

Jackiw-Teitelboim gravity with matter, generalized eigenstate thermalization hypothesis, and random matrices

Daniel Louis Jafferis¹, David K. Kolchmeyer^{1,2}, Baur Mukhametzhanov^{3,4}, and Julian Sonner⁵

¹*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*

²*Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

³*Institute for Advanced Study, Princeton, New Jersey 08540, USA*

⁴*Department of Physics, Cornell University, Ithaca, New York 14853, USA*

⁵*Department of Theoretical Physics, University of Geneva, Geneva, Switzerland*



(Received 3 January 2023; accepted 1 August 2023; published 20 September 2023)

We present evidence for a duality between Jackiw-Teitelboim gravity minimally coupled to a free massive scalar field and a single-trace two-matrix model. One matrix is the Hamiltonian H of a holographic disorder-averaged quantum mechanics, while the other matrix is the light operator \mathcal{O} dual to the bulk scalar field. The single-boundary observables of interest are thermal correlation functions of \mathcal{O} . We study the matching of the genus zero-, one- and two-boundary expectation values in the matrix model to the disk and cylinder Euclidean path integrals. The non-Gaussian statistics of the matrix elements of \mathcal{O} correspond to a generalization of the eigenstate thermalization hypothesis ansatz. We describe multiple ways to construct double-scaled matrix models that reproduce the gravitational disk correlators. One method involves imposing an operator equation obeyed by H and \mathcal{O} as a constraint on the two matrices. Separately, we design a model that reproduces certain double-scaled Sachdev-Ye-Kitaev correlators that may be scaled once more to obtain the disk correlators. We show that in any single-trace, two-matrix model, the genus zero two-boundary expectation value, with up to one \mathcal{O} insertion on each boundary, can be computed directly from all of the genus zero one-boundary correlators. Applied to the models of interest, we find that these cylinder observables depend on the details of the double-scaling limit. To the extent we have checked, it is possible to reproduce the gravitational double-trumpet, which is UV divergent, from a systematic classification of matrix model 't Hooft diagrams. The UV divergence indicates that the matrix integral saddle of interest is perturbatively unstable. A nonperturbative treatment of the matrix models discussed in this work is left for future investigations.

DOI: [10.1103/PhysRevD.108.066015](https://doi.org/10.1103/PhysRevD.108.066015)

I. INTRODUCTION

The eigenstate thermalization hypothesis (ETH) is an ansatz for the statistical properties of the matrix elements of simple operators between high-energy microstates in quantum chaotic and thermalizing systems [1,2]. In a given system that exhibits level repulsion of an exponentially dense spectrum, such matrix elements are expected to be pseudorandom, and an ensemble can be formed by averaging within narrow energy bands. Alternatively, the randomness may be understood as referring to an ensemble of systems sharing the same simple thermal correlation functions.

We will embed the ETH in a matrix model framework. This will make the ansatz more precise, extend its regime of validity, and clarify how to compute thermal higher-point functions. In the examples we will consider, the matrix model will reproduce correlation functions at time scales ranging from prethermalization to postscrambling. Matrix models of the type we define would not be expected to reproduce behavior at ultrashort times in theories with a free UV, or at exponentially long times that depend on the specific fine-grained microstates.

Perhaps surprisingly, the matrix models that describe thermal mean field theory are strongly non-Gaussian. Thus the standard Gaussian ETH ansatz does not agree with higher-point functions at leading order, even in systems for which the thermal correlators factorize into one and two-point functions. This observation resonates well with the non-Gaussian generalizations of the ETH ansatz put forward in [3] capable to accommodate nontrivial OTOCs [4–8]. Despite these complications, one of the

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

models we consider will turn out to be integrable in a certain sense.

A two-matrix model provides a unified framework for the ETH and the Wigner-type matrix model for the Hamiltonian H that defines the spectrum. In this paper, we will demonstrate this for the quantum mechanics dual to Jack-Teitelboim (JT) gravity with propagating matter. With \mathcal{O} the operator dual to a free massive scalar field in the bulk, we will define double-scaled two-matrix models for H and \mathcal{O} , generalizing the Saad-Shenker-Stanford (SSS) matrix model [9] and its deformations [10,11]. The genus expansion will now be decorated by \mathcal{O} graphs, which are associated to geodesic worldlines.

Unlike pure JT gravity, which is exactly solvable and finite, JT gravity with propagating matter can at best be regarded as an effective field theory due to its UV divergences. As was mentioned in Sec. 6.1 of [9] and Sec. 5 of [12], these divergences arise from long, thin wormholes. We find that it is possible to reproduce these divergences using matrix-model perturbation theory. Because a nonperturbative treatment of the matrix integral is not necessary to study the bulk effective field theory, we may refer to our matrix models as “effective matrix models.”

To be precise, we are interested in two-matrix models with a single-trace potential. These matrix integrals take the form

$$\int d\mathcal{O} dH e^{-\text{Tr } V(H, \mathcal{O})}, \quad (1.1)$$

where $V(H, \mathcal{O})$ is an arbitrary (and generically infinite) linear combination of words made with the letters \mathcal{O} and H . In contrast, the SSS matrix model is a single-trace model of a single matrix, H . Single-trace models (with any number of matrices) admit a genus expansion where each order in the expansion corresponds to a sum over ‘t Hooft double-line graphs with a given topology. Furthermore, each trace present in an expectation value corresponds to a boundary. In the SSS model, the leading-order result for the partition function of H is given by a sum over ‘t Hooft graphs with disk topology¹:

$$\langle \text{Tr } e^{-\beta H} \rangle_{\text{disk}} = \text{[Diagram: A circle containing a planar graph with multiple vertices and edges, representing a disk topology]} + \dots \quad (1.2)$$

¹ $\langle \cdot \rangle$ refers to the expectation value defined using the matrix integral. The “disk” subscript refers to the topology of the ‘t Hooft diagrams.

This matrix-model computation is equivalent to a disk gravitational path integral in pure JT gravity. In our two-matrix models, we can construct thermal n -point correlators using the matrices \mathcal{O} and H . To leading order in the genus expansion, the thermal two-point function in the matrix model is again given by a sum over planar diagrams:

$$\langle \text{Tr } e^{-\beta_1 H} \mathcal{O} e^{-\beta_2 H} \mathcal{O} \rangle_{\text{disk}} = \text{[Diagram: A circle containing a planar graph with two red double-line loops and black double-line segments]} + \dots \quad (1.3)$$

The red double-line is a propagator of the \mathcal{O} matrix, while the black double-line is a propagator of the H matrix. In JT gravity minimally coupled to a free massive scalar field,² the corresponding disk two-point function is computed using the extrapolate dictionary applied to a Euclidean black hole background.

The main technical objective of this paper is to explore the duality between single-trace, two-matrix models and JT gravity with matter. We are interested in correlators with additional \mathcal{O} insertions, additional boundaries, and additional handles. Our technical results may be summarized as follows:

- (i) In Sec. VII, we provide an algorithm for constructing a potential $V(H, \mathcal{O})$ for which the planar matrix model n -point functions [namely, the higher-point analogs of (1.3)] equal the disk n -point functions of \mathcal{O} in JT gravity. This establishes the duality at the level of the disk. Explicitly determining $V(H, \mathcal{O})$ is tedious in practice. The algorithm introduces a fictitious parameter ϵ and weights each of the various terms that contribute to a gravitational n -point correlation function by some power of ϵ , such that (roughly speaking) higher-point correlators are weighted by higher powers of ϵ . The potential $V(H, \mathcal{O})$ is organized as a series expansion in ϵ . The last step in the algorithm is to set $\epsilon \rightarrow 1$. We consider two specific schemes for ϵ -deforming the gravitational correlators, which we call the “Selberg regulator” and the “ q -deformed regulator.” We have not proven that the series representation of $V(H, \mathcal{O})$ is convergent, so strictly speaking the existence of these matrix models is conjectural.

²Going forward, we may refer to this theory as “JT gravity” or “JT gravity with matter” out of convenience. The theory without matter will always be called “pure JT gravity.”

- (ii) In Sec. VIII, we argue that the matrix model defined using the q -deformed regulator can be generalized into a three-parameter model (where ϵ is one of the parameters) for which the disk correlators equal the correlators that were studied in [13] in the double-scaled Sachdev-Ye-Kitaev (SYK) model. It was shown in [13] that these correlators can be scaled once more to obtain those of JT gravity with matter. To the extent that we checked, the three-parameter matrix model has a symmetry that may be predicted using the framework of [13]. In our calculations, the appearance of this symmetry is highly nontrivial and provides further support for the claim that the model is well-defined.
- (iii) In Sec. III, we provide another argument for the existence of two-matrix models that reproduce the desired gravitational disk correlators. This argument is more explicit about the form that $V(H, \mathcal{O})$ takes. We first point out that in any ensemble-averaged theory whose correlators are exactly given by the disk correlators of JT gravity with matter (such as the SYK model in the appropriate scaling regime), the operators \mathcal{O} and H obey an operator equation. In the semiclassical (or high-energy) limit,³ where the correlators become conformally invariant [with conformal group $PSL(2, \mathbb{R})$], this operator equation becomes the condition that the scaling dimensions of the primary operators appearing in the $\mathcal{O}\mathcal{O}$ operator product expansion (OPE) are in the set $\{1\} \cup \{2\Delta + 2n : n \in \mathbb{Z}_{\geq 0}\}$. Using the 1D conformal field theory (CFT) bootstrap, one may prove⁴ that this condition, together with associativity of the OPE and conformal invariance, guarantees that all of the n -point correlators of \mathcal{O} agree with those of a bosonic generalized free field (GFF). Although we are generally interested in JT gravity away from the semiclassical limit, this result nonetheless motivates us to construct a two-matrix ensemble by squaring our operator equation and adding it to the matrix potential with a large coefficient such that it is enforced as a constraint. Note that associativity of the OPE is guaranteed in our model because we are representing the operators \mathcal{O} and H using matrices, and matrix multiplication is associative. This model

is a single-trace matrix model and we conjecture that it correctly computes all of the disk correlators in JT gravity with matter. In Sec. VI, we support this conjecture by showing that a large class of Schwinger-Dyson equations in this model are solved by the gravitational n -point correlators. To verify the Schwinger-Dyson equations, we need to use certain integrability relations. One of these is only available in the double-scaling limit.

- (iv) Our next results concern the two-boundary, genus zero correlators in the matrix model. These are computed by summing over ‘t Hooft diagrams with cylinder topology. In Sec. IX, we show that in any single-trace, two-matrix model, the connected two-boundary correlators

$$\langle \text{Tr} e^{-\beta_L H} \text{Tr} e^{-\beta_R H} \rangle_{\text{cylinder}} \quad (1.4)$$

and

$$\langle \text{Tr} \mathcal{O} e^{-\beta_L H} \text{Tr} \mathcal{O} e^{-\beta_R H} \rangle_{\text{cylinder}} \quad (1.5)$$

may be determined directly from the disk correlators even if the matrix potential $V(H, \mathcal{O})$ is unknown. Roughly speaking, our strategy is to cut into pieces the ‘t Hooft diagrams that contribute to the disk amplitudes and reassemble the pieces to form diagrams with cylinder topology, which we systematically classify. Our result is reminiscent of recursion in one-matrix models, where higher-genus and higher-boundary correlators may be recursively computed from lower-genus and lower-boundary correlators.⁵ For a ‘t Hooft-scaled model (where the matrix potential is proportional to N , the number of eigenvalues of each matrix), our algorithm for computing the cylinder correlators yields unambiguous results. However, for the double-scaled matrix models that are dual to JT gravity at the level of the disk, our algorithm only returns an unambiguous answer once the scheme for taking the double-scaling limit is specified. We obtain results using the aforementioned Selberg and q -deformed regulators. As explained in the first point above, these regulators are part of the definition of the matrix models.

Using the Selberg regulator, we match the matrix-model cylinder computations to their analogous double-trumpet calculations in JT gravity (either the empty double-trumpet or a double-trumpet with

³By “semiclassical limit,” we are always referring to the $G_N \rightarrow 0$ limit. Equivalently, this is the limit in which the coefficient of the Schwarzian action, which we may call ϕ_b , goes to infinity. By dimensional analysis, ϕ_b has units of length, and is the only scale in the Schwarzian theory. Because all inverse temperatures are naturally measured in units of ϕ_b , the semiclassical limit may be thought of as a high-energy limit. Furthermore, in this limit, the wiggly AdS boundaries become rigid, so the conformal invariance of boundary correlation functions is restored.

⁴See Appendix D for the proof. See [14,15] for further 1D CFT bootstrap results.

⁵We are hopeful that there exists a general formulation of two-matrix model recursion, but finding it is beyond the scope of this work.

one \mathcal{O} inserted on each boundary). In JT gravity, these quantities receive contributions from the partition function of the scalar field on the rigid double-trumpet, which is given by

$$Z_{\text{scalar}}(b) = \sum_{n=0}^{\infty} \frac{e^{-n\Delta b}}{(1 - e^{-b})(1 - e^{-2b}) \dots (1 - e^{-nb})} \quad (1.6)$$

$$= \frac{1}{(e^{-\Delta b}; e^{-b})_{\infty}}, \quad (1.7)$$

where b is the length of the closed geodesic that wraps the double-trumpet and Δ encodes the mass of the scalar. We have explicitly matched the contributions from the first four terms above to the first four classes of ‘t Hooft diagrams that appear in our infinite classification of cylindrical ‘t Hooft diagrams. We conjecture that the sum over all the ‘t Hooft diagrams reproduces the full double-trumpet result. We further conjecture that our method for computing the sum over cylinder diagrams can be extended to compute the sum over diagrams with any topology, and we expect that the result will match the gravitational path integral with the same topology (including the matter-determinant factor).

Using the q -deformed regulator, the results for the cylinder correlators are the same as in the Selberg model, except the partition function is replaced by

$$Z(b) = \sum_{n=0}^{\infty} \frac{e^{-n\Delta b}}{(1 - e^{-b})^n} \quad (1.8)$$

$$= \frac{1 - e^{-b}}{1 - e^{-b} - e^{-\Delta b}}. \quad (1.9)$$

Curiously, this partition function has a Hagedorn temperature because the denominator goes to zero for sufficiently small b . Finding an independently defined bulk theory that reproduces this result is an interesting question which is beyond the scope of this paper.

- (v) Note that the gravitational duals of (1.4) and (1.5) are ill-defined because the gravitational path integral includes an integral over b , and this integral is either nonconvergent (in the Selberg model) or simply ill-defined (in the q -deformed model) due to the small- b behavior of the above partition functions. As mentioned above, we can still reproduce $Z_{\text{scalar}}(b)$ for all b using the Selberg matrix model by writing the result as an infinite series of terms that are individually well-defined. In Sec. IX D, we explain

that our two-matrix models are necessarily non-perturbatively unstable. To illustrate this point, we consider a single-matrix multitrace matrix model for H [which could arise by integrating out \mathcal{O} in (1.1)]. For simplicity, we consider a ‘t Hooft scaled matrix model, where the n -trace⁶ part of the potential is weighted by N^{2-n} . We review that the partition function $\langle \text{Tr} e^{-\beta H} \rangle$ to leading order in N is determined by the saddle point of the integral over the eigenvalues of H . We then show that the connected two-boundary correlator $\langle \text{Tr} e^{-\beta_L H} \text{Tr} e^{-\beta_R H} \rangle_c$ to leading order in N is directly determined by the Hessian of the matrix potential evaluated at the saddle. The bad behavior of the double-trumpet is directly linked to a perturbative instability of the saddle in the eigenvalue integral.

- (vi) Our final main result pertains only to JT gravity with matter. It was shown in [16,17] that the disk n -point correlators may be conveniently computed using a set of Feynman rules. A graph drawn on a disk represents a set of Wick contractions of the external operators. We show in Sec. II that these Feynman rules may be naturally extended to other topologies. In particular, we draw graphs on genus zero surfaces with two or three boundaries and show that the naively extended Feynman rules yield sensible results. This result supports many of the other results in this paper but may also be of independent interest to JT gravity experts.

We now outline the structure of the remainder of this paper. In Sec. II, we review JT gravity with matter with an emphasis on the disk and double-trumpet calculations that will be compared against the matrix models. We also conjecture how the gravitational Feynman rules which were originally developed for disk computations can be extended to arbitrary topologies. In Sec. III, we briefly step away from JT gravity to discuss the ETH and its relevance to holography in full generality. We then return to JT gravity to discuss the operator equation relating \mathcal{O} and H that is used to construct a constrained matrix ensemble. In Sec. IV, we introduce a toy model where $V(H, \mathcal{O})$ is Gaussian in \mathcal{O} . This model succeeds in computing the holographic two-point function but fails to correctly compute higher-point functions. The purpose of introducing this model is to make the reader comfortable with two-matrix models, and also to introduce notations and ideas that will be useful later in the paper. Section V summarizes the most important points of the remainder of the paper while avoiding most of the technical details. In particular, we outline the definitions of the Selberg and q -deformed matrix models, and we review the basic strategy of our double-trumpet

⁶“1-trace” corresponds to single-trace, “2-trace” corresponds to double-trace, etc.

computations. The reader who has no time can skip to the discussion section after Sec. V. In Sec. VI, we show how in the constrained matrix ensemble, a large class of Schwinger-Dyson equations is solved by the disk correlators of JT gravity with matter. In Sec. VII, we explain in more detail our algorithm that allows one to start from the answers for the holographic disk n -point functions and work backwards to determine the potential $V(H, \mathcal{O})$. In Sec. VIII we carefully define the Selberg and q -deformed matrix models and explain the connection between the q -deformed model and the double-scaled SYK model, using the results of [13]. In Sec. IX and Appendix F, we discuss our computation of (1.4) and (1.5) in full detail. We compare the matrix model results against their gravitational counterparts. We also discuss the perturbative instability in the matrix model that is dual to the UV divergence in the double-trumpet. In the discussion section we speculate on how our results may be generalized to higher dimensions, and we compare our results with other works that used the ETH to study JT gravity with matter. The appendixes mostly contain derivations of various special function identities as well as other technical calculations.

II. JACK-TEITELBOIM GRAVITY WITH MATTER

In this section we review JT gravity [9,18] minimally coupled to a free massive scalar. Correlation functions on the disk are determined by a set of Feynman rules [16]. Using the boundary particle formulation of JT gravity [17,19,20], we also compute the two-point function on the double-trumpet, including the matter 1-loop determinant. The result is again described by the Feynman rules of [16], suggesting that they generalize to correlation functions on all genus g Riemann surfaces with n boundaries $\mathcal{M}_{g,n}$.

We consider JT gravity coupled to a scalar with the Euclidean action

$$I[g, \phi, \varphi] = -S_0\chi + I_{JT}[g, \phi] + I_m[g, \varphi], \quad (2.1)$$

$$I_{JT}[g, \phi] = -\frac{1}{2} \int_{\mathcal{M}} \sqrt{g} \phi (R+2) - \int_{\partial\mathcal{M}} \sqrt{h} \phi (K-1), \quad (2.2)$$

$$I_m[g, \varphi] = \frac{1}{2} \int_{\mathcal{M}} \sqrt{g} (g^{ab} \partial_a \varphi \partial_b \varphi + m^2 \varphi^2), \quad (2.3)$$

where χ is the Euler characteristic of the two-dimensional manifold, S_0 is the entropy of an extremal black hole, ϕ is the dilaton and φ is a scalar field with mass m . To compute the path integral on $\mathcal{M}_{g,n}$, interpreted as $\langle Z(\beta_1) \dots Z(\beta_n) \rangle_{g,n}$ in the dual matrix model [9], we fix the regulated boundary lengths of n asymptotically anti-de Sitter (AdS) regions to be $\beta_1/\epsilon, \dots, \beta_n/\epsilon$ and the dilaton $\phi|_{\partial\mathcal{M}} = \gamma/\epsilon$. We follow the

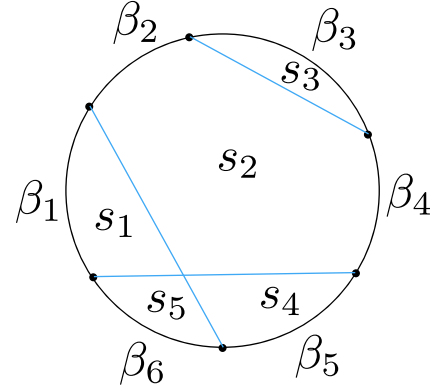


FIG. 1. An example of a Feynman diagram that contributes to the thermal six-point \mathcal{O} correlator. Each blue line is a Wick-contraction through the bulk of two boundary operators with scaling dimension Δ . The correlator is a sum over all ways of contracting the operators. Each disk-shaped region is labeled by an s parameter that is integrated in the range $s \in (0, \infty)$ to obtain the value of the correlator. The Euclidean-time separation between the external operators is indicated by the β_1, \dots, β_6 parameters. We adopt a convention where a blue bulk line always has scaling dimension Δ .

conventions of [9] and set $\gamma = 1/2$.⁷ For the scalar field φ we take either Dirichlet or Neumann boundary conditions, such that the dual boundary operator \mathcal{O} has the scaling dimension

$$\Delta = \frac{1}{2} \pm \sqrt{\frac{1}{4} + m^2}. \quad (2.4)$$

The choice of sign depends on which of the two standard boundary conditions we choose for φ . Given any $\Delta > 0$, there is a unique choice of the sign and $m^2 > -\frac{1}{4}$ such that (2.4) holds.

A. Correlation functions on the disk

Disk correlators of \mathcal{O} have been computed in [16,17,19–21], which we now review. In the absence of gravity, the correlators of \mathcal{O} are those of a GFF. The two-point function at zero temperature is $\langle \mathcal{O}(\tau) \mathcal{O}(0) \rangle_{\text{GFF}} = \tau^{-2\Delta}$ and higher correlators are computed by Wick contractions.

To include the gravitational corrections, one must reparametrize the GFF correlators and then integrate over all reparametrizations with the Schwarzian action,

$$\begin{aligned} & \langle \text{Tr} e^{-\beta H} \mathcal{O}(\tau_1) \dots \mathcal{O}(\tau_n) \rangle_{\text{disk}} \\ &= e^{S_0} \int \frac{\mathcal{D}f}{SL(2, \mathbb{R})} (f'(\tau_1))^\Delta \dots (f'(\tau_n))^\Delta \\ & \times \langle \mathcal{O}(f(\tau_1)) \dots \mathcal{O}(f(\tau_n)) \rangle_{\text{GFF}} e^{-I_{\text{Sch}}[f]}, \end{aligned} \quad (2.5)$$

⁷We use a different convention for the overall normalization of the density of states, see below.

where $f(\tau)$ is a reparametrization field that determines the embedding of the boundary curve into the hyperbolic disk. The notation on the left-hand side of (2.5) indicates that the gravitational correlator is not normalized by the partition function $\langle \text{Tre}^{-\beta H} \rangle$.

The disk correlator of an arbitrary number of \mathcal{O} operators can be computed exactly, and the result resums all the perturbative gravitational corrections. In particular, [16] devised an elegant set of Feynman rules that assigns a

finite-dimensional integral to each Wick contraction of the \mathcal{O} operators.⁸ For example, a Feynman diagram that represents a particular six-point Wick-contraction is depicted in Fig. 1.

The value of each Feynman diagram is set by the following rules:

- (i) For each boundary segment of length β , we include a factor of $e^{-\beta s^2}$, where s labels the disk-shaped region that is adjacent to the boundary segment.

- (ii) For each \mathcal{O} insertion, we include a factor of

$$\Delta \text{---} \begin{array}{c} s_1 \\ \text{---} \text{---} \text{---} \\ s_2 \end{array} = \left(\frac{\Gamma(\Delta \pm i s_1 \pm i s_2)}{\Gamma(2\Delta)} \right)^{1/2} \equiv (\Gamma_{12}^\Delta)^{1/2}, \quad (2.6)$$

where s_1 and s_2 are associated to the two regions adjacent to the operator, and \pm means that we take a product of gamma functions for all four choices of the signs.⁹ Here, the normalization of operators is chosen to be such that at short distances the two-point function is $\langle \text{Tre}^{-\beta H} \mathcal{O}(\tau) \mathcal{O}(0) \rangle_{\text{disk}} \approx \tau^{-2\Delta} \langle \text{Tre}^{-\beta H} \rangle_{\text{disk}}$, $\tau \rightarrow 0$.

- (iii) For each crossing of two lines, we include a factor of

$$\begin{array}{ccc} \Delta_1 & & \Delta_2 \\ & \searrow & \nearrow \\ & s_2 & \\ & \nearrow & \searrow \\ s_1 & & s_3 \\ & \searrow & \nearrow \\ & s_4 & \end{array} = \left\{ \begin{array}{ccc} \Delta_1 & s_1 & s_2 \\ \Delta_2 & s_3 & s_4 \end{array} \right\}, \quad (2.7)$$

which is the $6j$ -symbol of the $\mathfrak{sl}(2, \mathbb{R})$ algebra.¹⁰ The parameters Δ_1 and Δ_2 represent the scaling dimensions associated to the two crossing lines. The parameters s_1, \dots, s_4 represent the four disk-shaped regions that surround the crossing. In the theory with one scalar operator we set $\Delta_1 = \Delta_2 = \Delta$. More generally, if we had two different free scalar fields in the bulk, we would use (2.7). The symmetries of the $6j$ -symbol are the symmetries of its graphical representation, e.g. reflections across Δ_1 or Δ_2 lines and reflections across the horizontal or vertical axes. The definition of the $6j$ -symbol and its properties are reviewed in Appendix A.

- (iv) After including all of the appropriate factors as specified above, we integrate over each s parameter

with the Schwarzian density of states $\int_0^\infty ds \rho(s)$, where¹¹

$$\rho(s) = \frac{1}{2\pi\Gamma(\pm 2is)} = \frac{s}{\pi^2} \sinh(2\pi s). \quad (2.8)$$

Sometimes we use the energy basis $E = s^2$ and the corresponding density of states

$$\rho_0(E) = \frac{1}{2\pi^2} \sinh(2\pi\sqrt{E}) \quad (2.9)$$

such that $\rho(s)ds = \rho_0(E)dE$.

1. Examples

To demonstrate how the rules described above work in practice and also for later use, we explicitly write down a few correlation functions.

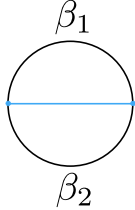
⁸An alternative derivation of these rules was given in [17], which is based on a formalism developed in [19,20].

⁹That is $\Gamma(\Delta \pm i s_1 \pm i s_2) = \Gamma(\Delta + i s_1 + i s_2) \Gamma(\Delta + i s_1 - i s_2) \Gamma(\Delta - i s_1 + i s_2) \Gamma(\Delta - i s_1 - i s_2)$.

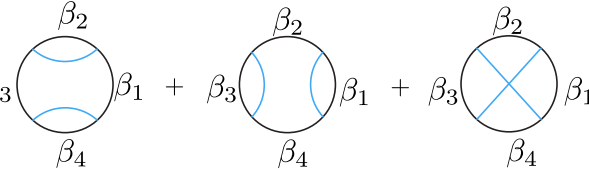
¹⁰ $6j$ symbols also appear in the computation of AdS amplitudes in general dimensions, as was shown in [22].

¹¹Our convention for the Schwarzian density of states is related to [9] by $\rho(s) = 2\rho_{\text{there}}(s)$. Equivalently, we use conventions of [9] with a rescaling $e^{S_0} \rightarrow 2e^{S_0}$. We find this normalization more convenient for computations involving $\mathfrak{sl}(2, \mathbb{R})$ $6j$ -symbols.

We start with the two-point function

$$\langle \text{Tr } e^{-\beta H} \mathcal{O}(\tau) \mathcal{O}(0) \rangle_{\text{disk}} = \text{Diagram} = e^{S_0} \int_0^\infty ds_1 ds_2 \rho(s_1) \rho(s_2) e^{-(\beta_1 s_1^2 + \beta_2 s_2^2)} \frac{\Gamma(\Delta \pm i s_1 \pm i s_2)}{\Gamma(2\Delta)}, \quad (2.10)$$


where $\beta_1 = \tau, \beta_2 = \beta - \tau$. The four-point function has three terms corresponding to three different GFF Wick contractions

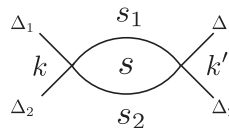
$$\langle \text{Tr } e^{-\beta H} \mathcal{O}(\tau_1) \dots \mathcal{O}(\tau_4) \rangle_{\text{disk}} = \text{Diagram 1} + \text{Diagram 2} + \text{Diagram 3} \quad (2.11)$$


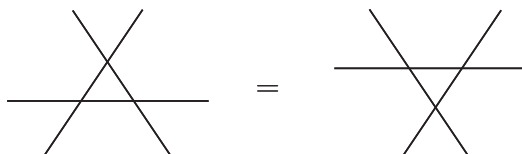
$$= e^{S_0} \int_0^\infty \prod_{i=1}^4 (ds_i \rho(s_i) e^{-\beta_i s_i^2}) (\Gamma_{12}^\Delta \Gamma_{23}^\Delta \Gamma_{34}^\Delta \Gamma_{41}^\Delta)^{1/2} \quad (2.12)$$

$$\times \left(\frac{\delta(s_1 - s_3)}{\rho(s_1)} + \frac{\delta(s_2 - s_4)}{\rho(s_2)} + \left\{ \begin{matrix} \Delta & s_1 & s_2 \\ \Delta & s_3 & s_4 \end{matrix} \right\} \right), \quad (2.13)$$

where β_i denote the time differences between operators. Assuming $0 < \tau_4 < \dots < \tau_1 < \beta$, we defined $\beta_1 = \beta - (\tau_1 - \tau_4), \beta_2 = \tau_1 - \tau_2, \beta_3 = \tau_2 - \tau_3, \beta_4 = \tau_3 - \tau_4$. For convenience, in the first two terms we introduced two energies for the same region and included corresponding delta functions to remove one of the energy integrals. In Appendix B we check that the relative coefficient between the first two terms and the third term is indeed as given in (2.13). Later, we will also need diagrams contributing to the six-point function and we will discuss them in due course.

For some Feynman diagrams we have a choice of where to put the intersections of the bulk lines. Once we fix a Wick-contraction, the result is independent of where we put the intersections. This is guaranteed by the orthogonality and Yang-Baxter equations. We express them pictorially as

$$\text{Diagram 1} = \frac{s_1}{s_2} \frac{k}{k'}, \quad (2.14)$$


$$\text{Diagram 2} = \text{Diagram 3}, \quad (2.15)$$


respectively. The expression for the orthogonality relation is

$$\int_0^\infty ds \rho(s) \left\{ \begin{matrix} \Delta_1 & s_1 & k \\ \Delta_2 & s_2 & s \end{matrix} \right\} \left\{ \begin{matrix} \Delta_1 & s_1 & k' \\ \Delta_2 & s_2 & s \end{matrix} \right\} = \frac{\delta(k - k')}{\rho(k)}. \quad (2.16)$$

Similarly, one can write the Yang-Baxter equation from (2.15). If we impose the condition that the bulk lines should not have any voluntary crossings, then only the Yang-Baxter equation is needed to ensure that the Feynman rules are unambiguous. More identities satisfied by the $6j$ -symbol are described in Appendix A. In Sec. VIII we will encounter a q -deformed set of Feynman rules where the $6j$ -symbol obeys the Yang-Baxter equation, but not the orthogonality equation.

