3}S_1^5-\frac{871}{3}S_1^4S_2\right.\\&+\frac{2636}{3}S_1^3S_2^2-\frac{2636}{3}S_1^2S_2^3\\&\left.+\frac{871}{3}S_1S_2^4-\frac{56}{3}S_2^5\right)+\cdots.\end{aligned}\tag{3.1}$$

Here the first term receives a contribution from the measure of the unbroken gauge group $U(M_1)\times U(M_2)$ [17], where each factor gives a standard term $S_i^2\log S_i$ that reproduces the Veneziano-Yankielowicz superpotential. The one-loop diagrams for Φ and the ghosts B, C account for the Δ dependence of the first two terms in Eq. (3.1)

$$\left(\frac{1}{2}S_1^2-2S_1S_2+\frac{1}{2}S_2^2\right)\log\Delta.$$

Finally, the Λ dependence reflects the ambiguity in the cutoff of the full $U(M_1+M_2)$ gauge group and should therefore multiply $(S_1+S_2)^2$. The higher order perturbative terms have the combinatorial meaning we explained in the previous section. For example, the terms that involve only S_1 or S_2 enumerate planar cubic diagrams and were computed in [27].

Note that the function $\mathcal{F}_0(S_1,S_2)$ is symmetric in S_1 and $-S_2$. This reflects the symmetry of the potential: we can exchange the stable and unstable critical points if we change the overall sign of the potential by $g_s\rightarrow -g_s$. Since $S_i=g_sM_i$ this gives $S_1\leftrightarrow -S_2$. From the combinatorial point of view this was explained in Sec. II D in terms with the connection to the $O(2)$ model on a random surface—it is an obvious property of the Feynman rules.

We should now extremize the effective glueball superpotential

$$W_{\rm eff}(S)=\sum_i\left(N_i\frac{\partial\mathcal{F}_0(S)}{\partial S_i}-2\pi i\tau_0S_i\right).\tag{3.2}$$

In the present case we have $N_1 = N_2 = 1$ and we will also set to zero the bare coupling τ_0.

The physical quantity to compute in this model is the matrix of the $U(1)\times U(1)$ couplings in the effective low-energy theory. It is given by the second derivatives of matrix model free energy

$$\tau_{ij} = \frac{\partial^2 \mathcal{F}_0(S)}{\partial S_i \partial S_j}. \tag{3.3}$$

Note that by a scaling argument the matrix τ_{ij} does not depend on the deformation parameter ϵ and therefore it should reproduce the coupling constant of the $\mathcal{N}=2$ Seiberg-Witten theory at the relevant point of the Coulomb branch. Minimizing the effective superpotential (3.2), that in this case simplifies to

$$W_{\rm eff}(S) = \sum_i \frac{\partial \mathcal{F}_0(S)}{\partial S_i},$$

gives the condition

$$\sum_i \tau_{ij} = 0.$$

So we see that at the extremum τ_{ij} takes the form

$$\begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{pmatrix} = \tau \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \tag{3.4}$$

where τ is the effective gauge coupling for the "off-diagonal" $U(1) \subset SU(2) \subset U(2)$. Note that we automatically managed to get rid of the diagonal $U(1)$ factor by setting the bare coupling constant to zero in Eq. (3.2).

The extremization of $W_{\rm eff}(S)$ we can do using the perturbative expansion of $\mathcal{F}_0$ (3.1). However, before we do this, let us recall that, in terms of the exact solutions, this extremization has a clear geometric interpretation [7,26]. The free energy $\mathcal{F}_0$ can be described in terms of the dual geometry (2.17) that in this case of a cubic superpotential takes the form of a genus one curve

$$y^2 = (x^2 - a^2)^2 + b_1 x + b_0. \tag{3.5}$$

Here the coefficients b_1, b_0 are determined by the 't Hooft couplings S_1, S_2. In particular one has the simple relation $b_1 = -4(S_1 + S_2)$. Minimizing $W_{\rm eff}(S)$ with respect to S_1 and S_2 gives the condition

$$S_1 = -S_2. \tag{3.6}$$

Therefore the algebraic curve (3.5) reduces to nothing but the Seiberg-Witten (SW) curve for $SU(2)$ theory [28],

$$y^2 = (x^2 - u)^2 + \Lambda^4,$$

where one has to make the identification of parameters (with $\Delta = 2a$)

$$u = \frac{1}{2}\langle \mathrm{tr}\,\Phi^2 \rangle = \frac{1}{4}\Delta^2. \tag{3.7}$$

So at the extremum the free energy $\mathcal{F}_0$ can be thought of as a function of only one variable $S = S_1 = -S_2$ that is determined by the parameter Δ (or u) of the SW curve.

Both in the matrix model and in the SW solution the exact expression for the coupling τ of the off-diagonal $U(1)$ that appears in Eq. (3.4) follows directly from this geometric interpretation as the modulus of an elliptic curve. Given the parametrization of this curve, we can expand τ in terms of the variable u or Δ and obtain the exact result

$$\tau(u) = 2\log\left(\frac{\Lambda}{\Delta}\right) + 20\left(\frac{\Lambda}{\Delta}\right)^4 + 538\left(\frac{\Lambda}{\Delta}\right)^8 + \frac{62048}{3}\left(\frac{\Lambda}{\Delta}\right)^{12} + \cdots. \tag{3.8}$$

We can now reconstruct this exact solution in perturbation theory by simply evaluating the second derivative of $\mathcal{F}_0$ at the critical point up to a fixed number of loops. Given the perturbative expansion (3.1) of $\mathcal{F}_0$ in terms of a loop expansion of planar diagrams, we should first compute $S = S_1 = -S_2$ at the extremum. This gives a series of the form

$$\frac{S}{\Delta^3} = \left(\frac{\Lambda}{\Delta}\right)^4 + 6\left(\frac{\Lambda}{\Delta}\right)^8 + 140\left(\frac{\Lambda}{\Delta}\right)^{12} + 4620\left(\frac{\Lambda}{\Delta}\right)^{16} + \cdots. \tag{3.9}$$

