# Patterns at finite $N$

Ji Hoon Lee

2109.02545 D. Gaiotto, JHL
2204.09286 JHL
work in progress

A recurring motif in our study of holography is the idea that the Feynman diagrams of a large $N$ gauge theory reorganize into a genus expansion of *some* string theory.

It is natural to ask to what extent the notion of gauge-string duality persists at finite $N$.

In this talk, I will describe an intriguing pattern that appears in the spectra of finite $N$ gauge theories at zero coupling. Namely, the exact finite $N$ spectrum organizes as a systematic set of $e^{-N}$ corrections to the large $N$ spectrum. These corrections turn out to have a transparent holographic interpretation in the bulk – they are contributions from giant graviton branes and open string excitations thereof.

After describing the pattern, I will explain why the pattern is there and what we can learn about the bulk space of states with this knowledge.

The attitude with which I'll proceed is to learn what the basic kinematic rules are in the bulk string theory by squeezing the dual gauge theory. The giant graviton expansion is an open-closed SFT formula which clearly suggests some rules in the bulk that we did not know before. The goal will be to understand better how the gauge-string dictionary could be extended to finite $N$.

Let us begin with simple observations about various finite $N$ spectra. We start with observations in the superconformal index of $U(N)$ $\mathcal{N}=4$ SYM for context, but I will later explain why these patterns persist in any theory of gauged $N \times N$ matrices.

The half-BPS index of $U(N)$ $\mathcal{N}=4$ SYM

$$Z_N = \prod_{n=1}^{N} \frac{1}{1-x^n},$$

counts states of the form $\prod_i (\operatorname{Tr} X^{m_i})|0\rangle$. A notable feature is that the product has a cutoff at $N$ due to trace relations that render the operator $\operatorname{Tr} X^{N+1}$ no longer independent of the lower traces.

Suppose we want to isolate the finite $N$ effects on the spectrum. A way to do this is to divide $Z_N$ by $Z_\infty$. Doing so for large values of $N$, one observes that

$$\begin{aligned}\frac{Z_N}{Z_\infty} &= 1 - x^{N+1}(1 + x + x^2 + x^3 + x^4 + x^5 + \cdots) \\ &\quad + x^{2N+3}(1 + x + 2x^2 + 2x^3 + 3x^4 + 3x^5 + \cdots) \\ &\quad - x^{3N+6}(1 + x + 2x^2 + 3x^3 + 4x^4 + 5x^5 + \cdots) + \cdots \\ &= \sum_{k=0}^{\infty} x^{kN} \frac{(-1)^k x^{k(k+1)/2}}{\prod_{m=1}^{k}(1 - x^m)}\end{aligned}$$

which is the so-called "giant graviton expansion" for the simplest case of half-BPS. The finite $N$ corrections enter in the spectrum roughly at every interval of $N$.

For the $\frac{1}{16}$-BPS index of $\mathcal{N}=4$ SYM, one finds a similar structure:

$$
\begin{aligned}
\frac{Z_N}{Z_\infty} =& 1 + w^{2N}\frac{1}{2!}\Big[(-N^2-5N-6)w^2 + 2(N^2+3N+2)w^3 - 3(N(N+1))w^4 \\
&\qquad + 6(N^2-N-2)w^5 + \cdots\Big] \\
&+ w^{4N}\frac{1}{5!}\Big[2(N^5+25N^4+245N^3+1175N^2+2754N+2520)w^8 \\
&\qquad - 2(8N^5+165N^4+1370N^3+5775N^2+12482N+11160)w^9 \\
&\qquad + 12(6N^5+105N^4+760N^3+2915N^2+6054N+5490)w^{10} \\
&\qquad + \cdots\Big] \\
&+ w^{6N}\Big[\frac{1}{8!}(-3N^8-180N^7-4662N^6-68040N^5-611667N^4 \\
&\qquad - 3466260N^3-12084468N^2-23681520N-19958400)w^{18} \\
&\qquad + \cdots\Big] + \cdots
\end{aligned}
$$

The $N$-dependence of coefficients appears because I've written a specialization of fugacities that is typically used to study $\frac{1}{16}$-BPS black holes. A fully-refined version of this formula has $N$-independent coefficients.