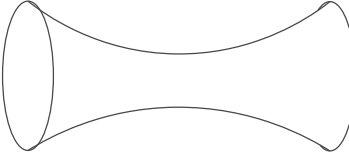
B. Correlation functions on the double-trumpet

We are also interested in correlation functions on the double-trumpet. Here, we discuss the two simplest cases: the double-trumpet with no \mathcal{O} insertions, and the two-point function on the double-trumpet, with one \mathcal{O} on each boundary. In both cases, we include the matter 1-loop determinant.

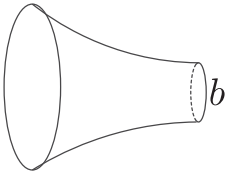
1. Double-trumpet with 1-loop determinant

The path integral on the double-trumpet without \mathcal{O} insertions is

$$\langle \text{Tr} e^{-\beta_L H} \text{Tr} e^{-\beta_R H} \rangle_{\text{cyl}} = \int_0^\infty db \, b \, Z_{\text{tr}}(\beta_L, b) Z_{\text{tr}}(\beta_R, b) Z_{\text{scalar}}(b) \quad (2.17)$$

$$= \beta_L \text{ (diagram) } \beta_R, \quad (2.18)$$


where cyl indicates the cylinder topology. The partition function of the scalar field on the double-trumpet is Z_{scalar} , which we explicitly define in (2.23). If we omit the Z_{scalar} term, we obtain the answer computed in [9] for pure JT gravity. It is given as an integral over the moduli space of hyperbolic manifolds with two boundaries, parametrized by the length of the closed geodesic b . The “trumpet” partition function $Z_{\text{tr}}(\beta, b)$ represents the integral over the extrinsic curvatures of a “wiggly” AdS boundary and is given by [9]

$$Z_{\text{tr}}(\beta, b) = \beta \text{ (diagram) } b \quad (2.19)$$


$$= \int_0^\infty ds \frac{\cos(bs)}{\pi} e^{-\beta s^2} \quad (2.20)$$

$$= \frac{1}{2\sqrt{\pi}} \beta^{-1/2} e^{-\frac{b^2}{4\beta}}. \quad (2.21)$$

To compute Z_{scalar} , we may canonically quantize the scalar field on the analytic continuation of the double-trumpet to Lorentzian signature with the metric

$$ds^2 = d\rho^2 - \cosh^2 \rho dt^2, \quad (2.22)$$

where $\rho \in (-\infty, \infty)$ is the spatial coordinate, and t is time. Each positive-frequency mode is labeled by $n \in \mathbb{Z}_{\geq 0}$ and has energy $\Delta + n$. The partition function, computed with periodicity b in the imaginary time direction, is thus

$$Z_{\text{scalar}}(b) = \prod_{n=0}^{\infty} \frac{1}{1 - e^{-b(\Delta+n)}} \quad (2.23)$$

$$= \sum_{n=0}^{\infty} \frac{e^{-n\Delta b}}{(1 - e^{-b})(1 - e^{-2b}) \dots (1 - e^{-nb})} \quad (2.24)$$

$$= \exp \left(\sum_{w=1}^{\infty} \frac{e^{-wb\Delta}}{w(1 - e^{-wb})} \right), \quad (2.25)$$

where we provided three equivalent expressions with different interpretations. The first formula makes the relation with the single particle spectrum manifest. In the second formula (2.24), we can separate out the counting of conformal primaries and descendants. For example, the $n = 1$ term

$$\frac{e^{-\Delta b}}{1 - e^{-b}} = \sum_{k=0}^{\infty} e^{-(\Delta+k)b}$$

counts the primary state Δ together with its descendants. The $n = 2$ term

$$\frac{e^{-2\Delta b}}{(1 - e^{-b})(1 - e^{-2b})} = \sum_{m=0}^{\infty} \frac{e^{-(2\Delta+2m)b}}{1 - e^{-b}} \quad (2.26)$$

counts the double-trace operators (states) $2\Delta + 2m$, where the factor

$$\frac{1}{1 - e^{-b}} = \sum_{k=0}^{\infty} e^{-kb}$$

again accounts for the descendants. The $n = 3$ term counts the primary states with dimensions $3\Delta + 2m_1 + 3m_2$ (for $m_1, m_2 \in \mathbb{Z}_{\geq 0}$) and their descendants. The pattern continues for higher-trace operators. Finally, the third expression (2.25) connects with the Selberg trace formula which we discuss next.¹²

An alternative way to compute the 1-loop determinant, which is simpler to generalize to an arbitrary hyperbolic manifold, is to use the Selberg trace formula. It relates the spectrum of the Laplace operator on a hyperbolic manifold \mathcal{M} to the spectrum of closed geodesics (e.g. see [23])

$$\text{tr } e^{\frac{\tau}{2} \nabla^2} = f(\tau) A + \frac{e^{-\tau/8}}{\sqrt{2\pi\tau}} \sum_{\gamma} \sum_{w=1}^{\infty} \frac{b_{\gamma}}{\sinh \frac{wb_{\gamma}}{2}} e^{-\frac{(wb_{\gamma})^2}{2\tau}}. \quad (2.27)$$

In the rhs the sum is over all closed geodesics γ and b_{γ} is the length of this geodesic.¹³ The sum over w can be thought of as a sum over multiple windings of the “primitive” (traversed once) geodesic γ . In the first term in the rhs A is the area of the manifold and $f(\tau)$ does not depend on the particular manifold and will not play a role in our discussion.¹⁴

To compute the determinant we use the standard trick to relate it to the heat kernel

¹²To relate (2.25) to (2.23) we first expand $\frac{1}{1 - e^{-wb}} = \sum_{n=0}^{\infty} e^{-nwb}$ and then sum over w .

¹³Here, our convention is that we sum over unoriented primitive geodesics.

¹⁴ $f(\tau) = \frac{e^{-\tau/8}}{(2\pi\tau)^{3/2}} \int_0^\infty dr \frac{r e^{-r^2/2\tau}}{\sinh^2 \frac{r}{2}}.$

$$Z_{\text{scalar}}(\mathcal{M}) = \det(-\nabla^2 + m^2)^{-1/2} \quad (2.28)$$

$$= \exp \frac{-1}{2} \text{tr} \log(-\nabla^2 + m^2) \quad (2.29)$$

$$= \exp \frac{1}{2} \text{tr} \int_{\epsilon}^{\infty} \frac{d\tau}{\tau} e^{-\frac{\tau}{2}(-\nabla^2 + m^2)} \quad (2.30)$$

up to a constant that diverges as we take $\epsilon \rightarrow 0$. After inserting the Selberg trace formula (2.27) and integrating over τ we have

$$Z_{\text{scalar}}(\mathcal{M}) = \exp \left(\#A + \frac{1}{2} \sum_{\gamma} \sum_{w=1}^{\infty} \frac{e^{-wb_{\gamma}(\Delta - \frac{1}{2})}}{w \sinh \frac{wb_{\gamma}}{2}} \right). \quad (2.31)$$

Strictly speaking, the Selberg trace formula applies to compact hyperbolic manifolds. On the other hand, we are interested in manifolds with infinite area. To deal with this we use Gauss-Bonnet theorem to write $A = -2\pi\chi + \int_{\partial M} K$ and absorb this term into the renormalization of S_0 and γ .

On the double-trumpet there is a single primitive geodesic and (2.31) agrees with (2.25). While on the disk there are no closed geodesics and the 1-loop determinant contributes only the area term that renormalizes S_0, γ .

In the presence of matter the integral (2.17) has a UV divergence at small¹⁵ b [9]. This can be seen by expanding the Pochhammer symbol $(e^{-\Delta b}, e^{-b})_{\infty} = \prod_{n=0}^{\infty} (1 - e^{-b(\Delta+n)})$ in (2.23) at small b . Alternatively we can note that in the UV limit $b \rightarrow 0$ the mass is not important and we can approximate by a massless free scalar on a strip with the partition function dominated by the vacuum $\sim e^{\frac{14\pi^2}{24b}} = e^{\frac{\pi^2}{6b}}$.

We consider this divergence as a reflection of the fact that JT with matter should be considered as a low-energy effective field theory. We will match this divergence in the matrix model. In other words, we consider correlation functions at fixed b and match them with a computation in the matrix model.

2. Two-point correlator

The next observable we consider is the double-trumpet with one \mathcal{O} inserted on each boundary

$$\langle \text{Tr} e^{-\beta_L H} \mathcal{O} \text{Tr} e^{-\beta_R H} \mathcal{O} \rangle_{\text{cyl}} = \text{Diagram} \quad (2.32)$$

To compute this quantity, we first compute the two-point function on the rigid double-trumpet as a function of b , integrate over boundary reparametrizations using the Schwarzian action, multiply by $Z_{\text{scalar}}(b)$, and finally integrate over b . The result may be expressed as an infinite sum (2.36), where each term in the sum captures the contribution to $Z_{\text{scalar}}(b)$ from a single term in (2.24). We use the boundary

particle formalism [19] to do this computation in Appendix E. The upshot is that the Feynman rules of [16] discussed in Sec. II A still apply, but they are generalized to allow for closed loops of a bulk line. The $6j$ -symbol governs the intersection of all bulk lines, regardless of whether they form closed loops or extend to the boundaries. For the two-point correlator on the double-trumpet we have

$$\langle \text{Tr} e^{-\beta_L H} \mathcal{O} \text{Tr} e^{-\beta_R H} \mathcal{O} \rangle_{\text{cyl}} \quad (2.33)$$

$$= \int_0^{\infty} ds_L ds_R \rho(s_L) \rho(s_R) e^{-(\beta_L s_L^2 + \beta_R s_R^2)} (\Gamma_{LL}^{\Delta} \Gamma_{RR}^{\Delta})^{1/2} \quad (2.34)$$

$$\times \left(\frac{\delta(s_L - s_R)}{\rho(s_L)} + \left\{ \begin{matrix} \Delta & s_L & s_R \\ \Delta & s_R & s_L \end{matrix} \right\} + \sum_{m=0}^{\infty} \left\{ \begin{matrix} 2\Delta + 2m & s_L & s_R \\ \Delta & s_R & s_L \end{matrix} \right\} + \dots \right) \quad (2.35)$$

$$= \text{Diagram 1} + \text{Diagram 2} + \sum_{m=0}^{\infty} \text{Diagram 3} + \dots \quad (2.36)$$

¹⁵This is analogous to the “tachyon” divergence in string theory.

Here, the gamma functions are as prescribed by the rule (2.6) with $\Gamma_{LL}^\Delta = \frac{\Gamma(\Delta)^2}{\Gamma(2\Delta)} \Gamma(\Delta \pm 2is_L)$ and similarly for s_R . The three explicitly shown terms in (2.36) correspond to the $n = 0, 1, 2$ terms in (2.24). In the first term, the identity operator (or ground state) propagates on the closed geodesic and the energies s_L, s_R are the same. In the second term, a primary operator \mathcal{O} and its descendants propagate around the closed geodesic. The closed geodesic is assigned a label equal to the scaling dimension of \mathcal{O} (which is Δ). The intersection of the closed geodesic with the geodesic connecting the AdS boundaries implies that we include a $6j$ -symbol. The third term corresponds to the double-trace operators propagating around the closed geodesic. They have scaling dimensions $2\Delta + 2m$ for $m \in \mathbb{Z}_{\geq 0}$. The general rule is that for each primary operator (aside from the identity, which is treated as a special case), we assign its scaling dimension to the closed geodesic line and include a $6j$ symbol for the intersection. We then sum over the terms we get for all the primary operators. The second term above captures the lone single-trace operator, while the third term captures the double-trace operators. In particular, the sum over m in the third term directly corresponds to the sum over m in (2.26). Similarly, we can include triple-trace and higher operators, which we denoted by ellipsis in (2.36).

C. Pair of pants and beyond

In the previous subsection we saw that the gravitational Feynman rules that compute disk correlators may be extended to compute the two-point function on the double-trumpet. Here, we compute an amplitude involving the pair of pants and find evidence that the Feynman rules may be extended even further.

We are interested in the pair of pants geometry with three AdS boundaries and a geodesic that winds in a figure-8 pattern contributing to the 1-loop determinant, see Fig. 2. In the previous examples we considered, the bulk lines divided the geometry into disk-shaped regions. Here, the bulk lines divide the geometry into three cylinders. We may extend the gravitational Feynman rules by declaring that for each cylinder we define two parameters s, s' , one near each boundary of the cylinder. And include a factor of $\langle \rho(s)\rho(s') \rangle_{\text{cyl}}$ in the integrand, the zero genus two-boundary density-density correlator in the SSS model [9]. The self-intersection of the bulk line corresponds to the $6j$ symbol as usual. Using these extended Feynman rules, the full amplitude for Fig. 2 is

$$e^{-S_0} \int_0^\infty \prod_{j=1}^3 (ds_j \langle Z(\beta_j) \rho(s_j) \rangle_{\text{cyl}}) \left\{ \begin{matrix} \Delta & s_1 & s_2 \\ \Delta & s_3 & s_2 \end{matrix} \right\}. \quad (2.37)$$

To show this is correct we relate it to the contribution of the closed geodesic to the 1-loop determinant (2.31). This is done using an integral representation of the $6j$ -symbol (A48) that we derive in Appendix A 5

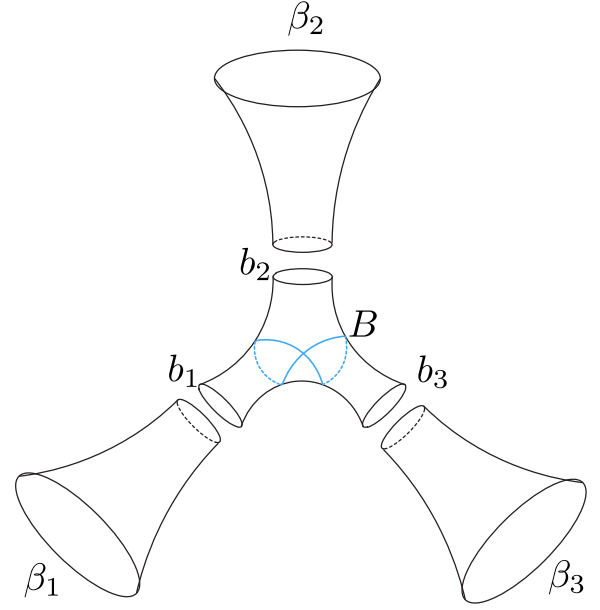


FIG. 2. The blue geodesic divides the pair of pants into three regions with cylinder topology. The length of the blue geodesic B is given in (2.39).

$$\left\{ \begin{matrix} \Delta & s_1 & s_2 \\ \Delta & s_3 & s_2 \end{matrix} \right\} = \frac{1}{2} \int_0^\infty \prod_{j=1}^3 \left(db_j b_j \frac{2 \cos(b_j s_j)}{b_j} \right) \frac{e^{-\Delta B}}{1 - e^{-B}}, \quad (2.38)$$

where $B = B(b_1, b_2, b_3)$ is determined by

$$\cosh \frac{B}{2} = \cosh \frac{b_2}{2} + 2 \cosh \frac{b_1}{2} \cosh \frac{b_3}{2}. \quad (2.39)$$

It turns out that B is also the length of the blue geodesic in Fig. 2 on the pair of pants with geodesic boundaries of lengths b_1, b_2, b_3 . Inserting this into (2.37) and exchanging integrals over b_j and s_j we have

$$\frac{1}{2} e^{-S_0} \int_0^\infty \prod_{j=1}^3 (db_j b_j \langle Z(\beta_j) x(b_j) \rangle_{\text{cyl}}) \frac{e^{-\Delta B}}{1 - e^{-B}}, \quad (2.40)$$

where we defined

$$x(b) = \frac{2}{b} \text{Tr} \cos(b\sqrt{H}). \quad (2.41)$$

Each factor $\langle Z(\beta) x(b) \rangle_{\text{cyl}} = Z_{\text{tr}}(\beta, b)$ is the trumpet partition function. We discuss this fact in more detail in Sec. IV B. Without the factor $\frac{e^{-\Delta B}}{1 - e^{-B}}$, this is the path integral on the pair of pants $\langle Z(\beta_1) Z(\beta_2) Z(\beta_3) \rangle_{g=0, n=3}$ computed in [9].¹⁶ Then the factor $\frac{e^{-\Delta B}}{1 - e^{-B}}$ corresponds to the $w = 1$

¹⁶We have an extra $\frac{1}{2}$ because our conventions for the density of states correspond to [9] with a rescaling $e^{S_0} \rightarrow 2e^{S_0}$, also see (2.8).

contribution to the determinant (2.31) on the pair of pants. We therefore showed that (2.37) is indeed the contribution of the “figure-8” geodesic in Fig. 2. Finally, it is easy to generalize the above computation to any multitrace operator propagating on the figure-8 geodesic. We simply substitute Δ by the dimension of the desired the multitrace operator.

It is natural to conjecture that the gravitational Feynman rules can be extended to compute amplitudes for arbitrary genus topologies with an arbitrary number of boundaries and \mathcal{O} insertions. A general diagram consists of bulk lines drawn on a Riemann surface with boundaries. The bulk lines can end on AdS boundaries where \mathcal{O} operators are inserted, or the bulk lines can form closed geodesics, which may represent contributions to the 1-loop determinant. These lines divide the surface into subregions, and each subregion is characterized by the number of boundaries n and the genus g . For every boundary of every subregion, we assign an s parameter. For each subregion, we get a factor of $\langle \rho(s_1) \dots \rho(s_n) \rangle_{g,n}$, which is the inverse Laplace transform of $Z_{g,n}(\beta_1, \dots, \beta_n)$ defined in Eq. (127) of [9]. Wherever two bulk lines intersect, we get a $6j$ symbol involving the four adjacent s parameters. Furthermore, wherever there is a closed geodesic that does not intersect other geodesics (including itself), we get a function of the two adjacent s parameters:

$$\int_0^\infty db \, b \frac{e^{-\Delta' b}}{1 - e^{-b}} \frac{2}{b} \cos(bs_1) \frac{2}{b} \cos(bs_2), \quad (2.42)$$

where Δ' is the dimension of the primary operator propagating on the closed geodesic. This ensures that the closed geodesic is weighted by $\frac{e^{-\Delta' b}}{1 - e^{-b}}$ in the moduli space integral. After integrating over all of the s parameters, the result is equal to the sum/integral over all the gravitational configurations (including the metric and the choice of non-homotopic geodesics) that have the topology of the diagram.

Another check of this conjecture comes from the two-point function on the disk with a handle where the geodesic connecting the two \mathcal{O} insertions is nonself-intersecting. This was computed in [24–26].

D. Gravitationally dressed OPE

The GFF correlators have $SL(2, \mathbb{R})$ conformal symmetry and can be decomposed into conformal blocks. For example, the four-point function is

$$\langle \mathcal{O}(\tau_1) \mathcal{O}(\tau_2) \mathcal{O}(\tau_3) \mathcal{O}(\tau_4) \rangle_{\text{GFF}} = \tau_{12}^{-2\Delta} \tau_{34}^{-2\Delta} + \tau_{14}^{-2\Delta} \tau_{23}^{-2\Delta} + \tau_{13}^{-2\Delta} \tau_{24}^{-2\Delta} \quad (2.43)$$

$$= \tau_{12}^{-2\Delta} \tau_{34}^{-2\Delta} \left(1 + \left(\frac{z}{1-z} \right)^{2\Delta} + z^{2\Delta} \right), \quad (2.44)$$

where $\tau_{ij} = \tau_i - \tau_j$ and the cross-ratio is $z = \frac{\tau_{12}\tau_{34}}{\tau_{13}\tau_{24}}$. In the OPE channel $\mathcal{O}(\tau_1) \mathcal{O}(\tau_2)$, the first term in (2.44) corresponds to the exchange of the unit operator, while the other two terms are due to double-trace operators. To see the latter, we decompose [e.g. see Equations (4.14) and (4.15) in [27]]

$$\left(\frac{z}{1-z} \right)^{2\Delta} = \sum_{n=0}^{\infty} \frac{(2\Delta)_n^2}{n!(4\Delta + n - 1)_n} g_{2\Delta+n}(z), \quad (2.45)$$

$$z^{2\Delta} = \sum_{n=0}^{\infty} (-1)^n \frac{(2\Delta)_n^2}{n!(4\Delta + n - 1)_n} g_{2\Delta+n}(z), \quad (2.46)$$

where the conformal block is defined by

$$g_h(z) \equiv z^h {}_2F_1(h, h, 2h, z). \quad (2.47)$$

Let us assume $\tau_1 > \dots > \tau_4$ as in (2.13). Now, if we dress (2.44) with the Schwarzian mode as in (2.5), the result is the three gravitational Feynman diagrams (2.13). Each of the three terms in (2.13) correspond to the three terms in (2.44). The first two are the uncrossed diagrams, while the last one is the crossed diagram.

One might ask what happens with the OPE expansion (2.45), (2.46) after we dress it with the Schwarzian mode. It turns out that explicit expressions may be derived for each of the individual terms on the rhs of (2.45), (2.46) after dressing with the Schwarzian. These are the last two terms in (2.13) expanded as

$$\frac{\delta(s_2 - s_4)}{\rho(s_2)} = \sum_{n=0}^{\infty} P_n^{\Delta, \Delta}(s_2; s_1, s_3) P_n^{\Delta, \Delta}(s_4; s_1, s_3), \quad (2.48)$$

$$\left\{ \begin{array}{ccc} \Delta & s_1 & s_2 \\ \Delta & s_3 & s_4 \end{array} \right\} = \sum_{n=0}^{\infty} (-1)^n P_n^{\Delta, \Delta}(s_2; s_1, s_3) P_n^{\Delta, \Delta}(s_4; s_1, s_3), \quad (2.49)$$

where P_n is proportional to a Wilson polynomial. The definition of P_n and a more detailed discussion of these identities can be found in Appendix A. These expressions can be further integrated over s_j as in (2.13) to obtain formulas in Euclidean time τ_j .

An important point that we will use later is that a system of orthonormal polynomials P_n diagonalizes the $6j$ symbol as a map between s_2, s_4 . In Appendix C 1 we take the semiclassical limit of (2.48) and (2.49) and show that it reduces to (2.45) and (2.46).

III. THE EIGENSTATE THERMALIZATION HYPOTHESIS AS A MATRIX MODEL

Holographic CFTs have a large- N expansion and a sparse spectrum of light operators [28]. By definition, the scaling dimensions of and OPE coefficients involving

the light operators only are well-defined in the large- N limit. In contrast, the OPE coefficients that involve heavy operators do not have a large- N limit. In particular, the spectrum of black hole microstates becomes arbitrarily dense for large N . For this reason, computing the exact spectrum and OPE coefficients at some fixed, large value of N is in general a hopeless task.

It is much easier to make a well-motivated guess for the CFT data. That is, it is natural to conjecture that the scaling dimensions of and OPE coefficients involving the heavy operators look like a typical draw from some ensemble.¹⁷ We will refer to this ensemble as the “ETH ensemble.” This conjecture amounts to applying the ETH to holographic theories, and it was previously studied in connection to wormholes and AdS/CFT in [29]. A key point is that self-averaging observables (such as the black hole partition function, or thermal correlators of an order one number of light operators at sufficiently early times) are only sensitive to the choice of the ensemble itself at large N . Thus, for a judiciously chosen ensemble, the ensemble’s predictions for self-averaging observables should be comparable to the results of gravitational path integral calculations. The ensemble should be chosen to correctly reproduce the results of all the gravitational calculations that we know how to perform at some given energy scale and some given level of precision.¹⁸ Note that there are certain calculations involving wormholes that we do not know how to perform.

To illustrate these points, consider the expectation value of a light operator in a black hole microstate, which we will write as $\langle E|\mathcal{O}|E\rangle$. Let $\overline{\langle E|\mathcal{O}|E\rangle}$ refer to the average value of this matrix element in a microcanonical window of width δE centered around E , in the limit of large N and small δE ,

$$\overline{\langle E|\mathcal{O}|E\rangle} \equiv \lim_{\delta E \rightarrow 0} \lim_{N \rightarrow \infty} e^{-S(E, \delta E)} \sum_{|E_a - E| < \frac{\delta E}{2}} \langle E_a|\mathcal{O}|E_a\rangle, \quad (3.1)$$

where $e^{S(E, \delta E)}$ is the number of states in the microcanonical window. Note that when N is large and δE is small, we have that $e^{S(E, \delta E)} = \delta E \rho(E)$, where $\rho(E)$ is the large- N density of black hole microstates, which may be computed from the black hole saddle in the Euclidean path integral. We can holographically determine $\overline{\langle E|\mathcal{O}|E\rangle}$ by taking the inverse Laplace transform of the thermal one-point function $\text{Tr} e^{-\beta H} \mathcal{O}$, again computed from the black hole saddle (we always assume that the temperature is above the Hawking-Page transition). The result is a smooth function

of E . In JT gravity coupled to a free massive scalar, this smooth function is zero. For the remainder of this discussion, let us for simplicity assume that $\overline{\langle E|\mathcal{O}|E\rangle}$ vanishes (this can be achieved by subtracting a multiple of the identity from the bulk operator that \mathcal{O} is dual to).

Computing the thermal two-point function $\text{Tr} e^{-\beta_1 H} \mathcal{O} e^{-\beta_2 H} \mathcal{O}$ in the same way, we can deduce another self-averaging quantity,

$$\begin{aligned} & \overline{|\langle E_1|\mathcal{O}|E_2\rangle|^2} \\ & \equiv \lim_{\delta E \rightarrow 0} \lim_{N \rightarrow \infty} e^{-S(E_1, \delta E) - S(E_2, \delta E)} \sum_{\substack{|E_a - E_1| < \frac{\delta E}{2} \\ |E_b - E_2| < \frac{\delta E}{2}}} |\langle E_a|\mathcal{O}|E_b\rangle|^2. \end{aligned} \quad (3.2)$$

Again, one may compute (3.2) by taking the inverse Laplace transform with respect to β_1 and β_2 of the two-point function evaluated from the black hole saddle, which takes the form

$$\int_{E_0}^{\infty} dE_1 \rho(E_1) dE_2 \rho(E_2) \overline{|\langle E_1|\mathcal{O}|E_2\rangle|^2} e^{-\beta_1 E_1 - \beta_2 E_2}, \quad (3.3)$$

where E_0 is the black hole threshold.

Another statistical quantity of interest is the variance,

$$\overline{\langle E|\mathcal{O}|E\rangle^2} \equiv \lim_{\delta E \rightarrow 0} \lim_{N \rightarrow \infty} e^{-S(E, \delta E)} \sum_{|E_a - E| < \frac{\delta E}{2}} \langle E_a|\mathcal{O}|E_a\rangle^2. \quad (3.4)$$

Equation (3.4) is not to be confused with (3.2) evaluated for $E_1 = E_2 = E$, which we will denote by

$$\overline{|\langle E_1|\mathcal{O}|E_2\rangle|^2}_{E_1=E_2=E}. \quad (3.5)$$

If the matrix elements of \mathcal{O} in a given microcanonical Hilbert space are independent and identically distributed random variables (up to the Hermiticity condition for \mathcal{O}), then (3.4) and (3.5) will be equal. More generally, (3.4) and (3.5) may differ, perhaps at a subleading order in N , due to nontrivial correlations among these variables. Because (3.5) and its finite N corrections should be computable from the bulk black hole two-point function, (3.4) and its subleading corrections must be computable from some other bulk observable. Starting from the black hole two-point function, one could try to obtain (3.4) by setting $\beta_1 = \beta + iT$ and $\beta_2 = \beta - iT$, integrating over $T \in (-\infty, \infty)$ to set the two energies equal, and then taking an inverse Laplace transform on β , but the result will be the same smooth function that one would have obtained by performing separate inverse Laplace transforms on β_1 and β_2 . The black hole two-point function does not have enough information for us to determine both (3.5) and (3.4) to all orders in N . The only other candidate for the holographic dual to (3.4) is a two-point function on a wormhole geometry that connects two asymptotically AdS

¹⁷To be clear, for the present discussion we are considering a single, nondisorder averaged theory.

¹⁸Different calculations in the bulk should correspond to different ensembles. For example, gravitational calculations that incorporate the effects of irrelevant operators in the Lagrangian should be more sensitive to the exact CFT spectrum. Hence, the corresponding ensemble should be more fine-grained. Of course, the effects of irrelevant operators are more important at greater energy scales and higher levels of precision.

boundaries, where one \mathcal{O} operator is inserted on each boundary.¹⁹

Euclidean wormhole amplitudes are challenging because the size of the wormhole can become small in the off-shell path integral. Where the wormhole is small, the local temperature of the bulk fields is large, leading to a UV divergence in the amplitude. For example, in JT gravity minimally coupled to a CFT, the double-trumpet matter partition function diverges as $e^{\frac{c^2}{6b}}$ for small b . Hence, the wormhole amplitude is undefined. This issue can be side-stepped when the path integral has a saddle. For instance, if the dimension of \mathcal{O} scales as $\frac{1}{G_N}$, then by inserting an \mathcal{O} on each AdS boundary we can stabilize the wormhole at a finite size. However, when the dimension is of order one, there is no saddle. Hence, wormhole amplitudes reflect a limitation on what the gravitational Euclidean path integral can teach us about the ETH ensemble.

As illustrated above, the Euclidean path integral alone is not able to entirely determine the ETH ensemble. Thus, some additional principle is needed to motivate which ensemble to use. The simplest guess is to follow the original ETH literature [1,2,30] and draw the matrix elements of \mathcal{O} from a Gaussian ensemble,

$$\langle E_a | \mathcal{O} | E_b \rangle = \overline{\langle E_a | \mathcal{O} | E_a \rangle} \delta_{ab} + \sqrt{|\langle E_a | \mathcal{O} | E_b \rangle|^2} R_{ab}, \quad (3.6)$$

where R_{ab} is a random GUE matrix with zero mean and unit variance. Unfortunately, a Gaussian ETH ensemble does not correctly compute all of the thermal correlators. For concreteness, let us specialize to the case of JT gravity coupled to a massive scalar and use the conventions of Sec. II. In the next section, we will find that the ansatz (3.6) computes the correct two-point function only. For the four-point function, the Gaussian ETH reproduces the two gravitational Feynman diagrams where the bulk lines do not cross [see Eq. (2.11)]. In the semiclassical limit at zero temperature, the corresponding correlator is

$$\begin{aligned} \langle \mathcal{O}(\tau_1) \mathcal{O}(\tau_2) \mathcal{O}(\tau_3) \mathcal{O}(\tau_4) \rangle \\ = (\tau_{12} \tau_{34})^{-2\Delta} + (\tau_{14} \tau_{23})^{-2\Delta}. \end{aligned} \quad (3.7)$$

The main issue with (3.7) is that it is inconsistent with crossing symmetry, which is the requirement that multipoint \mathcal{O} correlators ought to be invariant under permutations of the τ_j variables.²⁰ The crossing-symmetric four-point function that we wish to reproduce using the ETH ensemble is

$$\begin{aligned} \langle \mathcal{O}(\tau_1) \mathcal{O}(\tau_2) \mathcal{O}(\tau_3) \mathcal{O}(\tau_4) \rangle \\ = (\tau_{12} \tau_{34})^{-2\Delta} + (\tau_{14} \tau_{23})^{-2\Delta} + (\tau_{13} \tau_{24})^{-2\Delta}. \end{aligned} \quad (3.8)$$

This is the four-point function of a generalized free field (GFF). In 1D CFT, a simple condition that fixes the correlation functions of \mathcal{O} to be those of a GFF is that the only primary operators that appear in the $\mathcal{O}\mathcal{O}$ OPE aside from the identity have dimensions $2\Delta + 2n$ for $n \in \mathbb{Z}_{\geq 0}$.²¹ In this case, the OPE looks like

$$\mathcal{O}(\tau) \mathcal{O}(0) = \tau^{-2\Delta} + \sum_{n=0}^{\infty} \tau^{2n} [\mathcal{O}\mathcal{O}]_n + \text{descendants} \quad (3.9)$$

where $[\mathcal{O}\mathcal{O}]_n$ refers to a double-trace primary. Equation (3.9) is equivalent to the following operator equation:

$$\begin{aligned} [\mathcal{O}(t_1), \mathcal{O}(t_2)] &= \lim_{\epsilon \rightarrow 0} \left[\frac{1}{(it_{12} + \epsilon)^{2\Delta}} - \frac{1}{(it_{12} - \epsilon)^{2\Delta}} \right] \\ &= \begin{cases} -\frac{2i \sin(\pi\Delta)}{|t_1 - t_2|^{2\Delta}} \text{sign}(t_1 - t_2), & \Delta \notin \mathbb{Z}_{\geq 1} \\ \frac{2\pi i}{(2\Delta-1)!} (-1)^{\Delta-1} \delta^{(2\Delta-1)}(t_{12}), & \Delta \in \mathbb{Z}_{\geq 1} \end{cases}. \end{aligned} \quad (3.10)$$

That is, the commutator of two \mathcal{O} operators separated in Lorentzian time is proportional to the identity operator. It is straightforward to check that (3.9) implies (3.10). To see that (3.10) implies (3.9), note that if any primary operators with dimensions not in the set $\{1\} \cup \{2\Delta + 2n : n \in \mathbb{Z}_{\geq 0}\}$ appeared on the rhs of (3.9), then they would make additional contributions to the rhs of (3.10). Furthermore, it is simple to show that (3.10) is obeyed by the GFF correlators.