Plugging this into $\partial^2 \mathcal{F}_0 / \partial S^2$ gives us a systematic approximation of the effective coupling τ. It is instructive to see how the instanton expansion of τ computed from the n-loop free energy of the matrix model for various n gives a sequence of series expansions gradually converging to the exact result:

$$\begin{aligned}
\tau_{\text{1-loop}} &= 2\log\left(\frac{\Lambda}{\Delta}\right), \\
\tau_{\text{2-loop}} &= 2\log\left(\frac{\Lambda}{\Delta}\right) + 20\left(\frac{\Lambda}{\Delta}\right)^4 + 120\left(\frac{\Lambda}{\Delta}\right)^8 + 1080\left(\frac{\Lambda}{\Delta}\right)^{12} + \cdots, \\
\tau_{\text{3-loop}} &= 2\log\left(\frac{\Lambda}{\Delta}\right) + 20\left(\frac{\Lambda}{\Delta}\right)^4 + 538\left(\frac{\Lambda}{\Delta}\right)^8 + 7816\left(\frac{\Lambda}{\Delta}\right)^{12} + \cdots, \\
\tau_{\text{4-loop}} &= 2\log\left(\frac{\Lambda}{\Delta}\right) + 20\left(\frac{\Lambda}{\Delta}\right)^4 + 538\left(\frac{\Lambda}{\Delta}\right)^8 + \frac{62048}{3}\left(\frac{\Lambda}{\Delta}\right)^{12} + \cdots, \\
&\vdots \\
\tau_{\text{exact}} &= 2\log\left(\frac{\Lambda}{\Delta}\right) + 20\left(\frac{\Lambda}{\Delta}\right)^4 + 538\left(\frac{\Lambda}{\Delta}\right)^8 + \frac{62048}{3}\left(\frac{\Lambda}{\Delta}\right)^{12} + \cdots.
\end{aligned} \tag{3.10}$$

As an aside we point out that the condition $S_1 = -S_2$, which naturally emerges from minimizing the effective superpotential, means that from the point of view of the matrix model we are dealing with a symmetric filling of the two cuts. The exact solution to this model has interesting properties and is further analyzed in Appendix A. In particular there it is discussed that this model, as well as its generalization with asymmetric filling of the two cuts, has a nontrivial scaling limit in the universality class of the $c=1$ string.

Remembering the relation $S_i = g_s M_i$, we see that because of the minus sign in Eq. (3.6) in the symmetric filling the number of eigenvalues in the unstable cut (the maximum of the potential) is negative. This is clearly an unphysical solution and should be interpreted as obtained by analytic continuation. In fact, if we put a positive number of eigenvalues at an unstable critical point the eigenvalue cut will not lie on the real axis but the cut will rotate itself along the imaginary axis. (This can be seen by simply analytically continuing $\Phi \to i\Phi$ in the Gaussian approximation.) Instead of working with negative numbers it is perhaps better to think of this solutions in terms of ''eigenvalue holes'' obtained by filling the Dyson sea almost to the top of the potential.

Finally, let us point out that using matrix model results we could also obtain other holomorphic quantities, such as the SW periods a and a_D. For example, for the expectation values $\langle \mathrm{tr}\,\Phi^k \rangle$ one finds a nice expression,

$$\langle \mathrm{tr} h \Phi^k \rangle = \oint x^k h,$$

written in terms of the 1-form $h = W''(x)dx/y$, which can be interpreted as the smeared density of the eigenvalues of the adjoint field Φ [3,26]. In particular, the case $k=1$ gives rise to the SW period a.

B. $\mathcal{N}=1^*$ theory

As another illustration of the perturbative technique in the matrix model applied to nonperturbative gauge theory, we consider a massive deformation of $\mathcal{N}=4$ gauge theory, the so-called $\mathcal{N}=1^*$ theory. In $\mathcal{N}=1$ superspace this theory is described by $U(N)$ gauge theory with three adjoint chiral superfields and a tree-level superpotential,

$$W_{\mathrm{tree}} = \mathrm{tr}\left(g\Phi_1[\Phi_2,\Phi_3] + \frac{m}{2}\sum_{i=1}^{3} \Phi_i^2 \right). \tag{3.11}$$

For simplicity, we also assume that the eigenvalues of all the Higgs fields are in the same classical vacuum (perturbation theory around other vacua is discussed in Appendix B). Computing planar Feynman diagrams up to three loops in the corresponding matrix model we will be able to reproduce the leading terms in the (exact) effective superpotential of $\mathcal{N}=1^*$ theory.

From the topology of planar Feynman diagrams in this matrix model it is easy to see that the free energy, $\mathcal{F}_0(S)$, has the following structure:

$$\mathcal{F}_0(S) = \sum_k c_{k+1} \frac{g^{2k}}{m^{3k}} S^{k+2}. \tag{3.12}$$

Following the notations of [3], henceforth we set $g=1$.

Given the matrix model free energy $\mathcal{F}_0(S)$, one can compute the effective superpotential $W_{\mathrm{eff}}(S)$ using the relation (2.5). Furthermore, integrating out the field S in $W_{\mathrm{eff}}(S)$ gives the effective superpotential as a function of the coupling con-

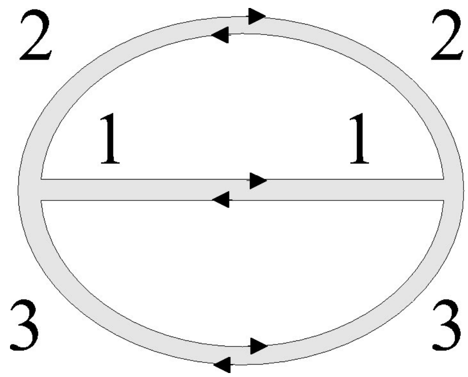

FIG. 5. Two-loop contribution to the free energy in the 3-matrix model corresponding to the $\mathcal{N}=1^*$ theory. The numbers next to propagators label the choice of one of the three matrix fields.

stants. For the $\mathcal{N}=1^*$ theory the answer can be computed explicitly [3] by using the matrix model techniques developed in [29]. Specifically, one obtains

$$W_{\mathrm{eff}} = -\frac{Nm^3}{24} E_2(\tau), \tag{3.13}$$

where $\tau = \tau_0/N$ and $E_2(\tau)$ is the Eisenstein series. This agrees with the analysis of [30] based on field theory dualities. Up to an additive constant, we can write the effective superpotential (3.13) as a power series in the variable $q = \exp(2\pi i\tau)$,

$$W_{\mathrm{eff}} = Nm^3(q + 3q^2 + 4q_3 + 7q^4 + 6q_5 + \cdots). \tag{3.14}$$

Our goal is to reproduce this result by the perturbative technique in the corresponding 3-matrix model

$$\int d\Phi \exp\left[-\mathrm{tr}\left(\Phi_1[\Phi_2,\Phi_3] + \frac{m}{2}\sum_{i=1}^{3} \Phi_i^2 \right)\right]. \tag{3.15}$$

Namely, computing the planar Feynman diagrams up to three loops we shall find numerical coefficients c_k in the perturbative series (3.12) and, in particular, to check the first few coefficients in Eq. (3.14).