The RHS of the expansions

$$Z_N^{\frac{1}{2}\mathrm{BPS}} = Z_\infty^{\frac{1}{2}\mathrm{BPS}} \sum_{k=0}^{\infty} x^{kN} \hat{Z}_k(x)$$

$$Z_N^{\frac{1}{16}\mathrm{BPS}} = Z_\infty^{\frac{1}{16}\mathrm{BPS}} \sum_{k_1,k_2,k_3=0}^{\infty} x^{k_1 N} y^{k_2 N} z^{k_3 N} \hat{Z}_{(k_1,k_2,k_3)}(x, y, z, p, q)$$

have independent definitions in terms of $\prod_i U(k_i)$ gauge integrals owing to their bulk interpretations. The terms in the sum correspond to $k$ coincident branes wrapping $S^3 \subset S^5$ and their open string modes. The prefactor $Z_\infty$ counts closed strings, or the tower of BPS Kaluza-Klein modes.

The integrand for the 1-loop open string modes $\hat{Z}_{k_i}$ can be found either by (1) counting the modifications of determinant operators in gauge theory [Gaiotto & JHL '21], or by (2) counting modes on probe D3 giant gravitons in the supergravity limit [Imamura '21]. These results agree in holographic examples.

The $\prod_i U(k_i)$ gauge theories living on bulk branes are defined with an unusual integration cycle, essentially because the usual $k$-torus contour fails to pick up residues from poles that correspond to their physical excitations. One needs to define the $\prod_i U(k_i)$ theory with a contour prescription taking residues from physical poles [JHL '22].

The derivation of the integrand and the integration cycle for general brane indices $\hat{Z}_{k_i}$ are interesting on their own right, but unfortunately I won't have time to explain them today.

In the half-BPS sector, the $k$-th term is the half-BPS partition function of the $U(k)$ gauge theory living on $k$ coincident giants. Recall that half-BPS chiral primaries are a purely-bosonic sector where the index equals the partition function.

$$\hat{Z}_k = \frac{1}{k!}\frac{1}{(1-x^{-1})^k}\oint\prod_{a=1}^{k}\frac{d\sigma_a}{2\pi i\sigma_a}\prod_{\substack{a,b=1\\a\neq b}}^{k}\frac{(\sigma_a-\sigma_b)}{(\sigma_a-x^{-1}\sigma_b)}$$

Why does the inverse $x^{-1}$ appear? The inverse $x^{-1}$ appears because $x$ is the fugacity for an R-charge and half-BPS excitations of maximal giants (i.e. those that wrap a maximal $S^3$ in $S^5$) can only take away R-charges.

$$\hat{Z}_k = \frac{1}{\prod_{m=1}^{k}(1-x^{-m})} = (-1)^k \frac{x^{k(k+1)/2}}{\prod_{m=1}^{k}(1-x^m)}$$

While a naive interpretation of the brane index $\hat{Z}_k$ at $|x| > 1$ would say that the open-string spectrum is unbounded below, we should remember that excitations on top of a given closed-string vacuum need to be understood in the domain $|x| < 1$ where the Kaluza-Klein spectrum $Z_\infty$ is defined.

Indeed, one important purpose of the modified $\prod_i U(k_i)$ integration cycle was to make the integral well-defined in the region $|x| < 1$. Therefore, the space of normalizable half-BPS open-string states has a spectrum that is given by the rightmost expression.

There is an important consequence to understanding the brane partition function $\hat{Z}(x)$ in the region $|x| < 1$. The analytic continuation from the region outside the unit disk to inside the unit disk gives an extra overall sign $(-1)^k$ for odd numbers of giants.

Let us observe the manner in which stacks of giant gravitons sum to give the half-BPS spectrum:

$$\prod_{n=1}^{N} \frac{1}{1-x^n} = \frac{1}{\prod_{n=1}^{\infty}(1-x^n)} \sum_{k=0}^{\infty} (-1)^k x^{kN} \frac{x^{k(k+1)/2}}{\prod_{m=1}^{k}(1-x^m)}$$

Via direct power expansion, one can see that each giant graviton term on the RHS, supplemented with Kaluza-Klein modes, overcounts the gauge theory partition function.

