

A homological approach to the Gaussian unitary ensemble

Owen Gwilliam, Alastair Hamilton, and Mahmoud Zeinalian

Abstract. We study the Gaussian unitary ensemble (GUE) using noncommutative geometry and the Batalin–Vilkovisky (BV) formalism, and we show how this homological approach provides canonical relations between correlation functions in the GUE. As applications of this method, we obtain new ways to prove generalizations of Wigner’s semicircle law, to compute all the large N statistical correlations for multi-trace functions as random variables in the GUE, and to determine the leading (and subleading) order behavior of the correlation functions with respect to the rank N . Along the way, and illuminating the connection with prior work, we develop an explicit dictionary between this homological approach and the well-known combinatorics of ribbon graphs that lead to counting problems for the corresponding surfaces.

1. Introduction

In order to model complicated quantum mechanical systems, Wigner hit upon the clever idea of studying random matrices [18]. The most basic version involves placing a Gaussian measure

$$\mu_N := \frac{1}{Z_N} e^{-\frac{1}{2}\mathrm{Tr}(X^2)} dX$$

on the space of $N \times N$ Hermitian matrices \mathfrak{h}_N , where Z_N normalizes the measure to have mass one. In the large N limit, he found that the expected values of trace functions such as

$$\int_{\mathfrak{h}_N} \mathrm{Tr}(X^{36}) \mu_N$$

exhibited remarkable behavior: to leading order in N , the expected value reduces to a simple one-dimensional integral. For the function above,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \int_{\mathfrak{h}_N} \mathrm{Tr} \left(\left(\frac{X}{\sqrt{N}} \right)^{36} \right) \mu_N = \frac{1}{2\pi} \int_{-2}^2 x^{36} \sqrt{4-x^2} dx;$$

an example of his semicircle law [19]. Out of Wigner’s work grew an area of mathematics where physicists, probabilists, operator algebraists, and others meet.

Mathematics Subject Classification 2020: 81T75 (primary); 15B52, 60B20, 60F99, 81T18, 81T32, 81T40, 81T70 (secondary).

Keywords: Gaussian unitary ensemble (GUE), large N asymptotics, Batalin–Vilkovisky (BV) formalism, Loday–Quillen–Tsygan theorem, ribbon graphs.

A central goal of this paper is to show how homological algebra offers a novel perspective that clarifies how and why these amazing results in probability theory are related to noncommutative geometry – specifically the noncommutative symplectic geometry of Kontsevich [11] – and hence to the topology of moduli spaces of Riemann surfaces. From the work of Harer [9], Mumford, Penner [15] and Thurston, it is known that the cohomology of the moduli space of Riemann surfaces may be described in terms of an orbifold complex generated by ribbon graphs. By the results of [11] and [7], the homological aspects of this complex may be recast in terms of the Batalin–Vilkovisky (BV) formalism and noncommutative geometry, and it is this perspective that we use in this paper.

The BV formalism is a homological way to encode the idea of the path integral in general; this framework has had great success in the study of gauge and string field theories. As the Gaussian unitary ensemble (GUE) can be interpreted as 0-dimensional quantum field theory, it is natural (at least to BV enthusiasts) to try it in this context. One appealing aspect of the BV approach is that it shifts the emphasis onto the algebra of observables rather than on constructing a measure, and this shift means that the large N limit is being explored in an algebraic setting. Thanks to the Loday–Quillen–Tsygan (LQT) Theorem [13, 17], this limit is well understood for the classical field theory.

In the paper [5] we explained how to quantize the LQT theorem, and we found a very simple differential graded algebra that encodes the GUE. (Our treatment here is essentially self-contained; we write so that the reader does not need to understand the LQT theorem in general, and instead we provide the explicit formulas relevant to the GUE.) In this paper, we focus on exploiting that presentation to find recurrence relations among multi-trace expected values and in particular to find expressions for the (sub-)leading order behavior that go beyond the original results of Wigner and his semicircle law (see Section 5 and 6 respectively). In brief, the noncommutative symplectic geometry encodes algebraically the combinatorics of ribbon graphs, allowing us to make efficient computations.

A quick consequence of this large N asymptotic analysis, performed using homological algebra and an understanding of the properties of the Catalan numbers, is a generalization of the semicircle law to multi-trace functions. In particular, the expected value of a multi-trace function “decouples” into a product of the expected values for each factor; see Section 6.4 for these results. As another application, we compute the large N statistical correlation coefficients for a certain infinite family of random variables in the GUE defined by multi-trace functions; see Section 6.3 for these results. These results are known to experts (see [1, 2, 4] as entry points among the relevant sources), and so the novelty is in the method and the new relationships it forges between disparate mathematical domains. (For the experts, note that the homological relations of the BV formalism encode the Dyson–Schwinger equations, so that our approach can be understood from that viewpoint. See, for instance, [6].)

Perhaps the most conceptually useful aspect of our method is that it directly connects ribbon graphs, and hence the topology of the moduli space of Riemann surfaces, to the integration-by-parts relations between multi-trace functions arising through the BV formalism. This connection arises because the LQT theorem relates cyclic cohomology –

loosely speaking, part of a closed string field theory – to Lie algebra cohomology – loosely speaking, part of a gauge theory; making this assertion precise is a central part of [5]. There is a deep reservoir of results about how closed string field theory relates to the topology of Riemann surfaces, and here we use but a handful of those insights.

Remark 1.1. Shortly after this article appeared on the arXiv we were contacted by Kontsevich, who informed us that he had just the day before delivered a talk on matrix models [12] that was closely related to ideas explored in this paper but introduced in [5]. In particular, in the talk he discussed a conjecture on “decoupling” that is a version of our Theorem 6.18 in the much less straightforward case of a multi-matrix model depending on a nonquadratic potential. We subsequently benefited from an interesting email exchange with Kontsevich, and we think there is much to explore in these directions.

As a guide for the reader, note that Section 2 is a lightning review of the BV formalism and the quantum LQT theorem of [5]. Section 3 then reviews how the GUE appears as an example in this framework. In Section 4, we begin on new material, reviewing ribbon graphs and explaining how they arise in the setting of the GUE and the quantum LQT theorem. Finally, we get to the applications: Section 5 proves that the multi-trace correlation functions are polynomial functions of the rank N with nonnegative integer coefficients, determines their degree and establishes a recurrence relation for them, while Section 6 computes their leading order behavior with respect to the rank N .

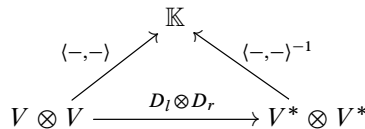
1.1. Notation and conventions

Throughout the paper our convention will be to work with differential graded symplectic vector spaces. We also follow the convention of working with cohomologically graded objects; hence the suspension ΣV of a graded vector space V is defined by $\Sigma V^i := V^{i+1}$. We will assume that our symplectic vector spaces V carry a symplectic form $\langle -, - \rangle$ of odd degree. The differential d on V is required to be compatible with the inner product in the sense that

$$\langle dx, y \rangle + (-1)^x \langle x, dy \rangle = 0. \tag{1.1}$$

We emphasize that, for the sake of brevity, we will refer to these spaces simply as *symplectic vector spaces*; with the understanding that they carry a differential (possibly zero) and that the symplectic form has odd degree.

We will define the inverse form $\langle -, - \rangle^{-1}$ on the dual space V^* by the commutative diagram



where $D_l(y) := \langle y, - \rangle$ and $D_r(y) := \langle -, y \rangle$. Note that while the form $\langle -, - \rangle$ is skew-symmetric, the Koszul sign rule implies that the inverse form $\langle -, - \rangle^{-1}$ is symmetric. More

generally, we will use the same formula for the inverse of *any* nondegenerate bilinear form $\langle -, - \rangle$.

We denote the symmetric group by \mathbb{S}_n . We follow the convention that coinvariants are indicated by a subscript. The graded symmetric algebra on a graded vector space V will be denoted by $S(V)$. Throughout the paper we work over a ground field \mathbb{K} of characteristic zero, usually $\mathbb{K} = \mathbb{C}$. The cardinality of a finite set X will be denoted by $|X|$.

2. Recollections on the quantum LQT theorem

In this section, we will recall the basic framework of the Batalin–Vilkovisky formalism [16], including its formulation in noncommutative geometry coming from the work of Kontsevich [11]. We then recall from [5] how the quantum LQT maps intertwine the commutative and noncommutative aspects of this framework.

2.1. The commutative geometry of the Batalin–Vilkovisky formalism

We begin by describing the classical setup for the Batalin–Vilkovisky formalism and its commutative geometry.

Definition 2.1. Given a symplectic graded vector space V we define

$$P[V] := S(V^*) = \bigoplus_{k=0}^{\infty} [(V^*)^{\otimes k}]_{\mathbb{S}_k}.$$

From the inverse form $\langle -, - \rangle^{-1}$ on V^* we define a Poisson bracket $\{ -, - \}$ on $P[V]$ of odd degree by extending the inverse form on V^* to $P[V]$ using the Leibniz rule;

$$\{a, bc\} = \{a, b\}c + (-1)^{(a+1)b}b\{a, c\}. \tag{2.1}$$

This structure is sometimes referred to as a (differential graded) *shifted Poisson algebra*. There is a natural quantization of this structure, in the Batalin–Vilkovisky sense, obtained by turning on a differential.

Definition 2.2. The *BV-Laplacian* Δ on $S(V^*)$ is the unique differential operator satisfying

$$\Delta(ab) = (\Delta a)b + (-1)^a a(\Delta b) + \{a, b\} \tag{2.2}$$

for all $a, b \in S(V^*)$ and such that Δ vanishes on all $v \in V^*$. These conditions ensure that $\Delta^2 = 0$. From (1.1) it follows that $(d + \Delta)$ is still a differential and more generally that

$$BV[V] := (S(V^*), (d + \Delta), \cdot, \{ -, - \})$$

is a *Batalin–Vilkovisky algebra*, where we have used \cdot to denote the commutative multiplication on the underlying graded-commutative algebra.

2.2. Noncommutative geometry in the Batalin–Vilkovisky formalism

We now recall the formulation of the Batalin–Vilkovisky formalism in the framework of noncommutative geometry. This begins with a construction of Kontsevich [11], with further input from the work of Movshev [14].

Definition 2.3. Given a symplectic graded vector space V , we define

$$H[V] := \bigoplus_{k=0}^{\infty} [(V^*)^{\otimes k}]_{\mathbb{Z}/k\mathbb{Z}}.$$

We define a Lie bracket of odd degree on $H[V]$ by the formula

$$\begin{aligned} & \{(a_1 \cdots a_m), (b_1 \cdots b_n)\} \\ & := \sum_{i=1}^m \sum_{j=1}^n \pm \langle a_i, b_j \rangle^{-1} (a_{i+1} \cdots a_m a_1 \cdots a_{i-1} b_{j+1} \cdots b_n b_1 \cdots b_{j-1}), \end{aligned}$$

where $a_1, \dots, a_m, b_1, \dots, b_n \in V^*$ and the sign is determined canonically by the Koszul sign rule. Note that when $m = 1 = n$,

$$\{(a), (b)\} = \langle a, b \rangle^{-1},$$

which is a constant.

A Lie cobracket

$$\nabla : H[V] \rightarrow (H[V] \otimes H[V])_{S_2}$$

of odd degree may also be defined by the formula

$$\nabla(a_1 \cdots a_n) := \sum_{1 \leq i < j \leq n} \pm \langle a_i, a_j \rangle^{-1} (a_{i+1} \cdots a_{j-1}) \otimes (a_{j+1} \cdots a_n a_1 \cdots a_{i-1}).$$

Again, the sign is determined canonically by the Koszul sign rule. These structures turn $H[V]$ into a Lie bialgebra.

Consider the subspace of $H[V]$ consisting of positive cyclic powers,

$$H_+[V] := \bigoplus_{k=1}^{\infty} [(V^*)^{\otimes k}]_{\mathbb{Z}/k\mathbb{Z}},$$

and denote the generator of $(V^*)^{\otimes 0}$ inside $H[V]$ by ν , which has degree zero. Then we may write

$$H[V] = \mathbb{K}\nu \oplus H_+[V]. \tag{2.3}$$

It is easily checked using the above definitions that the canonical quotient map

$$\sigma : H[V] \rightarrow P[V]$$

(which sends ν to 1) is a map of Lie algebras. The commutative algebra $S(\mathbb{H}[V])$ has the canonical structure of a shifted Poisson algebra in which the bracket on $\mathbb{H}[V]$ is extended to $S(\mathbb{H}[V])$ using the Leibniz rule (2.1). The above map then extends to a map

$$\sigma : S(\mathbb{H}[V]) \rightarrow P[V] \tag{2.4}$$

of shifted Poisson algebras in a unique way.

Note that using the decomposition (2.3) we may write

$$\begin{aligned} S(\mathbb{H}[V]) &= S(\mathbb{K}\nu \oplus \mathbb{H}_+[V]) \\ &= S(\mathbb{K}\nu) \otimes S(\mathbb{H}_+[V]) = \mathbb{K}[\nu] \otimes S(\mathbb{H}_+[V]), \end{aligned}$$

where we have identified the symmetric algebra on $\mathbb{K}\nu$ with polynomials in ν .

We may now define, following [5], the noncommutative counterparts of the Batalin–Vilkovisky algebra described by Definition 2.2.

Definition 2.4. Given a symplectic graded vector space V , set

$$\text{BV}_\nu^{\text{nc}}[V] := S(\mathbb{H}[V]) = \mathbb{K}[\nu] \otimes S(\mathbb{H}_+[V]).$$

We extend the shifted Poisson algebra structure on $S(\mathbb{H}[V])$ described above to the structure of a Batalin–Vilkovisky algebra by defining a BV-Laplacian using the formula

$$\Delta_\nu := \nabla + \delta,$$

where δ denotes the Chevalley–Eilenberg differential on $S(\mathbb{H}[V])$ determined by the Lie bracket on $\mathbb{H}[V]$ and ∇ denotes the cobracket on $\mathbb{H}[V]$, which is extended to $S(\mathbb{H}[V])$ using the Leibniz rule. This provides us with the structure of a Batalin–Vilkovisky algebra,

$$\text{BV}_\nu^{\text{nc}}[V] = (S(\mathbb{H}[V]), (d + \Delta_\nu), \cdot, \{-, -\}).$$

It is a consequence of equation (2.2) that the map (2.4) yields a map of Batalin–Vilkovisky algebras – that is to say, it commutes with the BV-Laplacians. The details are not difficult, and are spelled out in [5, Proposition 4.3]. We will denote this map of BV algebras by

$$\sigma_\nu : \text{BV}_\nu^{\text{nc}}[V] \rightarrow \text{BV}[V]. \tag{2.5}$$

We may refine the picture described above by including another deformation parameter, bringing us closer to the construction described in [7].

Definition 2.5. Given a symplectic graded vector space V , set

$$\text{BV}_{\gamma,\nu}^{\text{nc}}[V] := \mathbb{K}[\gamma] \otimes S(\mathbb{H}[V]) = \mathbb{K}[\gamma, \nu] \otimes S(\mathbb{H}_+[V]).$$

This also has the structure of a shifted Poisson algebra, where we extend those structures that we defined on $S(\mathbb{H}[V])$ above linearly with respect to the parameter γ (which has degree zero). We equip it with the BV-Laplacian

$$\Delta_{\gamma,\nu} := \nabla + \gamma \cdot \delta.$$

This defines a Batalin–Vilkovisky algebra

$$\text{BV}_{\gamma,\nu}^{\text{nc}}[V] = (\mathbb{K}[\gamma] \otimes S(\mathbb{H}[V]), (d + \Delta_{\gamma,\nu}), \cdot, \gamma\{-, -\}).$$

Note that in this case we must multiply the Poisson bracket $\{-, -\}$ by the parameter γ to retain the structure of a Batalin–Vilkovisky algebra.