In a typical instance of the Gaussian ETH ensemble, the matrix \mathcal{O} does not obey any simple operator equations. In an ETH ensemble that reproduces the disk correlators of JT gravity with matter, \mathcal{O} should be constrained to obey (3.10) in the semiclassical limit. We propose an ETH ensemble where the analog of (3.10) away from the semiclassical limit is manifestly obeyed. First, we must find the operator equation that generalizes (3.10) away from the semiclassical limit. We propose the following expression:

$$\begin{aligned} e^{-S_0} \sum_b \left\{ \begin{array}{ccc} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{array} \right\} \left[\frac{\mathcal{O}_{ab} \mathcal{O}_{bc}}{e^{-S_0} \sqrt{\Gamma_{ab}^{\Delta} \Gamma_{bc}^{\Delta}}} - \delta_{ac} \right] \\ = \left[\frac{\mathcal{O}_{ad} \mathcal{O}_{dc}}{e^{-S_0} \sqrt{\Gamma_{ad}^{\Delta} \Gamma_{dc}^{\Delta}}} - \delta_{ac} \right], \end{aligned} \quad (3.11)$$

where $\mathcal{O}_{ab} \equiv \langle E_a | \mathcal{O} | E_b \rangle$ refers to matrix elements of \mathcal{O} in the energy eigenbasis, and $s^2 = E$ relates the s parameters to the energies. The sum is over all of the energy

¹⁹See the discussion section of [24] for further comments.

²⁰Strictly speaking, this definition is correct when Δ is an integer. More generally, there are branch cuts in the complex τ plane that introduce an ambiguity when continuing one operator past another. A more general definition of crossing symmetry is that the correlators should be invariant under permutations up to the phases associated with continuing the operators past one another.

²¹See Appendix D for an explanation.

eigenvalues. Equation (3.11) holds within any correlator (both at the disk level and, conjecturally, at higher genus) in any ensemble-averaged theory that is dual to JT gravity minimally coupled to a scalar.

For example, let us insert (3.11) into a correlator with two other \mathcal{O} insertions. The ensemble average of

$$\sum_{a,b,c,d,e} e^{-\beta_1 E_a} e^{-\beta_2 E_e} e^{-S_0} \left\{ \begin{array}{ccc} \Delta & s_a & s_b \\ \Delta & s_c & s_e \end{array} \right\} \left[\frac{\mathcal{O}_{ab} \mathcal{O}_{bc}}{e^{-S_0} \sqrt{\Gamma_{ab}^\Delta \Gamma_{bc}^\Delta}} - \delta_{ac} \right] \times e^{-\beta_3 E_c} \mathcal{O}_{cd} e^{-\beta_4 E_d} \mathcal{O}_{da} \quad (3.12)$$

is, at disk level,

$$\int_0^\infty ds_e \rho(s_e) e^{-\beta_2 s_e^2} e^{2S_0} \int_0^\infty ds_a \rho(s_a) ds_b \rho(s_b) ds_c \rho(s_c) ds_d \times \rho(s_d) e^{-\beta_1 s_a^2} e^{-\beta_3 s_c^2} e^{-\beta_4 s_d^2} (\Gamma_{cd}^\Delta \Gamma_{da}^\Delta)^{1/2} \quad (3.13)$$

$$\times \left\{ \begin{array}{ccc} \Delta & s_a & s_b \\ \Delta & s_c & s_e \end{array} \right\} \left(\frac{\delta(s_b - s_d)}{\rho(s_b)} + \left\{ \begin{array}{ccc} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{array} \right\} \right), \quad (3.14)$$

$$= e^{2S_0} \int_0^\infty ds_a \rho(s_a) ds_e \rho(s_e) ds_c \rho(s_c) ds_d \rho(s_d) \times e^{-\beta_1 s_a^2} e^{-\beta_2 s_e^2} e^{-\beta_3 s_c^2} e^{-\beta_4 s_d^2} (\Gamma_{cd}^\Delta \Gamma_{da}^\Delta)^{1/2} \quad (3.15)$$

$$\times \left(\left\{ \begin{array}{ccc} \Delta & s_a & s_d \\ \Delta & s_c & s_e \end{array} \right\} + \frac{\delta(s_d - s_e)}{\rho(s_e)} \right), \quad (3.16)$$

where we have used the orthogonality of $6j$ -symbols (2.14). This is the disk level expression in the ensemble-averaged theory for

$$\sum_{a,c,d,e} e^{-\beta_1 E_a} e^{-\beta_2 E_e} \left[\frac{\mathcal{O}_{ae} \mathcal{O}_{ec}}{e^{-S_0} \sqrt{\Gamma_{ae}^\Delta \Gamma_{ec}^\Delta}} - \delta_{ac} \right] e^{-\beta_3 E_c} \mathcal{O}_{cd} e^{-\beta_4 E_d} \mathcal{O}_{da}. \quad (3.17)$$

It is straightforward to check that (3.11) holds in correlators with more \mathcal{O} insertions. Equation (3.11) is a natural generalization of (3.10) away from the semiclassical limit because it is the only operator equation that we are aware of that is quadratic in \mathcal{O} . Heuristically, the $6j$ -symbol exchanges the order of the operators.

If we integrate both sides of (3.11) with respect to $P_n^{\Delta\Delta}(s_d; s_a, s_c)$ defined in (A7), we may derive

$$\sum_b P_n^{\Delta\Delta}(s_b; s_a, s_c) \left[\frac{\mathcal{O}_{ab} \mathcal{O}_{bc}}{e^{-S_0} \sqrt{\Gamma_{ab}^\Delta \Gamma_{bc}^\Delta}} - \delta_{ac} \right] = 0, \quad n \in 2\mathbb{Z}_{\geq 0} + 1. \quad (3.18)$$

This can be interpreted as the statement that primaries with weights $2\Delta + n$ for odd n do not appear in the \mathcal{OO} OPE. Thus, in every member of the ensemble, $\mathcal{O}_{ab} \mathcal{O}_{bc}$ should be

equal to²² $\delta_{ac} e^{-S_0} (\Gamma_{ab}^\Delta \Gamma_{bc}^\Delta)^{1/2}$, which represents the identity operator contribution in the OPE, plus a linear combination of $(\Gamma_{ab}^\Delta \Gamma_{bc}^\Delta)^{1/2} P_n^{\Delta\Delta}(s_b; s_a, s_c)$ for even n , which represent the blocks associated to primaries with dimensions $2\Delta + n$ for n even.²³

We will construct a new ETH ensemble by imposing (3.11) as a constraint. That is, let us define

$$M_{ac}^d = \frac{1}{2} \sum_b \left(e^{-S_0} \left\{ \begin{array}{ccc} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{array} \right\} - \delta_{bd} \right) \times \left(\frac{\mathcal{O}_{ab} \mathcal{O}_{bc}}{e^{-S_0} \sqrt{\Gamma_{ab}^\Delta \Gamma_{bc}^\Delta}} - \delta_{ac} \right). \quad (3.19)$$

The ensemble is then defined by the following matrix integral:

$$\int dH d\mathcal{O} \exp \left(-\text{Tr} V(H) - \frac{\Lambda}{2} \sum_{a,c,d} |M_{ac}^d|^2 \right), \quad (3.20)$$

where Λ is a large parameter that enforces $M_{ac}^d = 0$, or (3.11), as a constraint. \mathcal{O} and H are Hermitian matrices with the usual measure. The role of $V(H)$ is to ensure that to leading order in S_0 , the spectrum of the Hamiltonian H agrees with $\rho(s)$. This is a two-matrix model with a single-trace potential. One can define a genus expansion in terms of 't Hooft ribbon diagrams (the diagrammatic rules are given in (6.7) and (6.10)). We propose that the disk correlators of this matrix model agree with the disk correlators of JT gravity minimally coupled to a scalar field. In fact, (3.20) is a solvable model in the sense that the disk correlators of JT gravity solve its planar Schwinger-Dyson equations (to the extent that we checked). Furthermore, the analytic expressions for 't Hooft diagrams greatly simplify thanks to the unlacing rules in (2.14) and (2.15). We will explain this in more detail in Sec. VI. Note that the size of the matrices in (3.20) is formally infinite, so (3.20) represents the end product of a double-scaling limit. We discuss how one can back away from the double-scaling limit in Sec. VIB. The manipulations in Sec. VI will suggest that any matrix model that looks like (3.20) in a double-scaling limit will compute the JT disk correlators. Away from the double-scaling limit, the model is not solvable because generically the unlacing rules would no longer hold.²⁴

Even though (3.20) is not solvable away from the double-scaling limit, we can argue that in the double-

²²After averaging the E_b energy in a microcanonical window.

²³Note that the $P_n^{\Delta\Delta}(s_b; s_a, s_c)$ functions obey a completeness relation. See (A30).

²⁴In a regulated model, the spectrum of H will have compact support, so (2.16) will be replaced by an integral over a finite domain. It is not possible for an integral over a finite domain to produce a delta function.

scaling limit it reproduces the desired gravitational correlators at disk level. It is natural to then ask what the multiboundary and higher genus correlators are. One of the technical results of this paper is that the double-trumpet correlators²⁵ (and most likely all multi-boundary correlators for any genus) of a single-trace model of two Hermitian matrices can be directly determined from the disk correlators without knowledge of the matrix potential itself. However, the result depends on how the double-scaling limit is taken. In particular, one needs to have regulated expressions for the gravitational correlators. Although in Sec. VIB we write an explicit matrix potential that represents a regulated version of (3.20), we have not determined the corresponding regulated expressions for the six- and higher-point disk correlators.²⁶ Thus, beginning in Sec. VII, we consider different ways of regulating the disk gravitational Feynman rules that *a priori* are independent from the regulator in Sec. VIB, and we argue that one can work backwards to determine the matrix potential. We will consider two specific regulators that lead to two specific models, and we will see in Sec. IX (at the level of the empty double-trumpet and the double-trumpet two-point function with one \mathcal{O} on each boundary) that the two bulk theories have the same correlators up to the matter determinant factor [which for a massive scalar on the double-trumpet is $Z_{\text{scalar}}(b)$, defined in (2.23)–(2.25)]. We will also see that the bulk 1-loop determinant is sensitive to how the matrix model is defined slightly away from the double-scaling limit.

Having an ETH ensemble allows us in principle to make sense of the otherwise ill-defined wormhole amplitudes. In Sec. IX D, we show that the empty double-trumpet in the matrix model is directly determined by the Hessian of the matrix potential evaluated at the saddle point that defines the perturbative expansion.²⁷ The double-trumpet becomes ill-defined precisely when the Hessian is not positive-definite. Hence, the problematic UV behavior of wormholes is linked to a perturbative instability in the matrix model. To make the wormholes well-defined, it would be interesting to find a stable saddle that the unstable saddle could decay to. We leave this question to future work.

IV. WARMING UP WITH A SOLVABLE TWO-MATRIX MODEL

To gain intuition for how a two-matrix model can compute the disk correlators of an arbitrary number of \mathcal{O} insertions, it is useful to first consider a solvable toy

²⁵Strictly speaking, in this work we only explicitly investigate the empty double-trumpet and the double-trumpet two-point function with one \mathcal{O} on each AdS boundary.

²⁶For the two- and four-point disk correlators, we conjecture explicit expressions that obey a Schwinger-Dyson equation in the double-scaling limit.

²⁷To be more precise, we integrate out \mathcal{O} and then work with the effective potential for H , which is multitrace.

model that only correctly computes the disk two-point function. Later, we will generalize the model described in this section to a more complicated model that correctly computes all of the disk correlators.

Consider a two-matrix model with Hermitian matrices H and \mathcal{O} and a single-trace matrix potential $V(H, \mathcal{O})$. The potential has a complicated dependence on H but is quadratic in \mathcal{O} . We use the standard flat measure for the real and imaginary parts of the matrix elements of H and \mathcal{O} . The matrix integral of this toy model is given by

$$Z_{\text{toy}} = \int dH d\mathcal{O} e^{-V_{\text{toy}}(H, \mathcal{O})}, \quad (4.1)$$

$$V_{\text{toy}}(H, \mathcal{O}) = \sum_a (V_{\text{SSS}}(E_a) + V_{c.t.}(E_a)) + \frac{1}{2} \sum_{a,b} F(E_a, E_b) \mathcal{O}_{ab} \mathcal{O}_{ba}, \quad (4.2)$$

where $F(E_a, E_b) = F(E_b, E_a)$ is a smooth function and E_a refers to an eigenvalue of H . We have chosen to write this single-trace potential in the eigenbasis of H , so the sums are over all of the eigenvalues of H . The $V_{\text{SSS}}(H)$ term is the matrix potential of the SSS model [9], which is dual to pure JT gravity.

We have included a counterterm potential $V_{c.t.}(H)$ that is chosen to ensure that after integrating out \mathcal{O} the disk density of states for H is still $e^{S_0} \rho_0(E)$, given in (2.9). This is because in JT gravity coupled to the free scalar, integrating out the scalar does not affect the disk partition function, except for the renormalization of S_0, γ . This was discussed after (2.31). We will determine $V_{c.t.}(H)$ momentarily.

The last term in (4.2) is single-trace because we can write an arbitrary function of two energies as $F(E_a, E_b) = \sum_{nm} c_{nm} E_a^n E_b^m$, with each term being single-trace $\text{Tr} H^n \mathcal{O} H^m \mathcal{O}$.

Note that [9] did not provide an explicit formula for the matrix potential in their model, because the details of how the double-scaling limit is taken do not affect their results. These details are also irrelevant for this section. In this section, the number of eigenvalues is infinity and we are working directly in the double-scaling limit. A more rigorous treatment is provided in Sec. VIII.

To determine $V_{c.t.}(H)$, we should first integrate out \mathcal{O} . The result is

$$Z_{\text{toy}} = \int dH e^{-\text{Tr} V_{\text{SSS}}(H) - \tilde{V}(H)}, \quad (4.3)$$

$$\tilde{V}(H) = \sum_a V_{c.t.}(E_a) + \frac{1}{2} \sum_{a,b} \log F(E_a, E_b). \quad (4.4)$$

The last term in (4.4) is a double-trace term, and it is represented by a double-line loop in 't Hooft diagrams.

After expanding the $e^{-\tilde{V}(H)}$ term, we may diagrammatically compute corrections to the disk density of states for H , as shown in Fig. 3. We should pick $V_{c.t.}$ such that $\tilde{V}(H)$ becomes

$$\tilde{V}(H) = \frac{1}{2} \int dE_1 dE_2 (\rho_H(E_1) - e^{S_0} \rho_0(E_1)) (\rho_H(E_2) - e^{S_0} \rho_0(E_2)) \log F(E_1, E_2), \quad (4.5)$$

where

$$\rho_H(E) = \text{Tr} \delta(E - H) = \sum_a \delta(E - E_a). \quad (4.6)$$

The variation of (4.5) with respect to $\rho_H(E)$ vanishes to first order, when evaluated for $\rho_H(E) = e^{S_0} \rho_0(E)$. Hence, the addition of \tilde{V} to V_{SSS} does not change the saddle-point density of states, as desired. Note that adding a single-trace counterterm $\text{Tr} V_{c.t.}(H)$ is enough to ensure this.

A. Correlation functions on the disk

We next consider the disk two-point function of \mathcal{O} in our toy model. It is given by

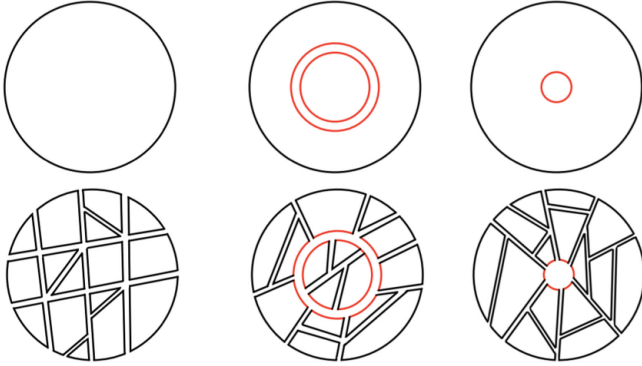


FIG. 3. Top row: The left disk represents the disk computation of $\langle \text{Tr} e^{-\beta H} \rangle$ in the SSS model [9]. One should imagine filling in the disk with all possible planar ‘t Hooft diagrams of H . The middle disk represents a correction from a single insertion of the double-trace term in $\tilde{V}(H)$ (in general, there could be arbitrarily many insertions, which are all summed over). One should imagine filling in the regions inside and outside the red double-line loop with the ‘t Hooft diagrams of H in the SSS model of disk and cylinder topology, respectively. In our terminology, the red double-line loop in the center is an example of an “ \mathcal{O} bubble diagram.” The right disk represents a correction from the single-trace counterterm potential that is designed to cancel the contribution from the \mathcal{O} bubble diagram. The counterterm is a single-trace term in the potential and hence is represented by a single loop. Bottom row: we provide an example of one of the infinitely many ways the diagrams in the top row can be filled in with planar ‘t Hooft diagrams involving the H matrix. A black double-line represents the propagator of the H matrix.

$$\begin{aligned} & \langle \text{Tr} e^{-\beta_1 H} \mathcal{O} e^{-\beta_2 H} \mathcal{O} \rangle_{\text{disk}} \\ &= e^{2S_0} \int_0^\infty dE_1 dE_2 e^{-\beta_1 E_1} e^{-\beta_2 E_2} \rho_0(E_1) \rho_0(E_2) F(E_1, E_2)^{-1}. \end{aligned} \quad (4.7)$$

Comparing this with the gravity answer (2.10), we can determine $F(E_1, E_2)$

$$F(E_1, E_2) = e^{S_0} \left(\frac{\Gamma(\Delta \pm i\sqrt{E_1} \pm i\sqrt{E_2})}{\Gamma(2\Delta)} \right)^{-1}. \quad (4.8)$$

We express this match between the matrix model and gravity computations of the two-point function pictorially

$$\text{(red double-line loop)} = \text{(blue single-line loop)}, \quad (4.9)$$

where the lhs represents the sum over ‘t Hooft diagrams computing the two-point function in the matrix model. The red double-line is the propagator of the \mathcal{O} matrix. Each single red line represents an energy, and the two closed red loops should be filled in like a disk (that is, with \mathcal{O} bubble diagrams and H double-lines in all possible ways). Filling in a loop like a disk corresponds to integrating over the corresponding energy with the disk density of states $e^{S_0} \rho_0(E)$. The rhs is the Feynman diagram in JT gravity coupled to a free scalar (2.10).

We can similarly compute the four-point function. In (4.10) to (4.14) only, we imagine that the spectrum of H has been fixed to some instance, and $\langle \dots \rangle$ refers to the expectation value in the ensemble defined by the \mathcal{O} matrix integral,

$$\begin{aligned} & \left\langle \text{Tr} \prod_{j=1}^4 e^{-\beta_j H} \mathcal{O} \right\rangle \\ &= \sum_{a_1, \dots, a_4} e^{-\sum_{j=1}^4 \beta_j E_{a_j}} \langle \mathcal{O}_{a_1 a_2} \mathcal{O}_{a_2 a_3} \mathcal{O}_{a_3 a_4} \mathcal{O}_{a_4 a_1} \rangle \end{aligned} \quad (4.10)$$

$$\begin{aligned} &= \sum_{a_1, \dots, a_4} e^{-\sum_{j=1}^4 \beta_j E_{a_j}} (\langle \mathcal{O}_{a_1 a_2} \mathcal{O}_{a_2 a_3} \rangle \langle \mathcal{O}_{a_3 a_4} \mathcal{O}_{a_4 a_1} \rangle \\ &+ \langle \mathcal{O}_{a_1 a_2} \mathcal{O}_{a_4 a_1} \rangle \langle \mathcal{O}_{a_2 a_3} \mathcal{O}_{a_3 a_4} \rangle) \end{aligned} \quad (4.11)$$

$$+ \langle \mathcal{O}_{a_1 a_2} \mathcal{O}_{a_3 a_4} \rangle \langle \mathcal{O}_{a_2 a_3} \mathcal{O}_{a_4 a_1} \rangle) \quad (4.12)$$

$$\begin{aligned} &= \sum_{a_1, \dots, a_4} e^{-\sum_{j=1}^4 \beta_j E_{a_j}} F(E_{a_1} E_{a_2})^{-1} \\ &\times F(E_{a_3} E_{a_4})^{-1} (\delta_{a_1 a_3} + \delta_{a_2 a_4}) \end{aligned} \quad (4.13)$$

$$+ \sum_a e^{-(\beta_1 + \dots + \beta_4) E_a} F(E_a, E_a)^{-2}. \quad (4.14)$$

We used Wick contractions above. The propagator in the Gaussian model is $\langle \mathcal{O}_{a_1 a_2} \mathcal{O}_{a_3 a_4} \rangle = \delta_{a_1 a_4} \delta_{a_2 a_3} F(E_{a_1}, E_{a_2})^{-1}$. The last term (4.14) is nonplanar and we neglect it.

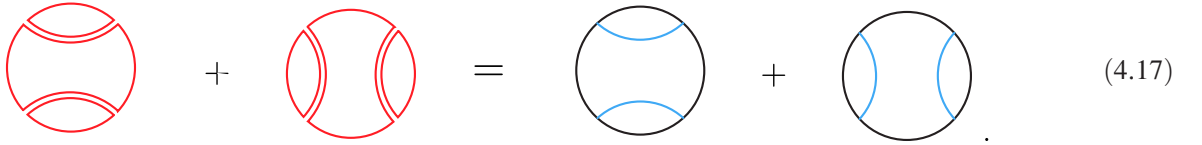
To get the disk four-point function in the double-scaled two-matrix model, we use (4.8), we substitute δ_{ab} with

$\frac{\delta(E_a - E_b)}{e^{S_0} \rho_0(E_a)}$, and we substitute \sum_a with $\int dE_a e^{S_0} \rho_0(E_a)$. The result is that we reproduce the first two terms of the gravity answer (2.11)

$$\left\langle \text{Tr} \prod_{j=1}^4 e^{-\beta_j H} \mathcal{O} \right\rangle_{\text{disk}} = e^{S_0} \int_0^\infty \prod_{j=1}^4 (dE_j e^{-\beta_j E_j} \rho_0(E_j)) \quad (4.15)$$

$$\frac{\Gamma(\Delta \pm i\sqrt{E_1} \pm i\sqrt{E_2})}{\Gamma(2\Delta)} \frac{\Gamma(\Delta \pm i\sqrt{E_3} \pm i\sqrt{E_4})}{\Gamma(2\Delta)} \left(\frac{\delta(E_1 - E_3)}{\rho_0(E_1)} + \frac{\delta(E_2 - E_4)}{\rho_0(E_2)} \right). \quad (4.16)$$

We therefore find a match between 't Hooft diagrams in the matrix model and gravitational Feynman diagrams in JT gravity



Each \mathcal{O} propagator in the matrix model may be interpreted as a bulk line in the gravitational Feynman rules.

Note that the third term in (2.11) is not captured by our toy model. This is because the third Wick contraction (4.14) is nonplanar, while in (2.11) the third term contributes at the same order e^{S_0} .

It should now be clear that the Gaussian in \mathcal{O} matrix model considered in this section captures correctly all gravitational Feynman diagrams that do not have bulk line intersections. While the Feynman diagrams with intersections (that depend on the $6j$ -symbols) are not reproduced. To deal with this issue we will eventually add interactions in the matrix model for \mathcal{O} in Sec. VII.

B. Double-trumpet

Before adding interactions, we would like to consider the double-trumpet in the Gaussian matrix model. This will facilitate a similar discussion in the interacting case later on.

We consider the connected correlator $\langle \text{Tr} e^{-\beta_L H} \text{Tr} e^{-\beta_R H} \rangle_c$ in the toy matrix model. To leading order in the genus expansion, this is computed by summing 't Hooft diagrams of cylinder topology. In the SSS model, it was shown [9] that that this sum correctly reproduces the gravity answer, i.e. Eq. (2.17) without $Z_{\text{scalar}}(b)$. Therefore, we naively want to show that adding $\tilde{V}(H)$ in (4.3) is equivalent to inserting the 1-loop determinant $Z_{\text{scalar}}(b)$ in (2.17). This will turn out not to be true, but it will be instructive to go through this computation.

We will compute the effect of $\tilde{V}(H)$ in (4.3) in a perturbative expansion in the SSS matrix model. In order

to do so, it is convenient to use an integral representation of \tilde{V} . We use that

$$\begin{aligned} \log \Gamma(\Delta \pm ik_1 \pm ik_2) \\ = \int_0^\infty db b \frac{e^{-\Delta b}}{1 - e^{-b}} \frac{2}{b} \cos(bk_1) \frac{2}{b} \cos(bk_2) + \text{const.} \end{aligned} \quad (4.18)$$

The additive constant here is divergent and cancels the $b \rightarrow 0$ divergence of the integral, but it is independent of k_1, k_2 . More precisely

$$\begin{aligned} \log \Gamma(z + 1) \\ = -\gamma z + \int_0^\infty \frac{db}{b(e^b - 1)} [e^{-zb} - (1 - zb)]. \end{aligned} \quad (4.19)$$

When we sum four such integrals to compute $\log \Gamma(\Delta \pm ik_1 \pm ik_2)$, the dependence on k_1, k_2 drops out everywhere except for the first term in the integral. The potential (4.5) and (4.8) can now be written

$$-\tilde{V}(H) = \frac{1}{2} \int_0^\infty db b \frac{e^{-\Delta b}}{1 - e^{-b}} X(b)^2 + \text{const}, \quad (4.20)$$

where we defined

$$X(b) = \frac{2}{b} \left(\text{Tr} \cos(b\sqrt{H}) - \int_0^\infty dE e^{S_0} \rho_0(E) \cos(b\sqrt{E}) \right) \quad (4.21)$$

$$= \frac{2}{b} (\text{Tr} \cos(b\sqrt{H}) - \langle \text{Tr} \cos(b\sqrt{H}) \rangle_{\text{disk}}). \quad (4.22)$$

The reader might worry that these expressions look divergent in the double-scaling limit. The important point is that divergences in the two terms (4.21) cancel and correlators of $X(b)$ are finite.

Now we explain correlation functions of $X(b)$ in the SSS matrix model. We call $X(b)$ a “geodesic loop,” the name that will be justified momentarily. This observable was also considered in [31].

Consider the correlators of resolvents in the SSS matrix model [9] related to Weil-Petersson volumes $V_{g,n}$

$$W(z) = \text{Tr} \frac{2z}{z^2 + H} = \text{Tr} \left(\frac{1}{z + i\sqrt{H}} + \frac{1}{z - i\sqrt{H}} \right), \quad (4.23)$$

$$\langle W(z_1) \dots W(z_n) \rangle_{g,n} = \int_0^\infty \prod_{j=1}^n (db_j b_j e^{-b_j z_j}) V_{g,n}(b_1, \dots, b_n). \quad (4.24)$$

To get $V_{g,n}$ directly as correlators in the matrix model, we can do the inverse Laplace transform

$$x(b) = \frac{1}{b} \int_{\epsilon-i\infty}^{\epsilon+i\infty} \frac{dz}{2\pi i} e^{bz} W(z) \quad (4.25)$$

$$= \frac{2}{b} \text{Tr} \cos(b\sqrt{H}), \quad (4.26)$$

where we closed the contour to the left and used (4.23) to compute by the residue theorem. From (4.24) we now find

$$\langle x(b_1) \dots x(b_n) \rangle_{g,n} = V_{g,n}(b_1, \dots, b_n). \quad (4.27)$$

The cases $g=0, n=1, 2$ (disk and double-trumpet) are special. For the double-trumpet, Eq. (4.24) still holds if we assume²⁸ $V_{g=0,n=2}(b, b') = \frac{1}{b} \delta(b - b')$. Therefore,

$$\langle x(b)x(b') \rangle_{\text{cyl}} = \frac{1}{b} \delta(b - b'). \quad (4.28)$$

The disk correlator $\langle x(b) \rangle_{\text{disk}}$ diverges and we subtract it explicitly in (4.21)

$$X(b) = x(b) - \langle x(b) \rangle_{\text{disk}}. \quad (4.29)$$

To properly define this difference, we imagine first computing correlators of $X(b)$ at large N , without the double-scaling limit. For example, in a 1-cut matrix model with finite support of the density of states $\langle x(b) \rangle_{\text{disk}}$ is finite. And then we take the double-scaling limit. The result is that the correlators of $X(b)$ are finite

²⁸In this case $\langle W(z_1)W(z_2) \rangle_{g=0,n=2} = \int_0^\infty db_1 db_2 b_1 b_2 \times e^{-b_1 z_2 - b_2 z_2} \frac{\delta(b_1 - b_2)}{b_1} = \frac{1}{(z_1 + z_2)^2}$.

$$\langle X(b) \rangle_{\text{disk}} = 0, \quad (4.30)$$

$$\langle X(b)X(b') \rangle_{\text{cyl}} = \frac{1}{b} \delta(b - b'), \quad (4.31)$$

$$\langle X(b_1) \dots X(b_n) \rangle_{g,n} = V_{g,n}(b_1, \dots, b_n). \quad (4.32)$$

Taking the inverse Laplace only on some of z_j in (4.24) we can compute mixed correlators of W and X . In particular

$$\langle W(z)X(b) \rangle_{\text{cyl}} = \int_0^\infty db' b' e^{-b'z} V_{g=0,n=2}(b', b) \quad (4.33)$$

$$= e^{-bz}, \quad (4.34)$$

$$\langle Z(\beta)X(b) \rangle_{\text{cyl}} = \oint \frac{dw}{2\pi i} e^{\beta w} \frac{1}{2\sqrt{w}} \langle W(\sqrt{w})X(b) \rangle_{\text{cyl}} \quad (4.35)$$

$$= \oint \frac{dw}{2\pi i} e^{\beta w - b\sqrt{w}} \frac{1}{2\sqrt{w}} \quad (4.36)$$

$$= Z_{\text{tr}}(\beta, b), \quad Z(\beta) = \text{Tr} e^{-\beta H}. \quad (4.37)$$

In the second computation we used $e^{-\beta H} = \oint \frac{dw}{2\pi i} \frac{e^{\beta w}}{w+H}$. The contour in (4.37) goes around the branch cut $w \in (-\infty, 0]$ of \sqrt{w} .