This overcounting in the bulk side of the formula is saved only by extra signs $(-1)^k$, which results from the analytic continuation of the R-symmetry fugacity $x$.

Odd stacks of giant gravitons effectively behave as fermions in the bulk half-BPS Hilbert space, even though the half-BPS Hilbert space of the dual CFT only involves gauge-invariant combinations of bosonic scalars $X$.

I want to emphasize that the sign does not come from the usual $(-1)^F$, because we are not computing an index in the half-BPS sector. The sign appears in the formula purely due to the analytic continuation.

If we wish to give a Hilbert space interpretation to the bulk side of the formula, the most natural postulate for the presence of $(-1)^k$ would be that there is a different $\mathbb{Z}_2$-grading in the bulk space of states that is not present in the CFT Hilbert space.

The extra signs from the analytic continuation appear in all known examples, on top of the usual $(-1)^F$ for the index.

Having stated the problem, i.e. of understanding the extra grading in the bulk, it is natural to ask if there is a way to understand the grading from the (more tractable) gauge theory side even if it does not appear at the level of its Hilbert space.

In the rest of the talk I will explain, starting from the data of a free $U(N)$ gauge theory, how to construct the Hilbert space associated with the bulk side of the giant graviton formula.

A surprise is that the reason why the formula exists is not associated with the theory having supersymmetry-protected quantities or the string dual admitting a weakly-curved gravity regime. Rather, the only relevant information is that we are computing invariants of a set of $N \times N$ matrices.

The notion of giant gravitons was conceived in order to provide a bulk explanation of the presence of finite $N$ trace relations, in terms of the polarization of large charge gravitons into branes. The goal of the remaining part of the talk will be to explain why giant gravitons implement trace relations as transparently as possible.

We need to develop basic algebraic tools along the way, but it will eventually be clear that these tools have physical significance.

Let us consider a free $U(N)$ gauge theory on $S^1 \times M$ for $M$ a compact manifold. If its Hilbert space splits into different topological sectors, we restrict attention to one sector at a time.

Consider the ring of all gauge invariant polynomials of fields and their derivatives

$$R = \mathbb{C}[\operatorname{Tr} X, \operatorname{Tr} XY, \operatorname{Tr} \psi F..., \operatorname{Tr} \partial ... \partial ... X, \cdots].$$

At $N = \infty$, the space of states is the free module $M_\infty = R|0\rangle$ built on the vacuum $|0\rangle$. $M_\infty$ is graded with respect to the quantum numbers of its generators.

To find the space of states a free gauge theory at finite $N$, one takes the quotient

$$M_N = M_\infty / \tilde{M}_N$$

of $M_\infty$ by the submodule $\tilde{M}_N = I_N|0\rangle$, where the ideal over $R$

$$I_N = \langle \text{ trace relations at } N \rangle_R$$

is generated by the set of all trace relations present at finite $N$.

For a single matrix $X$ at $N = 2$, these relations look like

$$\mathrm{Tr}(X)^3 - 3\,\mathrm{Tr}(X)\,\mathrm{Tr}(X^2) + 2\,\mathrm{Tr}(X^3) = 0$$
$$\mathrm{Tr}(X)^4 - 6\,\mathrm{Tr}(X)^2\,\mathrm{Tr}(X^2) + 3\,\mathrm{Tr}(X^2)^2 + 8\,\mathrm{Tr}(X)\,\mathrm{Tr}(X^3) - 6\,\mathrm{Tr}(X^4) = 0$$

and so on.

What makes the study of the quotient module $M_N$ difficult is that $M_N$ is in general not freely generated, i.e. it is not a Fock space. This means that we can only describe $M_N$ at best in terms of a set of generators that necessarily have relations among themselves.

Having these relations is problematic if we are interested in describing the space of states as a Fock space built from single trace oscillators.

We can address this problem by computing the free resolution of the quotient module $M_N$. A free resolution of $M_N$ is an exact sequence of free $R$-modules $V_k$ along with differential $\hat{Q}$

$$\cdots \to V_3 \xrightarrow{\hat{Q}} V_2 \xrightarrow{\hat{Q}} V_1 \xrightarrow{\hat{Q}} M_\infty \to M_N \to 0$$

with $\hat{Q}V_1 = \tilde{M}_N$, i.e. the image of $\hat{Q}_1$ is the submodule generated by the trace relations.