Remark 2.6. Later, we will use the parameters γ and ν to keep track of the genus and number of boundary components associated to a ribbon graph, cf. Section 4.3.

It is clear that by sending the deformation parameter γ to 1 we get a map of Batalin–Vilkovisky algebras

$$\text{BV}_{\gamma,\nu}^{\text{nc}}[V] \xrightarrow{\gamma=1} \text{BV}_{\nu}^{\text{nc}}[V].$$

Combining this map with the map (2.5) yields a map of Batalin–Vilkovisky algebras which we denote by

$$\sigma_{\gamma,\nu} : \text{BV}_{\gamma,\nu}^{\text{nc}}[V] \rightarrow \text{BV}[V]. \tag{2.6}$$

2.3. The Morita map

In this section, we will introduce the map that appears in the formulation of Morita invariance in Hochschild cohomology. It is defined by taking the trace of a product of matrices. It is the compatibility of this map with the framework of the Batalin–Vilkovisky formalism that will allow us to make contact with quantities in random matrix theory.

Let V be a symplectic vector space and consider the space of N -by- N matrices with entries in V , which we denote

$$\text{M}_N(V) = V \otimes \text{M}_N(\mathbb{K}).$$

It is also a (differential graded) symplectic vector space whose symplectic form of odd degree is defined by

$$\langle x \otimes A, y \otimes B \rangle := \langle x, y \rangle \text{Tr}(AB).$$

Consider the multilinear maps

$$t_k : \text{M}_N(\mathbb{K})^{\otimes k} \rightarrow \mathbb{K} \tag{2.7}$$

where

$$t_k(A_1, \dots, A_k) := \text{Tr}(A_1 \cdots A_k).$$

These are used to define the Morita map as follows.

Definition 2.7. Let V be a symplectic vector space and consider the map

$$(V^{\otimes k})^* \rightarrow (V^{\otimes k})^* \otimes (\text{M}_N(\mathbb{K})^{\otimes k})^* = (\text{M}_N(V)^{\otimes k})^*$$

sending ξ_k to $\xi_k \otimes t_k$. As the multilinear maps (2.7) are cyclically symmetric, the above determines a well-defined map

$$\mathcal{M} : \mathbb{H}[V] \rightarrow \mathbb{H}[\text{M}_N(V)]. \tag{2.8}$$

Note that since the identity matrix has trace $\text{Tr}(I_N) = N$ it follows that

$$\mathcal{M}(v) = Nv.$$

The following result, [5, Lemma 4.6], shows that this map respects the structures on $H[V]$ introduced by Definition 2.3.

Proposition 2.8. *The map (2.8) is a morphism of Lie bialgebras.*

This result has the following immediate consequence. The map (2.8) extends to a map

$$\mathcal{M} : S(H[V]) \rightarrow S(H[M_N(V)])$$

of shifted Poisson algebras. In fact, it also provides a map of BV algebras, which we denote by

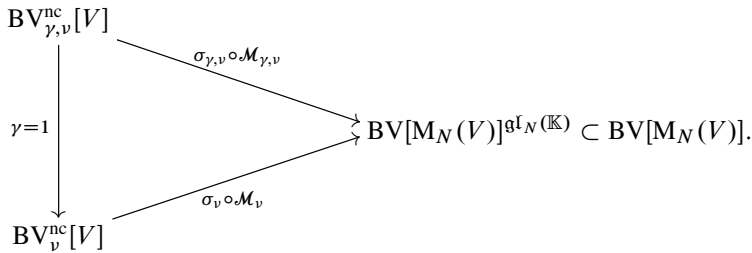
$$\mathcal{M}_v : \text{BV}_v^{\text{nc}}[V] \rightarrow \text{BV}_v^{\text{nc}}[M_N(V)].$$

Denote the γ -linear extension of \mathcal{M}_v to $\text{BV}_{\gamma,v}^{\text{nc}}[V] = \mathbb{K}[\gamma] \otimes S(H[V])$ by

$$\mathcal{M}_{\gamma,v} : \text{BV}_{\gamma,v}^{\text{nc}}[V] \rightarrow \text{BV}_{\gamma,v}^{\text{nc}}[M_N(V)].$$

It is also a map of BV algebras.

We can combine these maps with the morphisms (2.5) and (2.6) of BV algebras. Note that since the multilinear maps (2.7) are $\mathfrak{gl}_N(\mathbb{K})$ -invariant, the combined maps will land in the $\mathfrak{gl}_N(\mathbb{K})$ -invariants $\text{BV}[M_N(V)]^{\mathfrak{gl}_N(\mathbb{K})}$. All told, we have the following commutative diagram of Batalin–Vilkovisky algebras



Remark 2.9. The diagonal maps in this diagram are what we will sometimes refer to as the “quantized LQT maps”. This terminology is carried over from [5], where they correspond to the maps in the Loday–Quillen–Tsygan theorem in the special case of a vanishing algebra structure on the graded vector space V .

3. Matrix integrals in the Gaussian unitary ensemble

In this section, we recall from [5] how, as an application of the constructions just reviewed, a two-dimensional symplectic vector space encodes Hermitian matrix integrals and hence

describes certain correlation functions in the Gaussian unitary ensemble. We are then able to articulate a key construction of this paper, which will be the subject of the next section.

Let \mathfrak{h}_N denote the real subspace of $\mathfrak{gl}_N(\mathbb{C})$ consisting of all Hermitian matrices. We are interested in studying the k -trace correlation functions

$$I_{k_1, k_2, \dots, k_n}^N := \frac{\int_{\mathfrak{h}_N} \text{Tr}(X^{k_1}) \text{Tr}(X^{k_2}) \dots \text{Tr}(X^{k_n}) e^{-\frac{1}{2} \text{Tr}(X^2)} dX}{\int_{\mathfrak{h}_N} e^{-\frac{1}{2} \text{Tr}(X^2)} dX}. \tag{3.1}$$

Note that while the integrals in both the numerator and the denominator of (3.1) depend upon a choice of linear identification of \mathfrak{h}_N with \mathbb{R}^{N^2} , the ratio does not.

The symplectic vector space leading to these correlation functions is remarkably simple.

Definition 3.1. Let \mathcal{A} be the two-dimensional complex vector space with generators a and b of degrees zero and one respectively. Define the differential graded symplectic structure on \mathcal{A} by

$$da = b, \quad \langle b, a \rangle = 1 = -\langle a, b \rangle.$$

Note that \mathcal{A} is acyclic, so that many complexes we construct from it have simple cohomology. What will be of interest later is *how* various cocycles are related.

Let a^* and b^* denote the dual basis of \mathcal{A}^* and set

$$x := a^* \quad \text{and} \quad \xi := -b^*.$$

Then x has degree zero, ξ has degree minus-one, and the definitions above become

$$d\xi = -x, \quad \langle x, \xi \rangle = \langle x, \xi \rangle^{-1} = 1 = \langle \xi, x \rangle^{-1} = \langle \xi, x \rangle.$$

Observe that $H[\mathcal{A}]$, the space of cyclic tensor powers, is isomorphic to $\mathbb{C}[x]$ in degree zero, as cyclic words in one generator correspond to symmetric words in one generator.

Writing $\mathcal{A} = \mathbb{C} \oplus \Sigma^{-1}\mathbb{C}$ where the even generator a sits in the left-hand summand and the odd generator b sits in the right-hand summand, we have the decomposition

$$M_N(\mathcal{A}) = \mathcal{A} \otimes M_N(\mathbb{C}) = M_N(\mathbb{C}) \oplus \Sigma^{-1}M_N(\mathbb{C}).$$

The Hermitian matrices \mathfrak{h}_N sit inside the left-hand summand as a subspace and hence we may restrict any polynomial superfunction f in $\text{BV}[M_N(\mathcal{A})]$ to a complex-valued polynomial function on Hermitian matrices \mathfrak{h}_N . In this way we define the *expectation value*

$$\langle - \rangle : \text{BV}[M_N(\mathcal{A})] \rightarrow \mathbb{C} \tag{3.2}$$

by

$$\langle f \rangle := \frac{\int_{\mathfrak{h}_N} f(X) e^{-\frac{1}{2} \text{Tr}(X^2)} dX}{\int_{\mathfrak{h}_N} e^{-\frac{1}{2} \text{Tr}(X^2)} dX}.$$

The following is Proposition 5.2 of [5]. It follows from some standard arguments in the Batalin–Vilkovisky formalism, which encode the integration by parts relations for the Gaussian unitary ensemble.

Proposition 3.2. *The expectation value map (3.2) is a quasi-isomorphism of complexes whose one-sided inverse is the inclusion of \mathbb{C} inside $\text{BV}[M_N(\mathcal{A})]$ as the constant polynomials.*

It is helpful at this point to summarize the present situation by the following (incomplete) commutative diagram:

$$\begin{array}{ccc}
 \text{BV}_{\gamma,v}^{\text{nc}}[\mathcal{A}] & \xleftarrow{\exists?} & \mathbb{C}[\gamma, v] \\
 \downarrow \sigma_{\gamma,v} \circ \mathcal{M}_{\gamma,v} & \searrow \gamma=1 & \swarrow \gamma=1 \\
 & \text{BV}_v^{\text{nc}}[\mathcal{A}] & \xleftarrow{\exists?} \mathbb{C}[v] \\
 & \swarrow \sigma_v \circ \mathcal{M}_v & \searrow v=N \\
 \text{BV}[M_N(\mathcal{A})] & \xleftarrow{\langle - \rangle : f \mapsto \langle f \rangle} & \mathbb{C}
 \end{array} \tag{3.3}$$

The left vertical map is what we referred to earlier in Remark 2.9 as the quantum LQT map; in degree zero (the most important degree for us) this map sends a cyclic word (x^k) to the function $\text{Tr}(X^k)$ on matrices, and it sends a symmetric product of cyclic words $(x^{k_1}) \cdots (x^{k_n})$ to the multi-trace function $\text{Tr}(X^{k_1}) \cdots \text{Tr}(X^{k_n})$. All the horizontal arrows are quasi-isomorphisms; this fact follows from a standard spectral sequence argument that is spelled out in detail in [5, Proposition 5.2 and 5.3].

The goal of the next section is to construct a *noncommutative* analogue

$$\langle - \rangle_{\gamma,v} : \text{BV}_{\gamma,v}^{\text{nc}}[\mathcal{A}] \rightarrow \mathbb{C}[\gamma, v]$$

of the expectation value map (3.2). This map will fill in the dashed arrow atop diagram (3.3) and thus complete the above commutative diagram. That section culminates in Theorem 4.8.

Foreshadowing one of the punchlines of the paper, we claim that this dashed arrow $\langle - \rangle_{\gamma,v}$ will amount to taking a product of cyclic words, interpreting them as the vertices of a ribbon graph, and then returning a polynomial encoding the topology of those ribbon graphs that one can make from these vertices. Thanks to the commutativity of the diagram, we can evaluate that polynomial to compute the expected value of the multi-trace function determined by the product of cyclic words. In particular, if we use $p_{k_1, \dots, k_n}^{\gamma,v} \in \mathbb{C}[\gamma, v]$ to denote the “noncommutative expectation value”

$$\langle (x^{k_1}) \cdots (x^{k_n}) \rangle_{\gamma,v},$$

then we will show that evaluating $p_{k_1, \dots, k_n}^{\gamma,v}$ at $\gamma = 1$ and $v = N$ recovers the multi-trace expectation value

$$I_{k_1, \dots, k_n}^N = \frac{\int_{\mathfrak{h}_N} \text{Tr}(X^{k_1}) \cdots \text{Tr}(X^{k_n}) e^{-\frac{1}{2} \text{Tr}(X^2)} dX}{\int_{\mathfrak{h}_N} e^{-\frac{1}{2} \text{Tr}(X^2)} dX}.$$

This result is the content of Proposition 5.2.

4. The ribbon graph expansion of the expectation value

In this section, we will introduce some instances of *noncommutative* analogues of Feynman diagram expansions. The role of graphs (or Feynman diagrams), ubiquitous and familiar in the “commutative” setting due to Wick’s theorem, is replaced here by ribbon graphs in this noncommutative setting. The topology of the surfaces described by these ribbon graphs, determined by the genus and number of boundary components, is easily read off from the ribbon graph itself and encoded in a polynomial with two corresponding variables. These constructions will be used to complete diagram (3.3), which is the central result of this section, and thus relate the quantum LQT map to the topology of surfaces.

4.1. Ribbon graphs

We begin by recalling some basic material about ribbon graphs.

Definition 4.1. A ribbon graph Γ consists of the following data:

- A finite set $H(\Gamma)$ whose elements are called the *half-edges* of Γ .
- A partition $E(\Gamma)$ of $H(\Gamma)$ into pairs. Elements of $E(\Gamma)$ are called *edges*.
- A partition $V(\Gamma)$ of $H(\Gamma)$, whose elements are called *vertices*.
- A cyclic ordering of each vertex $v \in V(\Gamma)$.

This data may be equivalently described as follows. Note that a cyclic ordering of the half-edges of a vertex is the same thing as a cyclic permutation of all the half-edges at that vertex. Combining these permutations at every vertex gives us a permutation ζ_Γ of all the half-edges of the ribbon graph. The cyclic decomposition of the permutation ζ_Γ recovers the vertices of the ribbon graph Γ along with their cyclic structure. Similarly, we may define a permutation κ_Γ of the half-edges of Γ satisfying

- $\kappa_\Gamma^2 = \text{id}$,
- $\kappa_\Gamma h \neq h$, for all $h \in H(\Gamma)$;

simply by transposing the half-edges of each edge.

The correspondence

$$\Gamma \mapsto (H(\Gamma), \zeta_\Gamma, \kappa_\Gamma)$$

determines an isomorphism between the category of ribbon graphs and the category whose objects are 3-tuples consisting of a set H , a permutation ζ on H and a permutation κ on H satisfying the two conditions just listed above (both categories have an obvious notion of isomorphism). Frequently, it is more convenient to define certain constructions in the latter category.

Given any ribbon graph Γ we may construct a connected oriented surface S_Γ with boundary, as shown in Figure 1. (Our construction will differ slightly from the one that the reader may already be familiar with, e.g., from [10].) Given a ribbon graph Γ with n vertices

$$v_1, \dots, v_n$$

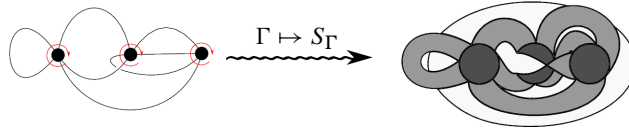


Figure 1. The surface associated to a ribbon graph. We note that the left-most central band is twisted in the picture in order to maintain the orientability of the surface.

of valency k_1, \dots, k_n , we take a sphere S^2 and remove n disks, creating k_i parameterized intervals around the boundary of each disk. We then use the cyclic structure at each vertex v_i to place the half-edges incident to v_i into one-to-one correspondence with the parameterized intervals of the i th disk. Now for every edge e of Γ we take a strip $I \times I$ and glue the ends of this strip to the two parameterized intervals corresponding to the two half-edges that form e . This gives a well-defined (up to homeomorphism) construction of an oriented surface \mathcal{S}_Γ , since we may always rotate and permute the parameterized boundary components by a homeomorphism of the sphere. The Euler characteristic of the surface \mathcal{S}_Γ is

$$\chi(\mathcal{S}_\Gamma) = 2 - |V(\Gamma)| - |E(\Gamma)|, \tag{4.1}$$

as can be checked directly.