The two-loop contribution to $\mathcal{F}_0$ comes from the Feynman diagrams of the type shown in Fig. 5. It is one of the diagrams that appears in a simple 1-matrix model with cubic potential, see Fig. 1. The second type of two-loop diagrams in Fig. 1 does not appear here due to the index structure of the cubic interaction. Thus, we obtain the two-loop coefficient $c_2 = -1$.

At the next order, i.e., at three loops, there are two types of diagrams which are presented in Fig. 6. Taking into account also the index structure of the diagrams one finds many different terms. Adding all of them together gives $c_3 = 7/2$.

Summarizing, up to three loops the perturbative expansion of $\mathcal{F}_0(S)$ has the form

$$\mathcal{F}_0(S) = -\frac{S^3}{m^3} + \frac{7}{2}\frac{S^4}{m^6} + \cdots. \tag{3.16}$$

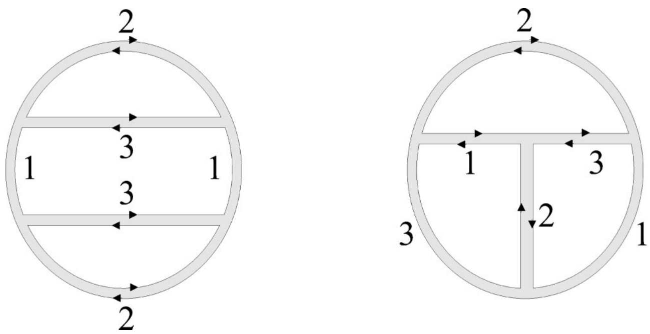

FIG. 6. Two types of three-loop diagrams that contribute to $\mathcal{F}_0$ with one of the possible labeling of the propagators.

Note that the expansion we find indeed has the general structure expected in Eq. (3.12). The relative minus signs in this expansion are due to the interaction vertices with both positive and negative weight arising from the commutator in Eq. (3.15).

Substituting Eq. (3.16) into Eq. (2.5) we obtain the leading behavior of the effective superpotential

$$W_{\mathrm{eff}}=NS\log(S/m^3)-2\pi i\tau_0 S-3N\frac{S^2}{m^3}+14N\frac{S^3}{m^6}+\cdots. \tag{3.17}$$

Now, integrating out the gluino field S we obtain the final expression for the effective superpotential

$$W_{\mathrm{eff}}=-Nm^3q-3Nm^3q^2-4Nm^3q^3+\cdots. \tag{3.18}$$

The leading coefficients in this expression agree with the first coefficients in the expansion of the exact answer (3.14), written in terms of the Eisenstein series $E_2(\tau)$.

Since we can do this calculation order by order, and since n-loop diagrams give rise to n-instanton terms in W_{eff}, it is instructive to look at the higher order terms and to see how the result depends on n. For example, if we keep only the leading term in the perturbative series $\mathcal{F}_0$, the superpotential (3.17) looks like

$$W_{\mathrm{eff}}=NS\log(S)-2\pi i\tau_0 S-3NS^2. \tag{3.19}$$

This leads to the effective superpotential

$$W_{\mathrm{eff}}=Nm^3(q+3q^2+18q^3+144q^4+1350q^5+\cdots), \tag{3.20}$$

where we retained the terms of higher order in q, most of which cannot be trusted in this approximation.

If we compute perturbative free energy $\mathcal{F}_0$ to three loops, as we did above, we obtain the effective superpotential (3.18), where one can trust three leading terms. Moreover, the values of the higher order terms in Eq. (3.18) are slightly "improved" compared to Eq. (3.20). One can continue and do a similar calculation up to four loops and so on. As a result, one finds a sequence of instanton expansions which gradually approach the exact answer (3.14):

$$W_{1\text{-loop}}=Nm^3\underline{q},$$

$$W_{2\text{-loop}}=Nm^3\left(\underline{q+3q^2}+18q^3+144q^4+1350q^5+\frac{69984}{5}q^6+\frac{777924}{5}q^7+\cdots\right),$$

$$W_{3\text{-loop}}=Nm^3\left(\underline{q+3q^2+4q^3}-108q^4-1548q^5-\frac{43416}{5}q^6+\frac{345744}{5}q^7+\cdots\right),$$

$$W_{4\text{-loop}}=Nm^3\left(\underline{q+3q^2+4q^3+7q^4}+1212q^5+\frac{108384}{5}q^6+\frac{874744}{5}q^7+\cdots\right),$$

$$W_{5\text{-loop}}=Nm^3\left(\underline{q+3q^2+4q^3+7q^4+6q^5}-\frac{72516}{5}q^6-\frac{1657856}{5}q^7+\cdots\right),$$

$$W_{6\text{-loop}}=Nm^3(\underline{q+3q^2+4q^3+7q^4+6q^5+12q^6}+190976q^7+\cdots),$$

$$\vdots$$

$$W_{\mathrm{exact}}=Nm^3(\underline{q+3q^2+4q^3+7q^4+6q^5+12q^6+8q^7+\cdots}). \tag{3.21}$$

Here, the underlined terms represent the exact terms in the instanton expansion whose coefficients "stabilize" beyond a certain order. It is curious to note that although all the numerical coefficients in the exact superpotential W_{exact} are integer numbers, it is not the case for the result obtained from a finite number of loops in matrix perturbation theory. Moreover, the n-loop approximation to W_{exact} is not a modular form, and one can see from the examples listed above that in the truncation to n loops the mistake in the $(n+1)$th coefficient is quite large. This emphasizes the fact that the Montonen-Olive duality is not put in by hand in this formalism, but rather is derived. In this sense, we are going one step beyond duality.

Note that we can express the S-duality of the $N=1^*$ theory succinctly as the statement that the effective glueball superpotential $W_{\mathrm{eff}}(S)$ is given by a *Legendre transform* of a modular form, in this given by $E_2(\tau)$ (with $\tau=\tau_0/N$).