We now return to the potential (4.20). To leading order in the genus expansion, we compute the connected two-boundary correlator in the toy model perturbatively in \tilde{V}

$$\langle Z(\beta_L)Z(\beta_R) \rangle_{\text{cyl}} = \sum_{n=0}^{\infty} \frac{1}{n!} \langle Z(\beta_L)(-\tilde{V})^n Z(\beta_R) \rangle \quad (4.38)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \int_0^\infty \prod_{j=1}^n \left(db_j b_j \frac{e^{-\Delta b_j}}{1 - e^{-b_j}} \right) \times \langle Z(\beta_L) \frac{X(b_1)^2}{2} \dots \frac{X(b_n)^2}{2} Z(\beta_R) \rangle \quad (4.39)$$

$$\approx \sum_{n=0}^{\infty} \int_0^\infty \prod_{j=1}^n \left(db_j b_j \frac{e^{-\Delta b_j}}{1 - e^{-b_j}} \right) \quad (4.40)$$

$$\langle Z(\beta_L)X(b_1) \rangle_{\text{cyl}} \langle X(b_1)X(b_2) \rangle_{\text{cyl}} \dots \langle X(b_{n-1})X(b_n) \rangle_{\text{cyl}} \times \langle X(b_n)Z(\beta_R) \rangle_{\text{cyl}} \quad (4.41)$$

$$= \int_0^\infty db b Z_{\text{tr}}(\beta_L, b) Z_{\text{tr}}(\beta_R, b) \sum_{n=0}^{\infty} \left(\frac{e^{-\Delta b}}{1 - e^{-b}} \right)^n. \quad (4.42)$$

Here, the correlator in the lhs $\langle Z(\beta_L)Z(\beta_R) \rangle_{\text{cyl}}$ is in the toy model (4.3), while all other correlators are in the SSS model. In the third line we retained only the leading order in genus expansion. The n th term in the sum corresponds to a

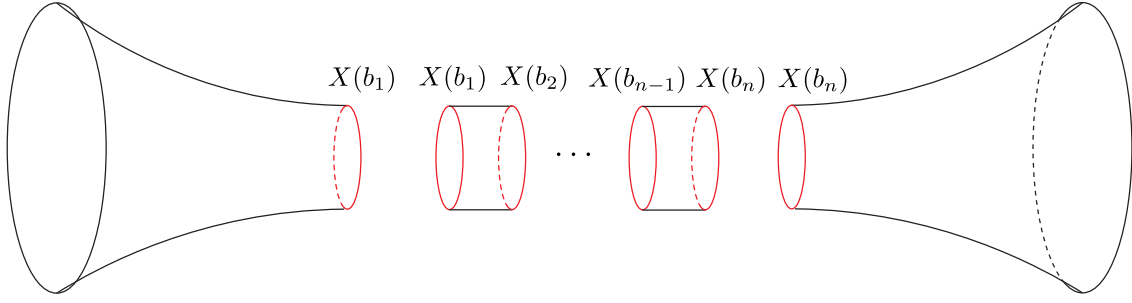


FIG. 4. Two adjacent red line loops $X(b_j)$ represents the double-trace term in (4.20). These double-lines separate the double-trumpet into $n+1$ regions, and one should imagine filling in these regions with diagrams in the SSS model with cylinder topology. Any diagrams that contain a contractible red double-line loop are canceled because $\langle X(b) \rangle_{\text{disk}} = 0$.

diagram that looks like the Fig. 4 with n red double-line loops.

Only the $n=0, 1$ terms in (4.42) agree with the corresponding terms in (2.24). We will correct the matrix model to reproduce (2.24) below.

C. Corrected effective potential

The double-trumpet in (4.42) does not agree with the result in (2.17) and (2.23) from JT minimally coupled to a scalar. However, having understood the computation (4.38)–(4.42), it is not difficult to find the potential that agrees with the gravity result. Instead of the potential $V_{\text{SSS}}(H) + \tilde{V}(H)$ consider

$$\int dH \exp \left(-V_{\text{SSS}}(H) + \frac{1}{2} \int_0^\infty db \, b [1 - (e^{-\Delta b}, e^{-b})_\infty] X(b)^2 \right). \quad (4.43)$$

A computation similar to (4.38)–(4.42) shows

$$\begin{aligned} \langle Z(\beta_L) Z(\beta_R) \rangle_{\text{cyl}} &= \int_0^\infty db \, b Z_{\text{tr}}(\beta_L, b) Z_{\text{tr}}(\beta_R, b) \sum_{n=0}^\infty [1 - (e^{-\Delta b}, e^{-b})_\infty]^n \\ & \quad (4.44) \end{aligned}$$

$$= \int_0^\infty db \, b Z_{\text{tr}}(\beta_L, b) Z_{\text{tr}}(\beta_R, b) \frac{1}{(e^{-\Delta b}, e^{-b})_\infty}, \quad (4.45)$$

where the Pochhammer symbol is $(e^{-\Delta b}, e^{-b})_\infty = \prod_{n=0}^\infty (1 - e^{-b(\Delta+n)})$. This agrees with the gravity result (2.17) and (2.23). In the next section, we will explain how one can incorporate the \mathcal{O} matrix into (4.43).

V. OUTLINE OF THE REST OF THE PAPER

In the last section we considered a toy model that is Gaussian in \mathcal{O} . This model is undesirable for two reasons:

- (1) The Gaussian model fails to correctly reproduce the disk $2n$ -point functions for $n \geq 2$. Instead, its disk

correlators are given by sums over gravitational Feynman diagrams with no intersections of the bulk lines. While the disk correlators in JT gravity minimally coupled to a scalar also include diagrams where the bulk lines cross.

- (2) The empty double-trumpet, given in (4.42), differs from (2.17). If we insist that the Gaussian model is dual to JT gravity minimally coupled to some matter theory, then the partition function of this theory in AdS_2 differs from that of a scalar field.

In this section, we summarize two strategies for finding matrix models that do not suffer from the above issues. We then comment on what these models can potentially teach us about UV divergences in wormhole amplitudes.

A. A well-motivated guess for the matrix potential

Our first strategy is to make a well-motivated guess for the matrix potential and then argue that it correctly reproduces the gravitational correlators. This was discussed in Sec. III. We identified an operator equation that is quadratic in \mathcal{O} (but has a complicated dependence on H) that holds in any disk correlator. In the eigenbasis of H , this expression is given in (3.11). It is reasonable to expect that in any instance of a matrix ensemble that reproduces the correct gravitational disk correlators, the two matrices should represent (to a good approximation) an abstract operator algebra where the operators \mathcal{O} and H obey (3.11). Hence, we wrote in (3.20) a model where (3.11) is manifestly obeyed in the limit $\Lambda \rightarrow \infty$. In Sec. VI, we write down rules for constructing ‘t Hooft diagrams in this model. Because these rules involve the $6j$ symbol, ‘t Hooft diagrams greatly simplify due to the unlacing rule in (2.14). We use (2.14) to show that a large class of Schwinger-Dyson equations is solved by the correct gravitational disk correlators.

B. Solving for the matrix potential given the gravitational correlators

Our second strategy is to start with the correct gravitational disk correlators and work backwards to find the corresponding matrix potential. To do this, we generalize the gravitational correlators by introducing a parameter ϵ

such that the original correlators are recovered as $\epsilon \rightarrow 1$, and the correlators of the Gaussian toy model are recovered as $\epsilon \rightarrow 0$. We think of ϵ as either a regulator or backing away from the double-scaling limit. There are many ways to define the regulated correlators for intermediate values of ϵ , and in Sec. VII we introduce two specific regulators; the Selberg and the q -deformed regulators.²⁹ Given a regulator, we explain in Sec. VII that one may algorithmically determine the matrix potential as a series expansion in ϵ .³⁰ After sending $\epsilon \rightarrow 1$, the matrix integral takes the form

$$\mathcal{Z} = \int dH d\mathcal{O} \exp \left(-\text{Tr}[V_{\text{SSS}}(H) + V_{c.t.}(H)] - e^{S_0} \sum_{ab} \frac{F_{ab}}{2} \mathcal{O}_{ab} \mathcal{O}_{ba} + \text{interactions} \right). \quad (5.1)$$

The interaction terms, which include higher powers of \mathcal{O} , are designed so that the matrix integral produces the correct gravitational disk correlators. As in the toy model, the role of the counter-term $V_{c.t.}(H)$ is to ensure that the density of states of H is the same as in the SSS model.³¹ In analogy with one-matrix models, we are fixing the model using only the disk data.

Next, we would like to compute correlation functions with more boundaries and/or handles and compare the

results with gravity. In Sec. IX A, we study the empty double-trumpet, or the connected two-boundary correlator

$$\langle \text{Tre}^{-\beta_L H} \text{Tre}^{-\beta_R H} \rangle_c \quad (5.2)$$

at leading order in the genus expansion. Surprisingly, we find that despite the presence of the interaction terms in (5.1), the empty double-trumpet agrees with the $\epsilon = 0$ answer (4.42). We could not find a simple reason for this, so the interested reader is encouraged to read Sec. IX A for the technical details. This result does not depend on whether the q -deformed or Selberg regulator was used.

In Sec. IX B, we compute the double-trumpet two-point function with one \mathcal{O} inserted on each AdS boundary. The result depends on whether the q -deformed or Selberg regulator is used. We now briefly sketch some of the ingredients that go into this computation, leaving the full computation for Sec. IX B. Readers who are only interested in the answer may jump to Sec. V B 2.

1. The double-trumpet two-point function

Consider the matrix model computation of the disk four-point function, which is designed to agree with the gravitational answer. In analogy to (4.17) in the toy model, we represent this computation as follows:

Diagrammatic equation (5.3) showing the decomposition of a connected four-point function into planar and non-planar components. On the left, three red diagrams are summed: a double-line with two red blobs, a double-line with two red blobs and a red blob labeled 'C', and a red blob labeled 'C'. On the right, three blue diagrams are summed: a circle with two vertical blue lines, a circle with two horizontal blue lines, and a circle with two intersecting blue lines forming an 'X'.

where a double-line with a red blob refers to the exact planar two-point function, and the blob labeled “C” refers to the sum over connected planar four-point diagrams. Each diagram on the left side corresponds to a diagram on the right side. It is also convenient to amputate the connected four-point function and define a new blob labeled “A” such that

Diagrammatic equation (5.4) defining the blob “A”. A red blob labeled “C” is equal to a red blob labeled “A” with four red lines extending from its corners.

The blobs labeled “C” and “A” refer to smooth functions of four energies, where each energy corresponds to one of the four single-lines shown in the above graphical representation.³² It is convenient to label an energy E with an s parameter, where $E = s^2$.

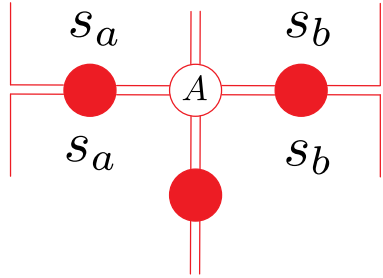
²⁹In Sec. VIII, we carefully define the regulators and show how they correspond to a ‘t Hooft-scaled matrix model. The double-scaling limit corresponds to removing the regulator.

³⁰We have not fully answered the question of whether this expansion converges for $|\epsilon| < 1$. We discuss this issue more in Secs. VII and VIII.

³¹The explicit formula for $V_{c.t.}(H)$ differs from that of the toy model due to the interaction terms, which create new \mathcal{O} bubble diagrams whose disk contributions need to be canceled.

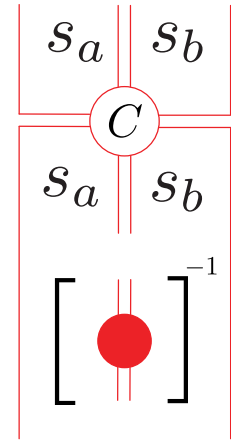
³²It looks like there are eight single-lines emanating from the “C” or “A” blobs, but in planar diagrams they are connected to each other in pairs through the blob. For example, the simplest contribution to each of the two blobs is a tree four-point vertex.

We can build a class of diagrams that contribute to the double-trumpet two-point function by gluing together two opposite propagators in (5.4) after removing one of the red blobs. Explicitly, we have



$$(5.5)$$

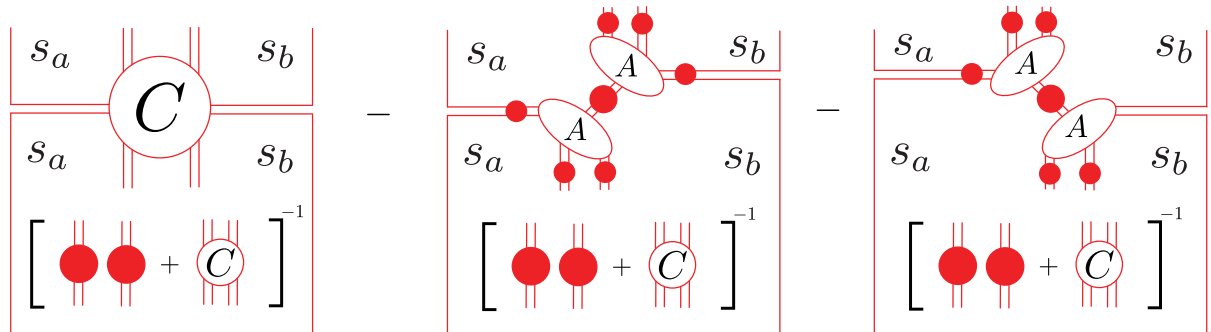
where the vertical single-lines on the far left and far right of the diagram represent the two traces in (5.2), and the top and bottom ends of this diagram are identified to obtain the cylinder topology. Removing the red blob is necessary to avoid overcounting diagrams. Another way to write this is as follows:



$$(5.6)$$

where raising the red blob to the -1 power is another way to express that a red blob should be removed from one of the external propagators of the “C” blob before making the identification. In practice, the red blob is a function of s_a and s_b , and we want to divide the “C” blob by this function. The result is a function of the two energies $E_a = s_a^2$ and $E_b = s_b^2$ that are associated to the two AdS boundaries.

So far, this discussion did not depend on whether the Selberg or q -deformed regulators were used. Now, we will build another class of diagrams starting from the connected six-point function. As above, we will identify opposite propagators to construct diagrams with cylinder topology. However, to avoid overcounting diagrams, we must remove certain subdiagrams from the connected six-point function. The appropriate generalization of (5.6) is



$$(5.7)$$

Without the $[\dots]^{-1}$ insertions above, the process of identifying the top two double-lines with the bottom two double-lines would result in an overcounting of diagrams, as the computation would not only sum over diagrams but also sum over the locations where the diagrams can be cut in two places to obtain connected planar six-point diagrams. The two subtractions are introduced so that (5.7) does not count any diagrams that were already counted in (5.6).³³ Unlike (5.6)

and (5.7) contains additional explicitly-shown closed single-line loops aside from those associated with s_a and s_b . This means that to remove the diagrams inside the brackets from the blob labeled “C,” we must integrate one of the energies of the “C” blob against the inverse of an appropriate “two-to-two propagator” that we will return to shortly.

Explicit formulas for the blobs in (5.7) are known because by construction, the matrix model disk correlators reproduce the gravitational disk correlators [for example, we may use (5.3) to obtain the explicit formula for the four-point “C” blob]. We can thus convert (5.7) into an expression that may be evaluated using the gravitational Feynman rules,

³³A more detailed explanation of our procedure for systematically classifying diagrams is provided in Sec. IX.

$$\left[\begin{array}{c} s_a \\ s_a \end{array} \left[\begin{array}{c|c|c} s_a & s_q & s_b \\ \hline s_a & s_p & s_b \end{array} + \begin{array}{c|c|c} s_a & s_q & s_b \\ \hline s_a & s_p & s_b \end{array} \right] \begin{array}{c} s_b \\ s_b \end{array} \right] \left[\begin{array}{c|c|c} s_a & s_p & s_b \\ \hline s_a & s_q & s_b \end{array} + \begin{array}{c|c|c} s_a & s_p & s_b \\ \hline s_a & s_q & s_b \end{array} \right]^{-1} \right] \quad (5.8)$$

The first set of brackets (without the -1 exponent) is meant to be expanded out into two terms, where each term is evaluated using the gravitational Feynman rules. The blue lines attached to the black boundaries correspond to insertions of $(\Gamma_{aa}^\Delta)^{1/2}$ and $(\Gamma_{bb}^\Delta)^{1/2}$, as per the Feynman rules. The diagrams contained in the second set of brackets (with the -1 exponent) become, using the Feynman rules,

$$s_a \left[\begin{array}{c|c} s_p & \\ \hline s_q & \end{array} \right] s_b + s_a \left[\begin{array}{c|c} s_p & \\ \hline s_q & \end{array} \right] s_b \leftrightarrow \frac{\delta(s_p - s_q)}{\rho(s_p)} + \left\{ \begin{array}{ccc} \Delta & s_a & s_p \\ \Delta & s_b & s_q \end{array} \right\} \quad (5.9)$$

We should think of (5.9) as a “two-to-two propagator” that acts on a function of s_q to produce a function of s_p . The first term (with the delta function, or the two parallel lines) acts as the identity operator. The two-to-two propagator is a projector because it squares to itself, thanks to the unlacing rule (2.14) and the fact that (5.9) is invariant under exchanging the two top (or two bottom) endpoints of the blue lines. The $[\cdot \cdot]^{-1}$ term in (5.8) corresponds to the inverse propagator. Having defined the two bracketed terms in (5.8), the parameters s_p and s_q should be integrated over using the density of states $\rho(s)$. The s_p, s_q integrals are associated to the additional internal single-line loops mentioned in the previous paragraph. The two vertically separated bracketed terms are conveniently interpreted as two operators that are multiplied together, and the identification of the top and bottom ends of the diagram corresponds to taking a trace.

A subtlety arises because the two-to-two propagator is not invertible, as it is a projector. This can lead to puzzles in the evaluation of (5.8). For instance, the first bracketed term is rescaled by a factor of two under the action of the two-to-two propagator, as it is invariant under exchanging the top two (or bottom two) endpoints. This naively suggests that the inverse propagator in (5.8) may be replaced by a factor of $\frac{1}{2}$,

$$\left[\begin{array}{c|c} \text{---} & \text{---} \\ \hline \text{---} & \text{---} \end{array} + \begin{array}{c|c} \text{---} & \text{---} \\ \hline \text{---} & \text{---} \end{array} \right] \stackrel{?}{=} \frac{1}{2} \left[\begin{array}{c|c} \text{---} & \text{---} \\ \hline \text{---} & \text{---} \end{array} + \begin{array}{c|c} \text{---} & \text{---} \\ \hline \text{---} & \text{---} \end{array} \right] \quad (5.10)$$

Another naive guess is that the inverse propagator undoes the sum over the crossed and uncrossed vertical lines that pass through the horizontal line in the first bracketed term,

$$\left[\begin{array}{c|c} \text{---} & \text{---} \\ \hline \text{---} & \text{---} \end{array} + \begin{array}{c|c} \text{---} & \text{---} \\ \hline \text{---} & \text{---} \end{array} \right] \stackrel{?}{=} \begin{array}{c|c} \text{---} & \text{---} \\ \hline \text{---} & \text{---} \end{array} \quad (5.11)$$

Using the regulators mentioned above, the two-to-two propagator becomes invertible and one can obtain definite answers. Equation (5.10) is correct using the Selberg regulator, while (5.11) is correct using the q -deformed regulator. The Selberg

result for (5.8) reproduces the third term in Eq. (2.36). The result for the q -deformed regulator is the same except the sum over dimensions $2\Delta + 2m$ for $m \in \mathbb{Z}_{\geq 0}$ becomes instead a sum over dimensions $2\Delta + m$ for $m \in \mathbb{Z}_{\geq 0}$.

In Appendix F 2, we explicitly compute one further class of diagrams. Using the Selberg regulator, the analog of (5.10) becomes

$$\frac{1}{3!} \left(\begin{array}{c} \text{Diagram 1} \\ \text{Diagram 2} \\ \text{Diagram 3} \\ \text{Diagram 4} \\ \text{Diagram 5} \\ \text{Diagram 6} \end{array} \right) \quad (5.12)$$

We further show in Appendix F that after identifying the top and bottom energies of (5.12), we obtain a contribution to the two-point function on the double-trumpet

$$\sum_{n,m=0}^{\infty} \int_0^{\infty} ds_a ds_b \rho(s_a) \rho(s_b) e^{-\beta_L s_a^2 - \beta_R s_b^2} (\Gamma_{aa}^{\Delta} \Gamma_{bb}^{\Delta})^{1/2} \times \left\{ \begin{array}{cc} 3\Delta + 2n + 3m & s_a \quad s_b \\ \Delta & s_b \quad s_a \end{array} \right\}. \quad (5.13)$$

This is indeed the gravity answer if we continued the expansion (2.36) to the next term. Seeing a pattern, we conjecture that using the Selberg regulator, the class of 't Hooft diagrams where the lowest number of double-line propagators crossed by a left-right path is n returns the result

$$\frac{1}{n!} \left(\begin{array}{c} n \text{ lines} \\ \text{Diagram} \end{array} + \text{permutations} \right), \quad (5.14)$$

and that this term reproduces the n th term in the sum in (2.36). Using the q -deformed regulator, the pattern is that only the trivial permutation in (5.14) appears, with no prefactor. This trivial permutation may be evaluated using successive applications of the pentagon identity (A32), which allows two vertical lines (crossing over a horizontal line) to be fused into a sum over single vertical lines with different dimensions, and the orthogonality relation of the Wilson polynomials (see Appendix A for an introduction to the relevant special functions and identities). The final result after removing the regulator (or taking the JT limit) is the n th term in the sum in (4.42).

2. Constructing the Selberg and q -deformed matrix models

The result of the computation sketched above is that the connected correlator

$$\langle \text{Tr} \mathcal{O} e^{-\beta_L H} \text{Tr} \mathcal{O} e^{-\beta_R H} \rangle_c \quad (5.15)$$

at leading order in the genus expansion is equal to a sum over infinitely many 't Hooft diagrams in the matrix model (5.1), which we have systematically classified. Each class of diagrams may be summed to obtain an explicit formula. The gravitational computation of the double-trumpet two-point function may be written as an infinite series as in (2.36), where the n th term computes the contribution from the n th term in (2.24) to the final answer.³⁴ We explicitly checked that using the Selberg regulator, the first four classes of 't Hooft diagrams reproduce the first four terms in (2.36). We conjecture that the full set of 't Hooft diagrams reproduces all of the terms in (2.36).

Thus, the model (5.1) defined using the Selberg regulator succeeds in reproducing the double-trumpet two-point function of JT gravity minimally coupled to a scalar field. However, the model does not reproduce the correct empty double-trumpet. We can amend the matrix potential in (5.1) to make the model compute the correct empty double-trumpet without affecting the computation of (5.15). The Selberg matrix model is defined by replacing (5.1) by

$$\mathcal{Z} = \int dH d\mathcal{O} \exp \left(-\text{Tr} [V_{\text{SSS}}(H) + V_{c.t.}(H)] \right) \quad (5.16)$$

$$+ \frac{1}{2} \int_0^{\infty} db b \left([1 - (e^{-\Delta b}, e^{-b})_{\infty}] - \frac{e^{-\Delta b}}{1 - e^{-b}} \right) X(b)^2 \quad (5.17)$$

$$- e^{S_0} \sum_{ab} \frac{F_{ab}}{2} \mathcal{O}_{ab} \mathcal{O}_{ba} + \text{interactions} \Big), \quad (5.18)$$

³⁴In (2.36) we only explicitly showed the $n = 0, 1, 2$ terms, leaving the rest in the \dots .

which is the same potential up to the addition of (5.17).³⁵ Because (5.17) does not depend on \mathcal{O} , the disk correlators of \mathcal{O} are the same as in (5.1). In particular, the saddle-point density of states for H is unaffected due to the subtractions involved in the definition of $X(b)$, (4.21). If we neglected the interactions above [so that the potential is that of the toy Gaussian model plus (5.17)] and integrated out \mathcal{O} , then from (4.3), (4.20), and (4.43), it is clear that the model would reproduce the correct empty double-trumpet. We pointed out above that the interactions in (5.1) do not change the empty double-trumpet result from that of the toy model. Hence, the Selberg model computes the correct empty double-trumpet (2.17).

We conjecture that the Selberg model correctly computes all higher genus and multiboundary correlators in JT gravity minimally coupled to a scalar. This conjecture is based on the success of the Selberg model in reproducing the scalar partition function on the double-trumpet (to the extent that we explicitly checked). We expect that the correlators on general topologies may also be computed from the disk correlators using similar methods to our double-trumpet computation. We make some comments on the pair of pants in Sec. IX C.

The double-trumpet two-point function using the q -deformed regulator takes a form that resembles (2.36), except the dimensions of the states/operators that propagate around the closed geodesic differ. The partition function that counts the states/operators that propagate in (2.36) is given in (2.24). If we instead use the q -deformed regulator, we find that the relevant partition function is

$$Z(b) = \sum_{n=0}^{\infty} \left(\frac{e^{-\Delta b}}{1 - e^{-b}} \right)^n = \frac{1 - e^{-b}}{1 - e^{-b} - e^{-\Delta b}}. \quad (5.19)$$

As in the Selberg-regulated computation, we found this by computing the $n = 0, 1, 2, 3$ terms explicitly and conjecturing that the pattern continues for arbitrary n . This partition function agrees with the partition function that appears in (4.42). Hence, the q -deformed regulator seems to compute amplitudes in JT gravity minimally coupled to a scalar field up to a modification of the 1-loop determinant. For this interpretation to hold, there is no need to add additional double trace terms to (5.1), so we will define the q -deformed matrix model to simply be (5.1), where the interaction terms and counterterms are determined using the q -deformed regulator.

An interesting feature of the bulk dual of the q -deformed model is the modification of the matter 1-loop determinant

³⁵The matrix model no longer appears to be single-trace because (5.17) is a double-trace term. However, if we rescale \mathcal{O}_{ab} by a function of E_a and E_b , the transformation of the measure $d\mathcal{O}$ causes a double-trace term to be added to the potential. Hence, we can change variables in the \mathcal{O} integral to make the potential single-trace again.

in global AdS_2 . In particular, (5.19) has a Hagedorn temperature, which is discussed more in Sec. IX D 1. This is suggestive of the physics of strings, which would then appear in the dual of the closely related double-scaled SYK model.³⁶ We will show that the q -deformed model can be naturally defined away from the double-scaling limit such that its disk correlators compute the SYK correlators studied in [13]. This connection is detailed in Sec. VIII.

C. Matrix model interpretation of UV divergences in wormhole amplitudes

The double-trumpet amplitudes in both the Selberg and q -deformed matrix models suffer from UV divergences. In the Selberg model, $Z_{\text{scalar}}(b)$ diverges as $e^{\frac{2}{b}}$ for small b . In the q -deformed model, the partition function has a Hagedorn temperature. Although these amplitudes are ill-defined, we can reproduce them from the matrix model because we can write each amplitude as an infinite sum [such as in (2.36)], and we know which 't Hooft diagrams reproduce each term in the sum. Hence, the sum over all 't Hooft diagrams with cylinder topology does not converge. This is a sign that the saddle point around which the perturbative genus expansion is defined is unstable. In Sec. IX D 2, we analyze the effective potential for H (after integrating out \mathcal{O}) and show that the empty double-trumpet is directly determined by the Hessian of the matrix potential evaluated at the saddle point of the eigenvalue integral, similarly to how the disk is directly determined by the location of the saddle point. In the q -deformed model, the Hessian has negative eigenvalues, which implies an instability. In the Selberg model, infinitely many eigenvalues become arbitrarily close to zero in the double-scaling limit, which effectively also amounts to an instability.³⁷

Having a matrix model description of the UV divergences might allow us to understand how the gravitational theory, viewed as an effective theory, can be nonperturbatively completed. By modifying the matrix potential far away from the location of the saddle point, it should be possible to make the model nonperturbatively well-defined. Then, it would be interesting to find a stable saddle that the unstable saddle can decay to. Note that the disk correlators computed using this new hypothetical saddle could be very different from the original disk correlators. Hence, a more physical approach to understanding the model might

³⁶It would be interesting to relate this idea to the work of [32] as well as [33–35].

³⁷The matrix model interpretation of these UV divergences is different from the analysis of [36], which studied UV divergences arising from dynamical end-of-the-world branes. Integrating out the branes only modified the single-trace matrix potential. To make the saddle unstable, one would need to modify the intereigenvalue repulsive force associated to the Vandermonde determinant.

involve introducing new bulk modes/interactions that could possibly render the original saddle stable.

Sufficiently far away from the double-scaling limit, the saddle in the q -deformed model becomes perturbatively stable. In Sec. VIII, we explain how the q -deformed model can be naturally generalized to a three-parameter model,³⁸ and in this model the double-trumpet can be explicitly computed [see (9.29)]. Having a bulk interpretation of this three-parameter model would allow us to understand the physical mechanism that determines whether there is a Hagedorn temperature or not. While there are infinitely many ways to define the q -deformed model (or any double-scaled matrix model) away from the double-scaling limit, the three-parameter model is canonically defined by its relation to the double-scaled SYK model computations in [13]. Unfortunately, we do not have a canonical way to define the Selberg model away from the double-scaling limit, so we do not have explicit regulated formulas that could admit a nice physical interpretation.

VI. CONSTRAINED MATRIX ENSEMBLE

In this section we analyze the model defined in (3.20) in more detail. We will show that in the double-scaling limit the disk amplitudes of JT gravity coupled to a scalar (discussed in Sec. II) are solutions to a large class of planar Schwinger-Dyson equations for this model. The orthogonality relation (2.14) will play a key role. This section does not contain any prerequisite material that is necessary for understanding Sec. VII and beyond.

We will first work directly in the double-scaling limit and give an argument that the JT disk correlators solve the Schwinger-Dyson equations. The argument is based on “unlacing” relations of the $6j$ -symbol. Some terms in the Schwinger-Dyson equation turn out to be divergent at high energies. These divergences arise because of the double-scaling (low-energy) limit, where the right edge of the spectrum is taken to infinity and the density of states is supported on a semi-infinite interval. To properly deal with these divergences, in the second part of this section we back away from the double-scaling limit and consider a regularized version of the matrix-model potential (3.20). The regularized matrix model is related to a certain q -deformation of JT correlators [13], that will be discussed in more detail in Sec. VIII. In the q -deformed model the density of states has a finite support and high-energy divergences are regulated.