Definition 4.2. Given a ribbon graph $\Gamma = (H(\Gamma), \zeta_\Gamma, \kappa_\Gamma)$, consider the permutation $\beta_\Gamma := \zeta_\Gamma \kappa_\Gamma$ and define the *dual ribbon graph* by

$$\Gamma^* := (H(\Gamma), \beta_\Gamma, \kappa_\Gamma).$$

The significance of the dual graph is the following. The boundary components of the surface \mathcal{S}_Γ are in one-to-one correspondence with the vertices of the dual graph Γ^* . Using this and formula (4.1) for the Euler characteristic of \mathcal{S}_Γ , we get a simple expression for the genus. Define $b(\Gamma) := |V(\Gamma^*)|$ to be the number of boundary components of \mathcal{S}_Γ and set

$$g(\Gamma) := \frac{|V(\Gamma)| + |E(\Gamma)| - b(\Gamma)}{2} \tag{4.2}$$

to be the genus of the surface \mathcal{S}_Γ .

We now consider the operation of contracting an edge in a ribbon graph. This is most conveniently described in terms of the dual graph.

Definition 4.3. Let Γ be a ribbon graph and let $e \in E(\Gamma)$ be an edge. To contract the edge e in Γ , we simply remove this edge from the dual graph Γ^* ; that is if Γ/e denotes the graph Γ with the edge e contracted then

$$(\Gamma/e)^* = (H(\Gamma) - e, \beta_{\Gamma/e}, \kappa_{\Gamma/e})$$

where ζ/e denotes the permutation ζ with the entries from e deleted from its cyclic decomposition.

This operation may be more directly described as follows. If the edge $e = \{h, h'\}$ joins two different vertices

$$(hh_1 \dots h_k) \quad \text{and} \quad (h'h'_1 \dots h'_{k'})$$

of Γ then contracting the edge e combines these vertices into a single vertex

$$(h_1 \dots h_k h'_1 \dots h'_{k'}). \tag{4.3}$$

If $e = \{h, h'\}$ is a loop on a vertex

$$v = (hh_1 \dots h_k h'h'_1 \dots h'_{k'})$$

of Γ then contracting e splits v into two separate vertices

$$(h_1 \dots h_k) \quad \text{and} \quad (h'_1 \dots h'_{k'}). \tag{4.4}$$

Since contracting the edge $e = \{h, h'\}$ simply deletes the half-edges from the dual graph Γ^* , this will not change the number of vertices in Γ^* , unless of course some of those vertices are left with no remaining half-edges. The reader may observe that this situation occurs precisely when an empty vertex is produced as a result in either (4.3) or (4.4).

If $k = k' = 0$ in (4.3) then

$$b(\Gamma/e) = b(\Gamma) - 1. \tag{4.5}$$

If either $k = 0$ or $k' = 0$ (but not both) in (4.4) then

$$b(\Gamma/e) = b(\Gamma) - 1, \tag{4.6}$$

and if both $k = k' = 0$ then

$$b(\Gamma/e) = b(\Gamma) - 2. \tag{4.7}$$

The effect on the genus of contracting an edge may be summarized as follows. Contracting an edge e that joins two different vertices, as in (4.3), typically decreases the number of vertices and edges by one each and hence in this case

$$g(\Gamma/e) = g(\Gamma) - 1. \tag{4.8}$$

If the edge e is instead a loop then similar reasoning shows that the genus is not affected

$$g(\Gamma/e) = g(\Gamma). \tag{4.9}$$

Note that a careful analysis shows that (4.8) and (4.9) still hold even if the number $b(\Gamma)$ is affected by the edge contraction as above, as in this case the other terms in (4.2) will compensate appropriately.

4.2. Chord diagrams

We will use chord diagrams in our construction of the maps that fill in diagram (3.3) and later in Section 6. We begin by recalling their definition.

Definition 4.4. Given a set X of even cardinality, a chord diagram on X is a partition of X into pairs. We denote the set of all chord diagrams on X by $\mathcal{C}(X)$. In particular, if X is the set of all integers between 1 and $2l$ then we denote this set by $\mathcal{C}(2l)$.

If V is a graded vector space equipped with a symmetric bilinear form B of even degree, then for every chord diagram

$$c = \{i_1, j_1\}, \{i_2, j_2\}, \dots, \{i_l, j_l\} \in \mathcal{C}(2l),$$

we may define a map $B_c : V^{\otimes 2l} \rightarrow \mathbb{K}$ by

$$B_c(v_1, v_2, \dots, v_{2l}) = \pm B(v_{i_1}, v_{j_1})B(v_{i_2}, v_{j_2}) \cdots B(v_{i_l}, v_{j_l})$$

where the sign is determined canonically by the Koszul sign rule. Note that

$$B_c(\zeta \cdot x) = B_{\zeta^{-1} \cdot c}(x) \tag{4.10}$$

for all $x \in V^{\otimes 2l}$, $c \in \mathcal{C}(2l)$, and $\zeta \in \mathbb{S}_{2l}$.

We may also use chord diagrams to define ribbon graphs as follows.

Definition 4.5. Given a chord diagram

$$c = \{i_1, j_1\}, \{i_2, j_2\}, \dots, \{i_l, j_l\} \in \mathcal{C}(2l)$$

and a list of positive integers k_1, \dots, k_n such that

$$k_1 + k_2 + \dots + k_n = 2l,$$

let $\Gamma_{k_1, \dots, k_n}(c)$ denote the ribbon graph with $2l$ half edges h_1, \dots, h_{2l} . The edges of $\Gamma_{k_1, \dots, k_n}(c)$ are

$$\{h_{i_1}, h_{j_1}\}, \{h_{i_2}, h_{j_2}\}, \dots, \{h_{i_l}, h_{j_l}\}$$

and the vertices are defined by the permutation

$$(h_1 \dots h_{k_1})(h_{k_1+1} \dots h_{k_1+k_2}) \cdots (h_{k_1+\dots+k_{n-1}+1} \dots h_{k_1+\dots+k_{n-1}+k_n}).$$

4.3. Expectation values in noncommutative geometry

We are now ready to proceed with our construction of the maps that complete diagram (3.3). As a gloss – and because our constructions should generalize to any contractible space, though there is no present need to work at such a level of generality – we will provide a ribbon graph version of the computation of Gaussian moments by pairing

off legs of a vertex. Hence, we begin by describing our “propagator”, then describe our Wick formula, and finally verify that this construction fits into diagram (3.3).

On our two-dimensional symplectic vector space \mathcal{A} , we form a degree zero symmetric pairing

$$B : \mathcal{A} \otimes \mathcal{A} \rightarrow \mathbb{C}$$

where

$$B(u, v) := \langle du, v \rangle$$

using the symplectic form $\langle -, - \rangle$ and differential d . This pairing is degenerate, but becomes nondegenerate on the subspace of \mathcal{A} concentrated in degree zero. Let

$$B^{-1} = a \otimes a \in \mathcal{A} \otimes \mathcal{A}$$

denote the inverse of this form restricted to this subspace.

Now, given a list of positive integers k_1, \dots, k_n , we define a map

$$\langle - \rangle_{\gamma, \nu}^{k_1, \dots, k_n} : (\mathcal{A}^*)^{\otimes k_1} \otimes \dots \otimes (\mathcal{A}^*)^{\otimes k_n} \rightarrow \mathbb{C}[\gamma, \nu].$$

The map is zero if the sum of the k_i is odd. If the sum is even and equal to $2l$, then we define

$$\langle w \rangle_{\gamma, \nu}^{k_1, \dots, k_n} := \sum_{c \in \mathbb{C}(2l)} B_c^{-1}(w) \gamma^g(\Gamma_{k_1, \dots, k_n}(c)) \nu^b(\Gamma_{k_1, \dots, k_n}(c)) \tag{4.11}$$

where $w \in (\mathcal{A}^*)^{\otimes k_1} \otimes \dots \otimes (\mathcal{A}^*)^{\otimes k_n} = (\mathcal{A}^*)^{\otimes 2l}$.

This map plays a central role in what follows, so we make a few orienting remarks. Note that it records the topology of those ribbon graphs that are built from chord diagrams using the prescribed data consisting of the list of the k_i s. The input w is a kind of “non-commutative function” on \mathcal{A} (i.e., an element of the tensor algebra of the dual). Those familiar with Feynman diagrams will see the resemblance to the usual Wick formula.

It follows from (4.10) that this map is cyclically symmetric in the variables. A permutation $\zeta \in \mathbb{Z}/k_1\mathbb{Z} \times \dots \times \mathbb{Z}/k_n\mathbb{Z}$ determines an isomorphism

$$\zeta : \Gamma_{k_1, \dots, k_n}(c) \cong \Gamma_{k_1, \dots, k_n}(\zeta \cdot c).$$

Hence $\langle \zeta \cdot w \rangle_{\gamma, \nu}^{k_1, \dots, k_n} = \langle w \rangle_{\gamma, \nu}^{k_1, \dots, k_n}$ and (4.11) extends to give a well-defined map on $H_+[\mathcal{A}]^{\otimes n}$. Similar reasoning shows that, by permuting the vertices of the ribbon graph, this map extends to give a well-defined map on $S(H_+[\mathcal{A}])$. (Note that on $S^0(H_+[\mathcal{A}]) = \mathbb{C}$, this map is just the inclusion of \mathbb{C} into $\mathbb{C}[\gamma, \nu]$.) If we extend this map linearly with respect to the variables γ and ν , we see that (4.11) provides a well-defined map

$$\langle - \rangle_{\gamma, \nu} : \text{BV}_{\gamma, \nu}^{\text{nc}}[\mathcal{A}] \rightarrow \mathbb{C}[\gamma, \nu]. \tag{4.12}$$

We call this map the *noncommutative expectation value*. This map will provide the desired dashed arrow that completes diagram (3.3), but first we must prove some of its basic properties.

Theorem 4.6. *The noncommutative expectation value map (4.12) is a quasi-isomorphism of complexes whose one-sided inverse is the inclusion of $\mathbb{C}[\gamma, \nu]$ into $BV_{\gamma, \nu}^{nc}[\mathcal{A}]$.*

Proof. We must show first that (4.12) is a map of complexes; that is, we must show that it vanishes on the boundaries of $BV_{\gamma, \nu}^{nc}[\mathcal{A}]$. To see that the map is a quasi-isomorphism, first note that the cohomology of $BV_{\gamma, \nu}^{nc}[\mathcal{A}]$ is $\mathbb{C}[\gamma, \nu]$ because there is a filtration of $BV_{\gamma, \nu}^{nc}[\mathcal{A}]$ by powers of \mathcal{A} and the associated spectral sequence collapses immediately as the complex \mathcal{A} is acyclic. Now apply the comparison theorem.

Recall that \mathcal{A} is \mathbb{Z} -graded and concentrated in positive degrees, so that \mathcal{A}^* and hence $BV_{\gamma, \nu}^{nc}[\mathcal{A}]$ are also \mathbb{Z} -graded and concentrated in nonpositive degrees. To show (4.12) is a cochain map, it thus suffices only to consider the boundaries of degree minus-one elements in $BV_{\gamma, \nu}^{nc}[\mathcal{A}]$. Such elements are linear combinations of terms of the form

$$(\xi x^i)(x^{k_1}) \dots (x^{k_r})$$

where $i, r \geq 0$ and $k_1, \dots, k_r \geq 1$.

The boundary of such a typical element is

$$\begin{aligned} (d + \Delta_{\gamma, \nu})[(\xi x^i)(x^{k_1}) \dots (x^{k_r})] &= -(x^{i+1})(x^{k_1}) \dots (x^{k_r}) \\ &\quad + \sum_{s=1}^i (x^{s-1})(x^{i-s})(x^{k_1}) \dots (x^{k_r}) \\ &\quad + \gamma \sum_{t=1}^r k_t (x^{i+k_t-1})(x^{k_1}) \dots \widehat{(x^{k_t})} \dots (x^{k_r}). \end{aligned}$$

Hence we must show that

$$\begin{aligned} \langle (x^{i+1})(x^{k_1}) \dots (x^{k_r}) \rangle_{\gamma, \nu} &= \sum_{s=1}^i \langle (x^{s-1})(x^{i-s})(x^{k_1}) \dots (x^{k_r}) \rangle_{\gamma, \nu} \\ &\quad + \gamma \sum_{t=1}^r k_t \langle (x^{i+k_t-1})(x^{k_1}) \dots \widehat{(x^{k_t})} \dots (x^{k_r}) \rangle_{\gamma, \nu}. \end{aligned} \tag{4.13}$$

Obviously we may assume that $i + k_1 + \dots + k_r = 2l - 1$ for some $l \geq 1$. We begin by calculating the left-hand side of (4.13). As a point of notation, note that any chord diagram $c \in \mathcal{C}(2l)$ defines in an obvious way a permutation in \mathbb{S}_{2l} , also denoted by c , satisfying $c^2 = \text{id}$.

Now

$$\begin{aligned} \langle (x^{i+1})(x^{k_1}) \dots (x^{k_r}) \rangle_{\gamma, \nu} &= \sum_{c \in \mathcal{C}(2l)} \gamma^{g(\Gamma_{i+1, k_1, \dots, k_r}(c))} \nu^{b(\Gamma_{i+1, k_1, \dots, k_r}(c))} \\ &= \sum_{s=2}^{2l} \sum_{\{c \in \mathcal{C}(2l) : c(1) = s\}} \gamma^{g(\Gamma_{i+1, k_1, \dots, k_r}(c))} \nu^{b(\Gamma_{i+1, k_1, \dots, k_r}(c))} \end{aligned}$$

$$\begin{aligned}
 &= \sum_{s=1}^i \sum_{\{c \in \mathcal{C}(2l): c(1)=s+1\}} \gamma^g(\Gamma_{i+1, k_1, \dots, k_r}(c)) \nu^b(\Gamma_{i+1, k_1, \dots, k_r}(c)) \\
 &+ \sum_{t=1}^r \sum_{u=1}^{k_t} \sum_{\substack{\{c \in \mathcal{C}(2l): \\ c(1)=i+1+k_1+\dots+k_{t-1}+u\}}} \gamma^g(\Gamma_{i+1, k_1, \dots, k_r}(c)) \nu^b(\Gamma_{i+1, k_1, \dots, k_r}(c)).
 \end{aligned}$$

Hence to prove (4.13), we will show that

$$\begin{aligned}
 &\langle (x^{s-1})(x^{i-s})(x^{k_1}) \dots (x^{k_r}) \rangle_{\gamma, \nu} \\
 &= \sum_{\{c \in \mathcal{C}(2l): c(1)=s+1\}} \gamma^g(\Gamma_{i+1, k_1, \dots, k_r}(c)) \nu^b(\Gamma_{i+1, k_1, \dots, k_r}(c))
 \end{aligned} \tag{4.14}$$

for $s = 1, \dots, i$ and that

$$\begin{aligned}
 &\gamma \langle (x^{i+k_t-1})(x^{k_1}) \dots \widehat{(x^{k_t})} \dots (x^{k_r}) \rangle_{\gamma, \nu} \\
 &= \sum_{\substack{\{c \in \mathcal{C}(2l): \\ c(1)=i+1+k_1+\dots+k_{t-1}+s\}}} \gamma^g(\Gamma_{i+1, k_1, \dots, k_r}(c)) \nu^b(\Gamma_{i+1, k_1, \dots, k_r}(c))
 \end{aligned} \tag{4.15}$$

for $t = 1, \dots, r$ and $s = 1, \dots, k_t$.