Each component of this complex is graded with respect to the quantum numbers induced from the generators of the ring $R$. The complex itself has a grading which I'll call $B$, under which $\hat{Q}$ has degree $-1$.

$$\cdots \to V_3 \xrightarrow{\hat{Q}} V_2 \xrightarrow{\hat{Q}} V_1 \xrightarrow{\hat{Q}} M_\infty \to M_N \to 0$$

Intuitively, a free resolution "resolves" a highly-constrained finite $N$ Hilbert space $M_N$ by replacing it with an (infinite) sequence of Fock spaces $V_k$ that better and better approximate $M_N$.

Since $M_\infty$ maps to $M_N$, the zeroth approximation to the finite $N$ module $M_N$ is the large $N$ module $V_0 \equiv M_\infty$.

The subsequent free modules $V_k$ can be understood as the space of relations among the generators of $V_{k-1}$. For example, $V_1$ is the space of relations among the large $N$ Fock space generators of $M_\infty$, i.e. the generators of $V_1$ map to trace relations under $\hat{Q}$. Then $V_2$ is the space of relations among trace relations, and so on.

Therefore, $V_k$ is the space of $k$-th order relations among the large $N$ Fock space generators.

One may be concerned that the free resolution of $M_N$ is not unique, and indeed in general there are inequivalent ways to resolve $M_N$.

However, there is a notion of uniqueness for the free resolution of $M_N$, because there are only a finite number of generators of $R$ for given values of the charges.

So at any particular values of the charges, we get a finite resolution and there is a unique "minimal" free resolution defined to be a resolution for which each free module has a minimal number of generators.

Any resolution that I mention in subsequent discussion will be the minimal one.

The generators of $V_k$ and the matrix-valued differentials $\hat{Q}$ are explicitly computable, given the set of trace relations at some $N$. Of course, in practice we must work with a truncation at some "energy" $L$ for the generators of $R$ and of trace relations.

The result of the computation stabilizes as $L$ is increased.

Given a graded $R$-module $M$, one can compute its Hilbert series $HS_M$. This is the free partition function of a theory with Hilbert space $M$, up to a possible overall shift in the energy.

A key property of a free resolution is that the Hilbert series of the module being resolved is equal to the alternating sum of the Hilbert series of free modules in the resolution. In our case,

$$HS_{M_N}(t_1, \cdots, t_n) = \sum_{k=0}^{\infty} (-1)^k HS_{V_k}(t_1, \cdots, t_n)$$

where $n$ is the number of charges with which we refine the series.

I would now like to argue the following:

$$V_k = \mathcal{H}_\infty^{\rm closed} \otimes \mathcal{H}_k^{\rm open}$$

the free module $V_k$ of $k$-th order relations over the ring $R$ resolving the free $U(N)$ gauge theory space of states $M_N$ should be interpreted in the string dual as the $\alpha' \to \infty$ limit of the space of normalizable open string states in a sector with $k$ branes of the giant graviton-type, sharing a closed-string background with spectrum given in terms of $\mathcal{H}_\infty^{\rm closed}$.

If there are multiple types of brane systems with total number $k$, $\mathcal{H}_k^{\rm open}$ decomposes into a direct sum of spaces associated with each brane system labelled by $k$.

A consequence of this is an equivalence at the level of the partition functions at each $k$:

$$(-1)^k HS_{V_k}(t_1, \cdots, t_n) = Z_\infty \sum_{\sum_a k_a = k} t_1^{k_1 N} \cdots t_n^{k_n N} \hat{Z}_{(k_1, \cdots, k_n)}(t_1, \cdots, t_n)$$

In words, the Hilbert series $HS_{V_k}$, counting the dimensions of the free module $V_k$ of $k$-th order relations in a free $U(N)$ gauge theory, should be interpreted as the $\alpha' \to \infty$ limit of an open-closed SFT partition function in a sector with $k$ branes of the giant graviton-type.

This is the $k$-th term that appears in the giant graviton expansion.