C. Massive deformation of the Leigh-Strassler model

So far we considered only examples for which an exact solution was already known. This was helpful for establishing some confidence in the perturbative technique since it did not rely on the existence of the exact results, which we used

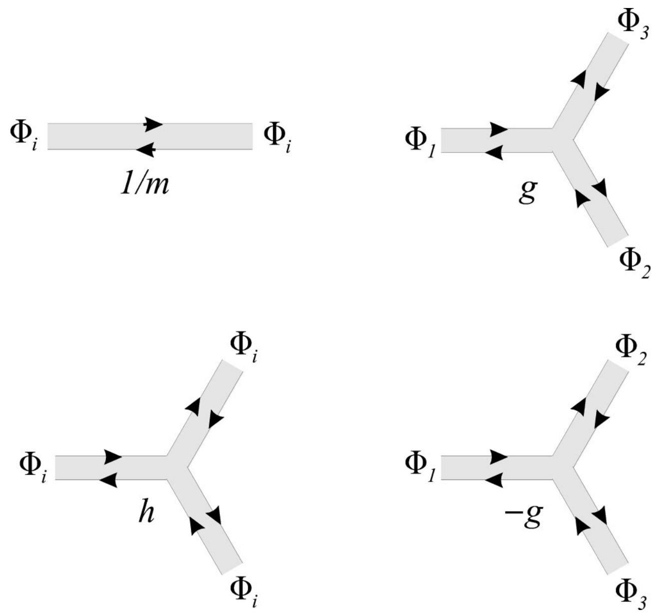

FIG. 7. The Feynman rules in the perturbative 3-matrix model corresponding to the massive deformation of the Leigh-Strassler model.

only to verify the perturbative answer. As we explained in the Introduction, in most of the models we do not have this luxury and, therefore, perturbative analysis remains as the only tool for obtaining nonperturbative results, such as instanton expansion of the effective superpotential. Here, we consider one such model.

Specifically, we consider a Leigh-Strassler deformation [25] of the model discussed in the previous section,

$$W_{\mathrm{tree}}=\mathrm{tr}\left(g\Phi_1[\Phi_2,\Phi_3]+\frac{h}{3}\sum_{i=1}^{3}\Phi_i^3+\frac{m}{2}\sum_{i=1}^{3}\Phi_i^2\right). \tag{3.22}$$

The corresponding 3-matrix model with action given by $W_{\mathrm{tree}}(\Phi_i)$ can be solved in the large M limit if either $g=0$ or $h=0$, but the exact solution is not known when both deformation parameters, g and h, are nonzero. On the other hand, perturbation theory is very simple, with the Feynman rules summarized in Fig. 7.

At the two-loop order, we find the following expression for the genus zero free energy:

$$\mathcal{F}_0=\frac{S^3}{m^3}(2h^2-g^2)+\cdots.$$

Substituting this into Eq. (2.5) gives the effective superpotential

$$W_{\mathrm{eff}}=NS\log(S/m^3)-2\pi i\tau_0 S+3(g^2-2h^2)N\frac{S^2}{m^3}+\cdots.$$

Finally, extremizing it with respect to S we obtain the value of the effective superpotential in the vacuum:

$$W_{\mathrm{eff}}=Nm^3[q+3(g^2-2h^2)q^2+\cdots].$$

The same technique applies to any $\mathcal{N}=1$ theory that admits a large N limit. In particular, one can systematically compute instanton corrections to the effective superpotential in a large class of $\mathcal{N}=1$ theories with any number of adjoint fields and generic tree-level superpotentials.

ACKNOWLEDGMENTS

We would like to thank M. Aganagic, R. Bousso, F. Cachazo, S. J. Gates, Jr., M. Mariño, A. Marshakov, H. Ooguri, S. Theisen, and K. Zarembo for valuable discussions. The research of R.D. is partly supported by FOM and the CMPA grant of the University of Amsterdam. S.G. is supported by the Clay Mathematics Institute, RFBR grants 01-01-00549 and 02-01-06322. V.A.K. is partly supported by European Union under the RTN contracts HPRN-CT-2000-00122 and -00131. C.V. is partly supported by NSF grants PHY-9802709 and DMS-0074329. We would like to thank the Max Planck Institute in Potsdam (V.A.K.), Ecole Normale Supérieure (S.G.), and Harvard University (R.D.) for kind hospitality during part of this work.

APPENDIX A: LARGE *M* SOLUTION OF THE TWO-CUT MATRIX MODEL

The results of perturbative expansions [7] used in this section can be reproduced, in accordance with the observations of [3], from the direct solution of the matrix model (2.1), as was shown in [1]. In this appendix we review, for the sake of completeness, both the matrix model derivation as well as the analytic form of the glueball superpotential in terms of elliptic functions. We take the cubic potential to be of the form

$$\begin{aligned}W(\Phi)&=\mathrm{tr}\left(\frac{1}{4}\Phi-\frac{1}{3}\Phi^3\right)\\&=\mathrm{tr}\left[\pm\frac{1}{2}\left(\Phi\pm\frac{1}{2}\right)^2-\frac{1}{3}\left(\Phi\pm\frac{1}{2}\right)^3\mp\frac{1}{12}\right].\end{aligned} \tag{A1}$$

The last line is the expansion around each of the two symmetric extrema of the potential. Note that we set here $\Delta=1$.

In terms of the eigenvalues, using Eq. (2.13), we write the usual for the one-matrix model saddle point equation (SPE) in the large M limit, in terms of the eigenvalues,

$$x^2-\frac{1}{4}=2\lambda\,⨍ du\,\rho(u)\frac{1}{x-u}, \tag{A2}$$

where $\lambda=g_sM$ is the overall 't Hooft coupling. The two-cut solution can be found in terms of the analytical function

$$\begin{aligned}G(x)&=2\left[\int_{x_1}^{x_2}+\int_{x_3}^{x_4}\right]du\,\rho(u)\frac{1}{x-u}\\&=\frac{1}{\lambda}\left[x^2-\frac{1}{4}-\sqrt{(x-x_1)(x-x_2)(x-x_3)(x-x_4)}\right]\end{aligned} \tag{A3}$$

having the large x asymptotics $G(x\to\infty)=2[(S_1+S_2)/\lambda x]$ and the corresponding couplings on each of the two intervals $S_j=g_sM_j$, $j=1,2$, finite in the limit $g_s\to 0$, M, $M_j\to\infty$.