We prove (4.14) first. Note that when $i = 0$ there is nothing to prove, so we begin by considering the case $i = 1 = s$. In this case, the left-hand side of (4.14) is $\nu^2 \langle (x^{k_1}) \dots (x^{k_r}) \rangle_{\gamma, \nu}$ and the right-hand side may be written as

$$\sum_{c \in \mathcal{C}(\{3, \dots, 2l\})} \gamma^g(\Gamma_{2, k_1, \dots, k_r}(c \cup \{1, 2\})) \nu^b(\Gamma_{2, k_1, \dots, k_r}(c \cup \{1, 2\})).$$

Contracting the edge $e = \{h_1, h_2\}$ in the above and applying (4.7) and (4.9), we see that this may be written as

$$\nu^2 \sum_{c \in \mathcal{C}(2l-2)} \gamma^g(\Gamma_{k_1, \dots, k_r}(c)) \nu^b(\Gamma_{k_1, \dots, k_r}(c)) = \nu^2 \langle (x^{k_1}) \dots (x^{k_r}) \rangle_{\gamma, \nu},$$

as desired.

Next we consider the case when $i \geq 2$ and $s = 1$ or $s = i$. For simplicity, assume $s = 1$. In this case, the left-hand side of (4.14) is $\nu \langle (x^{i-1})(x^{k_1}) \dots (x^{k_r}) \rangle_{\gamma, \nu}$ and the right-hand side may be written as

$$\sum_{c \in \mathcal{C}(\{3, \dots, 2l\})} \gamma^g(\Gamma_{i+1, k_1, \dots, k_r}(c \cup \{1, 2\})) \nu^b(\Gamma_{i+1, k_1, \dots, k_r}(c \cup \{1, 2\})).$$

Contracting the edge $e = \{h_1, h_2\}$ and applying (4.6) and (4.9), this becomes

$$\nu \sum_{c \in \mathcal{C}(2l-2)} \gamma^g(\Gamma_{i-1, k_1, \dots, k_r}(c)) \nu^b(\Gamma_{i-1, k_1, \dots, k_r}(c)) = \nu \langle (x^{i-1})(x^{k_1}) \dots (x^{k_r}) \rangle_{\gamma, \nu}.$$

Finally, consider the case when $1 < s < i$. In this case, we may write the right-hand side of (4.14) as

$$\sum_{c \in \mathcal{C}(\{1, \dots, 2l\} - \{1, s+1\})} \gamma^g(\Gamma_{i+1, k_1, \dots, k_r}(c \cup \{1, s+1\})) \nu^b(\Gamma_{i+1, k_1, \dots, k_r}(c \cup \{1, s+1\})).$$

Contracting the edge $e = \{h_1, h_{s+1}\}$, this becomes

$$\begin{aligned} & \sum_{c \in \mathcal{C}(2l-2)} \gamma^g(\Gamma_{s-1, i-s, k_1, \dots, k_r}(c)) \nu^b(\Gamma_{s-1, i-s, k_1, \dots, k_r}(c)) \\ &= \langle (x^{s-1})(x^{i-s})(x^{k_1}) \cdots (x^{k_r}) \rangle_{\gamma, \nu}. \end{aligned}$$

It remains to prove (4.15). We first consider the case $i = 0$ and $k_t = 1 = s$. In this case, the left-hand side of (4.15) is $\gamma \nu \langle (x^{k_1}) \cdots (x^{k_t}) \cdots (x^{k_r}) \rangle_{\gamma, \nu}$. Proceeding as above, but instead making use of (4.5) and (4.8), we may write the right-hand side of (4.15) as

$$\gamma \nu \sum_{c \in \mathcal{C}(2l-2)} \gamma^g(\Gamma_{k_1, \dots, \widehat{k_t}, \dots, k_r}(c)) \nu^b(\Gamma_{k_1, \dots, \widehat{k_t}, \dots, k_r}(c)),$$

as required.

Finally, we consider the case $i + k_t \geq 2$. Repeating the same arguments as above, but making use of (4.8), we may write the right-hand side of (4.15) as

$$\begin{aligned} & \gamma \sum_{c \in \mathcal{C}(2l-2)} \gamma^g(\Gamma_{i+k_t-1, k_1, \dots, \widehat{k_t}, \dots, k_r}(c)) \nu^b(\Gamma_{i+k_t-1, k_1, \dots, \widehat{k_t}, \dots, k_r}(c)) \\ &= \gamma \langle (x^{i+k_t-1})(x^{k_1}) \cdots (\widehat{x^{k_t}}) \cdots (x^{k_r}) \rangle_{\gamma, \nu}, \end{aligned}$$

as desired. ■

Since $\langle - \rangle_{\gamma, \nu}$ is γ -linear, it determines a unique map

$$\langle - \rangle_{\nu} : \text{BV}_{\nu}^{\text{nc}}[\mathcal{A}] \rightarrow \mathbb{C}[\nu] \tag{4.16}$$

that makes the upper quadrilateral of diagram (3.3) commute by setting $\gamma = 1$. This map may be described explicitly as the $\mathbb{C}[\nu]$ -linear map that is the identity on $S^0(\text{H}_+[\mathcal{A}]) = \mathbb{C}$ and such that

$$\langle w \rangle_{\nu}^{k_1, \dots, k_n} = \sum_{c \in \mathcal{C}(2l)} B_c^{-1}(w) \nu^{b(\Gamma_{k_1, \dots, k_n}(c))} \tag{4.17}$$

where $w \in (\mathcal{A}^*)^{\otimes k_1} \otimes \cdots \otimes (\mathcal{A}^*)^{\otimes k_n} = (\mathcal{A}^*)^{\otimes 2l}$. This observation leads us to the following corollary of Theorem 4.6.

Corollary 4.7. *The map (4.16) is a quasi-isomorphism of complexes whose one-sided inverse is the inclusion of $\mathbb{C}[\nu]$ into $\text{BV}_{\nu}^{\text{nc}}[\mathcal{A}]$.*

Proof. It follows tautologically from Theorem 4.6 that (4.16) is a map of complexes and the same spectral sequence argument used there applies to show that it is a quasi-isomorphism. ■

Finally, it remains to show that these maps fit into diagram (3.3).

Theorem 4.8. *Placing the noncommutative expectation value maps (4.12) and (4.16) in diagram (3.3) produces a commutative diagram in which all the horizontal arrows are quasi-isomorphisms.*

Proof. All that remains is to show that the resulting diagram is commutative. For this, it is sufficient to prove that the outside rectangle commutes. Since all the maps in this rectangle have degree zero it follows that in nonzero degrees the horizontal maps send everything to zero, and so in this case the diagram commutes trivially. Therefore the only interesting case is in degree zero because all four corners are nontrivial in this degree.

Note that we have already established that each row of this diagram describes a retract of the left-hand complex onto its cohomology: for $BV_{\gamma,v}^{nc}[\mathcal{A}]$ the arrows retract it onto $\mathbb{C}[\gamma, v]$. Now, any degree zero element f in $BV_{\gamma,v}^{nc}[\mathcal{A}]$ is a cocycle, since $BV_{\gamma,v}^{nc}[\mathcal{A}]$ is nonpositively graded and the differential has degree one. Using the retraction, we know that there is some polynomial $p_f \in \mathbb{C}[\gamma, v]$ cohomologous to f . Hence $f - p_f$ is a boundary in $BV_{\gamma,v}^{nc}[\mathcal{A}]$, and so it is sent to zero along either route around the diagram, as all the maps in diagram (3.3) are chain maps. As a consequence, when we follow the path around the northeast periphery of the diagram, both f and p_f will go to the same number in \mathbb{C} . The same will also be true when we take the other route. Now note that p_f is sent to the same number in \mathbb{C} along either of these two routes; here we have used the fact that the horizontal arrows are retractions. It now follows that the same is true of f . ■

5. Correlation functions in the Gaussian unitary ensemble

In this section, we apply our preceding results to describe the correlation functions (3.1) of the Gaussian unitary ensemble.

5.1. The correlation functions and a recursive formula

Definition 5.1. Consider the monomial

$$(x^{k_1})(x^{k_2}) \cdots (x^{k_n}) \in [H[\mathcal{A}]^{\otimes n}]_{\mathbb{S}_n} \tag{5.1}$$

given by a product of cyclic words on one generator. Define the polynomial $p_{k_1, \dots, k_n}^{\gamma, v} \in \mathbb{C}[\gamma, v]$ by

$$p_{k_1, \dots, k_n}^{\gamma, v} := \langle (x^{k_1}) \cdots (x^{k_n}) \rangle_{\gamma, v},$$

i.e., as the noncommutative expectation value of this monomial. Likewise, define the polynomial $p_{k_1, \dots, k_n}^v \in \mathbb{C}[v]$ by

$$p_{k_1, \dots, k_n}^v := \langle (x^{k_1}) \cdots (x^{k_n}) \rangle_v.$$

By Theorem 4.6 the polynomial $p_{k_1, \dots, k_n}^{\gamma, \nu}$ is the unique representative in $\mathbb{C}[\gamma, \nu]$ of the cohomology class (5.1) lying in $BV_{\gamma, \nu}^{nc}[\mathcal{A}]$. Likewise, by Corollary 4.7, $p_{k_1, \dots, k_n}^{\nu}$ is the unique representative in $\mathbb{C}[\nu]$ of the cohomology class (5.1) lying in $BV_{\nu}^{nc}[\mathcal{A}]$. Since diagram (3.3) commutes, the polynomial $p_{k_1, \dots, k_n}^{\nu}$ is obtained from $p_{k_1, \dots, k_n}^{\gamma, \nu}$ by setting $\gamma = 1$.

In what follows we will chiefly work with the polynomials $p_{k_1, \dots, k_n}^{\nu}$, since it will become apparent that the parameter γ has no role to play in analyzing the correlation functions of the Gaussian unitary ensemble; however, it makes sense to include the polynomials $p_{k_1, \dots, k_n}^{\gamma, \nu}$ in our discussion since they may be studied and computed using the same techniques that will be explained here – tracking the genus can be useful in closely related settings.

Note that the polynomials $p_{k_1, \dots, k_n}^{\gamma, \nu}$ and $p_{k_1, \dots, k_n}^{\nu}$ are well defined for nonnegative values of the indices k_r with

$$p_{0, k_1, \dots, k_n}^{\gamma, \nu} = \nu p_{k_1, \dots, k_n}^{\gamma, \nu} \quad \text{and} \quad p_{0, k_1, \dots, k_n}^{\nu} = \nu p_{k_1, \dots, k_n}^{\nu}.$$

This is because x^0 is precisely ν .

From formulas (4.11) and (4.17) for $\langle - \rangle_{\gamma, \nu}$ and $\langle - \rangle_{\nu}$, we know that the coefficients of the polynomials $p_{k_1, \dots, k_n}^{\gamma, \nu}$ and $p_{k_1, \dots, k_n}^{\nu}$ may be described by counting chord diagrams. Given positive integers k_1, \dots, k_n such that $k_1 + \dots + k_n = 2l$, set

$$\begin{aligned} \eta_{k_1, \dots, k_n}^{g, b} &:= \left| \{c \in \mathcal{C}(2l) : g(\Gamma_{k_1, \dots, k_n}(c)) = g \text{ and } b(\Gamma_{k_1, \dots, k_n}(c)) = b\} \right|, \\ \eta_{k_1, \dots, k_n}^{g, *} &:= \left| \{c \in \mathcal{C}(2l) : g(\Gamma_{k_1, \dots, k_n}(c)) = g\} \right|, \\ \eta_{k_1, \dots, k_n}^{*, b} &:= \left| \{c \in \mathcal{C}(2l) : b(\Gamma_{k_1, \dots, k_n}(c)) = b\} \right|. \end{aligned}$$

Then we may write

$$\begin{aligned} p_{k_1, \dots, k_n}^{\gamma, \nu} &= \sum_{g \geq 0, b \geq 1} \eta_{k_1, \dots, k_n}^{g, b} \gamma^g \nu^b, \\ p_{k_1, \dots, k_n}^{\nu} &= \sum_{b \geq 1} \eta_{k_1, \dots, k_n}^{*, b} \nu^b. \end{aligned} \tag{5.2}$$

Recall from (4.2) that for k_1, \dots, k_n as above

$$2g(\Gamma_{k_1, \dots, k_n}(c)) + b(\Gamma_{k_1, \dots, k_n}(c)) = n + l.$$

From this, it follows that when $2g + b = n + l$, we have

$$\eta_{k_1, \dots, k_n}^{g, b} = \eta_{k_1, \dots, k_n}^{g, *} = \eta_{k_1, \dots, k_n}^{*, b}$$

and zero otherwise. Hence we may write

$$\begin{aligned} p_{k_1, \dots, k_n}^{\gamma, \nu} &= \sum_{\substack{g \geq 0, b \geq 1: \\ 2g + b = n + l}} \eta_{k_1, \dots, k_n}^{g, b} \gamma^g \nu^b \\ &= \sum_{\substack{g \geq 0, b \geq 1: \\ 2g + b = n + l}} \eta_{k_1, \dots, k_n}^{g, *} \gamma^g \nu^b = \sum_{\substack{g \geq 0, b \geq 1: \\ 2g + b = n + l}} \eta_{k_1, \dots, k_n}^{*, b} \gamma^g \nu^b \end{aligned}$$

$$= \sum_{g=0}^{\lfloor \frac{n+l-1}{2} \rfloor} \eta_{k_1, \dots, k_n}^{g,*} \gamma^g \nu^{n+l-2g} = \sum_{g=0}^{\lfloor \frac{n+l-1}{2} \rfloor} \eta_{k_1, \dots, k_n}^{*,n+l-2g} \gamma^g \nu^{n+l-2g}$$

and

$$p_{k_1, \dots, k_n}^{\nu} = \sum_{g=0}^{\lfloor \frac{n+l-1}{2} \rfloor} \eta_{k_1, \dots, k_n}^{g,*} \nu^{n+l-2g} = \sum_{g=0}^{\lfloor \frac{n+l-1}{2} \rfloor} \eta_{k_1, \dots, k_n}^{*,n+l-2g} \nu^{n+l-2g}.$$

From this expression it is clear that the polynomials $p_{k_1, \dots, k_n}^{\nu}$ are even (respectively, odd) precisely when $(n + l)$ is even (respectively, odd).

The essential significance of the polynomials $p_{k_1, \dots, k_n}^{\nu}$ – of principal importance for this paper – is that they recover all the multi-trace correlation functions (3.1), thanks to [5, Proposition 5.4], which we now recall.

Proposition 5.2. *For every positive integer N ,*

$$p_{k_1, \dots, k_n}^{\nu}(N) = I_{k_1, \dots, k_n}^N = \frac{\int_{\mathfrak{h}_N} \text{Tr}(X^{k_1}) \dots \text{Tr}(X^{k_n}) e^{-\frac{1}{2} \text{Tr}(X^2)} dX}{\int_{\mathfrak{h}_N} e^{-\frac{1}{2} \text{Tr}(X^2)} dX}.$$

Proof. Although a proof is provided in [5], we mention for the sake of completeness that the result follows from the fact that diagram (3.3) commutes and that both σ_{ν} and \mathcal{M}_{ν} are maps of Batalin–Vilkovisky algebras, provided we recall that the image of $(x^k) \in H[\mathcal{A}]$ under $\sigma_{\nu} \circ \mathcal{M}_{\nu}$ is the polynomial function $\text{Tr}(X^k)$ for $X \in \mathfrak{gl}_N(\mathbb{C})$. ■

Our methods provide us with a recursive formula for computing these polynomials.