It makes sense that closed string spectrum $Z_\infty$ factors out of the expression, because we are computing the Hilbert series of a free module $V_k$ over $R$. $Z_\infty$ is simply the Hilbert series of $R$.

If $V_k$ were not free, the closed string spectrum would not factor out. So the open-string sector is the non-trivial part.

If we are computing the index with $(-1)^F$ in e.g. $\mathcal{N} = 4$ SYM dual to IIB in $\mathrm{AdS}_5 \times S^5$, rather than the free partition function, the RHS at fixed $k$ can be computed in the bulk in the $\alpha' \to 0$ limit from the DBI+CS action, as was done in various works of Imamura et al.

I checked the equivalence of the Hilbert series of $V_k$ with the $k$-th term in the giant graviton expansion (computed via the gauge theory prescription in Gaiotto-JHL) in a limited number of simple examples, e.g. $U(N)$-gauged boson $X$, fermion $\psi$, two bosons $X$ and $Y$, etc. at low orders.

There is computational difficulty in enumerating and resolving the finite $N$ trace relations over the large $N$ ring $R$.

More computations of free resolutions of $M_N$ at higher orders are required to understand properties such as the precise manner in which $HS_{V_k}$ depends on $N$.

For the case of the $U(N)$-gauged boson $X$, it is possible to show analytically that

$$(-1)^k HS_{V_k}(x) = \frac{1}{\prod_{n=1}^{\infty}(1-x^n)} \frac{(-1)^k x^{kN} x^{k(k+1)/2}}{\prod_{m=1}^{k}(1-x^m)}.$$

which I'll do in a moment.

Before I do so, I want to explain an earlier comment regarding an extra grading that appears in the giant graviton expansion.

The extra signs come from the $\mathbb{Z}$-grading $B$ in the free resolution, which is a grading that is independent of the other quantum numbers and that does not appear at the level of the gauge theory Hilbert space $M_N$.

That is, the bulk side of the giant graviton expansion is the refined Euler character of the free resolution I described:

$$HS_{M_N}(t_1, \cdots, t_n) = \sum_{k=0}^{\infty} (-1)^k HS_{V_k}(t_1, \cdots, t_n)$$

Having a free resolution consisting of modules $V_k$ allows us to say more, namely that $\hat{Q}$ has an interpretation as a differential relating string states in different open string vacua labelled by $k$.

The resolution however only gives a classical expression for the nilpotent matrix-valued differential $\hat{Q}$, but it would be very interesting if nonperturbative overlaps $O(e^{-N})$ due to instanton effects can be understood.

Now let us show analytically that

$$(-1)^k HS_{V_k}(x) = \frac{1}{\prod_{n=1}^{\infty}(1-x^n)} \frac{(-1)^k x^{kN} x^{k(k+1)/2}}{\prod_{m=1}^{k}(1-x^m)}.$$

for the $U(N)$-gauged boson $X$.

The ring of all gauge-invariant polynomials is

$$R = \mathbb{C}[\operatorname{Tr} X, \operatorname{Tr} X^2, \operatorname{Tr} X^3, \cdots].$$

At any value $N$, the space of trace relations is generated by

$$P_{N+1}, P_{N+2}, P_{N+3}, \cdots$$

where

$$\begin{aligned} P_1 &= \operatorname{Tr} X \\ P_2 &= \frac{1}{2}\left(\operatorname{Tr} X^2 - (\operatorname{Tr} X)^2\right) \\ P_3 &= \frac{1}{6}\left(\operatorname{Tr} X^3 - 3(\operatorname{Tr} X^2)(\operatorname{Tr} X) + 2(\operatorname{Tr} X)^3\right) \end{aligned}$$

and so on. The $P_n$ can be written down using the generating function

$$\exp\left(-\sum_{n=1}^{\infty} \frac{u^n}{n} \operatorname{Tr} X^n\right).$$

A special property of the $U(N)$-gauged boson $X$ is that the ring $R$ of all gauge-invariant polynomials can equivalently be expressed in terms of the trace relation generators $P_n$ themselves:

$$R = \mathbb{C}[P_1, P_2, P_3, \cdots],$$

a fact which leads to simplifications. The infinite $N$ module is $M_\infty = R|0\rangle$.