The limits x_i, $i=1,2,3,4$ are defined by the large x asymptotics,

$$\sum_i x_i=0,$$

$$\sum_i x_i^2=1, \tag{A4}$$

$$\sum_i x_i^3=12(S_1+S_2),$$

and by the normalization condition for the two intervals. The latter is given in terms of the elliptic integrals

$$S_1=\frac{1}{2\pi}\int_{x_1}^{x_2}dx\sqrt{(x_1-x)(x-x_2)(x-x_3)(x-x_4)},$$

$$S_2=\frac{1}{2\pi}\int_{x_3}^{x_4}dx\sqrt{(x-x_1)(x-x_2)(x-x_3)(x-x_4)}. \tag{A5}$$

Let us now compute the free energy $\mathcal{F}_0(S_1,S_2)=(1/M^2)\log Z$. From the eigenvalue representation of the matrix model we obtain the derivative of the free energy, amounting to the removal of the eigenvalue at the edge of each cut,

$$\partial_{S_1}\mathcal{F}_0(S_1,S_2)=g_s^{-1}[\mathcal{F}_0(S_1,S_2)-\mathcal{F}_0(S_1-g_s,S_2)]$$
$$=\frac{1}{\lambda^2}W(x_1)+\frac{2}{M\lambda}\sum_{j\neq 1}\log(x_1-x_j),$$

and a similar expression for $\partial_{S_2}\mathcal{F}_0(S_1,S_2)$. In terms of the eigenvalue density this gives

$$\lambda^2\partial_{S_1}\mathcal{F}_0(S_1,S_2)=W(x_1)+\frac{1}{2\pi i}\left[\oint_{x_1}^{x_2}+\oint_{x_3}^{x_4}\right]\times\sqrt{(x_1-x)(x-x_2)(x-x_3)(x-x_4)}\times\log(x_1-x)dx,$$

$$\lambda^2\partial_{S_2}\mathcal{F}_0(S_1,S_2)=W(x_4)+\frac{1}{2\pi i}\left[\oint_{x_1}^{x_2}+\oint_{x_3}^{x_4}\right]\times\sqrt{(x_1-x)(x-x_2)(x-x_3)(x-x_4)}\times\log(x_4-x)dx.$$

By expanding the contour of integration we pick up the contribution on the logarithmic cut (apart from singularities at $x=\infty$ which we have to subtract in the matrix model framework). This gives

$$\partial_{S_1}\mathcal{F}_0(S_1,S_2)=W(x_1)+\Pi_1+\text{subtractions for }\Lambda_0\to\infty,$$

$$\partial_{S_2}\mathcal{F}_0(S_1,S_2)=W(x_4)+\Pi_2+\text{subtractions for }\Lambda_0\to\infty, \tag{A6}$$

where $\Lambda_0\to\infty$ is a cutoff and

$$\Pi_1=\frac{1}{\pi}\int_{-\Lambda_0}^{x_1}\sqrt{(x-x_1)(x-x_2)(x-x_3)(x-x_4)}dx,$$

$$\Pi_2=-\frac{1}{\pi}\int_{x_4}^{\Lambda_0}\sqrt{(x-x_1)(x-x_2)(x-x_3)(x-x_4)}dx \tag{A7}$$

are the dual periods. Formulas of this type appeared in [31]; see also [1]. In [7] they follow from the analysis of the Calabi-Yau geometry with flux. Using Eqs. (A5), (A6), (A7), and (A4) one finds the small S_1, S_2 expansion for the free energy itself (3.1) from [7].

Let us note that the branch points are not necessarily placed on the real axis. For a general complex g_s, they will choose their positions according to the steepest decent in the eigenvalue integral. For a real g_s the stable cut will be on the real axis, whereas the unstable cut will cross the real axis, having the complex conjugated branch points. The situation when all branch points are on the real axis corresponds to the analytical continuation in the (originally positive) variables: $S_1>0$, $S_2<0$.

1. Symmetric filling of two intervals

Let us consider the case of the symmetric filling of two intervals $x\in(b,a)$ and $x\in(-a,-b)$. It corresponds to the "unphysical" filling parameters $\frac{1}{2}g_sM=S=S_1=-S_2>0$, but nevertheless it will reproduce the corresponding particular case of planar graph expansion considered in the previous section. One can say that the two intervals are filled by $M/2$ eigenvalues and $M/2$ "holes," respectively. As discussed in Sec. III A, this case describes the $SU(2)$ Seiberg-Witten (SW) solution. We will also see yet another way of obtaining a $c=1$ noncritical string at a self-dual radius from matrix models, when the end points of the cuts approach each other.

The function $G(x)$, having the large x asymptotics $G(x\to\infty)=O(1/x^2)$, can be represented as

$$\lambda G(x)=x^2-\frac{1}{4}-\sqrt{(x^2-a^2)(x^2-b^2)}.$$

The function $G(x)$, having the large x asymptotics $G(x\to\infty)=O(1/x^2)$, can be represented as

$$\lambda G(x)=x^2-\frac{1}{4}-\sqrt{(x^2-a^2)(x^2-b^2)}.$$

The large x asymptotics fixes one relation between a and b: $a^2+b^2=\frac{1}{2}$, and the normalization of the density $\int_b^a(dx/2\pi\lambda)\sqrt{(a^2-b^2)(x^2-b^2)}=1$ gives the relation (using [32],[2] 217.27 and 361.01)

$$\lambda S=\frac{1}{2\pi}\int_b^a dx\sqrt{(a^2-x^2)(x^2-b^2)}$$
$$=\frac{a^3}{6\pi}[(2-m)\mathbf{E}-2(1-m)\mathbf{K}], \tag{A8}$$

where $\mathbf{K}(m)$ and $\mathbf{E}(m)$ are the elliptic integrals of the first and second kind, $a=1/\sqrt{4-2m}$ and the elliptic nome is $m=1-b^2/a^2$. The derivative of the free energy (A6) can be calculated by the deformation of the contour to the dual period corresponding to the interval $(-b,b)$ as the complete elliptic integral

$$\partial_S\mathcal{F}(S,-S)=\frac{2}{\lambda}\int_{-b}^{b}\sqrt{(x^2-a^2)(x^2-b^2)}dx. \tag{A9}$$

However, the simplest quantity to calculate is actually the second derivative of the free energy, which is to be identified with the τ parameter of the SW curve. The latter can be seen already in the form of Eq. (A9). Indeed, by writing $(x^2-a^2)(x^2-b^2)=(x^2-1/4)^2-\Lambda^4$, where $\Lambda^4=m^2/16(2-m)^2$, we obtain

$$\partial_S^2\mathcal{F}_0(S,-S)=\frac{\partial_m\partial_S\mathcal{F}(S,-S)}{\partial_m S}=4\mathbf{K}(1-m)/\mathbf{K}(m)\equiv 4\tau. \tag{A10}$$

We found the explicit elliptic parametrization of the free energy: it is parametrized by m which can be expressed through S by means of Eq. (A8).