A. Working directly in the double-scaling limit

First, we rewrite the part of the potential that depends on \mathcal{O} . We find it more convenient to work with the rescaled matrix R_{ab} , defined by

$$\begin{aligned}\mathcal{O}_{ab} &:= R_{ab} \left(\frac{e^{-S_0} \Gamma(\Delta \pm i\sqrt{E_a} \pm i\sqrt{E_b})}{\Gamma(2\Delta)} \right)^{1/2} \\ &:= (e^{-S_0} \Gamma_{ab}^\Delta)^{1/2} R_{ab}.\end{aligned}\quad (6.1)$$

Note that the matrix elements of R depend on the eigenvalues of H . For added simplicity we set $S_0 = 0$ in this subsection.³⁹ We have that

$$\begin{aligned}\frac{\Lambda}{2} \sum_{nac} |M_{ac}^n|^2 &= \frac{\Lambda}{8} \sum_{nac} \sum_b \left(\left\{ \begin{matrix} \Delta & s_a & s_b \\ \Delta & s_c & s_n \end{matrix} \right\} - \delta_{bn} \right) \\ &\quad \times (R_{ab} R_{bc} - \delta_{ac}) \\ &\quad \times \sum_d \left(\left\{ \begin{matrix} \Delta & s_a & s_n \\ \Delta & s_c & s_d \end{matrix} \right\} - \delta_{dn} \right) \\ &\quad \times (R_{cd} R_{da} - \delta_{ac})\end{aligned}\quad (6.2)$$

$$\begin{aligned}&= \frac{\Lambda}{4} \sum_{abcd} \left(\delta_{bd} - \left\{ \begin{matrix} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{matrix} \right\} \right) \\ &\quad \times (R_{ab} R_{bc} - \delta_{ac}) (R_{cd} R_{da} - \delta_{ac})\end{aligned}\quad (6.3)$$

$$\begin{aligned}&:= \Lambda \left(\frac{1}{2} \sum_{ab} g_{ab} R_{ab} R_{ba} \right. \\ &\quad \left. + \frac{1}{4} \sum_{abcd} g_{abcd} R_{ab} R_{bc} R_{cd} R_{da} + \text{const} \right),\end{aligned}\quad (6.4)$$

where in the second equality we have used the orthogonality relation (2.14).⁴⁰ The couplings are

$$g_{ab} = -1 + \frac{1}{2} \sum_{cd} (\delta_{ac} + \delta_{bd}) \left\{ \begin{matrix} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{matrix} \right\},\quad (6.5)$$

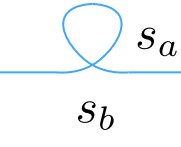
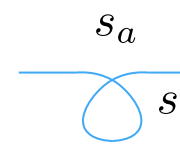
$$g_{abcd} = \frac{1}{2} (\delta_{ac} + \delta_{bd}) - \left\{ \begin{matrix} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{matrix} \right\}.\quad (6.6)$$

From (6.4), we may determine the propagator and four-point interaction vertex for the R matrix that is used to compute ‘t Hooft diagrams. The propagator is

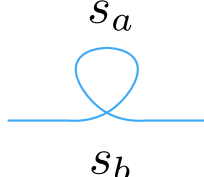
³⁹To restore the e^{S_0} factors, one should insert e^{-S_0} next to every $6j$ symbol, e^{S_0} next to every $\rho(s)$, and e^{-S_0} next to every $\delta(s_1 - s_2)$. This preserves the orthogonality relation (2.16).

⁴⁰Before we use (2.14), we replace the sum \sum_d by an integral $\int ds_d \rho(s_d)$. This is because we work at large N and consider only disk amplitudes in this section. We will make similar substitutions when we write down the diagrammatic rules of this model.

³⁸In Sec. VIII, we call these parameters q_A , q_B , and \tilde{q} .

$$\frac{s_a}{s_b} = \frac{1}{\Lambda} \left(-1 + \frac{1}{2} \text{diagram} + \frac{1}{2} \text{diagram} \right)^{-1} \quad (6.7)$$



where on the left-hand side, each index line in the double-line propagator corresponds to an energy (this convention was used in Sec. IV). On the right-hand side, the diagrams should be evaluated using the gravitational Feynman rules. That is,

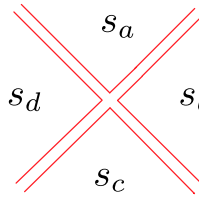
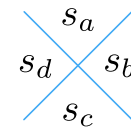
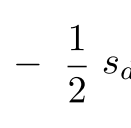
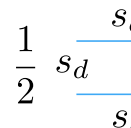
$$\text{diagram} = \int_0^\infty ds \rho(s) \left\{ \begin{matrix} \Delta & s_a & s_b \\ \Delta & s_a & s \end{matrix} \right\}. \quad (6.8)$$


This integral actually diverges because at high energies (see Appendix A)

$$\left\{ \begin{matrix} \Delta & s_a & s_b \\ \Delta & s_a & s \end{matrix} \right\} \sim e^{-\pi s}, \quad (s \rightarrow \infty), \quad (6.9)$$

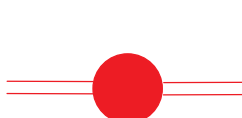

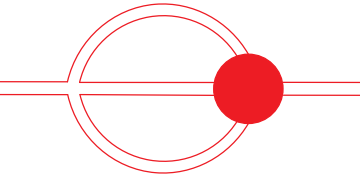
while $\rho(s) \sim e^{2\pi s}$ at large s . We will see however that such divergences will formally cancel out of our final expressions. We will deal with these divergences more carefully in a regularized model in the next subsection.

The quartic interaction vertex for R is given by

$$\text{diagram} = \Lambda \left(\text{diagram} - \frac{1}{2} \text{diagram} - \frac{1}{2} \text{diagram} \right) \quad (6.10)$$





where again the right-hand side should be evaluated using the gravitational Feynman rules. The first term on the right-hand side is the $6j$ symbol. The last two terms are delta functions $\frac{\delta(s_a - s_c)}{\rho(s_a)}$ and $\frac{\delta(s_b - s_d)}{\rho(s_b)}$ respectively.

From the diagrammatic rules above we may compute planar correlators of the R matrix. Any closed loop corresponds to an integral over the energy with measure $ds\rho(s)$. Next, we introduce a Schwinger-Dyson equation that relates the couplings in the matrix potential to the correlators. The equation is represented diagrammatically as follows:

$$\text{diagram} = \text{diagram} + \text{diagram} \quad (6.11)$$




The left-hand side represents the sum over all planar two-point diagrams and is equal to $\langle R_{ab} R_{ba} \rangle_{\text{disk}}$. The red blob on the right-hand side represents the sum over all planar (connected and disconnected) four-point diagrams and is given by $\langle R_{ab} R_{bc} R_{cd} R_{da} \rangle_{\text{disk}}$.⁴¹ To understand the above

equation, note that the leftmost double-line can either directly connect to the rightmost double-line, or it can connect to a vertex. The other three legs of this vertex and the rightmost double-line become the external lines of the four-point red blob, which accounts for all the remaining diagrams that contribute to the left-hand side. The Schwinger-Dyson equation (6.11) can also be derived from the identity

$$\int dR \frac{\partial}{\partial R_{ab}} (R_{ab} e^{-V(R)}) = 0 \quad (6.12)$$

⁴¹By $\langle R_{ab} R_{bc} R_{cd} R_{da} \rangle_{\text{disk}}$, we are referring to the smooth function of four energies that is obtained by performing inverse Laplace transforms on the four-point function in the planar limit of the matrix model. See Sec. VII for more comments.

and is expressed as

$$\langle R_{ab}R_{ba} \rangle_{\text{disk}} = \frac{1}{\Lambda} g_{ab}^{-1} - g_{ab}^{-1} \sum_{cd} g_{abcd} \langle R_{ab}R_{bc}R_{cd}R_{da} \rangle_{\text{disk}}, \quad (6.13)$$

where g_{ab}, g_{abcd} are defined in (6.5), (6.6).

We will now show that the gravitational correlators that correspond to $\langle R_{ab}R_{ba} \rangle_{\text{disk}}$ and $\langle R_{ab}R_{bc}R_{cd}R_{da} \rangle_{\text{disk}}$ solve

$$\begin{aligned} & \left(\text{---} \text{---} \text{---} - \frac{1}{2} \text{---} \text{---} \text{---} - \frac{1}{2} \text{---} \text{---} \text{---} \right) \left(\text{---} \text{---} \text{---} + \text{---} \text{---} \text{---} + \text{---} \text{---} \text{---} \right) \\ &= -1 + \frac{1}{2} \text{---} \text{---} \text{---} + \frac{1}{2} \text{---} \text{---} \text{---} \\ &+ \text{---} \text{---} \text{---} - \frac{1}{2} \text{---} \text{---} \text{---} - \frac{1}{2} \text{---} \text{---} \text{---} \end{aligned} \quad (6.14)$$

In the lhs, a product of two diagrams means that we connect the three lines in the middle, which corresponds to doing the sum \sum_{cd} . To get the rhs, we expanded the first line into nine terms and simplified the algebra.

Now we would like to argue that the last three terms in (6.14) cancel out. We know that 6j-symbols obey “unlacing” rules, such as the orthogonality relation (2.14). So heuristically, one might expect that more unlacing relations are obeyed, such that in the first term in the last line of (6.14) we can move the circle away from the horizontal line. Then the last three terms in (6.14) would cancel out. In practice, if we compute the integral corresponding to the first term in the last line of (6.14), we find a divergent result $\delta(0)$, delta

(6.11) or equivalently (6.13). Of the two terms on the right-hand side of (6.13), the first is order Λ^{-1} , while the second is order one in the large Λ expansion. Hence, we may drop the first term. To evaluate the second term, we replace the interaction vertex by the sum over the three diagrams in (6.10), and we replace the red blob by a sum over the three four-point gravitational Feynman diagrams (two uncrossed and one crossed), which correspond to the three terms in (5.3).⁴² We then find that $-g_{abcd} \langle R_{ab}R_{bc}R_{cd}R_{da} \rangle_{\text{disk}}$ is given by

function at zero argument. This can be seen from the orthogonality relation (2.14). The latter two terms are proportional to a square of the delta function and therefore contain $\delta(0)$ as well. For now we assume that such unlacing rules work and the last three terms cancel out. We will deal with this more carefully in the next subsection. The remaining terms in the second line of (6.14) are canceled by g_{ab}^{-1} in (6.13), giving one in total and Schwinger-Dyson equation becomes $\langle R_{ab}R_{ba} \rangle_{\text{disk}} = 1$, which is the correct result.

The next Schwinger-Dyson equation we consider is diagrammatically represented as follows:

$$\text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} = \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} - \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} - \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \quad (6.15)$$

where the blob with six external double-lines represents the sum over all planar (connected and disconnected) six-point diagrams. The blob labeled “C” on the left represents the sum over all connected four-point planar diagrams. To understand this equation, note that the leftmost double-line must be attached to a vertex, or else

the resulting diagram cannot be both connected and planar. The other three double-lines of the vertex together with the three double-lines on the right-hand side should be connected using additional propagators and vertices in all possible ways such that the entire diagram is connected. There are some contributions to the six-point blob that result in a disconnected diagram, and so these contributions are subtracted off as shown. Equation (6.15) can also be derived from the identity

⁴²The red four-point blob in (6.11) is by definition equal to the sum of the three terms on the left-hand side of (5.3).

$$\int dR \frac{\partial}{\partial R_{ad}} (R_{ab} R_{bc} R_{cd} e^{-V(R)}) = 0. \quad (6.16)$$

If we plug the gravitational correlators into the expressions for the matrix model correlators, the “C” blob becomes a $6j$ symbol, and the blob with six external double-lines becomes a sum over 15 gravitational Feynman diagrams, but 6 of those are subtracted off. The right-hand side of (6.15) is order one in the large N limit. Plugging in the diagrammatic rules (6.7) and (6.10), we may again verify that (6.15) holds, assuming unlacing rules. This can again be dealt with more carefully using the regularization in the next section.

We can continue to write down more Schwinger-Dyson equations where the left-hand side is analogous to the left-hand side of (6.15), except with additional double-lines emanating from the right side of the “C” blob. This entire class of Schwinger-Dyson equations is solved by the gravitational Feynman rules. This is evidence that (3.20) correctly reproduces JT gravity minimally coupled to a scalar at disk level.

B. Backing away from the double-scaling limit and q -deformation

In the previous subsection we considered the matrix model directly in the double-scaling limit and gave a suggestive argument that JT correlators solve the Schwinger-Dyson equations. However, some of the expressions were not completely well-defined due to high-energy divergences. This is to be expected in the double-scaling limit. To give a more precise matrix model description, in this section we consider a particular way to back away from the double-scaling limit. The model will depend on the size of matrices N and an extra parameter q . We will show that in the double-scaling limit $N \rightarrow \infty$, $q \rightarrow 1$, while keeping a particular combination (to be specified below) of N and q fixed, this matrix model is solved by JT correlators.

Let us consider a quartic matrix model of $N \times N$ matrices

$$V(R) = N \left(\frac{1}{2} \sum_{ab} g_{ab} |R_{ab}|^2 + \frac{1}{4} \sum_{abcd} g_{abcd} R_{ab} R_{bc} R_{cd} R_{da} \right). \quad (6.17)$$

To get ourselves oriented, let us first consider the usual large- N limit (no double-scaling). At large N we take the couplings to be order one; $g_{ab}, g_{abcd} \sim 1$, as it is common in matrix models. We can again derive the planar $N \rightarrow \infty$ Schwinger-Dyson equation from (6.12)

$$\frac{1}{N} = g_{ab} \langle |R_{ab}|^2 \rangle_{\text{disk}} + \sum_{cd} g_{abcd} \langle R_{ab} R_{bc} R_{cd} R_{da} \rangle_{\text{disk}}. \quad (6.18)$$

At large N , the correlators scale in the standard way

$$\langle |R_{ab}|^2 \rangle_{\text{disk}} \sim \frac{1}{N}, \quad (6.19)$$

$$\langle R_{ab} R_{bc} R_{cd} R_{da} \rangle_{\text{disk}} \sim \frac{1}{N^3}. \quad (6.20)$$

For the four-point function the disconnected part scales as $\langle R_{ab} R_{bc} \rangle \langle R_{cd} R_{da} \rangle_{\text{disk}} \sim \frac{1}{N^2} \delta_{ac}$, but the extra Kronecker δ_{ac} can be thought of as effectively $\frac{1}{N}$. The scaling is such, that if we sum over all indices the result is of order N : $\sum_{ab} \langle |R_{ab}|^2 \rangle_{\text{disk}} \sim N^2 \cdot \frac{1}{N} = N$ and $\sum_{abcd} \langle R_{ab} R_{bc} R_{cd} R_{da} \rangle_{\text{disk}} \sim N^4 \cdot \frac{1}{N^3} = N$. This is indeed what we expect for disk correlators.

Using the above scaling it is easy to check that all three terms in (6.18) are of the same order and must be kept in the large- N limit

$$\frac{1}{N} = \underbrace{g_{ab}}_{\sim 1} \underbrace{\langle |R_{ab}|^2 \rangle_{\text{disk}}}_{\sim \frac{1}{N}} + \sum_{\substack{cd \\ \sim N^2}} \underbrace{g_{abcd}}_{\sim 1} \underbrace{\langle R_{ab} R_{bc} R_{cd} R_{da} \rangle_{\text{disk}}}_{\sim \frac{1}{N^3}}. \quad (6.21)$$

Now we turn to the double-scaling limit. In this case we will see that the situation is different. In the limit that we define below, the lhs of (6.18) can be dropped, while the rest is solved by JT correlators.

To define the double-scaling limit we need to introduce a certain q -deformation of JT correlators. This will be described in much more detail in Sec. VIII A. Here we give a few results necessary for this section. First, instead of the Schwarzian density of states $\rho(s)$ we consider its q -deformation

$$\rho_q(s) = \frac{1}{2\pi\Gamma_q(\pm 2is)}, \quad \int_0^{\pi/|\log q|} ds e^{S_0} \rho_q(s) = e^{S_0} N_q, \quad (6.22)$$

where N_q is

$$N_q = \frac{1}{|\log q|(1-q)^2(q;q)_\infty^3}. \quad (6.23)$$

In particular, $N_q \rightarrow \infty$ as $q \rightarrow 1$. The density of states $e^{S_0} \rho_q(s)$ is supported on a finite interval $(0, \frac{\pi}{|\log q|})$, as in a 1-cut matrix model. It is also normalized as shown above. Therefore, it is natural to identify

$$N = e^{S_0} N_q. \quad (6.24)$$

The double-scaling limit that we will consider is defined by taking $N \rightarrow \infty$, $q \rightarrow 1$, while keeping $\frac{N}{N_q} = e^{S_0}$ finite.

In addition, we also deform the $6j$ -symbol to a q -deformed one

$$J_{abcd} \equiv \left\{ \begin{array}{ccc} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{array} \right\}_q. \quad (6.25)$$

The definition of the q -deformed $6j$ -symbol is given in Appendix A. For now, we only need to know that in the limit $q \rightarrow 1$ it gives back the classical $6j$ -symbol appearing in JT correlators. We take the exact correlators to be

$$\langle |R_{ab}|^2 \rangle_{\text{disk}} = \frac{N_q}{N} = e^{-S_0}, \quad (6.26)$$

$$\begin{aligned} \langle R_{ab} R_{bc} R_{cd} R_{da} \rangle_{\text{disk}} &= \left(\frac{N_q}{N} \right)^2 (\delta_{ac} + \delta_{bd}) \\ &+ \left(\frac{N_q}{N} \right)^3 \left\{ \begin{array}{ccc} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{array} \right\}_q \\ &= e^{-2S_0} \left(\delta_{ac} + \delta_{bd} + e^{-S_0} \left\{ \begin{array}{ccc} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{array} \right\}_q \right). \end{aligned} \quad (6.27)$$

$$(6.28)$$

It is interesting to note that the two-point function is enhanced by a factor N_q in comparison to the large- N

scaling (6.19). Similarly for four-point and higher correlators.

Now we would like to find the couplings g_{ab}, g_{abcd} such that in the double-scaling limit the Schwinger-Dyson equation (6.18) is solved by correlators (6.26)–(6.28). We first define a smeared delta function

$$\begin{aligned} \delta(k, k'; s_a, s_c | q) &:= \int_0^{\pi/|\log q|} ds \rho_q(s) \left\{ \begin{array}{ccc} \Delta & s_a & k \\ \Delta & s_c & s \end{array} \right\}_q \left\{ \begin{array}{ccc} \Delta & s_a & k' \\ \Delta & s_c & s \end{array} \right\}_q. \end{aligned} \quad (6.29)$$

In the limit $q \rightarrow 1$ this equation reduces to the orthogonality relation of the classical $6j$ -symbols (2.16) and

$$\lim_{q \rightarrow 1} \delta(k, k'; s_a, s_c | q) = \frac{\delta(k - k')}{\rho(k)}. \quad (6.30)$$

But for $0 < q < 1$ the function $\delta(k, k'; s_a, s_c | q)$ is smooth and finite. It is bell shaped and defines a smearing of the delta function.

Now we are ready to define the couplings. We choose them to be

$$g_{ab} = -e^{-S_0} \frac{1}{2} \sum_{cd} [\delta_{ac} \delta(s_b, s_d; s_a, s_c | q) + \delta_{bd} \delta(s_a, s_c; s_b, s_d | q)] \quad (6.31)$$

$$+ e^{-S_0} \sum_{cd} J_{abcd} \left(\delta_{ac} + \delta_{bd} - e^{-S_0} \frac{1}{2} [\delta(s_a, s_c; s_b, s_d | q) + \delta(s_b, s_d; s_a, s_c | q)] \right), \quad (6.32)$$

$$g_{abcd} = \frac{1}{2} \delta(s_a, s_c; s_b, s_d | q) + \frac{1}{2} \delta(s_b, s_d; s_a, s_c | q) - J_{abcd}. \quad (6.33)$$

Several comments are in order. The powers of e^{S_0} are chosen to be such that if we substitute $e^{S_0} = \frac{N}{N_q}$ and take $N \rightarrow \infty$, q —fixed limit, the couplings are of order 1. On the other hand, in the double-scaling limit $N \rightarrow \infty$, $q \rightarrow 1$ with $e^{S_0} = \frac{N}{N_q}$ —fixed we recover the couplings (6.5) and (6.6) where we now restored factors of e^{S_0} . To see the latter, the Kronecker delta and discrete sums are substituted in the double-scaling limit as

$$\delta_{ab} \rightarrow \frac{\delta(s_a - s_b)}{e^{S_0} \rho_q(s_a)}, \quad \sum_a \rightarrow \int_0^{\pi/|\log q|} ds e^{S_0} \rho_q(s). \quad (6.34)$$

Now we can check that the Schwinger-Dyson equation is satisfied in the double-scaling limit. First, the lhs of (6.18) can be dropped. Some of the individual terms in the rhs are in fact divergent in double-scaling limit, though of course all divergences cancel, as we now show. We compute

$$\sum_{cd} g_{abcd} \langle R_{ab} R_{bc} R_{cd} R_{da} \rangle_{\text{disk}} = \sum_{cd} \left(\frac{1}{2} \delta(s_a, s_c; s_b, s_d | q) + \frac{1}{2} \delta(s_b, s_d; s_a, s_c | q) - J_{abcd} \right) \quad (6.35)$$

$$e^{-2S_0} (\delta_{ac} + \delta_{bd} + e^{-S_0} J_{abcd}) \quad (6.36)$$

$$= -g_{ab} \langle |R_{ab}|^2 \rangle_{\text{disk}} \quad (6.37)$$

$$+ e^{-2S_0} \frac{1}{2} \sum_{cd} [\delta_{ac} \delta(s_a, s_c; s_b, s_d | q) + \delta_{bd} \delta(s_b, s_d; s_a, s_c | q)] \quad (6.38)$$

$$- e^{-3S_0} \sum_{cd} J_{abcd}^2 \quad (6.39)$$

$$= -g_{ab} \langle |R_{ab}|^2 \rangle_{\text{disk}}. \quad (6.40)$$

In the last equality we used that (6.38) and (6.39) cancel out. This follows from the definition of the smeared delta function (6.29), both give a smeared delta function at zero argument. We thus showed that the Schwinger-Dyson equation (6.18) is satisfied in the double-scaling limit with the lhs dropped.

In the double-scaling (IR) limit, we showed that we can drop the lhs of (6.18). This is equivalent to neglecting the bare propagator of the free theory, first term in the rhs of (6.11) and (6.13). It is interesting to note that this seems similar to the way Schwinger-Dyson equation is solved in the SYK model [37–40]. There in the low energy limit one neglects the bare UV part of the two-point function.

The next Schwinger-Dyson equation is derived from (6.16) and takes the form

$$\begin{aligned} & \frac{1}{N} (\delta_{ac} \langle |R_{ab}|^2 \rangle_{\text{disk}} + \delta_{bd} \langle |R_{cd}|^2 \rangle_{\text{disk}}) \\ &= g_{ad} \langle R_{ab} R_{bc} R_{cd} R_{da} \rangle_{\text{disk}} \\ &+ \sum_{ef} g_{adef} \langle R_{ab} R_{bc} R_{cd} R_{de} R_{ef} R_{fa} \rangle_{\text{disk}}. \end{aligned} \quad (6.41)$$

In Eq. (6.16) there is an extra term from differentiating R_{bc} giving $\delta_{ab} \delta_{cd} \langle R_{ab} R_{cd} \rangle$. However this contains too many deltas and, after summing over all energy indices, this would give a nonplanar contribution and is therefore suppressed at large N . The two-point coupling g_{ab} can be excluded using the first Schwinger-Dyson equation (6.18). After some algebra we obtain

$$\begin{aligned} & \sum_{ef} g_{adef} \langle R_{ab} R_{bc} R_{cd} \langle |R_{ad}|^2 \rangle_{\text{disk}} \\ & - |R_{da} \rangle \langle R_{ad}| \rangle R_{de} R_{ef} R_{fa} \rangle_{\text{disk}} \\ &= -\frac{1}{N} \langle R_{ab} R_{bc} R_{cd} R_{da} \rangle_{\text{disk, conn}}. \end{aligned} \quad (6.42)$$

In the lhs we have a six-point function up to subtractions. The subtraction is the product of two four-point functions. There are 15 chord diagrams contributing to the six-point function. However, if two R 's in $R_{ab} R_{bc} R_{cd}$ are connected by a chord, then such a term is canceled by the second term in the lhs. Therefore the only remaining contributions in the

lhs are chord diagrams where all three R 's in $R_{ab} R_{bc} R_{cd}$ are connected to one of the R 's in $R_{de} R_{ef} R_{fa}$. There are $3! = 6$ such chord diagrams

$$+ \quad (6.43)$$

We need to insert this and the coupling (6.33) into the Schwinger-Dyson equation (6.42). In the double-scaling limit we can drop the rhs in (6.42). The computation is somewhat involved and is easier to do pictorially. We checked that in the double-scaling limit $q \rightarrow 1$ it is indeed satisfied. Crucially, one has to use the Yang-Baxter equation for the (q -deformed) $6j$ -symbol.

One can consider more general Schwinger-Dyson equations that involve higher-point correlators. They can be derived similarly to (6.12) and (6.16).

VII. MATRIX MODEL POTENTIAL

We now return to the discussion of the matrix-model potential and argue that it is determined by the disk correlators. This is analogous to the statement that in a one-matrix model, the disk density of states determines the matrix potential and vice versa.

We will outline a systematic procedure for computing non-Gaussian corrections to the potential of \mathcal{O} , such that the matrix model correctly computes disk correlators with an arbitrary number of \mathcal{O} insertions. In Sec. VIII F, we explicitly compute the leading correction to the potential in a regulated model.

We write the matrix integral as follows:

$$\begin{aligned} \mathcal{Z} &= \int dH d\mathcal{O} e^{-V(H, \mathcal{O})} \\ &= \int dH d\mathcal{O} \exp \left(- \sum_a [V_{\text{SSS}}(E_a) + V_{c.t.}(E_a)] \right. \\ &\quad \left. - e^{S_0} \sum_{a,b} \mathcal{O}_{ab} \mathcal{O}_{ba} \frac{F_{ab}}{2} \right. \\ &\quad \left. + e^{S_0} \sum_{a,b,c,d} \mathcal{O}_{ab} \mathcal{O}_{bc} \mathcal{O}_{cd} \mathcal{O}_{da} \frac{G_{abcd}^{(4)}}{4} + \dots \right), \end{aligned} \quad (7.1)$$

where as usual we have chosen to work in the eigenbasis of H . We define F_{ab} and $G_{a_1 a_2 \dots a_{2n}}^{(2n)}$, $n \in \mathbb{Z}_{\geq 2}$, to be smooth real functions of the eigenvalues of H . That is, $F_{ab} = F(E_a, E_b)$ and $G_{abcd}^{(4)} = G^{(4)}(E_a, E_b, E_c, E_d)$. The \dots includes terms that are sixth and higher order in \mathcal{O} (we include even powers of \mathcal{O} only so that the model has a $\mathcal{O} \rightarrow -\mathcal{O}$ symmetry). The full set of coupling constants is specified by V_{SSS} and $V_{c.t.}$ together with the functions F_{ab} and $G_{a_1 a_2 \dots a_{2n}}^{(2n)}$ for $n \in \mathbb{Z}_{\geq 2}$. These functions are invariant under cyclic shifts of their indices (e.g. $G_{abcd}^{(4)} = G_{dabc}^{(4)}$) as well as reversals [e.g. $G_{abcd}^{(4)} = G_{dcba}^{(4)}$]. As in the previous section, we are working directly in the double-scaled limit, so the number of eigenvalues is infinite. The factors of e^{S_0} in the action ensure that 't Hooft diagrams have the correct factors of e^{S_0} according to their topology. In Sec. VIII C we will carefully consider a regulated model where the number of eigenvalues is large but finite.

As in the toy model, the requirement that the matrix integral (7.1) correctly computes the correlator $\langle \text{Tr} e^{-\beta H} \rangle$ at the level of the disk determines the counterterm potential $V_{c.t.}$ in terms of the other coupling constants. If we integrate out \mathcal{O} , the resulting matrix potential for H becomes a sum of multitrace terms. In analogy to Fig. 3, these multitrace terms generate \mathcal{O} bubble diagrams that correct the disk partition function in the SSS model. We choose the single-trace counterterm potential to cancel all of these corrections.

Next, we consider the disk two-point function. Some of the 't Hooft diagrams that contribute are shown below:

$$\text{Disk with red dot and lines} = \text{Disk with lines} + \text{Disk with red dot and lines} + \dots \quad (7.2)$$

The red blob is defined to be the sum over all planar 't Hooft diagrams with two external double-lines. As in (4.9) and (4.17), we have declined to draw any \mathcal{O} bubble diagrams or double-lines of the H matrix. These are automatically accounted for by the rule that each single-line loop corresponds to an integral over the corresponding energy with the measure $dE e^{S_0} \rho_0(E)$. Mathematically, (7.2) is represented by

$$\begin{aligned} &\langle \text{Tr} e^{-\beta_1 H} \mathcal{O} e^{-\beta_2 H} \mathcal{O} \rangle_{\text{disk}} \\ &= \int ds_a ds_b e^{S_0} \rho(s_a) e^{S_0} \rho(s_b) e^{-\beta_1 s_a^2} e^{-\beta_2 s_b^2} \langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}}, \end{aligned} \quad (7.3)$$

where $\rho(s)$ was defined in (2.8)⁴³ and $\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}}$ is a smooth function of s_a and s_b that is interpreted as a microcanonical two-point function.⁴⁴ To be precise, we define

$$\begin{aligned} &\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}} \\ &:= \frac{1}{\delta s_a e^{S_0} \rho(s_a)} \frac{1}{\delta s_b e^{S_0} \rho(s_b)} \langle \text{Tr} P(s_a) \mathcal{O} P(s_b) \mathcal{O} \rangle_{\text{disk}}, \end{aligned} \quad (7.4)$$

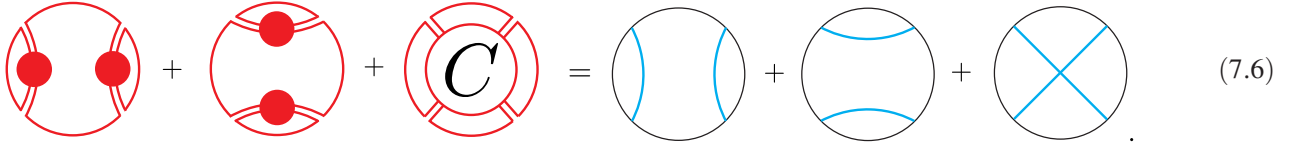
where $P(s_a)$ is a projection onto a microcanonical window centered around energy $E_a = s_a^2$ that has $\delta s_a e^{S_0} \rho(s_a)$ eigenvalues (to leading order in e^{S_0}). Taking inverse Laplace transforms after the large- N limit results in microcanonical-averaged correlators because the information about the fine-grained details of the spectrum is washed away at large N . We can only deduce the microcanonical-averaged correlators from the gravitational path integral. The requirement that the matrix model computes the correct two-point function implies that

$$\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}} = e^{-S_0} \Gamma_{ab}^\Delta. \quad (7.5)$$

We next consider the disk four-point function, which may be computed by summing 't Hooft diagrams with four external \mathcal{O} double-lines. These diagrams may be organized into connected and disconnected diagrams, as shown in (5.3), which we reproduce here,

⁴³For comparisons with gravitational amplitudes, it is more convenient to work with the s variables rather than the energies E . They are simply related by $E = s^2$.