Proposition 5.3. *For all $n \geq 1$ and for all $k_1 \geq 1$ and $k_2, \dots, k_n \geq 0$,*

$$p_{k_1, \dots, k_n}^{\gamma, \nu} = \sum_{r=1}^{k_1-1} p_{r-1, k_1-r-1, k_2, \dots, k_n}^{\gamma, \nu} + \gamma \sum_{r=2}^n k_r p_{k_1+k_r-2, k_2, \dots, \widehat{k}_r, \dots, k_n}^{\gamma, \nu}. \tag{5.3}$$

In particular,

$$p_{k_1, \dots, k_n}^{\nu} = \sum_{r=1}^{k_1-1} p_{r-1, k_1-r-1, k_2, \dots, k_n}^{\nu} + \sum_{r=2}^n k_r p_{k_1+k_r-2, k_2, \dots, \widehat{k}_r, \dots, k_n}^{\nu}. \tag{5.4}$$

Proof. Observe that (5.3) follows immediately from equation (4.13) and that (5.4) is obtained from (5.3) simply by setting $\gamma = 1$. ■

Remark 5.4. The preceding equations correspond to the Dyson–Schwinger equations for the Gaussian unitary ensemble, see for example [6, equation (2.5)].

5.2. Examples

Here we will provide some examples of how equation (5.4) can be used to compute the polynomials p_{k_1, \dots, k_n}^v and hence, by Proposition 5.2, the n -trace correlation functions.

Example 5.5. As an example of a very simple computation, we compute

$$\begin{aligned} p_2^v &= p_{0,0}^v = v^2, \\ p_{1,3}^v &= 3p_2^v = 3v^2, \\ p_{2,2}^v &= p_{0,0,2}^v + 2p_2^v = (v^2 + 2)p_2^v = (v^2 + 2)v^2. \end{aligned}$$

In principle, we may compute any polynomial p_{k_1, \dots, k_n}^v through the repeated application of the recursion relation (5.4). As a practical matter, however, the number of terms increases rapidly at each stage. Nonetheless, in some simple cases we obtain some general formulas.

Example 5.6. This example was presented in [5], but since we will require it later we present it here now as well. By the recursion relation (5.3) we have

$$p_{\underbrace{1,1, \dots, 1}_{2n \text{ terms}}}^{\gamma, v} = (2n - 1)\gamma \cdot p_{\underbrace{0, 1, 1, \dots, 1}_{2n-2 \text{ terms}}}^{\gamma, v} = (2n - 1)\gamma v p_{\underbrace{1, 1, \dots, 1}_{2n-2 \text{ terms}}}^{\gamma, v}.$$

Therefore,

$$p_{\underbrace{1, 1, \dots, 1}_{2n \text{ terms}}}^{\gamma, v} = (2n - 1)!! \gamma^n v^n \quad \text{and} \quad p_{\underbrace{1, 1, \dots, 1}_{2n \text{ terms}}}^v = (2n - 1)!! v^n$$

and, by Proposition 5.2, we have

$$\frac{\int_{\mathfrak{h}_N} \text{Tr}(X)^{2n} e^{-\frac{1}{2}\text{Tr}(X^2)} dX}{\int_{\mathfrak{h}_N} e^{-\frac{1}{2}\text{Tr}(X^2)} dX} = (2n - 1)!! N^n.$$

Example 5.7. By the recursion relation (5.3) we have

$$p_{\underbrace{2, 2, \dots, 2}_n}^{\gamma, v} = p_{\underbrace{0, 0, 2, 2, \dots, 2}_{n-1 \text{ terms}}}^{\gamma, v} + 2(n - 1)\gamma p_{\underbrace{2, 2, \dots, 2}_{n-1 \text{ terms}}}^{\gamma, v} = (v^2 + 2(n - 1)\gamma) p_{\underbrace{2, 2, \dots, 2}_{n-1 \text{ terms}}}^{\gamma, v}.$$

Therefore,

$$p_{\underbrace{2, 2, \dots, 2}_n}^{\gamma, v} = \prod_{j=0}^{n-1} (v^2 + 2j\gamma) \quad \text{and} \quad p_{\underbrace{2, 2, \dots, 2}_n}^v = \prod_{j=0}^{n-1} (v^2 + 2j).$$

More generally, one can show that

$$p_{\underbrace{2, 2, \dots, 2, k_1, \dots, k_m}_n}^{\gamma, v} = p_{k_1, \dots, k_m}^{\gamma, v} \prod_{j=0}^{n-1} \left(v^2 + \left(2j + \sum_{i=1}^m k_i \right) \gamma \right).$$

5.3. Order of the correlation functions

Finally, we will demonstrate how equation (5.4) may be used to compute the degree of the polynomial p_{k_1, \dots, k_n}^v .

Theorem 5.8. *Consider the polynomial p_{k_1, \dots, k_n}^v where $k_1, \dots, k_n \geq 0$ have even sum. Let*

$$m := \frac{1}{2} \sum_{i=1}^n k_i,$$

and let

$$q := |\{i : k_i \text{ is even}\}|$$

count the number of indices that are even. Then p_{k_1, \dots, k_n}^v is a polynomial of degree $(m + q)$.

Proof. The proof proceeds by induction on m . We will use the fact that the coefficients of the polynomials (5.2) are nonnegative. Note that if $m = 0$, then all the indices k_i are zero and the statement is trivial. If $m > 0$ then we consider two cases:

- *Case 1:* there is an index which is odd. There must then be at least two indices which are odd, and we may assume that it is both k_1 and k_2 that are odd. From (5.4) we have

$$\begin{aligned} p_{k_1, \dots, k_n}^v &= \sum_{r=1}^{k_1-1} p_{r-1, k_1-r-1, k_2, \dots, k_n}^v \\ &\quad + k_2 p_{k_1+k_2-2, k_3, \dots, k_n}^v + \sum_{r=3}^n k_r p_{k_1+k_r-2, k_2, \dots, \widehat{k}_r, \dots, k_n}^v. \end{aligned}$$

Applying the inductive hypothesis, we have

- The polynomial $p_{r-1, k_1-r-1, k_2, \dots, k_n}^v$ has degree $(m + q)$.
- The polynomial $p_{k_1+k_2-2, k_3, \dots, k_n}^v$ has degree $(m + q)$.
- The polynomial $p_{k_1+k_r-2, k_2, \dots, \widehat{k}_r, \dots, k_n}^v$ has degree either
 - * $(m + q)$ if k_r is odd, or
 - * $(m + q - 2)$ if k_r is even.

It follows that p_{k_1, \dots, k_n}^v has degree $(m + q)$.

- *Case 2:* there is an index which is even and positive. We assume that it is k_1 that is both even and positive. From (5.4) we have

$$\begin{aligned} p_{k_1, \dots, k_n}^v &= v p_{k_1-2, k_2, \dots, k_n}^v + \sum_{r=2}^{k_1-1} p_{r-1, k_1-r-1, k_2, \dots, k_n}^v \\ &\quad + \sum_{r=2}^n k_r p_{k_1+k_r-2, k_2, \dots, \widehat{k}_r, \dots, k_n}^v. \end{aligned}$$

Again, applying the inductive hypothesis, we have

- The polynomial $\nu p_{k_1-2, k_2, \dots, k_n}^\nu$ has degree $(m + q)$.
- The polynomial $p_{r-1, k_1-r-1, k_2, \dots, k_n}^\nu$ has degree either
 - * $(m + q)$ if r is odd, or
 - * $(m + q - 2)$ if r is even.
- The polynomial $p_{k_1+k_r-2, k_2, \dots, \widehat{k_r}, \dots, k_n}^\nu$ has degree $(m + q - 2)$.

Again, it follows that p_{k_1, \dots, k_n}^ν has degree $(m + q)$.

This completes our computation of the degree of p_{k_1, \dots, k_n}^ν . ■

Having now computed the degree of the polynomial p_{k_1, \dots, k_n}^ν , we will look at describing some of the large N asymptotic behavior of the polynomials in the next section.

6. Large N asymptotics of multi-trace correlation functions

In this section, we will explain how to compute the leading, as well as some of the sub-leading coefficients of the polynomials $p_{k_1, \dots, k_n}^\nu \in \mathbb{C}[\nu]$. By Proposition 5.2 this will allow us to describe some of the large N asymptotic behavior of the multi-trace correlation functions (3.1). In particular, we will use these results to compute the large N statistical correlation coefficients between random variables in the Gaussian unitary ensemble and to prove a generalization of Wigner’s semicircle law [19].

6.1. Computation of the leading coefficients

We begin by separating the cases of odd and even indices, using Theorem 5.8 to compute the degrees of the polynomials.

Definition 6.1. Consider the polynomial

$$p_{2i_1, 2i_2, \dots, 2i_k}^\nu, \quad \text{with } i_1, \dots, i_k \geq 0$$

having all even powers. It has degree $n_{i_1, \dots, i_k} := k + \sum_{r=1}^k i_r$. We use C_{i_1, \dots, i_k} to denote the coefficient of $\nu^{n_{i_1, \dots, i_k}}$ for this polynomial.

Now consider the polynomial

$$p_{2j_1+1, 2j_2+1, \dots, 2j_{2l}-1+1, 2j_{2l}+1}^\nu, \quad \text{with } j_1, \dots, j_{2l} \geq 0$$

having all odd powers. It has degree $m_{j_1, \dots, j_{2l}} := l + \sum_{r=1}^{2l} j_r$. We use $A_{j_1, \dots, j_{2l}}$ to denote the coefficient of $\nu^{m_{j_1, \dots, j_{2l}}}$ for this polynomial.

Now note that by Theorem 5.8 the polynomial

$$p_{2i_1, \dots, 2i_k, 2j_1+1, \dots, 2j_{2l}+1}^\nu, \quad \text{with } i_1, \dots, i_k, j_1, \dots, j_{2l} \geq 0 \tag{6.1}$$

has degree $n_{i_1, \dots, i_k} + m_{j_1, \dots, j_{2l}}$. The following result allows us to separate the cases of odd and even indices.

Theorem 6.2. *The coefficient of $v^{n_{i_1, \dots, i_k} + m_{j_1, \dots, j_{2l}}}$ for the polynomial (6.1) is*

$$C_{i_1, \dots, i_k} A_{j_1, \dots, j_{2l}}.$$

Corollary 6.3. *For all $i_1, \dots, i_k, j_1, \dots, j_{2l} \geq 0$,*

$$\begin{aligned} \lim_{N \rightarrow \infty} & \left[\frac{\int_{\mathfrak{h}_N} \text{Tr}(X^{2i_1}) \dots \text{Tr}(X^{2i_k}) \text{Tr}(X^{2j_1+1}) \dots \text{Tr}(X^{2j_{2l}+1}) e^{-\frac{1}{2} \text{Tr}(X^2)} dX}{N^{n_{i_1, \dots, i_k} + m_{j_1, \dots, j_{2l}}} \int_{\mathfrak{h}_N} e^{-\frac{1}{2} \text{Tr}(X^2)} dX} \right] \\ & = C_{i_1, \dots, i_k} A_{j_1, \dots, j_{2l}}. \end{aligned}$$

The proof of Theorem 6.2 will have to wait until we have at least computed the coefficients C_{i_1, \dots, i_k} . Note that Corollary 6.3 is a consequence of Proposition 5.2.

Recall that the n th Catalan number is

$$C_n = \frac{1}{n+1} \binom{2n}{n} = \frac{(2n)!}{n!(n+1)!}.$$

The following result thus explains our choice of notation for the leading coefficient of the polynomial $p_{2i_1, \dots, 2i_k}^v$.

Theorem 6.4. *For all $i_1, \dots, i_k \geq 0$, the leading coefficient of $p_{2i_1, \dots, 2i_k}^v$ is a product of Catalan numbers,*

$$C_{i_1, \dots, i_k} = \prod_{r=1}^k \left[\frac{1}{i_r + 1} \binom{2i_r}{i_r} \right] = \prod_{r=1}^k C_{i_r}. \tag{6.2}$$

In particular, the leading coefficient of p_{2n}^v is the n th Catalan number C_n .

Proof. The proof is by induction on $\sum_{r=1}^k i_r$. The base case is trivial and occurs when all the indices i_r are equal to zero.

Now, applying (5.4) and assuming that $i_1 \geq 1$ we have

$$\begin{aligned} p_{2i_1, \dots, 2i_k}^v & = \sum_{r=1}^{2i_1-1} p_{r-1, 2i_1-r-1, 2i_2, \dots, 2i_k}^v + \sum_{r=2}^k 2i_r p_{2(i_1+i_r-1), 2i_2, \dots, \widehat{2i_r}, \dots, 2i_k}^v \\ & = \sum_{r=0}^{i_1-1} p_{2r, 2(i_1-1-r), 2i_2, \dots, 2i_k}^v + \sum_{r=1}^{i_1-1} p_{2r-1, 2(i_1-r)-1, 2i_2, \dots, 2i_k}^v \\ & \quad + \sum_{r=2}^k 2i_r p_{2(i_1+i_r-1), 2i_2, \dots, \widehat{2i_r}, \dots, 2i_k}^v. \end{aligned} \tag{6.3}$$

Note that by Theorem 5.8, only the first sum contains polynomials of the same degree as $p_{2i_1, \dots, 2i_k}^v$, with the remaining sums containing polynomials of lower degrees. This yields the recursion relation

$$C_{i_1, \dots, i_k} = \sum_{r=0}^{i_1-1} C_{r, i_1-1-r, i_2, \dots, i_k}.$$

Upon substituting (6.2), this recursion relation becomes the well-known and canonical recursion relation

$$C_{n+1} = \sum_{i=0}^n C_i C_{n-i} \tag{6.4}$$

for the Catalan numbers. ■

We may now prove Theorem 6.2.

Proof of Theorem 6.2. To prove Theorem 6.2 we proceed by induction on

$$l + \sum_{r=1}^{2l} j_r + \sum_{r=1}^k i_r \geq 1.$$

The base case is $i_1 = \dots = i_k = j_1 = \dots = j_{2l} = 0$ and $l = 1$, in which case the polynomial (6.1) is $v^k p_{1,1}^v = v^{k+1}$; hence the base case is clearly satisfied.

Applying (5.4) to the polynomial $p_{2j_1+1, \dots, 2j_{2l}+1, 2i_1, \dots, 2i_k}^v$ we get

$$\begin{aligned} p_{2j_1+1, \dots, 2j_{2l}+1, 2i_1, \dots, 2i_k}^v &= 2 \sum_{r=1}^{j_1} p_{2(r-1), 2(j_1-r)+1, 2j_2+1, \dots, 2j_{2l}+1, 2i_1, \dots, 2i_k}^v \\ &\quad + \sum_{r=2}^{2l} (2j_r + 1) p_{2(j_1+j_r), 2j_2+1, \dots, \widehat{2j_r+1}, \dots, 2j_{2l}+1, 2i_1, \dots, 2i_k}^v \\ &\quad + \sum_{r=1}^k 2i_r p_{2(j_1+i_r)-1, 2j_2+1, \dots, 2j_{2l}+1, 2i_1, \dots, \widehat{2i_r}, \dots, 2i_k}^v \end{aligned} \tag{6.5}$$

Here we have split up the first sum in (5.4) into sums over even and odd indices and then identified these two sums. Note that by Theorem 5.8 the last sum in (6.5) contains polynomials of strictly smaller degree and hence may be ignored for our purposes.