Expanding Eqs. (A8) and (A10) in powers of $\Lambda^4=m^2/16(2-m)^2$, we get

$$S=\left(\frac{\Lambda}{\sqrt{2}}\right)^4+6\left(\frac{\Lambda}{\sqrt{2}}\right)^8+140\left(\frac{\Lambda}{\sqrt{2}}\right)^{12}+4620\left(\frac{\Lambda}{\sqrt{2}}\right)^{16}+\cdots, \tag{A11}$$

$$\tau=-\frac{i}{\pi}\left(2\log(\Lambda^2/8)+\frac{5}{2^3}\Lambda^4+\frac{269}{2^{10}}\Lambda^8+\frac{1939}{3\cdot 2^{12}}\Lambda^{12}+\frac{922253}{2^{23}}\Lambda^{16}+\cdots\right). \tag{A12}$$

The last is precisely the instanton expansion of the SW coupling constant; see, e.g., [33]. It is not surprising since the numerator and denominator of Eq. (A10) coincide (up to the same factor) with ω_D and ω from the formula (2.2) of [33], whose ratio defines τ of course. Restoring the modulus u and rescaling $\Lambda^2\to 2\Lambda^2$, one can write this result in conventions[3] of Ref. [34], which also agree with our conventions used in Sec. III A. Specifically, one finds (up to an overall numerical factor):

$$\tau=\log(\Lambda^2/4u)+\frac{5}{4}\frac{\Lambda^4}{u^2}+\frac{269}{2^7}\frac{\Lambda^8}{u^4}+\frac{1939}{3\cdot 2^7}\frac{\Lambda^{12}}{u^8}+\frac{922253}{2^{16}}\frac{\Lambda^{16}}{u^8}+\cdots. \tag{A13}$$

This is in agreement with Eq. (3.8), as follows directly from the identification (3.7) of the u variable with $\Delta^2/4$.

It is not surprising that inverting the series for S plugging it into Eq. (A10) and expanding in S we obtain

$$\partial_S^2\mathcal{F}_0(S,-S)=2\log S+68S+1500S^2+\frac{142520}{3}S^3+O(S^4), \tag{A14}$$

which coincides in the particular case $S=S_1=-S_2$ with the expansion from [7], quoted in Sec. III A (we put the dimensionful coupling $\lambda=1$).

Using Eqs. (A8) and (A10) we could also expand $\mathcal{F}$ itself in terms of the variable $q=e^{-\pi\tau}=e^{-\pi\mathbf{K}(1-m)/\mathbf{K}(m)}$, which will be the instanton expansion for the corresponding $\mathcal{N}=1$ symmetry theory with the $U(2)\to U(1)\times U(1)$ symmetry breaking cubic tree potential, according to the recipe of [3].

2. $c=1$ critical regime

In the context of the cubic potential matrix model, there are two distinct ways of getting a $c=1$ noncritical string: As noted in [35] $c=1$ at a self-dual radius is equivalent to the topological B-model on the deformed conifold, which in turn has been shown to be equivalent to the matrix model with quadratic potential [1]. Thus in the theory we are dealing with, if we zoom to the region near the critical points of the potential we obtain a $c=1$ system at a self-dual radius. However, there is another way of obtaining $c=1$ as well: We can consider the limit where the two ends of the cuts touch each other, which again leads to a conifold geometry but now the vanishing cycle is "magnetic" relative to the original "electric" cycle of the matrix model.

Let us now look at this regime in more detail. This corresponds to $b\to 0$, when $m_1=1-m\to 0$ and the two cuts in $F(x)$ merge into one. From Eq. (A8) we obtain in this limit

$$\lambda S\simeq\frac{1}{6\pi}-\frac{1}{8\pi}m_1\log(1/m_1)+O(m')$$

which shows that this transition happens at $S_c=1/6\pi\lambda$. The free energy (A10) in this limit is

$$\partial_S^2\mathcal{F}_0(S,-S)\simeq\frac{2}{\log 16/m_1}+O(m_1)$$

[2] Beware of a mistake there: $g\to a$.

[3] See also the footnote on page 3 in [34].

and has, as a function of $\delta=\lambda(S_c-S)$, the typical behavior of the $c=1$ matter coupled to the 2D gravity [36],

$$\mathcal{F}_0(S)-\mathcal{F}_0(S_c)\simeq\lambda^2\frac{\delta^2}{\log(1/6)}.$$

APPENDIX B: MASSIVE VACUA OF $\mathcal{N}=1^*$ THEORY

In this appendix we discuss matrix perturbation theory for nontrivial massive vacua of the $\mathcal{N}=1^*$ theory, corresponding to higher spin representations of $SU(2)$. As we shall see, there are some novelties in perturbation theory, which make these vacua conceptually similar to multicut matrix models. In fact, the vacua we are going to discuss also correspond to a multicut matrix model [10]. In both cases one finds (partial) gauge symmetry breaking which leads to new fermionic ghost degrees of freedom.

In order to describe this more specifically, let us rewrite the tree-level superpotential (3.11) in $\mathcal{N}=1^*$ theory in the following form:

$$W_{\mathrm{tree}}=\mathrm{tr}\left(i[\Phi_1,\Phi_2]\Phi_3+\sum_{i=1}^{3}\Phi_i^2\right). \tag{B1}$$

Supersymmetric vacua of the gauge theory correspond to the critical points of this superpotential. Thus, extremizing Eq. (B1) we find

$$[\Phi_1,\Phi_2]=2i\Phi_3 \tag{B2}$$

plus two similar equations obtained by permutation of indices 1, 2, 3. One obvious solution corresponds to $\Phi_i=0$. However, there are also some nontrivial solutions, corresponding to p-dimensional representations of $SU(2)$. In fact, suppose we start with a $U(N)$ gauge theory, with $N=pn$. Then, we can take n copies of such p-dimensional representations. This leads to a partial breaking of gauge symmetry,

$$U(pn)\to U(n). \tag{B3}$$

Note that the rank of the gauge group has been reduced in this case due to the fact that the irreducible representation we have taken for vacuum configurations are not one dimensional. The exact effective superpotential for all values of p is known [30,37,10], and can be written in terms of the Eisenstein series $E_2(\tau)$,

$$W_{\mathrm{eff}}=-\frac{Np^2}{12}E_2(\tau), \tag{B4}$$

very much like the superpotential in the for trivial vacuum, $p=1$. The only novelty here is the relation between τ and the bare coupling constant,