⁴⁴Note that $\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}}$ is the same object as (6.80) of [19] up to factors of the density of states.



On the right-hand side we specify the gravitational four-point function that the ‘t Hooft diagrams on the left-hand side must reproduce. We may write

$$\begin{aligned}
 & \langle \mathcal{O}_{ab} \mathcal{O}_{bc} \mathcal{O}_{cd} \mathcal{O}_{da} \rangle_{\text{disk}} \\
 &= \langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}} \langle \mathcal{O}_{bc} \mathcal{O}_{cb} \rangle_{\text{disk}} \frac{\delta(s_b - s_d)}{e^{S_0} \rho(s_b)} \\
 &+ \langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}} \langle \mathcal{O}_{ad} \mathcal{O}_{da} \rangle_{\text{disk}} \frac{\delta(s_a - s_c)}{e^{S_0} \rho(s_a)} \\
 &+ \langle \mathcal{O}_{ab} \mathcal{O}_{bc} \mathcal{O}_{cd} \mathcal{O}_{da} \rangle_{\text{disk}, c}. \quad (7.7)
 \end{aligned}$$

A correlator with the subscript disk, c refers to a function of the energies that represents the sum over connected planar ‘t Hooft diagrams only.⁴⁵ Because the two-point function has already been fixed in (7.5), the first two terms on the left side of (7.6) reproduce the first two terms on the right-hand side. It follows that

$$\begin{aligned}
 & \langle \mathcal{O}_{ab} \mathcal{O}_{bc} \mathcal{O}_{cd} \mathcal{O}_{da} \rangle_{\text{disk}, c} \\
 &= e^{-3S_0} (\Gamma_{ab}^\Delta \Gamma_{bc}^\Delta \Gamma_{cd}^\Delta \Gamma_{da}^\Delta)^{1/2} \left\{ \begin{array}{cc} \Delta & s_a \quad s_b \\ \Delta & s_c \quad s_d \end{array} \right\}. \quad (7.8)
 \end{aligned}$$

Continuing in a similar fashion, we can compare the matrix model ‘t Hooft diagrams with gravitational Feynman diagrams as above and derive expressions for the connected $2n$ -point functions $\langle \mathcal{O}_{a_1 a_2} \cdots \mathcal{O}_{a_{2n} a_1} \rangle_{\text{disk}, c}$. It follows from the Feynman rules in Sec. II that the connected $2n$ -point function is given by a sum over all connected gravitational $2n$ -point Feynman diagrams (defined such that the bulk lines of the Feynman diagrams form a connected graph).

Until now, we have assumed that for a suitable choice of the couplings, the matrix integral correctly computes all of the gravitational disk correlators. It follows that the sum over all connected planar $2n$ -point ‘t Hooft diagrams is completely determined by the gravitational Feynman rules, as illustrated above for $n = 1, 2$. We now show how this data can be used to systematically determine the couplings.⁴⁶ First, we modify the gravitational Feynman rules such that

⁴⁵Here, we mean diagrams where the external \mathcal{O} double-lines are connected through the bulk.

⁴⁶We will soon emphasize that there should be multiple ways to choose the couplings such that the disk correlators of the matrix model agree with the gravitational answers in the appropriate scaling limit. Here, we just specify one way.

each crossing of two blue lines comes with an additional factor of ϵ , where $0 < \epsilon < 1$. We may determine the matrix potential order by order in ϵ . In particular, we write

$$V_{c.t.}(H) = \sum_{m=0}^{\infty} \epsilon^m V_{c.t.}^{(m)}, \quad (7.9)$$

$$F = \sum_{m=0}^{\infty} \epsilon^m F^{(m)}, \quad (7.10)$$

$$G^{(2n)} = \sum_{m=1}^{\infty} \epsilon^m G^{(2n), (m)}, \quad (7.11)$$

where we have omitted the indices on F_{ab} and $G_{a_1 \cdots a_{2n}}^{(2n)}$ for convenience. Note that $\langle \mathcal{O}_{a_1 a_2} \cdots \mathcal{O}_{a_{2n} a_1} \rangle_{\text{disk}, c}$ is $O(\epsilon^{n-1})$ in the ϵ expansion. That is, any connected gravitational $2n$ -point Feynman diagram must have at least $n - 1$ crossings. When $\epsilon = 0$, the matrix potential must reduce to \mathcal{Z}_{toy} in (4.3), which is why the sum in (7.11) starts from $m = 1$.

Working to first order in ϵ , we set $G^{(2n), (1)} = 0$ for $n \geq 3$. Then, we choose $G^{(4), (1)}$ to ensure that the matrix model computes the correct connected four-point function, which is order ϵ . The only contribution is from a tree diagram,

$$\begin{aligned}
 & \frac{\epsilon G_{a_1 a_2 a_3 a_4}^{(4), (1)}}{F_{a_1 a_2}^{(0)} F_{a_2 a_3}^{(0)} F_{a_3 a_4}^{(0)} F_{a_4 a_1}^{(0)}} \\
 &= \langle \mathcal{O}_{a_1 a_2} \mathcal{O}_{a_2 a_3} \mathcal{O}_{a_3 a_4} \mathcal{O}_{a_4 a_1} \rangle_{\text{disk}, c}. \quad (7.12)
 \end{aligned}$$

After choosing $G^{(4), (1)}$, we note that the four-point vertex appears in loop diagrams that contribute to the two-point function at order ϵ . We choose $F^{(1)}$ to cancel these contributions. Next, note that both the four-point vertex and the propagator contribute to \mathcal{O} bubble diagrams at order ϵ [and the result depends on $F^{(1)}$ and $G^{(4), (1)}$]. As in Fig. 3, we choose $V_{c.t.}^{(1)}$ to cancel the order ϵ contributions of \mathcal{O} bubble diagrams to the disk density of states. This defines the matrix potential to order ϵ . See Fig. 5 for a list of the diagrams that appear in this calculation to order ϵ .⁴⁷

⁴⁷Note that the loop integrations in these diagrams may diverge. In Sec. VIII, we will show that the gravitational Feynman rules can be further modified by a q -deformation such that the q -deformed density of states has compact support, which implies that the energies running in the loops are integrated over a finite range so that all the loop integrations converge. We will properly address this point in Sec. VIII. Until then, we ignore the issue of divergent loop integrals.

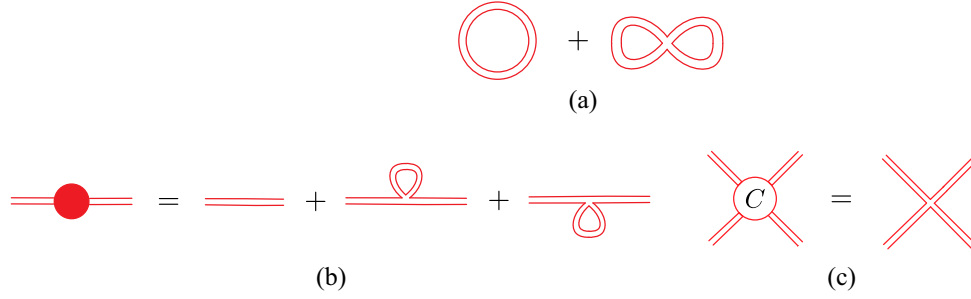


FIG. 5. To compute the correct gravitational observables to order ϵ , it suffices to let the quartic \mathcal{O} coupling be the only interaction to order ϵ . (a) \mathcal{O} bubble diagrams that contribute to the disk partition function at order ϵ . As in Fig. 3, these are canceled by counterterms in $V_{c.t.}$. (b) The ‘t Hooft diagrams that contribute to the two-point function. The order ϵ correction to the quadratic \mathcal{O} term is chosen to ensure that these diagrams sum to the known gravitational answer for the two-point function. (c) The value of the quartic \mathcal{O} coupling is determined by matching the four-point tree diagram to the known gravitational answer for the connected four-point function.

We now outline the general procedure to determine the matrix potential to any order in ϵ . The basic idea is that to any given order in ϵ , there are a finite number of nonzero connected planar $2n$ -point functions. Hence, we only need to adjust finitely many terms in (7.11) to make the matrix model correctly compute the gravitational correlators to a given order in ϵ . Thus, it is possible to systematically determine the matrix potential order by order in ϵ .

Let us explain the procedure in more detail. First, assume that the matrix potential has been chosen to order ϵ^{p-1} such that the matrix model correctly computes all of the connected correlators to order ϵ^{p-1} . Assume also that the couplings obey $G^{(2n),(m)} = 0$ for $n > 1 + m$. As shown above, these assumptions are true for $p = 2$, which we take as the base case of an inductive argument. To determine the couplings at order ϵ^p , we first choose $G^{(2n),(p)} = 0$ for $n > p + 1$. Next, note that the only contribution from $G^{(2p+2),(p)}$ to the connected $(2p+2)$ -point function at order ϵ^p is in a tree-level connected $(2p+2)$ -point diagram with a single $(2p+2)$ -point vertex,

$$\frac{\epsilon^p G_{a_1 \dots a_{2p+2}}^{(2p+2),(p)}}{F_{a_1 a_2}^{(0)} \dots F_{a_{2p+2} a_1}^{(0)}}. \quad (7.13)$$

We choose $G^{(2p+2),(p)}$ such that the sum of (7.13) and all of the other diagrams that contribute (which involve lower-point vertices only) yield the correct order ϵ^p result for $\langle \mathcal{O}_{a_1 a_2} \dots \mathcal{O}_{a_{2p+2} a_1} \rangle_{\text{disk}, c}$. Next, we consider the connected $2p$ -point function at order ϵ^p . This cannot depend on $G^{(2n),(p)}$ for $n < p$ because any diagram with a $(2n)$ -point vertex with $n < p$ must have at least one other vertex, and every vertex is at least order ϵ . Hence, the only order ϵ^p terms in the matrix potential that contribute to the $(2p)$ -point function at order ϵ^p are $G^{(2p),(p)}$ and $G^{(2p+2),(p)}$, and $G^{(2p),(p)}$ only contributes as part of a tree diagram. We choose $G^{(2p),(p)}$ such that the sum of this tree diagram and all the other diagrams yields the correct result for the $(2p)$ -point function. We may continue to choose $G^{(2n),(p)}$ for

successively lower values of n together with $F^{(p)}$ to ensure that the matrix model correctly computes all of the connected $2n$ -point functions at order ϵ^p . Finally, we choose $V_{c.t.}^{(p)}$ to cancel the corrections from \mathcal{O} bubble diagrams to the disk density of states. Thus, it is possible to choose the couplings to all orders in ϵ such that the matrix model agrees with the gravitational answers at disk level. After determining the matrix potential to all orders in ϵ , we can take $\epsilon \rightarrow 1$ at the end of any calculation.

We expect that there are many ways to define a double-scaled matrix model whose disk correlators agree with those of the gravitational theory. While we have described above a specific procedure for determining a suitable set of couplings, there could be other ways to determine the couplings. For instance, in the previous paragraph we modified the gravitational Feynman rules by including an additional factor of ϵ whenever two blue lines cross. We will refer to the associated matrix model as the “ q -deformed matrix model” for reasons that will become clear in the next section. Another way to modify the gravitational Feynman rules is to weight each connected $2n$ -point function with a factor of ϵ^{n-1} . The inductive argument outlined above can still be used to write the matrix potential in an ϵ expansion. The associated matrix model will be referred to as the “Selberg matrix model.” The $\epsilon \rightarrow 1$ limit corresponds to the double-scaling limit of the matrix model. We use the term “regulated matrix model” to refer to either of these models away from the double-scaling limit. The q -deformed and Selberg matrix models are double-scaled in different ways, but their disk correlators agree. However, as we will see later, their connected two-boundary (or double-trumpet) correlators disagree. Thus, it is important to distinguish between the two different double-scaling limits.

An important question for either the q -deformed or Selberg matrix models is whether the potential, written as a power series expansion in ϵ , actually converges for $0 < \epsilon < 1$. We want the matrix potential to be defined for ϵ in this range, so that the $\epsilon \rightarrow 1$ limit defines a scaling limit

of a well-defined matrix model. We do not have a rigorous proof that the sums in (7.9)–(7.11) converge for $0 < \epsilon < 1$. However, in the next section, we will provide highly nontrivial evidence in favor of the conclusion that the sums converge for the q -deformed matrix model. This evidence follows from the relationship between the q -deformed matrix model and the double-scaled SYK model, which was studied in [13]. We do not have an analogous argument for why the matrix potential converges in the Selberg matrix model. However, in the remainder of this paper, we will assume that the Selberg matrix model potential is well-defined for $0 < \epsilon < 1$ so that we can compute the double-trumpet in this model.

Even if the matrix potential is well-defined in the regulated models, these models may be nonperturbatively ill-defined. As an analogy, note that the potential $V(x) = x^2 - x^4$ has a local minimum at $x = 0$, which implies that the associated single-matrix model has a perturbatively stable saddle with a single-cut density of states centered on $x = 0$. However, the matrix integral itself is nonperturbatively ill-defined because the potential is unbounded below. We note here that the matrix model in [9] is nonperturbatively ill-defined (at least when the eigenvalue contour is \mathbb{R}), and the regulated matrix models we consider also appear to be nonperturbatively ill-defined. It would be interesting to find nonperturbatively well-defined models with the same genus expansion as the regulated models, in analogy to [41–49] but for two-matrix models.

VIII. THE REGULATED TWO-MATRIX MODELS

In the previous section we showed that it is possible to systematically determine the coupling constants of single-trace, two-matrix models that compute the disk correlators of JT gravity minimally coupled to a scalar field. When computing ‘t Hooft diagrams in these models, one encounters divergent loop integrals due to the noncompact support of the disk density of states $\rho(s)$. In order for the correlators to be finite in the double-scaling limit, the coupling constants cannot be well-defined in the double-scaling limit. This is simply because counterterms are needed to cancel the loop divergences.

We introduced two different regulated matrix models whose double-scaling limits reproduce the gravitational disk correlators. We referred to these models as the q -deformed model and the Selberg model. However, we did not explain how the divergent loop integrals should be regulated in these models. In this section, we begin by carefully defining the q -deformed matrix model. We will explain how the construction of the Selberg model differs. Then, we present a nontrivial calculation in support of the conclusion that the matrix potential of the q -deformed model is well-defined.

The reader who is mainly interested in the ‘t Hooft diagram computations in Sec. IX can skip most of this section without loss of continuity. To understand the results

in Sec. IX (aside from part of Secs. IX A 1 and IX D 1), the only important point from this section is that the special functions appearing in the gravitational Feynman rules may be deformed by a parameter $q \in [0, 1]$ (where $q = 1$ corresponds to the original, undeformed Feynman rules). We use these q -deformed special functions to regulate the aforementioned loop divergences. These special functions are defined in Appendix A.

A. Summary of chord diagram combinatorics

The q -deformed model has close connections with the results of [13], which we review in this subsection. The authors of [13] studied the SYK model in the double-scaled limit. If we write the SYK Hamiltonian as

$$H_{\text{SYK}} = i^{p/2} \sum_{i \leq i_1 < \dots < i_p \leq N} J_{i_1 i_2 \dots i_p} \psi_{i_1} \dots \psi_{i_p}, \quad (8.1)$$

then the double-scaled limit is a large- N limit where

$$\lambda \equiv \frac{2p^2}{N} \quad (8.2)$$

is held fixed. If we define

$$q = e^{-\lambda}, \quad (8.3)$$

then [13] showed that the moments of H_{SYK} only depend on q and may be computed by summing chord diagrams. A $2n$ -point chord diagram is defined to be a circle with $2n$ labeled points on the circumference and n chords. Each chord connects two points and each point is attached to exactly one chord. An example of a chord diagram is given in Fig. 6. Each chord diagram is assigned a value of q^n , where n is the number of involuntary chord crossings [we always assume that $q \in [0, 1]$]. The authors of [13] computed an elegant formula for the sum over all $2n$ -point chord diagrams:

$$\begin{aligned} & \text{sum over all } 2n\text{-point chord diagrams} \\ &= \int_0^\pi \frac{d\theta}{2\pi} (q, e^{\pm 2i\theta}, q)_\infty \left(\frac{2 \cos \theta}{\sqrt{1-q}} \right)^{2n}. \end{aligned} \quad (8.4)$$

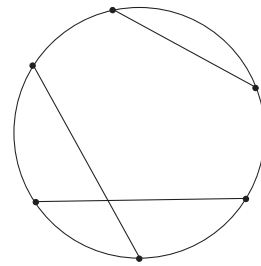


FIG. 6. An example of a six-point chord diagram. Pairs of points are connected with chords. This diagram is assigned the value q because there is one crossing.

Note that if one replaces $2n$ in (8.4) with an odd number, then (8.4) vanishes.

The authors of [13] also considered chord diagrams with two chord species. A two-species chord diagram is defined to be a circle with some number of labeled points on the circumference, and each point is either of type A or type B. All of the points are attached to a chord, and each chord connects exactly two points of the same type. If there are an odd number of either type-A or type-B points, then it is not possible to pair up all of the points and the diagram is assigned a value of zero. Otherwise, the diagram is assigned a value of $q_A^{n_1} q_B^{n_2} \tilde{q}^{n_3}$, where n_1 is the number of involuntary crossings of A-type chords, n_2 is the number of involuntary crossings of B-type chords, and n_3 is the number of involuntary crossings between an A-type and B-type chords. Another result of [13] is that the sum over all two-species chord diagrams with a fixed configuration of points on the circle and a fixed configuration of B-type chords may be computed using a set of Feynman rules that mirror the rules described in Sec. II A.⁴⁸ Figure 7 represents a sum over such chord diagrams.

We will often set $q_A = q$, $\tilde{q} = q^\Delta$. The chord-diagram Feynman rules are as follows (special functions are explained in Appendix A):

- (i) For each boundary segment labeled by n , we include a factor of

$$\left(\frac{2 \cos \theta}{\sqrt{1 - q}} \right)^n, \quad (8.5)$$

where θ labels the disk-shaped region that is adjacent to the boundary segment.

- (ii) For each B-type point (given by a blue dot in Fig. 7), we include a factor of

$$\begin{aligned} & \left(\frac{(q^{2\Delta}; q)_\infty}{(q^\Delta e^{i(\pm\theta_1 \pm \theta_2)}; q)_\infty} \right)^{1/2} \\ &= (1 - q)^\Delta \left(\frac{|\log q|}{1 - q} N_q \right)^{1/2} \left(\frac{\Gamma_q(\Delta \pm is_1 \pm is_2)}{\Gamma_q(2\Delta)} \right)^{1/2} \end{aligned} \quad (8.6)$$

$$\equiv (1 - q)^\Delta \left(\frac{|\log q|}{1 - q} N_q \right)^{1/2} (\Gamma_{12,q}^\Delta)^{1/2}, \quad (8.7)$$

where $\theta_j = s_j |\log q|$ are associated to the two disk-shaped regions adjacent to the blue dot and we introduced for later convenience

$$N_q = \frac{1}{|\log q| (1 - q)^2 (q; q)_\infty^3}. \quad (8.8)$$

⁴⁸Technically, [13] only proved this result for diagrams with up to one intersection of B-type chords. For general two-species chord diagrams, this result is a well-motivated conjecture.

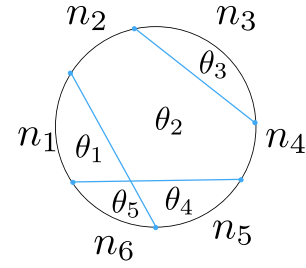


FIG. 7. The B-type points of this two-species chord diagram are labeled in blue. The labels n_1, n_2, \dots indicate how many A-type points (not shown) are between two B-type points. The sum over all ways of drawing the A-type chords is equal to a finite-dimensional integral which is determined by a set of Feynman rules. The Feynman rules require that each disk-shaped region is labeled with a θ parameter.

We expressed the rhs of (8.7) in a way that will be convenient for taking the limit $q \rightarrow 1$ below.

- (iii) For each involuntary crossing of two B-type chords, we include a factor ($q_A \equiv q$)

$$\begin{aligned} S_{\tilde{q}_1, \tilde{q}_2}^{q_A, q_B}(e^{i\theta_1}, e^{i\theta_2}, e^{i\theta_3}, e^{i\theta_4}) &\equiv q_B N_q \left\{ \begin{matrix} \Delta_1 & s_1 & s_2 \\ \Delta_2 & s_3 & s_4 \end{matrix} \right\}_q, \\ \theta_j &= s_j |\log q|, \end{aligned} \quad (8.9)$$

where $\tilde{q}_1 = q^{\Delta_1}$, $\tilde{q}_2 = q^{\Delta_2}$ and we need to set $\tilde{q}_1 = \tilde{q}_2 = \tilde{q} = q^\Delta$. Equation (8.9) for $\tilde{q}_1 \neq \tilde{q}_2$ is needed if one wants to introduce yet another chord species. The parameters $\theta_1, \dots, \theta_4$ represent the four disk-shaped regions that surround the crossing.

- (iv) After including all of the appropriate factors as specified above, we integrate over each $\theta \in (0, \pi)$ with the measure

$$\frac{d\theta}{2\pi}(q; q)_\infty(e^{\pm 2i\theta}; q)_\infty = \frac{1}{N_q} \frac{ds}{2\pi \Gamma_q(\pm 2is)} \quad (8.10)$$

$$\equiv \frac{1}{N_q} ds \rho_q(s), \quad (8.11)$$

where as usual $\theta = s |\log q|$ and $s \in (0, \pi/|\log q|)$. This measure is normalized

$$\int_0^\pi \frac{d\theta}{2\pi}(q; q)_\infty(e^{\pm 2i\theta}; q)_\infty = 1. \quad (8.12)$$

For example, the two-point correlator of B 's described by these rules is

$$\langle \text{Tr} B e^{-\beta_1 A} B e^{-\beta_2 A} \rangle_{\text{disk}, q} \quad (8.13)$$

$$= (1-q)^{2\Delta} \frac{|\log q|}{1-q} \frac{1}{N_q} \int_0^{\frac{\pi}{|\log q|}} \prod_{j=1}^2 \left[ds_j \rho_q(s_j) \right. \\ \left. \times \exp \left(-\beta_j \frac{2 \cos \theta_j}{\sqrt{1-q}} \right) \right] \frac{\Gamma_q(\Delta \pm is_1 \pm is_2)}{\Gamma_q(2\Delta)}. \quad (8.14)$$

B. Scaling limit

We now show that in the limit $q \rightarrow 1$ the q -deformed correlators reduce to JT correlators described in Sec. II A. The density of states $\rho_q(s)$ and the $6j$ -symbol (8.9) in the limit $q \rightarrow 1$ simply reduce to the corresponding JT values

$$\rho_q(s) \rightarrow \rho(s), \quad \left\{ \begin{matrix} \Delta_1 & s_1 & s_2 \\ \Delta_2 & s_3 & s_4 \end{matrix} \right\}_q \rightarrow \left\{ \begin{matrix} \Delta_1 & s_1 & s_2 \\ \Delta_2 & s_3 & s_4 \end{matrix} \right\}, \\ (q \rightarrow 1). \quad (8.15)$$

The counting of the factors N_q is as follows. Consider a chord diagram with $2n$ operators B . One can show that factors of N_q from all crossings (8.9) and densities of states (8.11) combine to give $N_q^{-(n+1)}$. Further, each operator B contributes $N_q^{1/2}$ (8.7), giving in total N_q^n from $2n$ operators. Combining these factors we have for each chord diagram $N_q^{-(n+1)} N_q^n = \frac{1}{N_q}$. To absorb the remaining factors in (8.7) we define an operator \mathcal{O}

$$B = (1-q)^\Delta \mathcal{O}. \quad (8.16)$$

The operator \mathcal{O} will reduce to the same \mathcal{O} considered in JT gravity in Sec. II A.

Now let's discuss the spectrum of A . Near the right edge in the limit $q \rightarrow 1$ it is

$$\frac{2 \cos(\theta)}{\sqrt{1-q}} = \frac{2 \cos(s|\log q|)}{\sqrt{1-q}} \quad (8.17)$$

$$\approx \frac{2}{\sqrt{1-q}} \left(1 - \frac{(1-q)^2}{2} s^2 + \dots \right), \quad (q \rightarrow 1). \quad (8.18)$$

To zoom into the right edge of the spectrum we define an operator H

$$A = \frac{2}{\sqrt{1-q}} \left(1 - \frac{(1-q)^2}{2} H \right) \quad (8.19)$$

such that in the limit $q \rightarrow 1$ the spectrum of H is s^2 with $s \in (0, \infty)$. This is the energy in JT limit. Combining everything together we find

$$\lim_{\substack{q \rightarrow 1 \\ q_B \rightarrow 1}} N \langle \text{Tr} \mathcal{O} e^{-\beta_1 H} \dots \mathcal{O} e^{-\beta_{2n} H} \rangle_{\text{disk}, q} \\ = \langle \text{Tr} \mathcal{O} e^{-\beta_1 H} \dots \mathcal{O} e^{-\beta_{2n} H} \rangle_{\text{disk}}, \quad (8.20)$$

where we also defined $N = e^{S_0} N_q$.⁴⁹ Here, the rhs is defined by the Feynman rules in JT gravity from Sec. II A. Later, in the matrix model N will be the size of matrices.

The parameter q_B is equivalent to ϵ in Sec. VII. The parameter q was not introduced in Sec. VII but is needed to carefully define the matrix model below. In the remainder of this paper, we refer to the limit $q \rightarrow 1, q_B \rightarrow 1$ as the “JT limit” of the q -deformed model.

C. Two two-matrix models regulating JT correlators

As we explained earlier, we expect JT gravity coupled to a free scalar to be dual to a double-scaled two-matrix model. To properly define the matrix model, we would like to describe it away from the double-scaled limit. We will consider two different ways to move away from the double-scaled limit: “ q -deformed matrix model” and “Selberg matrix model”.⁵⁰ The reason for considering two different regulators is the following. In the double-scaling limit both models will agree with the disk JT correlators. However, on the double-trumpet in the double-scaling limit they give rise to different results. The Selberg matrix model will agree with JT correlators on the double-trumpet, while the q -deformed matrix model will lead to a different answer.

We define the regulated two-matrix models by first specifying their regulated disk correlation functions. The matrix potential is then obtained by applying the algorithm in Sec. VII which produces a potential that computes a given set of disk correlators.

1. q -deformed matrix model

We define the “ q -deformed matrix model” such that its disk correlation functions are those of the double-scaled SYK model, which are described by the Feynman rules in Sec. VIII A. It depends on the parameters $q_A \equiv q$, q_B , and $\tilde{q} \equiv q^\Delta$. The potential may be found using the algorithm of Sec. VII.

In particular, the q -deformed model is a large N , ‘t Hooft-scaled two-matrix model with a single-trace potential, and we refer to the two Hermitian matrices as A and B . A single-trace, planar correlator of an arbitrary string of matrices, such as

⁴⁹Not to be confused with N in SYK.

⁵⁰The name “Selberg matrix model” is due to the fact that this model will reproduce the contribution of the 1-loop determinant on the double-trumpet, which in turn is computed by Selberg trace formula.

$$\lim_{N \rightarrow \infty} \frac{1}{N} \langle \text{Tr} A A B B A B A B A \rangle, \quad (8.21)$$

is computed by placing the same string of A - and B -type points on the circle of a two-species chord diagram and summing over all A - and B -type chord configurations.

2. Selberg matrix model

The Selberg regulator differs from the q -deformed regulator in two ways. First, we set $q_B = 1$. Second, we introduce a weight ϵ^{n-1} for each connected (in energy basis)⁵¹ $2n$ -point correlator. In other words, the blue chords of a Feynman diagram form a graph. For each connected subgraph with $2n$ external points, we include an additional factor of ϵ^{n-1} . For example, the connected four-point function using the Selberg regulator has a factor ϵ .

The q -deformed and Selberg regulators return the same results for the two and four-point functions if we identify q_B with ϵ . But the $2n$ -point correlators with $n \geq 3$ are different between the two regulators. The q -deformed regulator gives weights q_B according to the number of crossings in the Feynman diagram, while the Selberg model gives weights ϵ according to the number of external operators in each connected subgraph.

With these regulated gravitational Feynman rules, we may use the systematic procedure described toward the end of Sec. VII to construct a matrix potential to all orders in ϵ . The result accounts for (5.16) and (5.18). The double-scaling limit of the Selberg model corresponds to taking $q \rightarrow 1$ and then $\epsilon \rightarrow 1$ (the order is important, as we will see in Sec. IX). The matrix potential of the Selberg model also includes a double-trace term (5.17) that is added by hand; as explained around (5.17), its presence is required for the Selberg model to compute the correct empty double-trumpet.

In the remainder of this section, we focus only on the q -deformed model. We will revisit the Selberg model when we discuss the double-trumpet.

D. A single-matrix model warmup

Before we explain why the q -deformed two-matrix model should exist, we will first consider the easier problem of finding a one-matrix model whose single-trace, planar correlators are computed by sums over single-species chord diagrams. This is possible because the density of states (8.11) is supported on a finite interval $s \in (0, \frac{\pi}{|\log q|})$, just like in a single-cut matrix model.

We let M refer to a $N \times N$ Hermitian matrix, and the matrix integral is given by

$$\mathcal{Z}_q = \int dM e^{-N \text{Tr} V_q(M)}. \quad (8.22)$$

Our goal is to choose $V_q(M)$ such that the correlators of M obey

$$\lim_{N \rightarrow \infty} \frac{1}{N} \langle \text{Tr} M^k \rangle = \int_0^\pi \frac{d\theta}{2\pi} (q, e^{\pm 2i\theta}; q)_\infty \left(\frac{2 \cos \theta}{\sqrt{1-q}} \right)^k, \quad k \in \mathbb{Z}_{\geq 0}. \quad (8.23)$$

Let $\rho_{q,0}(E)$ be the normalized tree-level eigenvalue distribution of M , which obeys the saddle-point equation [9]

$$\int_{a_-}^{a_+} d\lambda \frac{\rho_{q,0}(\lambda)}{E - \lambda} = \frac{V'_q(E)}{2}, \quad E \in (a_-, a_+), \quad (8.24)$$

where the integral is a principal value integral, and the endpoints of the single-cut eigenvalue spectrum are

$$a_\pm = \pm \frac{2}{\sqrt{1-q}}. \quad (8.25)$$

Next, we define the rescaled quantities $\tilde{\rho}_{q,0}(x)$ and $\tilde{V}_q(x)$ such that

$$d\lambda \rho_{q,0}(\lambda) = dx \tilde{\rho}_{q,0}(x), \quad \tilde{V}_q(x) = V_q(\lambda), \quad \lambda = \frac{2x}{\sqrt{1-q}}. \quad (8.26)$$

Then, (8.24) becomes

$$\int_{-1}^1 dx \frac{\tilde{\rho}_{q,0}(x)}{y - x} = \frac{\tilde{V}'_q(y)}{2}. \quad (8.27)$$

Next, note that

$$(q, e^{\pm 2i\theta}; q)_\infty = \sum_{n \in \mathbb{Z}} (-1)^n q^{\frac{n^2}{2}} [q^{-\frac{n}{2}} + q^{\frac{n}{2}}] T_{2n}(\cos \theta), \quad (8.28)$$

where $T_n(\cos \theta) \equiv \cos n\theta$ is a Chebyshev polynomial of the first kind. Then, (8.23) implies that

$$\tilde{\rho}_{q,0}(x) = \frac{1}{2\pi \sqrt{1-x^2}} \left[\sum_{n \in \mathbb{Z}} (-1)^n q^{\frac{n^2}{2}} [q^{-\frac{n}{2}} + q^{\frac{n}{2}}] T_{2n}(x) \right]. \quad (8.29)$$

We will need to use the mathematical fact that

$$\int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} \frac{T_{2n}(x)}{y-x} = -\pi U_{2n-1}(y), \quad n \geq 0, \quad y \in (-1, 1), \quad (8.30)$$

where $U_n(\cos \theta) \equiv \frac{\sin((n+1)\theta)}{\sin \theta}$ is the Chebyshev polynomial of the second kind and the integral is a principal value

⁵¹That is before integrating over parameters s_j .

integral. We can then use (8.27) to directly compute $\tilde{V}'_q(y)$, and the result is

$$\frac{\tilde{V}'_q(y)}{2} = - \sum_{n=1}^{\infty} (-1)^n q^{\frac{n^2}{2}} [q^{-\frac{n}{2}} + q^{\frac{n}{2}}] U_{2n-1}(y). \quad (8.31)$$

Integrating the above with respect to y , we obtain

$$\tilde{V}_q(y) = - \sum_{n=1}^{\infty} (-1)^n q^{\frac{n^2}{2}} \frac{[q^{-\frac{n}{2}} + q^{\frac{n}{2}}]}{n} T_{2n}(y). \quad (8.32)$$

The expression for $V_q(\lambda)$ follows from (8.32) and (8.26):

$$V_q(\lambda) = - \sum_{n=1}^{\infty} (-1)^n q^{\frac{n^2}{2}} \frac{[q^{-\frac{n}{2}} + q^{\frac{n}{2}}]}{n} T_{2n}\left(\frac{\sqrt{1-q}}{2}\lambda\right). \quad (8.33)$$

We will refer to (8.33) in the next subsection. This single-matrix model is a regulated version of the SSS model [9].