Applying the inductive hypothesis to (6.5) we conclude that the leading coefficient of (6.1) is

$$\begin{aligned} &2 \sum_{r=1}^{j_1} C_{r-1, i_1, \dots, i_k} A_{j_1-r, j_2, \dots, j_{2l}} + \sum_{r=2}^{2l} (2j_r + 1) C_{j_1+j_r, i_1, \dots, i_k} A_{j_2, \dots, \widehat{j_r}, \dots, j_{2l}} \\ &= C_{i_1, \dots, i_k} \left(2 \sum_{r=1}^{j_1} C_{r-1} A_{j_1-r, j_2, \dots, j_{2l}} + \sum_{r=2}^{2l} (2j_r + 1) C_{j_1+j_r} A_{j_2, \dots, \widehat{j_r}, \dots, j_{2l}} \right) \end{aligned} \tag{6.6}$$

where on the last line of (6.6) we have applied equation (6.2).

In particular, setting $k = 0$ in (6.5) and applying the inductive hypothesis yields the identity

$$A_{j_1, \dots, j_{2l}} = 2 \sum_{r=1}^{j_1} C_{r-1} A_{j_1-r, j_2, \dots, j_{2l}} + \sum_{r=2}^{2l} (2j_r + 1) C_{j_1+j_r} A_{j_2, \dots, \widehat{j_r}, \dots, j_{2l}}. \tag{6.7}$$

Finally, substituting (6.7) into (6.6) we conclude that the leading coefficient of (6.1) is $C_{i_1, \dots, i_k} A_{j_1, \dots, j_{2l}}$. ■

It remains to compute the coefficient $A_{j_1, \dots, j_{2l}}$, which we will do using the recurrence relation (6.7). We begin by examining the simplest possible case and derive an expression in terms of the Catalan numbers C_n .

Lemma 6.5. *For all $j_1, j_2 \geq 0$, we have*

$$A_{j_1, j_2} = (2j_2 + 1) \sum_{r=0}^{j_1} (r + 1) C_r C_{j_1 + j_2 - r}. \tag{6.8}$$

Proof. Consider the generating function $c(t) = \sum_{n=0}^{\infty} C_n t^n$ for the Catalan numbers and recall that the canonical recursion relation (6.4) is encoded as

$$c(t) = 1 + tc(t)^2. \tag{6.9}$$

Now introduce the generating functions

$$a_{j_2}(t) := \sum_{n=0}^{\infty} A_{n, j_2} t^n \quad \text{and} \quad \vec{c}_{j_2}(t) := \sum_{n=0}^{\infty} C_{n+j_2} t^n.$$

From Theorem 6.2 and the recursion relation (6.7), we obtain the recursion

$$A_{j_1, j_2} = 2 \sum_{r=0}^{j_1-1} C_r A_{j_1-1-r, j_2} + (2j_2 + 1) C_{j_1+j_2}. \tag{6.10}$$

In terms of these generating functions, the recursion relation (6.10) can be written as

$$a_{j_2}(t) = (2j_2 + 1)\vec{c}_{j_2}(t) + 2tc(t)a_{j_2}(t)$$

from which we get

$$a_{j_2}(t) = (2j_2 + 1)(1 - 2tc(t))^{-1}\vec{c}_{j_2}(t). \tag{6.11}$$

One approach to computing the power series $(1 - 2tc(t))^{-1}$ is to differentiate (6.9); however, we choose to proceed as follows. Completing the square in (6.9) we obtain

$$\begin{aligned} (1 - 2tc(t))^2 &= 1 - 4t, \\ (1 - 2tc(t))^{-1} &= (1 - 4t)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n} (-4t)^n \\ &= \sum_{n=0}^{\infty} \binom{2n}{n} t^n = \sum_{n=0}^{\infty} (n + 1) C_n t^n, \end{aligned} \tag{6.12}$$

which one may observe is the derivative of $tc(t)$. Substituting the above into (6.11) we obtain equation (6.8). ■

Remark 6.6. Note that the coefficient A_{j_1, j_2} is symmetric in j_1 and j_2 , whilst there appears to be no (alternative) a priori explanation for why the expression on the right-hand side of (6.8) possesses this symmetry. One might view this as an indication that the formula (6.8) is not canonical, and we will indeed derive a much cleaner expression for this coefficient in Theorem 6.13.

The final step we must take is to derive an expression for $A_{j_1, \dots, j_{2l}}$ in terms of these simpler coefficients. Given a chord diagram

$$c = \{r_1, s_1\}, \{r_2, s_2\}, \dots, \{r_l, s_l\} \in \mathcal{C}(2l)$$

define

$$A_{j_1, \dots, j_{2l}}(c) := A_{j_{r_1}, j_{s_1}} A_{j_{r_2}, j_{s_2}} \cdots A_{j_{r_l}, j_{s_l}}.$$

Proposition 6.7. For all $j_1, \dots, j_{2l} \geq 0$,

$$A_{j_1, \dots, j_{2l}} = \sum_{c \in \mathcal{C}(2l)} A_{j_1, \dots, j_{2l}}(c).$$

Proof. The proof proceeds by induction on $\sum_{r=1}^{2l} j_r$. The base case occurs when all the indices j_r are equal to zero and hence follows from the calculation performed in Example 5.6 and the fact that there are $(2l - 1)!!$ chord diagrams in $\mathcal{C}(2l)$.

Now, assuming that $j_1 \geq 1$, we apply the recurrence relation (6.7) and the inductive hypothesis to conclude that

$$\begin{aligned} A_{j_1, \dots, j_{2l}} &= 2 \sum_{i=1}^{j_1} \sum_{c \in \mathcal{C}(2l)} C_{i-1} A_{j_1-i, j_2, \dots, j_{2l}}(c) \\ &\quad + \sum_{k=2}^{2l} \sum_{c \in \mathcal{C}(l-1)} (2j_k + 1) C_{j_1+j_k} A_{j_2, \dots, \widehat{j}_k, \dots, j_{2l}}(c). \end{aligned}$$

Using the identity

$$\sum_{c \in \mathcal{C}(2l)} A_{j_1, \dots, j_{2l}}(c) = \sum_{k=2}^{2l} \sum_{c \in \mathcal{C}(l-1)} A_{j_1, j_k} A_{j_2, \dots, \widehat{j}_k, \dots, j_{2l}}(c) \tag{6.13}$$

and the recurrence relation (6.10), we obtain

$$\begin{aligned} \sum_{c \in \mathcal{C}(2l)} A_{j_1, \dots, j_{2l}}(c) &= 2 \sum_{k=2}^{2l} \sum_{c \in \mathcal{C}(l-1)} \sum_{i=1}^{j_1} C_{i-1} A_{j_1-i, j_k} A_{j_2, \dots, \widehat{j}_k, \dots, j_{2l}}(c) \\ &\quad + \sum_{k=2}^{2l} \sum_{c \in \mathcal{C}(l-1)} (2j_k + 1) C_{j_1+j_k} A_{j_2, \dots, \widehat{j}_k, \dots, j_{2l}}(c) \end{aligned}$$

$$\begin{aligned}
 &= 2 \sum_{i=1}^{j_1} \sum_{c \in \mathcal{C}(2l)} C_{i-1} A_{j_1-i, j_2, \dots, j_{2l}}(c) \\
 &\quad + \sum_{k=2}^{2l} \sum_{c \in \mathcal{C}(l-1)} (2j_k + 1) C_{j_1+j_k} A_{j_2, \dots, \widehat{j_k}, \dots, j_{2l}}(c) = A_{j_1, \dots, j_{2l}}. \blacksquare
 \end{aligned}$$

Remark 6.8. From Proposition 6.7 and equation (6.13) we obtain the useful recurrence relation

$$A_{j_1, \dots, j_{2l}} = \sum_{k=2}^{2l} A_{j_1, j_k} A_{j_2, \dots, \widehat{j_k}, \dots, j_{2l}}.$$

This concludes the first step of calculating the leading coefficient of (6.1). In the meantime, as a simple demonstration of our results, we provide the following example.

Example 6.9. Using Corollary 6.3, Theorem 6.4, Lemma 6.5 and Proposition 6.7 one may calculate

$$\begin{aligned}
 &\lim_{N \rightarrow \infty} \left[\frac{\int_{\mathfrak{h}_N} \text{Tr}(X^{10}) \text{Tr}(X^{42}) \text{Tr}(X^{15}) \text{Tr}(X^{43}) \text{Tr}(X^{47}) \text{Tr}(X^{63}) e^{-\frac{1}{2} \text{Tr}(X^2)} dX}{N^{112} \int_{\mathfrak{h}_N} e^{-\frac{1}{2} \text{Tr}(X^2)} dX} \right] \\
 &= C_{5,21} A_{7,21,23,31} \\
 &= 25081904924688737847061935982290890890757044619026344345600000.
 \end{aligned}$$

The details are left, of course, to the enthusiastic reader.

6.2. Subleading coefficients

In order to compute the large N statistical correlations between certain random variables in the Gaussian unitary ensemble in Section 6.3, it is necessary to supplement our knowledge of the large N asymptotic behavior of the correlation functions (3.1) that we obtained in the preceding section by computing the coefficient of the subleading term in $p_{2i_1, \dots, 2i_k}^v$. This will also lead us to a cleaner formula for the coefficient $A_{j_1, \dots, j_{2l}}$.

Recall that the polynomial $p_{2i_1, \dots, 2i_k}^v$ is a polynomial of degree

$$n := k + \sum_{s=1}^k i_s.$$

Denote the coefficient of v^{n-2r} in $p_{2i_1, \dots, 2i_k}^v$ by $C_{i_1, \dots, i_k}(r)$ so that in the notation of the preceding section we have

$$C_{i_1, \dots, i_k} = C_{i_1, \dots, i_k}(0).$$

We will compute $C_{i_1, \dots, i_k}(1)$. To do so we make the auxiliary definition

$$\tilde{C}_{i,j}(1) := C_{i,j}(1) - C_i(1)C_j - C_j(1)C_i \tag{6.14}$$

which is symmetric with respect to i and j .

Proposition 6.10. For all $i_1, \dots, i_k \geq 0$,

$$C_{i_1, \dots, i_k}(1) = \sum_{r=1}^k C_{i_r}(1) C_{i_1, \dots, \widehat{i_r}, \dots, i_k} + \sum_{1 \leq r < s \leq k} \tilde{C}_{i_r, i_s}(1) C_{i_1, \dots, \widehat{i_r}, \dots, \widehat{i_s}, \dots, i_k}. \tag{6.15}$$

Remark 6.11. It follows from equation (6.15) that

$$\begin{aligned} C_{i_1, \dots, i_k, i'_1, \dots, i'_{k'}}(1) &= C_{i_1, \dots, i_k}(1) C_{i'_1, \dots, i'_{k'}} + C_{i'_1, \dots, i'_{k'}}(1) C_{i_1, \dots, i_k} \\ &\quad + \sum_{r=1}^k \sum_{s=1}^{k'} \tilde{C}_{i_r, i'_s}(1) C_{i_1, \dots, \widehat{i_r}, \dots, i_k} C_{i'_1, \dots, \widehat{i'_s}, \dots, i'_{k'}} \end{aligned} \tag{6.16}$$

for all $i_1, \dots, i_k; i'_1, \dots, i'_{k'} \geq 0$.

Proof of Proposition 6.10. The proof is by induction on $\sum_{r=1}^k i_r$. The base case occurs when all the indices i_r are equal to zero, in which case both sides of (6.15) clearly vanish.

We note that from equation (6.3), Theorem 5.8 and Theorem 6.2 we obtain the recursion relation

$$\begin{aligned} C_{i_1, \dots, i_k}(1) &= \sum_{r=0}^{i_1-1} C_{r, i_1-1-r, i_2, \dots, i_k}(1) + \sum_{r=0}^{i_1-2} A_{r, i_1-2-r} C_{i_2, \dots, i_k} \\ &\quad + 2 \sum_{r=2}^k i_r C_{i_1+i_r-1, i_2, \dots, \widehat{i_r}, \dots, i_k} \end{aligned} \tag{6.17}$$

which is valid for all $i_1 \geq 1$ and $i_2, \dots, i_k \geq 0$. In particular, setting $k = 1$ in (6.17) we obtain

$$C_i(1) = \sum_{r=0}^{i-1} C_{r, i-1-r}(1) + \sum_{r=0}^{i-2} A_{r, i-2-r} \tag{6.18}$$

which is valid for all $i \geq 1$. Combining (6.18) with (6.17) yields the recursion

$$\begin{aligned} C_{i_1, \dots, i_k}(1) &= \sum_{r=0}^{i_1-1} C_{r, i_1-1-r, i_2, \dots, i_k}(1) + C_{i_2, \dots, i_k} \left(C_{i_1}(1) - \sum_{r=0}^{i_1-1} C_{r, i_1-1-r}(1) \right) \\ &\quad + 2 \sum_{r=2}^k i_r C_{i_1+i_r-1, i_2, \dots, \widehat{i_r}, \dots, i_k} \end{aligned} \tag{6.19}$$

which is valid for all $i_1 \geq 1$ and $i_2, \dots, i_k \geq 0$.

Applying the inductive hypothesis and making use of equation (6.16), we find that

$$\begin{aligned} C_{r, i_1-1-r, i_2, \dots, i_k}(1) &= C_{r, i_1-1-r}(1) C_{i_2, \dots, i_k} + C_{i_2, \dots, i_k}(1) C_{r, i_1-1-r} \\ &\quad + \sum_{s=2}^k \tilde{C}_{r, i_s}(1) C_{i_1-1-r, i_2, \dots, \widehat{i_s}, \dots, i_k} \\ &\quad + \sum_{s=2}^k \tilde{C}_{i_1-1-r, i_s}(1) C_{r, i_2, \dots, \widehat{i_s}, \dots, i_k}. \end{aligned}$$

Substituting the above into (6.19) whilst making use of (6.4) and (6.2), we obtain

$$\begin{aligned}
 C_{i_1, \dots, i_k}(1) &= C_{i_2, \dots, i_k}(1) \sum_{r=0}^{i_1-1} C_r C_{i_1-1-r} + 2 \sum_{s=2}^k \sum_{r=0}^{i_1-1} \tilde{C}_{r, i_s}(1) C_{i_1-1-r, i_2, \dots, \widehat{i_s}, \dots, i_k} \\
 &\quad + C_{i_2, \dots, i_k} C_{i_1}(1) + 2 \sum_{s=2}^k i_s C_{i_1+i_s-1, i_2, \dots, \widehat{i_s}, \dots, i_k} \\
 &= C_{i_2, \dots, i_k}(1) C_{i_1} + C_{i_1}(1) C_{i_2, \dots, i_k} \\
 &\quad + 2 \sum_{s=2}^k C_{i_2, \dots, \widehat{i_s}, \dots, i_k} \left(i_s C_{i_1+i_s-1} + \sum_{r=0}^{i_1-1} \tilde{C}_{r, i_s}(1) C_{i_1-1-r} \right). \tag{6.20}
 \end{aligned}$$

Repeating the argument above once more, we obtain the formula

$$C_{i_1, i_s}(1) = C_{i_1}(1) C_{i_s} + C_{i_s}(1) C_{i_1} + 2 \left(i_s C_{i_1+i_s-1} + \sum_{r=0}^{i_1-1} \tilde{C}_{r, i_s}(1) C_{i_1-1-r} \right).$$

Substituting the above into (6.20) and using (6.14), we arrive at the equation

$$C_{i_1, \dots, i_k}(1) = C_{i_2, \dots, i_k}(1) C_{i_1} + C_{i_1}(1) C_{i_2, \dots, i_k} + \sum_{s=2}^k C_{i_2, \dots, \widehat{i_s}, \dots, i_k} \tilde{C}_{i_1, i_s}(1).$$

Applying the inductive hypothesis once more to the first term in this equation finally establishes equation (6.15). ■

Remark 6.12. A formula for the coefficient $C_i(1)$ may be found using the Harer–Zagier recursion relation [8]. Since we will not actually require such a formula for our purposes, we will not provide the full details, but it may be computed to be

$$C_i(1) = \begin{cases} \binom{2i-1}{3} C_{i-2}, & i \geq 2, \\ 0, & i = 0, 1 \end{cases}.$$

It therefore remains to compute the terms $\tilde{C}_{i,j}(1)$. This is done in a similar fashion to our computation of the coefficients $A_{i,j}$ in Lemma 6.5. One consequence of this analysis will be that we will also obtain a very clean formula for the coefficients $A_{i,j}$.