$$\tau=p(p\tau_0+k)/N.$$

In the effective field theory, this relation is set by the tree-level term and the one-loop anomaly term in the superpotential. The functional dependence on τ, on the other hand, is determined by matrix perturbative expansion $\mathcal{F}_0$ (around the corresponding vacuum). Since for all values of p we have the same functional dependence on τ—given by the Eisenstein series—we conclude that $\mathcal{F}_0$ should be the same for all vacua, i.e., for all values of p,

$$\mathcal{F}_0=-S^3+\frac{7}{2}S^4+\cdots. \tag{B5}$$

In order to reproduce this result directly by perturbative techniques in the matrix model, we have to expand the superpotential (B1) near a vacuum,

$$\Phi_1\to X+\Phi_1,\quad \Phi_1\to Y+\Phi_2,\quad \Phi_1\to Z+\Phi_3,$$

where X, Y, and Z solve Eq. (B2):

$$[X,Y]=2iZ,\quad \text{etc.} \tag{B6}$$

Substituting this into Eq. (B1) we find

$$W_{\mathrm{tree}}=\mathrm{tr}\Bigg(i[\Phi_1,\Phi_2]\Phi_3+\sum_{i=1}^{3}\Phi_i^2+iX[\Phi_2,\Phi_3] +iY[\Phi_3,\Phi_1]+iZ[\Phi_1,\Phi_2]\Bigg). \tag{B7}$$

Let us consider a specific case, corresponding to $p=2$. In this case, we have the following gauge symmetry breaking pattern:

$$U(2M)\to U(M). \tag{B8}$$

Hence, it is convenient to write all the matrix variables in terms of $M\times M$ blocks. Specifically, we take [it is easy to check that this is indeed a solution to Eq. (B2)]

$$X=\begin{pmatrix}0 & \mathbf{1}\\ \mathbf{1} & 0\end{pmatrix},\quad Y=i\begin{pmatrix}0 & -\mathbf{1}\\ \mathbf{1} & 0\end{pmatrix},\quad Z=\begin{pmatrix}\mathbf{1} & 0\\ 0 & -\mathbf{1}\end{pmatrix},$$

and for each Hermitian matrix Φ_i we introduce the notation

$$\Phi=\frac{1}{2}\begin{pmatrix}A^++A^- & D+iF\\ D-iF & A^+-A^-\end{pmatrix}, \tag{B9}$$

where $A^\pm$, D, and F are $M\times M$ matrices. Using this decomposition for all of the three matrix fields Φ_i, we get in total $3\times4=12$ matrices of size $M\times M$:

$$A_1^\pm,\quad D_1,\quad F_1,\quad A_2^\pm,\quad D_2,\quad F_2,\quad A_3^\pm,\quad D_3,\quad F_3. \tag{B10}$$

However, the gauge symmetry breaking (B8) suggests that $3M^2$ degrees of freedom can be gauge fixed to zero, so that effectively we should end up only with nine matrix fields. This is precisely what one finds.

Rewriting Eq. (B7) in terms of $M\times M$ matrices gives the following quadratic (mass) terms:

$$W_{\rm quadr}=\mathrm{tr}\Bigg(\frac{1}{2}\sum_i (A_i^+)^2+\frac{1}{2}D_1^2+\frac{1}{2}F_2^2+\frac{1}{2}(A_3^-)^2+D_1F_2 +F_2A_3^- -D_1A_3^- +\frac{1}{2}(D_2-F_1)^2+\frac{1}{2}(A_2^- -F_3)^2 +\frac{1}{2}(A_1^- +D_3)^2\Bigg). \tag{B11}$$

Here, the fields in the first and second lines have nondegenerate mass matrix. However, the fields in the third line appear only in certain linear combinations. Hence, their orthogonal combinations,

$$D_2+F_1,\quad A_2^- +F_3,\quad A_1^- -D_3, \tag{B12}$$

represent massless directions and can be potentially dangerous in the matrix integral. In fact, these are simply the usual Goldstone zero modes which can be removed by gauge fixing. We choose the following gauge, suggested by Eq. (B12):

$$D_2=-F_1,\quad A_2^-=-F_3,\quad A_1^-=D_3. \tag{B13}$$

This eliminates three out of twelve $M\times M$ matrices. For example, if we choose to eliminate D_2, A_2^-, and A_1^-, we end up with nine bosonic matrices:

$$A_1^+,\quad D_1,\quad F_1,\quad A_2^+,\quad F_2,\quad A_3^+,\quad A_3^-,\quad D_3,\quad F_3. \tag{B14}$$

Next, we should introduce fermionic ghost fields B,C. In order to do this, we note that under $SU(2M)$ gauge transformation the matrix fields Φ_i transform as

$$\delta\Phi\sim[\Phi,C].$$

Again, we write C in the 2×2 block form, similar to Eq. (B9):

$$C=\frac{1}{2}\begin{pmatrix} C_A & C_D+iC_F \\ C_D-iC_F & -C_A\end{pmatrix}. \tag{B15}$$

Applying the gauge transformation to Eq. (B13) and using the standard Faddeev-Popov method, one finds the action for the ghost fields B_α, C_α, where we introduced a new index notation $\alpha=A,D,F$. A straightforward, but slightly technical, calculation gives

$$W_{\rm ghost}=\mathrm{tr}\Bigg\{8iB_AC_A-4iB_DC_D-4iB_FC_F+\frac{1}{2}B_A[2iC_A(D_1-F_2)+2i(D_1-F_2)C_A+C_D(-A_2^+-iD_3)+C_F(-A_1^+-iF_3) +(A_2^+-iD_3)C_D+(A_2^+-iF_3)C_F]+\frac{1}{2}B_D[2C_A(iD_3-A_2^+)+2(A_2^++iD_3)C_A+C_D(iF_2-iA_3^-)+C_F(-A_3^++iF_1) +(iF_2-iA_3^-)C_D+(iF_1+A_3^+)C_F]+\frac{1}{2}B_F[2C_A(iF_3-A_1^+)+2(A_1^++iF_3)C_A+C_D(iF_1+A_3^+)+C_F(-iD_1-iA_3^-) +(iF_1-A_3^+)C_D+(-iD_1-iA_3^-)C_F]\Bigg\}. \tag{B16}$$