E. Solvable limits of the q -deformed model

In Sec. VIII B, we showed that in a certain limit, the q -deformed two-matrix model computes the disk correlators of JT gravity minimally coupled to a scalar field. In this subsection, we consider other limits of the q -deformed model, in which the model is solvable. Although we cannot write a closed-form expression for the entire matrix potential of the model, we can write the potential in these solvable limits. In the next subsection, we provide non-trivial evidence that the model also exists away from these solvable limits.

We remind the reader that the circle of a two-species chord diagram represents a trace in the matrix model, and each A- or B-type point on the circle corresponds to an insertion of the matrix A or B into the trace. In the q -deformed two-matrix model, the expectation value of this trace at planar order (normalized by N) is given by a sum over all the chord configurations that connect the boundary points. Two intersecting A-type chords come with a factor of q_A , two intersecting B-type chords come with a factor of q_B , and an intersection of an A-type and B-type chord comes with a factor of \tilde{q} . We let $\mathcal{Z}_{q_A, q_B, \tilde{q}}$ denote the matrix integral of the regulated two-matrix model.

The first solvable limit that we consider is $\tilde{q} = 0$. In this limit, an A-type chord cannot intersect a B-type chord. In the matrix model, this means that every connected planar 't Hooft diagram contains either A double-lines or B double-lines, but not both. Thus, there can be no interaction terms between A and B in the potential. The matrix integral must take the form

$$\mathcal{Z}_{q_A, q_B, \tilde{q}=0} = \int dA dB e^{-N(\text{Tr } V_{q_A}(A) + V_{q_B}(B))}, \quad (8.34)$$

where V_q was defined in (8.33).

The next solvable limit of interest is $q_B = 0$. In this case, two B-type chords may not intersect. Given that a B-type chord may be interpreted as a bulk line in the JT limit and that the toy matrix model \mathcal{Z}_{toy} does not allow two bulk lines to cross, it is clear that the q -deformed model with $q_B = 0$ is an appropriate deformation of \mathcal{Z}_{toy} . In particular, the potential of the q -deformed model at $q_B = 0$ must be quadratic in B . The matrix integral for $q_B = 0$ is

$$\begin{aligned} \mathcal{Z}_{q_A, q_B=0, \tilde{q}} &= \int dA dB \exp\left(-N \text{Tr } V_{q_A}(A) \right. \\ &\quad \left. - N \left[\sum_{a,b} B_{ab} B_{ba} \frac{F_{ab}^{q_A, \tilde{q}}}{2} \right] - N \text{Tr } V_{c.t.}^{q_A, \tilde{q}}(A) \right), \end{aligned} \quad (8.35)$$

where V_{q_A} is defined in (8.33), and

$$\begin{aligned} F_{ab}^{q, \tilde{q}} &\equiv F^{q, \tilde{q}}(\lambda_a, \lambda_b) \equiv \tilde{F}^{q, \tilde{q}}\left(\frac{\sqrt{1-q}}{2}\lambda_a, \frac{\sqrt{1-q}}{2}\lambda_b\right), \\ \tilde{F}^{q, \tilde{q}}(\cos \theta_1, \cos \theta_2) &\equiv \frac{(\tilde{q} e^{i(\pm\theta_1 \pm \theta_2)}; q)_{\infty}}{(\tilde{q}^2; q)_{\infty}}, \end{aligned} \quad (8.36)$$

where λ_a refers to an eigenvalue of A , and

$$V_{c.t.}^{q, \tilde{q}}(E) = - \int_{a_-}^{a_+} d\lambda \rho_{q,0}(\lambda) \log F^{q, \tilde{q}}(\lambda, E), \quad (8.37)$$

where a_{\pm} is defined in (8.25). The smooth function $F^{q, \tilde{q}}(\lambda_a, \lambda_b)$ is chosen so that the two-point function of B agrees with the chord-diagram Feynman rule (8.7). As in Sec. IV, the counterterm potential is chosen so that the tree-level eigenvalue distribution of A is $\rho_{q,0}$. To evaluate (8.37), note that

$$-\log(\tilde{q} e^{\pm i\theta}; q)_{\infty} = 2 \sum_{n=1}^{\infty} \frac{\cos n\theta}{(1-q^n)n} \tilde{q}^n \quad (8.38)$$

and thus

$$-\log(\tilde{q} e^{\pm i\theta_1 \pm i\theta_2}; q)_{\infty} = 4 \sum_{n=1}^{\infty} \frac{\cos n\theta_1 \cos n\theta_2}{(1-q^n)n} \tilde{q}^n. \quad (8.39)$$

It then follows from (8.37) that up to an unimportant additive constant,

$$V_{c.t.}^{q, \tilde{q}}(E) = \tilde{V}_{c.t.}^{q, \tilde{q}}\left(\frac{\sqrt{1-q}}{2}E\right) \quad (8.40)$$

$$\tilde{V}_{c.t.}^{q,\tilde{q}}(x) = \sum_{n=1}^{\infty} (-1)^n q^{\frac{n^2}{2}} \frac{q^{-\frac{n}{2}} + q^{\frac{n}{2}}}{(1 - q^{2n})n} T_{2n}(x) \tilde{q}^{2n}, \quad (8.41)$$

and thus,

$$V_q(E) + V_{c.t.}^{q,\tilde{q}}(E) = \sum_{n=1}^{\infty} (-1)^n q^{\frac{n^2}{2}} T_{2n}\left(\frac{\sqrt{1-q}}{2} E\right) \frac{q^{-\frac{n}{2}} + q^{\frac{n}{2}}}{n} \times \left(-1 + \frac{1}{(1 - q^{2n})} \tilde{q}^{2n}\right). \quad (8.42)$$

The last solvable limit we consider is $q_A = 0$. This limit is the same as the previous limit up to exchanging the labels A and B . This is because the matrix potential of the q -deformed two-matrix model must have a symmetry that exchanges the A and B labels everywhere (including the A , B matrices themselves as well as the q_A , q_B parameters). This symmetry follows from the simple fact that the two types of chords in a two-species chord diagram are treated on an equal footing. Of course, the JT limit defined in Sec. VIII B breaks this symmetry.

F. Perturbation theory around a solvable limit

In the previous subsection, we described three solvable limits of the q -deformed matrix model, corresponding to the cases where either q_A , q_B , or \tilde{q} are set to zero. To test our conjecture that the q -deformed model exists, one can compute perturbative corrections to the potential around each of these limits and check that they are mutually consistent. In this subsection, we show that the $O(q_B)$ correction to the $q_B = 0$ potential in (8.35) is consistent with the other two solvable limits. In particular, we will first compute this correction when $\tilde{q} = 0$, and the result will

agree with the $O(q_B)$ term of the potential in (8.34). Then, we will compute the $O(1)$ and $O(q_A)$ parts of this correction in the small- q_A expansion (for arbitrary \tilde{q}) and find that the terms in the potential we obtain are symmetric under relabeling A and B .

Our strategy to compute corrections to the $q_B = 0$ potential is the same as in Sec. VII, where we explained how one could systematically compute corrections to the matrix potential in \mathcal{Z}_{toy} order by order in ϵ (see Fig. 5 for a description of this procedure to first order). Here, q_B plays the role of ϵ . We write the matrix integral of the regulated model as follows:

$$\begin{aligned} \mathcal{Z}_{q_A, q_B, \tilde{q}} &= \int dA dB \exp(-V_{q_A, q_B, \tilde{q}}), \\ V_{q_A, q_B, \tilde{q}} &= N \text{Tr}(V_{q_A}(A) + V_{c.t.}^{q_A, \tilde{q}}(A) + q_B \Delta V^{q_A, \tilde{q}}(A)) \\ &\quad + N \sum_{a,b} B_{ab} B_{ba} \frac{(F_{ab}^{q_A, \tilde{q}} + q_B \Delta F_{ab}^{q_A, \tilde{q}})}{2} \\ &\quad - N q_B \sum_{a,b,c,d} B_{ab} B_{bc} B_{cd} B_{da} \frac{G_{abcd}^{q_A, \tilde{q}}}{4} \\ &\quad + O(q_B^2) \end{aligned} \quad (8.43)$$

We allow for a generic $O(q_B)$ correction to the terms that are already present in (8.35), and we also include a new term that is quartic in B . This quartic term makes the planar connected four-point function of B nontrivial, which allows B-type chords to intersect each other. In order for the single-trace, planar correlators of this matrix model to agree to order q_B with the corresponding sums over two-species chord diagrams, we must set

$$\begin{aligned} G_{abcd}^{q,\tilde{q}} &= \tilde{G}^{q,\tilde{q}}\left(\frac{\sqrt{1-q}}{2} \lambda_a, \frac{\sqrt{1-q}}{2} \lambda_b, \frac{\sqrt{1-q}}{2} \lambda_c, \frac{\sqrt{1-q}}{2} \lambda_d\right), \\ \tilde{G}^{q,\tilde{q}}(\cos(\theta_1), \cos(\theta_2), \cos(\theta_3), \cos(\theta_4)) \\ &= \sqrt{\tilde{F}^{q,\tilde{q}}(\cos \theta_1, \cos \theta_2) \tilde{F}^{q,\tilde{q}}(\cos \theta_2, \cos \theta_3) \tilde{F}^{q,\tilde{q}}(\cos \theta_3, \cos \theta_4) \tilde{F}^{q,\tilde{q}}(\cos \theta_4, \cos \theta_1)} S_{\tilde{q}, \tilde{q}}^{q,1}(e^{i\theta_1}, e^{i\theta_2}, e^{i\theta_3}, e^{i\theta_4}). \end{aligned} \quad (8.44)$$

This is because in any planar 't Hooft diagram with a quartic B vertex, the propagators can be set to the uncorrected propagator $(NF_{ab}^{q_A, \tilde{q}})^{-1}$ (since the $O(q_B)$ correction to the propagator becomes an $O(q_B^2)$ correction to the entire diagram, which we are not interested in). Thus the quartic interaction in (8.43) is entirely determined by the chord-diagram Feynman rule for two crossing B-type chords, (8.9).

To determine the $O(q_B)$ correction to the two-point coupling $\Delta F_{ab}^{q,\tilde{q}}$, we choose $\Delta F_{ab}^{q_A, \tilde{q}}$ to cancel the corrections to $\langle B_{ab} B_{ba} \rangle_{\text{disk}}$ from loop diagrams involving the quartic coupling, such that

$$\langle B_{ab} B_{ba} \rangle_{\text{disk}} = (NF_{ab}^{q_A, \tilde{q}})^{-1}. \quad (8.45)$$

We obtain

$$\Delta F_{ab}^{q,\tilde{q}} = \int d\lambda_c \rho_{q_A,0}(\lambda_c) \frac{G_{abcb}^{q_A, \tilde{q}}}{F_{bc}^{q_A, \tilde{q}}} + \int d\lambda_c \rho_{q_A,0}(\lambda_c) \frac{G_{abac}^{q_A, \tilde{q}}}{F_{ac}^{q_A, \tilde{q}}}. \quad (8.46)$$

To determine the correction to the B -independent part of the potential $\Delta V^{q_A, \tilde{q}}$, we should integrate out B in (8.43) to leading order in q_B . The result is

$$\begin{aligned} \mathcal{Z}_{q_A, q_B, \tilde{q}} = \int dAdB \exp \Big(& -N \text{Tr} (V_{q_A}(A) + V_{c.t.}^{q_A, \tilde{q}}(A) + q_B \Delta V^{q_A, \tilde{q}}(A)) - \frac{1}{2} \sum_{a,b} \log F_{ab}^{q_A, \tilde{q}} - \frac{q_B}{2} \sum_{a,b} \frac{\Delta F_{ab}^{q_A, \tilde{q}}}{F_{ab}^{q_A, \tilde{q}}} \\ & + N^{-1} \frac{q_B}{2} \sum_{a,b,c} \frac{G_{abac}^{q_A, \tilde{q}}}{F_{ab}^{q_A, \tilde{q}} F_{ac}^{q_A, \tilde{q}}} + N^{-1} \frac{q_B}{4} \sum_a \frac{G_{aaaa}^{q_A, \tilde{q}}}{F_{aa}^{q_A, \tilde{q}} F_{aa}^{q_A, \tilde{q}}} + O(q_B^2) \Big). \end{aligned} \quad (8.47)$$

We must choose $\Delta V^{q_A, \tilde{q}}$ to cancel the contributions from the other $O(q_B)$ terms to the saddle-point equation for the tree-level eigenvalue distribution of A . Thus,

$$\begin{aligned} \Delta V^{q_A, \tilde{q}}(\lambda_a) = & - \int d\lambda_b \rho_{q_A, 0}(\lambda_b) \frac{\Delta F_{ab}^{q_A, \tilde{q}}}{F_{ab}^{q_A, \tilde{q}}} \\ & + \frac{1}{2} \int d\lambda_b \rho_{q_A, 0}(\lambda_b) d\lambda_c \rho_{q_A, 0}(\lambda_c) \frac{G_{abac}^{q_A, \tilde{q}}}{F_{ab}^{q_A, \tilde{q}} F_{ac}^{q_A, \tilde{q}}} \\ & + \int d\lambda_b \rho_{q_A, 0}(\lambda_b) d\lambda_c \rho_{q_A, 0}(\lambda_c) \frac{G_{babc}^{q_A, \tilde{q}}}{F_{ba}^{q_A, \tilde{q}} F_{bc}^{q_A, \tilde{q}}}, \end{aligned} \quad (8.48)$$

and if we use (8.46), this becomes

$$\Delta V^{q_A, \tilde{q}}(\lambda_a) = -\frac{1}{2} \int d\lambda_b \rho_{q_A, 0}(\lambda_b) d\lambda_c \rho_{q_A, 0}(\lambda_c) \frac{G_{abac}^{q_A, \tilde{q}}}{F_{ab}^{q_A, \tilde{q}} F_{ac}^{q_A, \tilde{q}}}. \quad (8.49)$$

We now consider the $\tilde{q} = 0$ limit. From (8.36), (8.46), (8.44), (8.9), and (A70), we find that

$$G_{abcd}^{q_A, 0} = 1, \quad \Delta F_{ab}^{q_A, 0} = 2, \quad \Delta V^{q_A, 0}(A) = -\frac{1}{2}, \quad (8.50)$$

and in particular we may ignore $\Delta V^{q_A, 0}(A)$ because it is a constant. Then, (8.43) becomes

$$\begin{aligned} \mathcal{Z}_{q_A, q_B, 0} = \int dAdB \exp \Big(& -N \text{Tr} \left[V_{q_A}(A) + B^2 \left(\frac{1}{2} + q_B \right) \right. \\ & \left. - \frac{q_B}{4} B^4 + O(q_B^2) \right] \Big). \end{aligned} \quad (8.51)$$

Using (8.33), we see that this agrees with (8.34).

We now consider the computation of (8.44), (8.46), and (8.49) to order q_A for arbitrary \tilde{q} . Because (8.44) depends on the q -deformed $6j$ symbol, we use (A82) to write the q -deformed $6j$ symbol as an infinite sum. To a finite order in q_A , the sum truncates to a finite number of terms, and collecting the $O(1)$ and $O(q_A)$ terms is straightforward (but tedious). After determining $G_{abcd}^{q_A, \tilde{q}}$, the integrals in (8.46) and (8.49) are straightforward to evaluate to $O(q_A)$. We now present an explicit formula for the potential in (8.43) to $O(q_A)$:

$$\begin{aligned} \frac{1}{N} V_{q_A, q_B, \tilde{q}} = & \text{Tr} \left[\frac{(1 - \tilde{q}^2)}{2} A^2 + \frac{(1 - \tilde{q}^2)}{2} B^2 + \frac{\tilde{q}^2}{1 - \tilde{q}^2} B^2 A^2 - \frac{\tilde{q}(1 + \tilde{q}^2)}{2(1 - \tilde{q}^2)} ABAB \right] \\ & + q_A \text{Tr} \left[(1 - \tilde{q}^4) A^2 + \frac{(\tilde{q}^4 - 1)}{4} A^4 + \frac{\tilde{q}^2(1 - \tilde{q}^2)}{2} B^2 - \tilde{q}^2 A^2 B^2 \right. \\ & \left. + \tilde{q}^3 ABAB - \frac{\tilde{q}^3}{1 - \tilde{q}^2} A^3 BAB + \frac{\tilde{q}^2(1 + \tilde{q}^2)}{2(1 - \tilde{q}^2)} A^2 BA^2 B \right] \\ & + q_B \text{Tr} \left[\frac{\tilde{q}^2(1 - \tilde{q}^2)}{2} A^2 + (1 - \tilde{q}^4) B^2 - \tilde{q}^2 A^2 B^2 + \tilde{q}^3 ABAB \right. \\ & \left. + \frac{(\tilde{q}^4 - 1)}{4} B^4 - \frac{\tilde{q}^3}{(1 - \tilde{q}^2)} ABAB^3 + \frac{\tilde{q}^2(1 + \tilde{q}^2)}{2(1 - \tilde{q}^2)} AB^2 AB^2 \right] \\ & + q_B q_A \text{Tr} \left[\tilde{q}^4 (1 - \tilde{q}^2) A^2 - \frac{\tilde{q}^4(1 - \tilde{q}^2)}{2} A^4 + \tilde{q}^4 (1 - \tilde{q}^2) B^2 - \frac{\tilde{q}^4(2 - 5\tilde{q}^2 + \tilde{q}^4)}{1 - \tilde{q}^2} B^2 A^2 + \tilde{q}^4 B^2 A^4 \right. \\ & - \frac{2\tilde{q}^7}{1 - \tilde{q}^2} ABAB - \tilde{q}^5 A^3 BAB - \frac{\tilde{q}^4(1 - \tilde{q}^2)}{2} B^4 + \tilde{q}^4 A^2 B^4 - \tilde{q}^5 ABAB^3 - \frac{\tilde{q}^4}{2(1 - \tilde{q}^2)^2} A^2 B^2 A^2 B^2 \\ & \left. + \frac{\tilde{q}^5}{(1 - \tilde{q}^2)^2} A^2 BABAB^2 - \frac{\tilde{q}^6}{(1 - \tilde{q}^2)^2} A^2 BAB^2 AB + \frac{\tilde{q}^5}{(1 - \tilde{q}^2)^2} A^2 B^2 ABAB + \frac{\tilde{q}^4(\tilde{q}^4 - 2\tilde{q}^2 - 1)}{4(1 - \tilde{q}^2)^2} (AB)^4 \right] \\ & + \dots \end{aligned} \quad (8.52)$$

By inspection, the terms in (8.52) are symmetric under $A \leftrightarrow B$. At the level of this perturbative calculation, this result is highly nontrivial, and would be a coincidence if the disk observables of the q -deformed model did not have a chord diagram interpretation.

The above calculation is the first step of the systematic procedure described in Sec. VII for determining the couplings of the q -deformed model. We expect that if we continue to determine the couplings to higher order in q_B , we will continue to see the symmetry that interchanges A and B . This is strong evidence in favor of the conclusion that the matrix potential of the q -deformed model is well-defined. If (8.52) did not have the $A \leftrightarrow B$ symmetry, we would have to conclude that the power series representation of the matrix potential⁵² is merely a formal power series that does not converge (if it did converge, it would converge to a potential without the $A \leftrightarrow B$ symmetry, which contradicts the fact that the chord diagram Feynman rules treat A and B on equal footing). Although we lack a rigorous proof, we believe that the potential of the q -deformed model is well-defined. We will learn more about the q -deformed model in the next section, where we compute the double-trumpet. In particular, we will compute the empty double-trumpet as well as the double-trumpet with one \mathcal{O} inserted on each boundary. From either of these calculations, we can read off the partition function of the bulk matter theory on the double-trumpet. The fact that these two separate calculations return the same result for the matter partition function provides further nontrivial evidence that the q -deformed model exists.

We do not have analogous arguments for why the Selberg model exists because we have no *a priori* reasons to expect a symmetry in the Selberg model. Furthermore, in the Selberg model, we view the empty double-trumpet as part of the data that defines the model, as opposed to something that is nontrivially determined by the defining data.

IX. THE DOUBLE-TRUMPET

Previously, we argued that there exist single-trace two-matrix models which, in the double-scaling limit, correctly compute the disk amplitudes of JT gravity minimally coupled to a scalar field. Although we explained in Sec. VII that there is a systematic way to determine all of the terms in the matrix potential, actually computing these terms in practice quickly becomes tedious. The purpose of this section is to demonstrate that one does not need to know the detailed form of the matrix potential to compute the double-trumpet in a single-trace, two-matrix model. It is sufficient to know the (regulated) disk correlators. In both the q -deformed and Selberg models, we will compute contributions to the double-trumpet that can be directly compared against our calculations in Sec. II B. We

will consider the empty double-trumpet as well as the double-trumpet with an \mathcal{O} operator inserted on each boundary. For simplicity, we set $S_0 = 0$ in this section.

Let us briefly summarize the technical results of this section. In the q -deformed model, our result for the empty double trumpet in the JT limit is in agreement with (4.42), which we reproduce here,

$$\langle \text{Tr} e^{-\beta_L H} \text{Tr} e^{-\beta_R H} \rangle_{\text{cyl}} = \int_0^\infty db b Z_{\text{tr}}(\beta_L, b) Z_{\text{tr}}(\beta_R, b) Z(b), \quad (9.1)$$

$$Z(b) = \sum_{n=0}^{\infty} \left(\frac{e^{-\Delta b}}{1 - e^{-b}} \right)^n = \frac{1 - e^{-b}}{1 - e^{-b} - e^{-\Delta b}}. \quad (9.2)$$

We may interpret $Z(b)$ as a partition function of some matter theory with inverse temperature b . We found this result by classifying all of the 't Hooft diagrams that contribute to the empty double-trumpet, and we explicitly computed the first few classes. The first class of diagrams agrees with the $\epsilon = 0$ result (4.42). The next few classes of diagrams, which represent finite ϵ corrections, actually make a vanishing contribution to the empty double-trumpet. We conjecture that the sum over all the 't Hooft diagrams agrees with the $\epsilon = 0$ result, so that the matter partition function $Z(b)$ computed using the q -deformed regulator is the same as in the toy model of Sec. IV that is Gaussian in \mathcal{O} . Our computation of the double-trumpet with one \mathcal{O} inserted on each boundary provides more evidence in support of this result. As shown in Appendix E, the gravitational computation of this leads to an integral over the closed geodesic length b . The integrand is a product of the integral over the “boundary wiggles” at fixed b (which includes an $e^{-\Delta \ell}$ weighting factor, where ℓ is the renormalized geodesic length between the two \mathcal{O} insertions) and a partition function $Z(b)$. In the q -deformed model, we classify all of the 't Hooft diagrams, and we compute analytic formulas for the sums over the first four classes of diagrams. We express the results in a way that allows us to read off what $Z(b)$ is. We find that the first four classes of diagrams reproduce the $n = 0, 1, 2, 3$ terms in the sum in (9.2). We conjecture that the remaining classes of diagrams that we do not explicitly compute continue to match the remaining terms in the sum. Thus, both the empty double-trumpet and the two-point function return the same result (9.2) for the matter partition function. We comment on the Hagedorn behavior of $Z(b)$ in Sec. IX D 1.

Using the Selberg regulator, the empty double-trumpet is the same as in the q -deformed model. Thus, as explained in Sec. V, we add a double-trace term (5.17) to the matrix potential of the Selberg model to ensure that the empty double-trumpet becomes (9.1) except with $Z(b)$ replaced by $Z_{\text{scalar}}(b)$, which was defined in (2.23). Our freedom to add this term arises from an ambiguity in defining the integration measure for the matrix integral. We have

⁵²We are referring to (7.9)–(7.11), where q_B plays the role of ϵ .

identified two reasons for this ambiguity. First, by rescaling the matrix elements \mathcal{O}_{ab} by a function of E_a and E_b and transforming the measure $d\mathcal{O}$, we can add an arbitrary double-trace term to the potential that only depends on H . In other words, if we let $d\mathcal{O}$ in (1.1) refer to the standard flat measure for the rescaled matrix elements, then the measure written in terms of the original matrix elements includes a Jacobian factor that adds a double-trace term to the matrix potential that only depends on H . Unlike the q -deformed model, we have no way to canonically specify the measure for \mathcal{O} .⁵³ Second, if we view this model as an effective matrix model, then we should consider additional matrices other than \mathcal{O} that can couple to H . Upon integrating these other matrices out, we obtain the matrix potential of the two-matrix model, but perhaps with a modified measure for H . Note that the single-trace potential was obtained from a disk-level matching calculation (analogous to how one determines the effective Lagrangian of an EFT from matching). This matching calculation cannot rule out the presence of a double-trace term like (5.17) that does not change the disk answers. This double-trace term can be determined by matching to the gravitational double-trumpet.

Hence, in the Selberg model, the empty double-trumpet is viewed as part of the data that defines the model, while in the q -deformed model, the empty double-trumpet is a nontrivial prediction of the model. In the Selberg model, we compute the double-trumpet two-point function using similar techniques as in the q -deformed model, and we find that the first four classes of 't Hooft diagrams reproduce the $n = 0, 1, 2, 3$ terms in the sum in (2.24). We then conjecture that the sum over all 't Hooft diagrams in the Selberg model reproduces, in the double-scaling limit, the double-trumpet two-point function in JT gravity minimally coupled to a massive scalar field.

A. Empty double-trumpet

In this section, we study the double-trumpet with no \mathcal{O} insertions by explicitly analyzing 't Hooft diagrams. We illustrate our strategy by working directly in the double-scaling limit (as in Secs. IV and VII), although we will eventually write down explicit regulated expressions.

We are interested in the connected double-trace correlator

$$\langle \text{Tr} e^{-\beta_L H} \text{Tr} e^{-\beta_R H} \rangle_c \equiv \langle \text{Tr} e^{-\beta_L H} \text{Tr} e^{-\beta_R H} \rangle - \langle \text{Tr} e^{-\beta_L H} \rangle \langle \text{Tr} e^{-\beta_R H} \rangle. \quad (9.3)$$

Let $\langle \text{Tr} e^{-\beta_L H} \text{Tr} e^{-\beta_R H} \rangle_{\text{cyl}}$ denote the leading contribution to (9.3) in the genus expansion. An example of a 't Hooft diagram that contributes to $\langle \text{Tr} e^{-\beta_L H} \text{Tr} e^{-\beta_R H} \rangle_{\text{cyl}}$ is given in Fig. 8. To compute the sum over all such diagrams, it is

⁵³If we wish to preserve the $A \leftrightarrow B$ symmetry in the q -deformed model, we should not redefine the A matrix in a B -dependent way. As explained in Sec. VIII, in the double-scaling limit A becomes H and B becomes \mathcal{O} up to trivial rescalings.

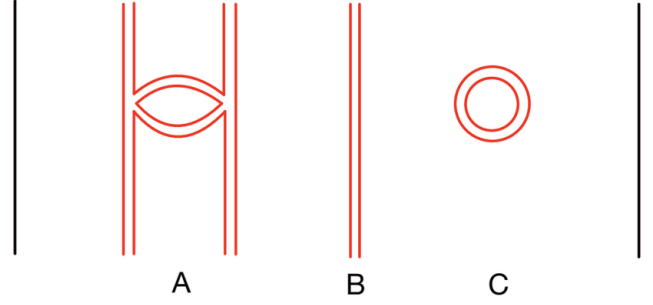


FIG. 8. A 't Hooft diagram that contributes to the connected double-trace correlator (9.3) at leading order in the genus expansion. The top and bottom ends of the diagram are identified to obtain the cylinder topology. Each of the two traces in (9.3) is represented by a black line. Each red double-line corresponds to the \mathcal{O} matrix. There can be many additional black double-lines which we have not explicitly drawn (these correspond to the H matrix). Ignoring the H double-lines, a general diagram can be separated into connected diagrams. Each of the three connected diagrams above is labeled by a letter. Diagrams A and B wrap the double-trumpet. Diagram C is contractible to a point. The contractible diagrams (namely, the \mathcal{O} bubble diagrams) will be cancelled by counterterms (just as in Fig. 3), so we will no longer draw them in subsequent figures. Note that diagrams B and C both arise from the same double-trace term that appears in the action for H after integrating out \mathcal{O} [in the toy model, this term is the second term in (4.4)].

convenient to first consider the sum over all the connected \mathcal{O} diagrams that wrap the double-trumpet (such as diagrams A and B in Fig. 8). Each diagram contains two noncontractible index loops and hence may be thought of as a function of the two energies running in these loops.⁵⁴ Let $\mathcal{D}(s_a, s_b)$ denote a smooth function that represents the sum over all connected \mathcal{O} diagrams that wrap the double-trumpet. For convenience, we will often write \mathcal{D}_{ab} in place of $\mathcal{D}(s_a, s_b)$.⁵⁵ A general 't Hooft diagram (such as Fig. 8) is obtained by connecting together the diagrams that contribute to \mathcal{D}_{ab} with H double-lines and contractible \mathcal{O} bubble diagrams. Because the contractible \mathcal{O} bubble diagrams are canceled by counterterms, we can use the double-trumpet of SSS to connect the diagrams that appear in \mathcal{D}_{ab} . Explicitly, we have that

⁵⁴The contractible index loops are also associated with energies, but these energies should be integrated over with the disk density of states for H . That is, these loops should be filled in like a disk.