Theorem 6.13. For all $i, j \geq 0$,

$$\tilde{C}_{i,j}(1) = \begin{cases} \frac{ij}{(i+j)} \binom{2i}{i} \binom{2j}{j}, & i+j > 0, \\ 0, & i+j = 0 \end{cases}, \tag{6.21}$$

$$A_{i,j} = \frac{(2i+1)!(2j+1)!}{(i+j+1)(i!j!)^2} = \frac{(2i+1)(2j+1)}{(i+j+1)} \binom{2i}{i} \binom{2j}{j}. \tag{6.22}$$

Proof. From the recursion relation (6.19) and formula (6.16) we obtain equation (6.20), which for two indices $i \geq 1$ and $j \geq 0$ provides us with the recursion relation

$$\tilde{C}_{i,j}(1) = 2 \left(jC_{i+j-1} + \sum_{r=0}^{i-1} \tilde{C}_{r,j}(1)C_{i-1-r} \right). \tag{6.23}$$

Now consider the generating functions

$$c(t) := \sum_{n=0}^{\infty} C_n t^n, \quad \tilde{c}_j(t) := \sum_{n=0}^{\infty} \tilde{C}_{n,j}(1)t^n, \quad \text{and} \quad \bar{c}_j(t) := \sum_{n=0}^{\infty} C_{n+j} t^n.$$

The recursion relation (6.23) may be written in terms of these generating functions as

$$\tilde{c}_j(t) = 2t(c(t)\tilde{c}_j(t) + j\bar{c}_j(t)).$$

Rearranging and using our expression (6.12) for $(1 - 2tc(t))^{-1}$, we obtain

$$\begin{aligned} \tilde{c}_j(t) &= 2j(1 - 2tc(t))^{-1}t\bar{c}_j(t) \\ &= 2j \sum_{n=1}^{\infty} \left(\sum_{r=0}^{n-1} (r+1)C_r C_{n+j-r-1} \right) t^n. \end{aligned}$$

Therefore,

$$\tilde{C}_{i,j}(1) = 2j \sum_{r=0}^{i-1} (r+1)C_r C_{i+j-r-1} = \frac{2j}{2j+1} A_{i-1,j} \tag{6.24}$$

where at the last step we have used equation (6.8).

Using the fact that the coefficients $A_{i,j}$ and $\tilde{C}_{i,j}(1)$ are symmetric in i and j , we obtain from (6.24) the recursion relation

$$\begin{aligned} A_{i,j-1} &= \frac{j(2i+1)}{i(2j+1)} A_{i-1,j}, & i, j \geq 1, \\ A_{0,j} &= (2j+1)C_j, & j \geq 0, \end{aligned} \tag{6.25}$$

where we have used equation (6.8) to compute $A_{0,j}$. The recursive computation (6.25) may be solved in a straightforward manner and it is easily verified that the solution is provided by equation (6.22). Equation (6.21) then follows from (6.22) and (6.24). ■

With this new formula for the coefficients $A_{i,j}$, we can use Proposition 6.7 to give a more convenient expression for the coefficients $A_{j_1, \dots, j_{2l}}$. Given a chord diagram

$$c = \{r_1, s_1\}, \{r_2, s_2\}, \dots, \{r_l, s_l\} \in \mathcal{C}(2l)$$

define

$$\mu_{j_1, \dots, j_{2l}}(c) := \prod_{k=1}^l \left[\frac{1}{j_{r_k} + j_{s_k} + 1} \right]$$

and

$$\mu_{j_1, \dots, j_{2l}} := \sum_{c \in \mathcal{C}(2l)} \mu_{j_1, \dots, j_{2l}}(c) \tag{6.26}$$

for all $j_1, \dots, j_{2l} \geq 0$. Note again the convenient formula

$$\mu_{j_1, \dots, j_{2l}} = \sum_{k=2}^{2l} \mu_{j_1, j_k} \mu_{j_2, \dots, \widehat{j_k}, \dots, j_{2l}}.$$

Theorem 6.14. For all $j_1, \dots, j_{2l} \geq 0$,

$$A_{j_1, \dots, j_{2l}} = \mu_{j_1, \dots, j_{2l}} \prod_{k=1}^{2l} \left[\frac{(2j_k + 1)!}{(j_k!)^2} \right]. \tag{6.27}$$

Proof. This is an immediate consequence of Proposition 6.7 and Theorem 6.13. ■

6.3. Large N statistical correlation coefficients in the Gaussian unitary ensemble

Consider the algebra of conjugation-invariant *polynomial* functions on matrices,

$$S(M_N(\mathbb{C})^*)^{\mathfrak{gl}_n(\mathbb{C})}.$$

By the Fundamental Theorem of Invariant Theory for $GL_N(\mathbb{C})$, we know that this algebra is generated by traces $\text{Tr}(X^k)$ of powers of the matrix X . We can see these functions, of course, as random variables in the Gaussian unitary ensemble. In terms of this paper, we note that

$$S(M_N(\mathbb{C})^*)^{\mathfrak{gl}_n(\mathbb{C})} \subset \text{BV}[M_N(\mathcal{A})],$$

so that our approach has full access to this class of random variables.

Let f and g be two conjugation-invariant, polynomial random variables. Recall that the variance, covariance and statistical correlation between the random variables f and g are defined by

$$\begin{aligned} \text{Var}(f) &:= \langle f^2 \rangle - \langle f \rangle^2, \\ \text{Cov}(f, g) &:= \langle fg \rangle - \langle f \rangle \langle g \rangle, \\ \rho(f, g) &:= \frac{\text{Cov}(f, g)}{\sqrt{\text{Var}(f)\text{Var}(g)}} \end{aligned}$$

where we remind the reader that the expected value $\langle f \rangle$ of the random variable f on the sample space of Hermitian matrices is defined by the expectation value map (3.2).

By Proposition 5.2 the expected value of the random variable

$$f(X) := \text{Tr}(X^{k_1})\text{Tr}(X^{k_2}) \dots \text{Tr}(X^{k_n}), \quad X \in \mathfrak{h}_N \tag{6.28}$$

is

$$\langle f \rangle = p_{k_1, \dots, k_n}^v(N).$$

In this section, we will compute the large N statistical correlations between the random variables (6.28); i.e., between multi-trace functions. We will find that the resulting expressions are dominated by the presence of odd powers in (6.28).

Theorem 6.15. *Fix nonnegative integers*

$$i_1, \dots, i_k; j_1, \dots, j_l; i'_1, \dots, i'_{k'}; j'_1, \dots, j'_{l'} \geq 0 \quad \text{with } k, l, k', l' \geq 0.$$

These determine sequences of random variables (f_N) and (g_N) , where for each $N \in \mathbb{N}$, we have functions on the rank N Gaussian unitary ensemble: for $X \in \mathfrak{h}_N$,

$$\begin{aligned} f_N(X) &:= \text{Tr}(X^{2i_1}) \dots \text{Tr}(X^{2i_k}) \text{Tr}(X^{2j_1+1}) \dots \text{Tr}(X^{2j_l+1}), \\ g_N(X) &:= \text{Tr}(X^{2i'_1}) \dots \text{Tr}(X^{2i'_{k'}}) \text{Tr}(X^{2j'_1+1}) \dots \text{Tr}(X^{2j'_{l'}+1}). \end{aligned}$$

Then the following are true:

- (1) *If l and l' have opposite parities, then*

$$\rho(f_N, g_N) = 0$$

for all $N \in \mathbb{N}$.

- (2) *If l and l' are both odd, then*

$$\lim_{N \rightarrow \infty} \rho(f_N, g_N) = \frac{\mu_{j_1, \dots, j_l, j'_1, \dots, j'_{l'}}}{\sqrt{\mu_{j_1, \dots, j_l, j_1, \dots, j_l} \mu_{j'_1, \dots, j'_{l'}, j'_1, \dots, j'_{l'}}}}. \tag{6.29}$$

- (3) *If l and l' are both even and $l, l' > 0$, then*

$$\lim_{N \rightarrow \infty} \rho(f_N, g_N) = \frac{\mu_{j_1, \dots, j_l, j'_1, \dots, j'_{l'}} - \mu_{j_1, \dots, j_l} \mu_{j'_1, \dots, j'_{l'}}}{\sqrt{\mu_{j_1, \dots, j_l, j_1, \dots, j_l} - \mu_{j_1, \dots, j_l}^2} \sqrt{\mu_{j'_1, \dots, j'_{l'}, j'_1, \dots, j'_{l'}} - \mu_{j'_1, \dots, j'_{l'}}^2}}. \tag{6.30}$$

- (4) *If l is even, $l > 0$, and $l' = 0$, then (assuming not all $i'_r = 0$)*

$$\lim_{N \rightarrow \infty} \rho(f_N, g_N) = 0. \tag{6.31}$$

- (5) *If $l = l' = 0$, then (assuming that $i_1, \dots, i_k, i'_1, \dots, i'_{k'} > 0$)*

$$\lim_{N \rightarrow \infty} \rho(f_N, g_N) = \frac{\sum_{r=1}^k \sum_{s=1}^{k'} \frac{i_r i'_s (i_r+1)(i'_s+1)}{(i_r+i'_s)}}{\sqrt{\sum_{r,s=1}^k \frac{i_r i_s (i_r+1)(i_s+1)}{(i_r+i_s)}} \sqrt{\sum_{r,s=1}^{k'} \frac{i'_r i'_s (i'_r+1)(i'_s+1)}{(i'_r+i'_s)}}}. \tag{6.32}$$

Remark 6.16. Note that for equation (6.32) the assumption that the indices i_1, \dots, i_k and $i'_1, \dots, i'_{k'}$ are all positive is not a real constraint, since the inclusion of indices that are equal to zero will not affect the correlation coefficient. This is because the statistical correlation between random variables does not change after multiplying the random variables by a scalar.

Remark 6.17. For convenience, we point out that in the simplest instances the above formulas provide the following nice expressions:

$$\begin{aligned} \lim_{N \rightarrow \infty} \rho(\text{Tr}(X^{2i+1}), \text{Tr}(X^{2j+1})) &= \frac{\sqrt{(2i+1)(2j+1)}}{(i+j+1)}, \\ \lim_{N \rightarrow \infty} \rho(\text{Tr}(X^{2i}), \text{Tr}(X^{2j})) &= \frac{2\sqrt{ij}}{(i+j)}. \end{aligned}$$

Proof of Theorem 6.15. (1) In this case $\langle f_N g_N \rangle = \langle f_N \rangle \langle g_N \rangle = 0$ for all positive integers N and hence the covariance $\text{Cov}(f_N, g_N)$ vanishes for all N .

(2) In this case we have $\langle f_N \rangle = \langle g_N \rangle = 0$ for all N . Therefore, using Theorem 6.2 and Theorem 5.8 we find that

$$\begin{aligned} \text{Cov}(f_N, g_N) &= \langle f_N g_N \rangle \\ &= p_{2i_1, \dots, 2i_k, 2i'_1, \dots, 2i'_{k'}, 2j_1+1, \dots, 2j_l+1, 2j'_1+1, \dots, 2j'_{l'}+1}(N) \\ &= C_{i_1, \dots, i_k, i'_1, \dots, i'_{k'}} A_{j_1, \dots, j_l, j'_1, \dots, j'_{l'}} N^n + O(N^{n-2}) \end{aligned}$$

where

$$n := k + k' + \frac{(l+l')}{2} + \sum_{r=1}^k i_r + \sum_{r=1}^{k'} i'_r + \sum_{r=1}^l j_r + \sum_{r=1}^{l'} j'_r. \tag{6.33}$$

We find similar expressions for the variances:

$$\begin{aligned} \text{Var}(f_N) &= \text{Cov}(f_N, f_N) \\ &= C_{i_1, \dots, i_k, i_1, \dots, i_k} A_{j_1, \dots, j_l, j_1, \dots, j_l} N^{2n_f} + O(N^{2(n_f-1)}), \\ \text{Var}(g_N) &= C_{i'_1, \dots, i'_{k'}, i'_1, \dots, i'_{k'}} A_{j'_1, \dots, j'_{l'}, j'_1, \dots, j'_{l'}} N^{2n_g} + O(N^{2(n_g-1)}); \end{aligned}$$

where

$$n_f := k + \frac{l}{2} + \sum_{r=1}^k i_r + \sum_{r=1}^l j_r \quad \text{and} \quad n_g := k' + \frac{l'}{2} + \sum_{r=1}^{k'} i'_r + \sum_{r=1}^{l'} j'_r \tag{6.34}$$

so that $n = n_f + n_g$.

Putting these together, we find that

$$\rho(f_N, g_N) = \frac{C_{i_1, \dots, i_k} C_{i'_1, \dots, i'_{k'}} A_{j_1, \dots, j_l, j'_1, \dots, j'_{l'}} + O(\frac{1}{N^2})}{\sqrt{C_{i_1, \dots, i_k}^2 A_{j_1, \dots, j_l, j_1, \dots, j_l} + O(\frac{1}{N^2})} \sqrt{C_{i'_1, \dots, i'_{k'}}^2 A_{j'_1, \dots, j'_{l'}, j'_1, \dots, j'_{l'}} + O(\frac{1}{N^2})}}$$

and hence that

$$\lim_{N \rightarrow \infty} \rho(f_N, g_N) = \frac{A_{j_1, \dots, j_l, j'_1, \dots, j'_{l'}}}{\sqrt{A_{j_1, \dots, j_l, j_1, \dots, j_l}} \sqrt{A_{j'_1, \dots, j'_{l'}, j'_1, \dots, j'_{l'}}}}.$$

Using formula (6.27) for the coefficients in the numerator and the denominator of the above expression we find that all the factorials cancel, leaving us with equation (6.29).