Summarizing, in the case of $p=2$ we find a $(9+6)$-matrix model, that is a matrix model with nine bosonic and six fermionic (ghost) fields,

$$\text{bosonic:}\quad A_1^+,\ D_1,\ F_1,\ A_2^+,\ F_2,\ A_3^+,\ A_3^-,\ D_3,\ F_3,$$
$$\text{fermionic:}\quad B_A,\ B_D,\ B_F,\ C_A,\ C_D,\ C_F, \tag{B17}$$

and with the following action:

$$W_{\rm tree}=W_{\rm quadr}+W_{\rm cubic}+W_{\rm ghost}, \tag{B18}$$

where the ghost action is given by Eq. (B16). The quadratic terms of the bosonic action are given by Eq. (B11):

$$W_{\rm quadr}=\mathrm{tr}\Bigg(\frac{1}{2}\sum_i (A_i^+)^2+2F_1^2+2F_3^2+2D_3^2+\frac{1}{2}D_1^2+\frac{1}{2}F_2^2+\frac{1}{2}(A_3^-)^2+D_1F_2+F_2A_3^- -D_1A_3^-\Bigg),$$

while the cubic interactions read

$$W_{\mathrm{cubic}}=\mathrm{tr}\bigg(\frac{i}{4}\{[A_1^+,A_2^+]A_3^+ +([F_1,F_2]-[D_1,F_1]-[D_3,F_3])A_3^+ +([D_3,D_1]+[F_3,F_1])A_2^+ \\ +([D_3,F_1]+[F_2,F_3])A_1^+ -[A_1^+,F_3]A_3^- +[D_3,A_2^+]A_3^- +i(-2F_1^2-D_1F_2-F_2D_1)A_3^- \\ +i(-2F_3^2D_1+F_1F_3D_3+F_1D_3F_3+2F_2D_3^2+F_1F_3D_3+F_1D_3F_3)\}\bigg). \tag{B19}$$

Computation of the planar Feynman diagrams in this matrix model is expected to reproduce the perturbative expansion of the free energy (B5). We will not pursue it further in this paper.

[1] R. Dijkgraaf and C. Vafa, Nucl. Phys. **B644**, 3 (2002).

[2] R. Dijkgraaf and C. Vafa, Nucl. Phys. **B644**, 21 (2002).

[3] R. Dijkgraaf and C. Vafa, ''A Perturbative Window into Non-Perturbative Physics,'' hep-th/0208048.

[4] M. Bershadsky, S. Cecotti, H. Ooguri, and C. Vafa, Commun. Math. Phys. **165**, 311 (1994).

[5] R. Gopakumar and C. Vafa, Adv. Theor. Math. Phys. **3**, 1415 (1999).

[6] C. Vafa, J. Math. Phys. **42**, 2798 (2001).

[7] F. Cachazo, K. A. Intriligator, and C. Vafa, Nucl. Phys. **B603**, 3 (2001).

[8] L. Chekhov and A. Mironov, Phys. Lett. B **552**, 293 (2003).

[9] N. Dorey, T. J. Hollowood, S. Prem Kumar, and A. Sinkovics, J. High Energy Phys. **11**, 039 (2002).

[10] N. Dorey, T. J. Hollowood, S. P. Kumar, and A. Sinkovics, J. High Energy Phys. **11**, 040 (2002).

[11] M. Aganagic and C. Vafa, ''Perturbative derivation of mirror symmetry,'' hep-th/0209138.

[12] F. Ferrari, Nucl. Phys. **B648**, 161 (2003).

[13] H. Fuji and Y. Ookouchi, J. High Energy Phys. **12**, 067 (2002).

[14] D. Berenstein, Phys. Lett. B **552**, 255 (2003).

[15] G. Bonnet, F. David, and B. Eynard, J. Phys. A **33**, 6739 (2000).

[16] I. Kostov, Nucl. Phys. **B326**, 583 (1989); M. Gaudin and I. Kostov, Phys. Lett. B **220**, 200 (1989).

[17] H. Ooguri and C. Vafa, Nucl. Phys. **B641**, 3 (2002).

[18] E. Witten, hep-th/9207094.

[19] S. J. Gates, Jr., M. T. Grisaru, M. Rocek, and W. Siegel, *Superspace, or One Thousand and One Lessons in Supersymmetry*, Frontiers in Physics Vol. 58 (Benjamin Cummings, Reading, MA, 1983), hep-th/0108200.

[20] P. Ginsparg, ''Matrix models of 2d gravity,'' Trieste Lectures, hep-th/9112013.

[21] A. D'adda, Class. Quantum Grav. **9**, L21 (1992); **9**, L77 (1992).

[22] S. Kharchev, A. Marshakov, A. Mironov, A. Morozov, and S. Pakuliak, Nucl. Phys. **B404**, 717 (1993); I. K. Kostov, Phys. Lett. B **297**, 74 (1992).

[23] I. K. Kostov, Nucl. Phys. **B376**, 539 (1992).

[24] C. Bachas, J. de Boer, R. Dijkgraaf, and H. Ooguri, J. High Energy Phys. **06**, 027 (2002).

[25] R. G. Leigh and M. J. Strassler, Nucl. Phys. **B447**, 95 (1995).

[26] F. Cachazo and C. Vafa, ''$N=1$ and $N=2$ Geometry from Fluxes,'' hep-th/0206017.

[27] E. Brezin, C. Itzykson, G. Parisi, and J. B. Zuber, Commun. Math. Phys. **59**, 35 (1978).

[28] N. Seiberg and E. Witten, Nucl. Phys. **B426**, 19 (1994); **B430**, 485(E) (1994).

[29] V. A. Kazakov, I. K. Kostov, and N. Nekrasov, Nucl. Phys. **B557**, 413 (1999).

[30] N. Dorey, J. High Energy Phys. **07**, 021 (1999); N. Dorey and S. P. Kumar, *ibid.* **02**, 006 (2000).

[31] F. David, Phys. Lett. B **302**, 403 (1993).

[32] P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals for Engineers and Physicists* (Springer-Verlag, Berlin, 1954).

[33] A. Klemm, W. Lerche, and S. Theisen, Int. J. Mod. Phys. A **11**, 1929 (1996).

[34] N. Dorey, V. V. Khoze, and M. P. Mattis, Phys. Lett. B **390**, 205 (1997).

[35] D. Ghoshal and C. Vafa, Nucl. Phys. **B453**, 121 (1995).

[36] V. A. Kazakov and A. A. Migdal, Nucl. Phys. **B311**, 171 (1988).

[37] O. Aharony, N. Dorey, and S. P. Kumar, J. High Energy Phys. **06**, 026 (2000).