We want to find the free resolution of the quotient module

$$M_N = M_\infty / I_N |0\rangle,$$

where $I_N$ is the ideal

$$I_N = \langle P_{N+1}, P_{N+2}, \cdots \rangle$$

generated by the trace relations.

A gauge theory perspective would be to simply set to zero the generators $P_{N+1}, P_{N+2}, \cdots$ of $I_N$. Then the partition function

$$Z_N = \prod_{n=1}^{N} \frac{1}{1 - x^n}$$

is computed from the Fock space of oscillators $P_1, P_2, \cdots, P_N$.

The bulk perspective would be to consider the free resolution of $M_N$:

$$\cdots \to V_3 \xrightarrow{\hat{Q}} V_2 \xrightarrow{\hat{Q}} V_1 \xrightarrow{\hat{Q}} M_\infty \to M_N \to 0$$

The zeroth component of this resolution is $M_\infty$, which is the half-BPS partition function in a closed-string background with no half-BPS giant gravitons:

$$HS_{M_\infty}(x) = \prod_{n=1}^{\infty} \frac{1}{1 - x^n}$$

The first component $V_1$ is composed of generators that map to trace relations under $\hat{Q}$

$$\hat{Q}\psi_{-N-a} = P_{N+a}$$

where $a = 1, 2, 3, \cdots$. The differential $\hat{Q}$ preserves the R-charge grading, so $\psi_{-N-a}$ has energy/R-charge $N + a$.

By working out the way in which the matrix-valued differential $\hat{Q}$ acts on generators of $V_2$, one finds

$$\hat{Q}(\psi_{-N-a}\psi_{-N-b}) = P_{-N-a}\psi_{-N-b} - P_{-N-b}\psi_{-N-a}$$

where $(\psi_{-N-a}\psi_{-N-b})$ denotes a two component generator of $V_2$, with $a < b$ and $a, b = 1, 2, 3, \cdots$.

The pattern continues for $V_k$:

$$\hat{Q}(\psi_{-N-a_{\mathbf{1}}} \cdots \psi_{-N-a_k}) = \sum_{i=1}^{k} (-1)^{k+1} P_{-N-a_i}(\psi_{-N-a_{\mathbf{1}}} \cdots \widehat{\psi_{-N-a_i}} \cdots \psi_{-N-a_k})$$

where $a_1 < \cdots < a_k$ and $a_1, \cdots, a_k = 1, 2, 3, \cdots$.

It is clear that, for $V_k$, we should count the number of fermionic oscillators starting at energies $N+1$:

$$\sum_{a_1 < \cdots < a_k} x^{N+a_1} x^{N+a_2} \cdots x^{N+a_k} = \frac{x^{kN} x^{k(k+1)/2}}{\prod_{m=1}^{k}(1-x^m)}$$

In the above, we've counted the generators of $V_k$ but we should remember that the generators have coefficients in the large $N$ ring $R$:

$$HS_{V_k}(x) = \frac{1}{\prod_{n=1}^{\infty}(1-x^n)} \frac{x^{kN} x^{k(k+1)/2}}{\prod_{m=1}^{k}(1-x^m)}.$$

This was the desired expression for the half-BPS partition function around $k$ giant gravitons.

Computing the refined Euler character of the free resolution we find

$$\prod_{n=1}^{N} \frac{1}{1-x^n} = \frac{1}{\prod_{n=1}^{\infty}(1-x^n)} \sum_{k=0}^{\infty} (-1)^k x^{kN} \frac{x^{k(k+1)/2}}{\prod_{m=1}^{k}(1-x^m)}$$

the giant graviton expansion for a single $U(N)$ bosonic matrix.

I hope now it is clear why the effective theory giant gravitons is fermionic. The generators of the space of $k$-th order relations were a $k$-collection of modes of a single fermionic field with $B = 1$

$$\psi(z) = \sum_{n \in \mathbb{Z}} \frac{\psi_n}{z^{n+N+1}}.$$

of dimension $N + 1$ and $B = 1$.

In this talk, I've only been able to explain basic kinematical aspects of what of a dual string theory can be squeezed from the spectrum of a free $U(N)$ gauge theory.

Much remains to be learned and computed at finite $N$...

# Thank you