(3) Proceeding as above, we find that

$$\begin{aligned} \text{Cov}(f_N, g_N) &= p_{2i_1, \dots, 2i_k, 2i'_1, \dots, 2i'_{k'}, 2j_1+1, \dots, 2j_l+1, 2j'_1+1, \dots, 2j'_{l'}+1}(N) \\ &\quad - p_{2i_1, \dots, 2i_k, 2j_1+1, \dots, 2j_l+1}(N) p_{2i'_1, \dots, 2i'_{k'}, 2j'_1+1, \dots, 2j'_{l'}+1}(N) \\ &= C_{i_1, \dots, i_k, i'_1, \dots, i'_{k'}} A_{j_1, \dots, j_l, j'_1, \dots, j'_{l'}} N^n + O(N^{n-2}) \\ &\quad - (C_{i_1, \dots, i_k} A_{j_1, \dots, j_l} N^{n_f} + O(N^{n_f-2})) \\ &\quad \cdot (C_{i'_1, \dots, i'_{k'}} A_{j'_1, \dots, j'_{l'}} N^{n_g} + O(N^{n_g-2})) \\ &= C_{i_1, \dots, i_k} C_{i'_1, \dots, i'_{k'}} (A_{j_1, \dots, j_l, j'_1, \dots, j'_{l'}} - A_{j_1, \dots, j_l} A_{j'_1, \dots, j'_{l'}}) N^n + O(N^{n-2}) \end{aligned}$$

where n , n_f and n_g are defined exactly as above in (6.33) and (6.34).

As before, in the limit the C_{i_1, \dots, i_k} and $C_{i'_1, \dots, i'_{k'}}$ terms cancel along with the factorials, leaving us with the expression (6.30) for the large N correlation. Note that both the numerator and the two factors in the denominator of (6.30) are positive when $l, l' > 0$. This follows from (6.26).

(4) Defining n , n_f and n_g according to (6.33) and (6.34) and applying Theorem 6.2 and equation (6.16), we compute

$$\begin{aligned} \text{Cov}(f_N, g_N) &= C_{i_1, \dots, i_k, i'_1, \dots, i'_{k'}} A_{j_1, \dots, j_l} N^n + O(N^{n-2}) \\ &\quad - (C_{i_1, \dots, i_k} A_{j_1, \dots, j_l} N^{n_f} + O(N^{n_f-2})) (C_{i'_1, \dots, i'_{k'}} N^{n_g} + O(N^{n_g-2})) \\ &= O(N^{n-2}), \\ \text{Var}(f_N) &= C_{i_1, \dots, i_k}^2 (A_{j_1, \dots, j_l, j_1, \dots, j_l} - A_{j_1, \dots, j_l}^2) N^{2n_f} + O(N^{2(n_f-1)}), \\ \text{Var}(g_N) &= C_{i'_1, \dots, i'_{k'}}^2 N^{2n_g} + C_{i'_1, \dots, i'_{k'}, i'_1, \dots, i'_{k'}}(1) N^{2(n_g-1)} + O(N^{2(n_g-2)}) \\ &\quad - (C_{i'_1, \dots, i'_{k'}} N^{n_g} + C_{i'_1, \dots, i'_{k'}}(1) N^{n_g-2} + O(N^{n_g-4}))^2 \\ &= (C_{i'_1, \dots, i'_{k'}, i'_1, \dots, i'_{k'}}(1) - 2C_{i'_1, \dots, i'_{k'}}(1) C_{i'_1, \dots, i'_{k'}}) N^{2(n_g-1)} + O(N^{2(n_g-2)}) \\ &= \left(\sum_{r,s=1}^{k'} \tilde{C}_{i'_r, i'_s}(1) C_{i'_1, \dots, \widehat{i'_r}, \dots, i'_{k'}} C_{i'_1, \dots, \widehat{i'_s}, \dots, i'_{k'}} \right) N^{2(n_g-1)} + O(N^{2(n_g-2)}). \end{aligned}$$

Putting these together, we get

$$\begin{aligned} \rho(f_N, g_N) &= \frac{O(\frac{1}{N})}{\sqrt{C_{i_1, \dots, i_k}^2 (A_{j_1, \dots, j_l, j_1, \dots, j_l} - A_{j_1, \dots, j_l}^2) \sum_{r,s=1}^{k'} \tilde{C}_{i'_r, i'_s}(1) C_{i'_1, \dots, \widehat{i'_r}, \dots, i'_{k'}} C_{i'_1, \dots, \widehat{i'_s}, \dots, i'_{k'}} + O(\frac{1}{N^2})}} \end{aligned}$$

from which (6.31) follows.

(5) Proceeding in the same manner as above using equation (6.16), we compute

$$\begin{aligned} & \text{Cov}(f_N, g_N) \\ &= \left(\sum_{r=1}^k \sum_{s=1}^{k'} \tilde{C}_{i_r, i'_s}(1) C_{i_1, \dots, \widehat{i_r}, \dots, i_k} C_{i'_1, \dots, \widehat{i'_s}, \dots, i'_{k'}} \right) N^{n-2} + O(N^{n-4}), \end{aligned}$$

where n is of course defined by (6.33). From this we get

$$\begin{aligned} & \rho(f_N, g_N) \\ &= \frac{\sum_{r=1}^k \sum_{s=1}^{k'} \tilde{C}_{i_r, i'_s}(1) C_{i_1, \dots, \widehat{i_r}, \dots, i_k} C_{i'_1, \dots, \widehat{i'_s}, \dots, i'_{k'}} + O(\frac{1}{N^2})}{\sqrt{\sum_{r,s=1}^k \tilde{C}_{i_r, i_s}(1) C_{i_1, \dots, \widehat{i_r}, \dots, i_k} C_{i_1, \dots, \widehat{i_s}, \dots, i_k} \sum_{r,s=1}^{k'} \tilde{C}_{i'_r, i'_s}(1) C_{i'_1, \dots, \widehat{i'_r}, \dots, i'_{k'}} C_{i'_1, \dots, \widehat{i'_s}, \dots, i'_{k'}} + O(\frac{1}{N^2})}}. \end{aligned}$$

Taking the limit and using equation (6.21), we see that the coefficients C_{i_1, \dots, i_k} and $C_{i'_1, \dots, i'_{k'}}$ in the numerator and the denominator cancel leaving us with expression (6.32). ■

6.4. A generalization of Wigner’s semicircle law and asymptotically decoupled random variables

As a final application of our results on the large N asymptotic behavior of the correlation functions (3.1), we will prove a generalization of Wigner’s law to multi-trace functions. We thank Gaëtan Borot for pointing out to us that the following theorem may be obtained using the results of [3].

Theorem 6.18. *For any $m \geq 1$ and polynomials $q_1, \dots, q_m \in \mathbb{C}[x]$,*

$$\begin{aligned} & \lim_{N \rightarrow \infty} \left(\frac{\int_{\mathfrak{h}_N} \prod_{i=1}^m \left[\frac{1}{N} \text{Tr} \left(q_i \left(\frac{X}{\sqrt{N}} \right) \right) \right] e^{-\frac{1}{2} \text{Tr}(X^2)} dX}{\int_{\mathfrak{h}_N} e^{-\frac{1}{2} \text{Tr}(X^2)} dX} \right) \\ &= \prod_{i=1}^m \left[\frac{1}{2\pi} \int_{-2}^2 q_i(x) \sqrt{4 - x^2} dx \right]. \end{aligned} \tag{6.35}$$

Of course, the case $m = 1$ is Wigner’s original semicircle law. In this sense, equation (6.35) may be restated as an equation amongst expectation values,

$$\lim_{N \rightarrow \infty} \left\langle \prod_{i=1}^m \left[\frac{1}{N} \text{Tr} \left(q_i \left(\frac{X}{\sqrt{N}} \right) \right) \right] \right\rangle = \prod_{i=1}^m \left[\lim_{N \rightarrow \infty} \left\langle \frac{1}{N} \text{Tr} \left(q_i \left(\frac{X}{\sqrt{N}} \right) \right) \right\rangle \right].$$

In other words, we can swap the limit with the product; we may say that the expectation value of the product becomes the product of the expectation values in the large N limit. Thus, in the large N limit, we observe the vanishing of certain variances and covariances; that is we see a “decoupling” between the random variables

$$\frac{1}{N} \text{Tr} \left(q \left(\frac{X}{\sqrt{N}} \right) \right), \quad X \in \mathfrak{h}_N$$

in the Gaussian unitary ensemble that are formed by polynomials $q \in \mathbb{C}[x]$.

Proof of Theorem 6.18. We begin by noting that both sides of (6.35) are multilinear functions of the polynomials q_1, \dots, q_m . It is therefore sufficient to assume that q_1, \dots, q_m are monomials of the form $q_r(x) = x^{k_r}$.

Obviously, if there is an odd number of indices k_r having odd parity, then both sides of (6.35) will vanish before even taking the limit. Therefore, we may begin by assuming that the number of indices k_r having odd parity is $2l > 0$. In this case, the right-hand side of (6.35) clearly vanishes, whilst it follows from Theorem 5.8 that the left-hand side is $O(\frac{1}{N^{2l}})$ and hence vanishes in the limit.

It remains to consider the case when $l = 0$ so that all the indices $k_r = 2i_r$ are even. In this case, it follows from Theorem 5.8 that the left-hand side is

$$\lim_{N \rightarrow \infty} \left(C_{i_1, \dots, i_m} + O\left(\frac{1}{N^2}\right) \right) = C_{i_1, \dots, i_m}$$

and the result now follows from Theorem 6.4 and the following integral expression for the Catalan numbers:

$$C_i = \frac{1}{2\pi} \int_{-2}^2 x^{2i} \sqrt{4 - x^2} \, dx, \quad i \geq 0;$$

which is well known. ■

Remark 6.19. Using the results of Section 6.1 and 6.2 we may compute the following large N covariance, refining ever so slightly the above result. Let $f, g \in \mathbb{C}[x]$ be two polynomials. Then

$$\begin{aligned} & \lim_{N \rightarrow \infty} \left[\text{Cov} \left(\text{Tr} \left(f \left(\frac{X}{\sqrt{N}} \right) \right), \text{Tr} \left(g \left(\frac{X}{\sqrt{N}} \right) \right) \right) \right] \\ &= \sum_{i, j=1}^{\infty} \frac{f^{(2i)}(0)g^{(2j)}(0)}{(i+j)!i!(i-1)!j!(j-1)!} + \sum_{i, j=0}^{\infty} \frac{f^{(2i+1)}(0)g^{(2j+1)}(0)}{(i+j+1)(i!)^2(j!)^2}. \end{aligned} \tag{6.36}$$

We leave the details to the reader. If there is a more intrinsic expression for the quantity on the right-hand side of (6.36), it is not readily apparent to the authors of this article.

Acknowledgments. We thank Gaëtan Borot and James Mingo for email exchanges about the literature on GUE and for suggestions of questions to pursue. In particular, Gaëtan gave us much insight into the current state of the art and alternative formulations of the kinds of results developed here.

Funding. Owen Gwilliam gratefully acknowledges the support of the National Science Foundation, grant 2042052.

References

[1] J. Ambjørn, L. Chekhov, and Y. Makeenko, [Higher genus correlators from the Hermitian one-matrix model](#). *Phys. Lett. B* **282** (1992), no. 3-4, 341–348 MR [1169645](#)

- [2] J. Ambjørn and Y. M. Makeenko, [Properties of loop equations for the Hermitian matrix model and for two-dimensional quantum gravity](#). *Modern Phys. Lett. A* **5** (1990), no. 22, 1753–1763 Zbl [1020.81806](#) MR [1075933](#)
- [3] G. Borot and A. Guionnet, [Asymptotic expansion of \$\beta\$ matrix models in the one-cut regime](#). *Comm. Math. Phys.* **317** (2013), no. 2, 447–483 Zbl [1344.60012](#) MR [3010191](#)
- [4] B. Eynard, [Counting surfaces](#). Prog. Math. Phys. 70, Birkhäuser/Springer, [Cham], 2016 Zbl [1338.81005](#) MR [3468847](#)
- [5] G. Ginot, O. Gwilliam, A. Hamilton, and M. Zeinalian, [Large \$N\$ phenomena and quantization of the Loday–Quillen–Tsygan theorem](#). *Adv. Math.* **409** (2022), no. part A, article no. 108631 Zbl [1502.18032](#) MR [4470611](#)
- [6] A. Guionnet, [Asymptotics of random matrices and related models](#). CBMS Reg. Conf. Ser. Math. 130, American Mathematical Society, Providence, RI, 2019 Zbl [1464.60004](#) MR [3931304](#)
- [7] A. Hamilton, [Noncommutative geometry and compactifications of the moduli space of curves](#). *J. Noncommut. Geom.* **4** (2010), no. 2, 157–188 Zbl [1193.14016](#) MR [2595312](#)
- [8] J. Harer and D. Zagier, [The Euler characteristic of the moduli space of curves](#). *Invent. Math.* **85** (1986), no. 3, 457–485 Zbl [0616.14017](#) MR [0848681](#)
- [9] J. L. Harer, [The virtual cohomological dimension of the mapping class group of an orientable surface](#). *Invent. Math.* **84** (1986), no. 1, 157–176 Zbl [0592.57009](#) MR [0830043](#)
- [10] M. Kontsevich, [Intersection theory on the moduli space of curves and the matrix Airy function](#). *Comm. Math. Phys.* **147** (1992), no. 1, 1–23 Zbl [0756.35081](#) MR [1171758](#)
- [11] M. Kontsevich, [Formal \(non\)commutative symplectic geometry](#). In *The Gel'fand Mathematical Seminars, 1990–1992*, pp. 173–187, Birkhäuser Boston, Boston, MA, 1993 Zbl [0821.58018](#) MR [1247289](#)
- [12] M. Kontsevich, [Over-determined Schwinger–Dyson equations for matrix models](#). Mathematical Physics Seminar, Rutgers University, Jun 8th 2022
- [13] J.-L. Loday and D. Quillen, [Cyclic homology and the Lie algebra homology of matrices](#). *Comment. Math. Helv.* **59** (1984), no. 4, 569–591 Zbl [0565.17006](#) MR [0780077](#)
- [14] M. V. Movshev, [Fukaya category with curves of higher genus](#). 1999, arXiv:[math/9911123v1](#)
- [15] R. C. Penner, [The decorated Teichmüller space of punctured surfaces](#). *Comm. Math. Phys.* **113** (1987), no. 2, 299–339 Zbl [0642.32012](#) MR [0919235](#)
- [16] A. Schwarz, [Geometry of Batalin–Vilkovisky quantization](#). *Comm. Math. Phys.* **155** (1993), no. 2, 249–260 Zbl [0786.58017](#) MR [1230027](#)
- [17] B. L. Tsygan, [Homology of matrix Lie algebras over rings and the Hochschild homology](#). *Uspekhi Mat. Nauk* **38** (1983), no. 2(230), 217–218 Zbl [0518.17002](#) MR [0695483](#)
- [18] E. P. Wigner, [Characteristic vectors of bordered matrices with infinite dimensions](#). *Ann. of Math. (2)* **62** (1955), 548–564 Zbl [0067.08403](#) MR [0077805](#)
- [19] E. P. Wigner, [On the distribution of the roots of certain symmetric matrices](#). *Ann. of Math. (2)* **67** (1958), 325–327 Zbl [0085.13203](#) MR [0095527](#)

Received 8 May 2023.

Owen Gwilliam

Department of Mathematics and Statistics, Lederle Graduate Research Tower, University of Massachusetts Amherst, 710 N. Pleasant Street, Amherst, MA 01003-9305, USA;
ogwilliam@umass.edu

Alastair Hamilton

Department of Mathematics and Statistics, Texas Tech University, Broadway & Boston, Lubbock, TX 79409-1042, USA; alastair.hamilton@ttu.edu

Mahmoud Zeinalian

Department of Mathematics, Lehman College, 250 Bedford Park Blvd W, Bronx, New York, NY 10468, USA; mahmoud.zeinalian@lehman.cuny.edu