⁵⁵As always, s is related to energy by $s^2 = E$. The two noncontractible index loops of a diagram are averaged over microcanonical windows centered on energies s_a^2 and s_b^2 respectively to obtain the contribution to \mathcal{D}_{ab} . In the toy model of Sec. IV, the only contribution to \mathcal{D}_{ab} comes from diagram B in Fig. 8 and is given by $\log \Gamma(\Delta \pm is_a \pm is_b)$. Note that \mathcal{D}_{ab} may be defined to include or not include a symmetry factor of 2 associated with the left-right symmetry. Our conventions for \mathcal{D}_{ab} are clear from (9.4).

$$\begin{aligned}
\langle \text{Tr} e^{-\beta_L H} \text{Tr} e^{-\beta_R H} \rangle_{\text{cyl}} &= \int_0^\infty dE_L dE_R e^{-\beta_L E_L} e^{-\beta_R E_R} \left[\langle \rho_H(E_L) \rho_H(E_R) \rangle_{\text{cyl,SSS}} \right. \\
&+ \int_0^\infty dE_1 dE_2 \langle \rho_H(E_L) \rho_H(E_1) \rangle_{\text{cyl,SSS}} \mathcal{D}(\sqrt{E_1}, \sqrt{E_2}) \langle \rho_H(E_2) \rho_H(E_R) \rangle_{\text{cyl,SSS}} \\
&+ \int_0^\infty dE_1 dE_2 dE_3 dE_4 \langle \rho_H(E_L) \rho_H(E_1) \rangle_{\text{cyl,SSS}} \mathcal{D}(\sqrt{E_1}, \sqrt{E_2}) \\
&\times \langle \rho_H(E_2) \rho_H(E_3) \rangle_{\text{cyl,SSS}} \mathcal{D}(\sqrt{E_3}, \sqrt{E_4}) \langle \rho_H(E_4) \rho_H(E_R) \rangle_{\text{cyl,SSS}} + \dots \left. \right], \quad (9.4)
\end{aligned}$$

where ρ_H was defined in (4.6), and $\langle \rho_H(E_L) \rho_H(E_R) \rangle_{\text{cyl,SSS}}$ denotes the leading order connected correlator in the SSS model and may be computed from the inverse Laplace transform of the SSS double-trumpet. Using the expression for $\langle \rho_H(E_L) \rho_H(E_R) \rangle_{\text{cyl,SSS}}$ from [9], Eq. (9.4) is equal to

$$\begin{aligned}
&\langle \text{Tr} e^{-\beta_L H} \text{Tr} e^{-\beta_R H} \rangle_c \\
&= \int_0^\infty b_L db_L b_R db_R Z_{\text{tr}}(\beta_L, b_L) Z_{\text{tr}}(\beta_R, b_R) \\
&\times \left[\frac{1}{b_R} \delta(b_R - b_L) + \tilde{\mathcal{D}}(b_L, b_R) \right. \\
&\left. + \int_0^\infty b_1 db_1 \tilde{\mathcal{D}}(b_L, b_1) \tilde{\mathcal{D}}(b_1, b_R) + \dots \right], \quad (9.5)
\end{aligned}$$

where

$$\tilde{\mathcal{D}}(b_1, b_2) \equiv \int_0^\infty ds_1 ds_2 \frac{\cos(s_1 b_1)}{\pi} \mathcal{D}(s_1, s_2) \frac{\cos(s_2 b_2)}{\pi}, \quad (9.6)$$

and $Z_{\text{tr}}(\beta, b)$ was defined in (2.21). Note that adding an energy-independent constant to \mathcal{D} does not affect the connected double-trace correlator because

$$\int_0^\infty dE_1 \langle \rho_H(E_1) \rho_H(E_2) \rangle_{\text{cyl,SSS}} = 0. \quad (9.7)$$

The integral of the density of states is equal to the number of eigenvalues, which does not fluctuate in the ensemble.

By deriving (9.4), we have reduced the task of computing the empty double-trumpet to the task of computing \mathcal{D}_{ab} , which is the sum over connected \mathcal{O} diagrams that wrap the double-trumpet. Our next step is to classify these diagrams by studying paths from the left to the right side of a diagram which are only allowed to cross over \mathcal{O} double-lines (and not vertices). Let $\mathcal{D}_{ab}^{(n)}$ denote the contribution to \mathcal{D}_{ab} from all the diagrams obeying the condition that the lowest number of \mathcal{O} double-lines that are crossed by a path from the left to the right side is n . For example, diagram B in Fig. 8 belongs to $\mathcal{D}_{ab}^{(1)}$ because every path from the left to the right side crosses exactly one \mathcal{O} double-line, while

diagram A belongs to $\mathcal{D}_{ab}^{(2)}$ because one can draw left-to-right paths that cross over two \mathcal{O} double-lines but not one. It follows that

$$\mathcal{D}_{ab} = \sum_{n=1}^{\infty} \mathcal{D}_{ab}^{(n)}. \quad (9.8)$$

In the next subsection, we compute $\mathcal{D}_{ab}^{(1)}$, $\mathcal{D}_{ab}^{(2)}$, and $\mathcal{D}_{ab}^{(3)}$ using the q -deformed and Selberg regulators and then claim that there is a pattern that generalizes to arbitrary n .

As mentioned above, by redefining the measure for \mathcal{O} , we can obtain different answers for the empty double-trumpet. In particular, adding the double-trace term

$$-\frac{1}{2} \sum_{a,b} \log f(E_a, E_b) \quad (9.9)$$

to the matrix potential is equivalent to adding $\log f(E_1, E_2)$ to $\mathcal{D}(\sqrt{E_1}, \sqrt{E_2})$. To add (9.9) to the matrix potential, we may define $\tilde{\mathcal{O}}_{ab}$ by

$$\tilde{\mathcal{O}}_{ab} := [f(E_a, E_b)]^{-1/2} \mathcal{O}_{ab} \quad (9.10)$$

and then define the measure to be $d\tilde{\mathcal{O}}$, or the standard flat measure for the $\tilde{\mathcal{O}}$ matrix elements. In this way, we can set $\mathcal{D}(\sqrt{E_1}, \sqrt{E_2})$ to whatever we want. In the Selberg model, we fix the ambiguity in the definition of the measure for \mathcal{O} by demanding that the empty double-trumpet agrees with (2.17). On the other hand, in the q -deformed model, we explicitly chose the standard measure for the A and B matrices. We should not modify the measure for B because the q -deformed model is supposed to treat the A and B matrices symmetrically.

1. Computing $\mathcal{D}^{(1)}$

To compute $\mathcal{D}_{ab}^{(1)}$, we consider diagrams that allow for left-right paths that cross over one \mathcal{O} double-line. We enumerate all of the possible diagrams in Fig. 10.

Let \mathcal{X}_{ab} refer to the 1PI two-point function (see Fig. 9). Referring to the matrix integral in (7.1), we find that \mathcal{X}_{ab} obeys

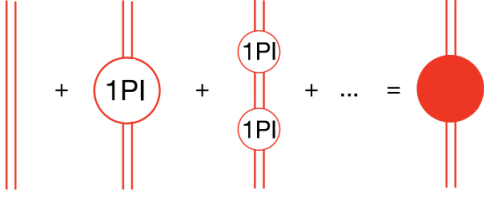


FIG. 9. The 1PI two-point function is related to the two-point function as shown. The red blob corresponds to $\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}}$. The top and bottom ends of the diagrams are *not* identified. We define \mathcal{X}_{ab} to be the function of two energies that represents the 1PI blob above.

$$\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}} = \frac{F_{ab}^{-1}}{1 - F_{ab}^{-1} \mathcal{X}_{ab}}, \quad (9.11)$$

which implies that

$$\begin{aligned} \mathcal{D}_{ab}^{(1)} &= -\log F_{ab} + \sum_{n=1}^{\infty} \frac{(F_{ab}^{-1} \mathcal{X}_{ab})^n}{n} \\ &= -\log F_{ab} - \log (1 - F_{ab}^{-1} \mathcal{X}_{ab}), \\ &= \log \langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}} \\ &= \log \frac{\Gamma(\Delta \pm i s_a \pm i s_b)}{\Gamma(2\Delta)}. \end{aligned} \quad (9.12)$$

The $\frac{1}{n}$ factor is a symmetry factor arising from the \mathbb{Z}_n symmetry of the diagrams in Fig. 10. Using (4.18) we then find that

$$\tilde{\mathcal{D}}^{(1)}(b, b') = \frac{1}{b} \delta(b - b') \frac{e^{-\Delta b}}{1 - e^{-b}}. \quad (9.13)$$

This contribution to the double-trumpet (9.5) precisely equals the double-trumpet result (4.42) in the toy model of Sec. IV. Changing the normalization of \mathcal{O} changes $\mathcal{D}^{(1)}$ by an additive constant that does not affect the double-trumpet, as explained above.

Before continuing on to $\mathcal{D}^{(2)}$, we will compute the contribution from $\mathcal{D}^{(1)}$ to the double-trumpet using (9.4) in the q -deformed model. To be precise, we will compute

$$\langle \text{Tr} e^{-\beta_L A} \text{Tr} e^{-\beta_R A} \rangle_{\text{cyl}}. \quad (9.14)$$

In Figs. 9 and 10, the red double-lines now correspond to the B matrix, and we will let the E dummy variable in (9.4) refer to an eigenvalue of the A matrix. Using the chord-diagram Feynman rules beginning from (8.5), we have that

$$\begin{aligned} &\langle \text{Tr} B A^{n_1} B A^{n_2} \rangle_{\text{disk}} \\ &= N_{q_A} \int_0^\pi \prod_{j=1}^2 \left[\frac{d\theta_j}{2\pi} (q_A; q_A)_\infty (e^{\pm 2i\theta_j}; q_A)_\infty \left(\frac{2 \cos \theta_j}{\sqrt{1 - q_A}} \right)^{n_j} \right] \\ &\quad \times \frac{(\tilde{q}^2; q_A)_\infty}{(\tilde{q} e^{i(\pm\theta_1 \pm \theta_2)}; q_A)_\infty}, \end{aligned} \quad (9.15)$$

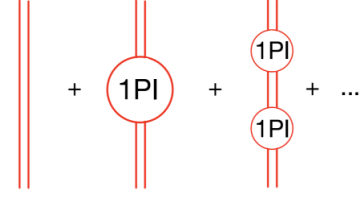


FIG. 10. The diagrams that contribute to $\mathcal{D}_{ab}^{(1)}$ consist of 1PI two-point functions threaded together. The top and bottom ends of the diagrams are identified.

where N_{q_A} , defined in (8.8), is the number of eigenvalues of each matrix. This implies that

$$\begin{aligned} &\frac{1}{\delta E_1 N_{q_A} \rho_{q_A,0}(E_1)} \frac{1}{\delta E_2 N_{q_A} \rho_{q_A,0}(E_2)} \langle \text{Tr} B P(E_1) B P(E_2) \rangle_{\text{disk}} \\ &= N_{q_A}^{-1} \frac{(\tilde{q}^2; q_A)_\infty}{(\tilde{q} e^{i(\pm\theta_1 \pm \theta_2)}; q_A)_\infty}, \\ &E_j = \frac{2 \cos \theta_j}{\sqrt{1 - q_A}}, \quad j = 1, 2, \end{aligned} \quad (9.16)$$

where $\rho_{q,0}$ was defined in (8.23) and (8.24). Equation (9.16) is defined analogously to (7.4), so $P(E_1)$ is a projector onto a microcanonical window of width δE_1 centered around E_1 , with $\delta E_1 N_{q_A} \rho_{q_A,0}(E_1)$ eigenvalues to leading order in N_{q_A} . Equation (9.16) is the analog of $\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}}$ in the q -deformed model, so $\mathcal{D}^{(1)}$ immediately follows from (9.12).

Next, we need to compute

$$\langle \rho_A(E_1) \rho_A(E_2) \rangle_{\text{cyl}, 1\text{-matrix}} \quad (9.17)$$

in the q -deformed model, where ρ_A is the density of states of A . Equation (9.17) refers to the cylinder in the single-matrix model of Sec. VIII D. That is, (9.17) is the analog of $\langle \rho_H(E_1) \rho_H(E_2) \rangle_{\text{cyl}, \text{SSS}}$ in the q -deformed model. The density-density correlator in the q -deformed 1-cut matrix model takes the universal form [50,51]

$$\begin{aligned} &\langle \rho_A(E_1) \rho_A(E_2) \rangle_{\text{cyl}, 1\text{-matrix}} \\ &= \frac{1}{(2\pi)^2} \frac{1}{(E_1 - E_2)^2} \frac{2 \left(\frac{E_1 E_2}{a^2} - 1 \right)}{\sqrt{\left(1 - \left(\frac{E_1}{a} \right)^2 \right) \left(1 - \left(\frac{E_2}{a} \right)^2 \right)}}, \end{aligned} \quad (9.18)$$

where

$$a = \frac{2}{\sqrt{1 - q_A}}. \quad (9.19)$$

Next, note that the Chebyshev polynomials of the first kind have the following generating function⁵⁶ for $|t| < 1$

$$\sum_{n=1}^{\infty} n T_n(x) T_n(y) t^n = \frac{(x^2 + y^2)(4t^3(t^2 + 1)xy - 4t^2(t^4 + 1)) + (t^6 + 7t^4 + 7t^2 + 1)txy + 2(1 - t^2)^2 t^2 - 16t^4 x^2 y^2}{(4t^2(x^2 + y^2) - 4(t^2 + 1)txy + (1 - t^2)^2)^2}. \quad (9.20)$$

Taking $t \rightarrow 1$, we formally obtain that

$$\sum_{n=1}^{\infty} n 2T_n(x) 2T_n(y) = \frac{2(xy - 1)}{(x - y)^2}, \quad (9.21)$$

which implies that

$$\begin{aligned} & \langle \rho_A(E_1) \rho_A(E_2) \rangle_{\text{cyl, 1-matrix}} \\ &= \frac{1}{(2\pi a)^2} \sum_{n=1}^{\infty} \frac{n 2T_n\left(\frac{E_1}{a}\right) 2T_n\left(\frac{E_2}{a}\right)}{\sqrt{\left(1 - \left(\frac{E_1}{a}\right)^2\right) \left(1 - \left(\frac{E_2}{a}\right)^2\right)}}. \end{aligned} \quad (9.22)$$

Meanwhile, from (8.39), (9.12), and (9.16), we have

$$\mathcal{D}^{(1)}(\sqrt{E_1}, \sqrt{E_2}) = \sum_{n=1}^{\infty} \frac{\tilde{q}^n}{n(1 - q_A^n)} 2T_n\left(\frac{E_1}{a}\right) 2T_n\left(\frac{E_2}{a}\right). \quad (9.23)$$

Noting that the Chebyshev polynomials are orthogonal,

$$\int_{-1}^1 \frac{dx}{2\pi\sqrt{1-x^2}} 2T_n(x) 2T_m(x) = \delta_{nm}, \quad n, m \geq 1, \quad (9.24)$$

we may plug (9.22) and (9.23) into (9.4) to obtain

$$\begin{aligned} & \langle \text{Tre}^{-\beta_L A} \text{Tre}^{-\beta_R A} \rangle_{\text{cyl}} \\ & \supset \sum_{n=1}^{\infty} n \frac{1 - q_A^n}{1 - q_A^n - \tilde{q}^n} \int_{-a}^a \prod_{i \in \{L, R\}} \left[\frac{dE_i}{2\pi a} e^{-\beta_i E_i} \frac{2T_n\left(\frac{E_i}{a}\right)}{\sqrt{1 - \left(\frac{E_i}{a}\right)^2}} \right]. \end{aligned} \quad (9.25)$$

The \supset symbol indicates that (9.25) only contains the contribution from $\mathcal{D}^{(1)}$. However, below we will see that neither $\mathcal{D}^{(2)}$ nor $\mathcal{D}^{(3)}$ make any nontrivial contributions to \mathcal{D} , and we will conclude that in general, only $\mathcal{D}^{(1)}$ contributes nontrivially to \mathcal{D} in the q -deformed model. Hence, we may replace the \supset by an $=$ above. Note that (9.4) only converges when⁵⁷

$$q_A^n + \tilde{q}^n < 1, \quad \forall n \geq 1. \quad (9.26)$$

⁵⁶This may be derived from the results of [52].

⁵⁷Given that q_A and q are in the interval (0,1), it is enough to write (9.26) for $n = 1$ only.

The integrals over E_i in (9.25) can be computed

$$\int_{-1}^1 \frac{dx}{2\pi\sqrt{1-x^2}} e^{-\beta x} 2T_n(x) = (-1)^n I_n(\beta). \quad (9.27)$$

So we get

$$\langle \text{Tre}^{-\beta_L A} \text{Tre}^{-\beta_R A} \rangle_{\text{cyl}} = \sum_{n=1}^{\infty} n \frac{1 - q_A^n}{1 - q_A^n - \tilde{q}^n} I_n(\beta_L a) I_n(\beta_R a). \quad (9.28)$$

Without the factor $\frac{1 - q_A^n}{1 - q_A^n - \tilde{q}^n}$ this is Eq. (50) from [53].⁵⁸ Our formula (9.28) is a generalization of that to the 2-matrix model.

In Sec. IX D 1, we will interpret the matrix model in the regime where (9.26) is violated. A simpler correlator is

$$\left\langle \text{Tr} 2T_n\left(\frac{A}{a}\right) \text{Tr} 2T_m\left(\frac{A}{a}\right) \right\rangle_{\text{cyl}} = n \frac{1 - q_A^n}{1 - q_A^n - \tilde{q}^n} \delta_{nm}. \quad (9.29)$$

It would be interesting to reproduce this result for the double-trumpet in q -deformed theory from the transfer matrix approach in the double-scaled SYK [13,55]. This computation can be thought of as the bulk dual of our matrix model result above. One can construct the bulk Hilbert space by slicing open the chord diagrams [55] and use it to compute the double-trumpet.

We can also take the JT limit of (9.29). We set $q_A = q$, $\tilde{q} = q^\Delta$, and we use (8.19) to relate H and A . Then define b_1 and b_2 via

$$q_A^n = e^{-b_1}, \quad q_A^m = e^{-b_2}. \quad (9.30)$$

Taking $q \rightarrow 1$ for fixed b_1, b_2 , we obtain

$$\begin{aligned} & \left\langle \text{Tr} \frac{2}{b_1} \cos(b_1 \sqrt{H}) \text{Tr} \frac{2}{b_2} \cos(b_2 \sqrt{H}) \right\rangle_{\text{cyl}} \\ &= \frac{1 - e^{-b_1}}{1 - e^{-b_1} - e^{-\Delta b_1}} \frac{1}{b_1} \delta(b_1 - b_2), \end{aligned} \quad (9.31)$$

which is in agreement with (4.42).

⁵⁸See also Eq. (2.34) in [54].

2. Computing $\mathcal{D}^{(2)}$

To compute $\mathcal{D}^{(2)}$, we consider diagrams that allow for left-right paths that cross over two \mathcal{O} double-lines, but we exclude those diagrams that were already counted in $\mathcal{D}^{(1)}$. It is convenient to think of each diagram as an operator that acts on functions of a single energy. Let $\mathcal{X}_{ab}^{(2)}(s_c, s_d)$ refer to a smooth function of s_a, s_b, s_c , and s_d that is defined in Fig. 11. We may think of $\mathcal{X}_{ab}^{(2)}(s_c, s_d)$ as an operator (which depends on the parameters s_a and s_b) that acts on a function of s_d to produce a function of s_c .

We enumerate all of the diagrams in $\mathcal{D}^{(2)}$ in Fig. 12, where we explicitly show the locations where a left-right path may be drawn to cross over two double lines. The remaining parts of the diagrams are encapsulated in $\mathcal{X}_{ab}^{(2)}(s_c, s_d)$.

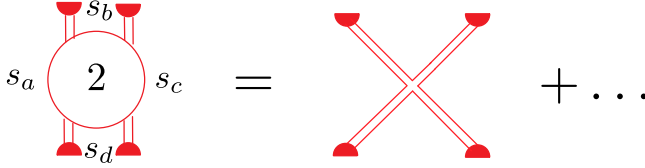


FIG. 11. The blob labeled “2” is the sum over two-particle irreducible diagrams, i.e. all connected planar four-point diagrams such that any left-right path through the blob crosses more than two double-lines. The simplest contribution to this blob comes from a tree four-point diagram (in this case, there are no left-right paths that can go through the blob, because we only let left-right paths cross over double-lines, and all of the double-lines are external. So it is vacuously true that all left-right paths through the blob cross over more than two double-lines). The blob labeled “2” is a function of the four energies shown above. We define $\mathcal{X}_{ac}^{(2)}(s_b, s_d)$ to be the product of this function and $(\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}} \langle \mathcal{O}_{bc} \mathcal{O}_{cb} \rangle_{\text{disk}} \langle \mathcal{O}_{cd} \mathcal{O}_{dc} \rangle_{\text{disk}} \langle \mathcal{O}_{ad} \mathcal{O}_{da} \rangle_{\text{disk}})^{1/2}$, which is represented by the four red semicircles. Including these factors ensures that $\mathcal{X}_{ac}^{(2)}(s_b, s_d)$ does not depend on the normalization of \mathcal{O} .

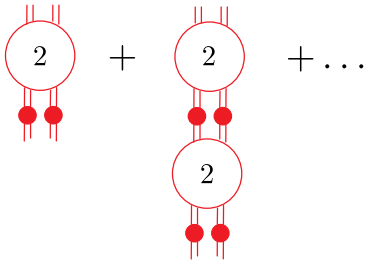


FIG. 12. The diagrams that contribute to $\mathcal{D}^{(2)}$. The top and bottom ends of these diagrams are identified. The red blob represents the exact two-point function $\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}}$.

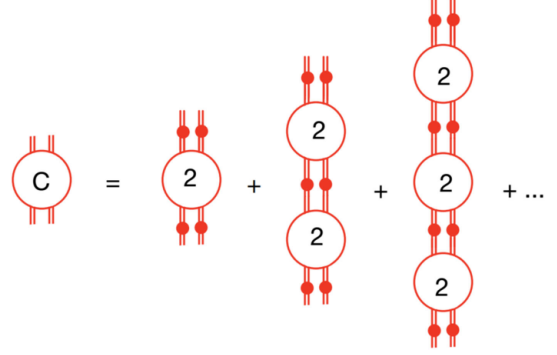


FIG. 13. A geometric series involving $\mathcal{X}_{ab}^{(2)}(s_c, s_d)$ computes the connected planar four-point function, which is represented by a blob labeled “C.” The top and bottom ends of these diagrams are *not* identified.

The sum over the diagrams in Fig. 12 is given by

$$\begin{aligned} \mathcal{D}_{ab}^{(2)} &= \int_0^\infty ds \rho(s) \mathcal{X}_{ab}^{(2)}(s, s) \\ &+ \frac{1}{2} \int_0^\infty ds_1 ds_2 \rho(s_1) \rho(s_2) \mathcal{X}_{ab}^{(2)}(s_1, s_2) \mathcal{X}_{ab}^{(2)}(s_2, s_1) \\ &+ \frac{1}{3} \int_0^\infty ds_1 ds_2 ds_3 \rho(s_1) \rho(s_2) \rho(s_3) \mathcal{X}_{ab}^{(2)}(s_1, s_2) \\ &\times \mathcal{X}_{ab}^{(2)}(s_2, s_3) \mathcal{X}_{ab}^{(2)}(s_3, s_1) + \dots, \end{aligned} \quad (9.32)$$

where the integrals refer to the closed index loops that are explicitly depicted in Fig. 12. The n th term in the sum has a symmetry factor of $\frac{1}{n}$ due to a \mathbb{Z}_n symmetry.

The function $\mathcal{X}_{ab}^{(2)}(s_c, s_d)$ may be determined from the disk four-point function. See Fig. 13. To determine $\mathcal{X}_{ac}^{(2)}(s_b, s_d)$, we write

$$\begin{aligned} &\frac{\langle \mathcal{O}_{ab} \mathcal{O}_{bc} \mathcal{O}_{cd} \mathcal{O}_{da} \rangle_{\text{disk}, c}}{\sqrt{\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}} \langle \mathcal{O}_{bc} \mathcal{O}_{cb} \rangle_{\text{disk}} \langle \mathcal{O}_{cd} \mathcal{O}_{dc} \rangle_{\text{disk}} \langle \mathcal{O}_{da} \mathcal{O}_{ad} \rangle_{\text{disk}}}} \\ &= \mathcal{X}_{ac}^{(2)}(s_b, s_d) + \int_0^\infty ds \rho(s) \mathcal{X}_{ac}^{(2)}(s_b, s) \mathcal{X}_{ac}^{(2)}(s, s_d) + \dots \\ &= \mathcal{X}_{ac}^{(2)} + \mathcal{X}_{ac}^{(2)} \mathcal{X}_{ac}^{(2)} + \dots, \\ &= \mathcal{X}_{ac}^{(2)} [1 - \mathcal{X}_{ac}^{(2)}]^{-1}, \end{aligned} \quad (9.33)$$

where in the third line above we have suppressed some of the s parameters and the integration to simplify the notation,⁵⁹ and in the fourth line we summed the geometric series (which is depicted in Fig. 13). We have

⁵⁹We will continue to suppress these s parameters as long as doing so does not cause confusion.

$$\frac{\langle \mathcal{O}_{ab} \mathcal{O}_{bc} \mathcal{O}_{cd} \mathcal{O}_{da} \rangle_{\text{disk},c}}{\sqrt{\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}} \langle \mathcal{O}_{bc} \mathcal{O}_{cb} \rangle_{\text{disk}} \langle \mathcal{O}_{cd} \mathcal{O}_{dc} \rangle_{\text{disk}} \langle \mathcal{O}_{da} \mathcal{O}_{ad} \rangle_{\text{disk}}}} = \left\{ \begin{array}{ccc} \Delta & s_b & s_c \\ \Delta & s_d & s_a \end{array} \right\} \equiv \mathcal{S}_{ac}(s_b, s_d), \quad (9.34)$$

from which it follows that

$$\mathcal{X}_{ac}^{(2)} = [1 + \mathcal{S}_{ac}]^{-1} \mathcal{S}_{ac}, \quad (9.35)$$

where we have formally inverted the geometric series in (9.33). Note however that $1 + \mathcal{S}_{ac}$, viewed as an operator acting on functions of a single s variable, is not invertible because

$$\begin{aligned} [1 + \mathcal{S}_{ac}](s_b, s_d) &= \frac{\delta(s_b - s_d)}{\rho(s_b)} + \left\{ \begin{array}{ccc} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{array} \right\} \\ &= 2 \sum_{n \in \mathbb{Z}_{\geq 0}} P_n^{\Delta, \Delta}(s_b; s_a, s_c) P_n^{\Delta, \Delta}(s_d; s_a, s_c), \end{aligned} \quad (9.36)$$

where we have used (A28), (A30). Because the sum runs over even n only, it follows that any $P_n^{\Delta, \Delta}(s_d; s_a, s_c)$ function with n odd is annihilated by $1 + \mathcal{S}_{ac}$. The q -deformed and Selberg models regulate these expressions. We define \mathcal{S}_{ac}^q as in (9.34) except with the q -deformed $6j$ symbol. Using either of the two regulators, (9.36) becomes

$$\begin{aligned} [1 + \epsilon \mathcal{S}_{ac}^q](s_b, s_d) &= \frac{\delta(s_b - s_d)}{\rho_q(s_b)} + \epsilon \left\{ \begin{array}{ccc} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{array} \right\}_q \\ &= \sum_{n \in \mathbb{Z}_{\geq 0}} P_n^{\Delta, \Delta}(s_b; s_c, s_a | q) \\ &\quad \times P_n^{\Delta, \Delta}(s_d; s_c, s_a | q) \\ &\quad \times \left(1 + \epsilon (-1)^n q^{\frac{n(n-1)}{2}} q^{2\Delta n} \right), \end{aligned} \quad (9.37)$$

which reduces to (9.36) as $q \rightarrow 1$ and $\epsilon \rightarrow 1$. In the q -deformed model, we have set $q_A = q$ and $q_B = \epsilon$. From (9.37) we see that $1 + \epsilon \mathcal{S}_{ac}^q$ is invertible. A simpler way to write (9.32) is

$$\mathcal{D}_{ab}^{(2)} = \text{Tr} \sum_{n=1}^{\infty} \frac{1}{n} [\mathcal{X}_{ab}^{(2)}]^n. \quad (9.38)$$

Using (9.35), we finally have

$$\begin{aligned} \mathcal{D}_{ab}^{(2)} &= -\text{Tr} \log(1 - \mathcal{X}_{ab}^{(2)}) = -\text{Tr} \log \left(1 - \frac{\epsilon \mathcal{S}_{ab}^q}{1 + \epsilon \mathcal{S}_{ab}^q} \right) \\ &= \text{Tr} \log(1 + \epsilon \mathcal{S}_{ab}^q). \end{aligned} \quad (9.39)$$

It follows that $\mathcal{D}_{ab}^{(2)}$ only depends on the eigenvalues of $1 + \epsilon \mathcal{S}_{ab}^q$, but from (9.37) the eigenvalues do not depend on

s_a or s_b . Hence, using the q -deformed or Selberg regulators, we find that $\mathcal{D}_{ab}^{(2)}$ is a constant that does not contribute to the empty double-trumpet.

The computation of $\mathcal{D}_{ab}^{(3)}$ is presented in Appendix F 1.

B. Two-point function on the double-trumpet

In this section, we consider the double-trumpet two-point function where a single \mathcal{O} operator is inserted on each boundary,

$$\langle \text{Tr} e^{-\beta_L H} \mathcal{O} \text{Tr} e^{-\beta_R H} \mathcal{O} \rangle_{\text{cyl}}, \quad (9.40)$$

where as before cyl refers to the leading contribution in the genus expansion. As in the previous section, we can systematically classify and compute the 't Hooft diagrams that contribute. We obtain results for both the q -deformed and Selberg models and explain for the q -deformed model why these results are consistent with our results for the empty double-trumpet.

The simplest class of diagrams that contribute to (9.40) is depicted in Fig. 14. In either the q -deformed model or the Selberg model, their total contribution in the JT limit is

$$\begin{aligned} &\int_0^\infty ds_a \rho(s_a) e^{-(\beta_L + \beta_R) s_a^2} \langle \mathcal{O}_{aa} \mathcal{O}_{aa} \rangle_{\text{disk},c} \\ &= \int_0^\infty ds_a \rho(s_a) e^{-(\beta_L + \beta_R) s_a^2} \Gamma_{aa}^\Delta, \end{aligned} \quad (9.41)$$

which reproduces the first term in (2.36).

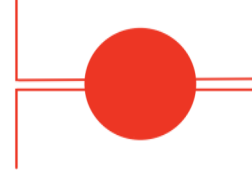


FIG. 14. The simplest class of 't Hooft diagrams that contribute to the two-point function on the double-trumpet consists of diagrams that admit a path from the left to the right boundary that does not cross over any \mathcal{O} double-lines. The red blob represents a sum over connected planar two-point diagrams. The top and bottom ends of this diagram are identified.

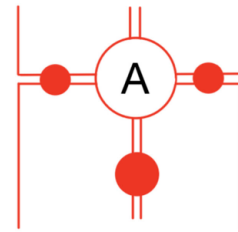


FIG. 15. We enumerate diagrams where a left-right path can cross a single \mathcal{O} double-line. The top and bottom ends of these diagrams are identified. As in Fig. 22, a blob with an "A" represents a sum over planar connected amputated diagrams.

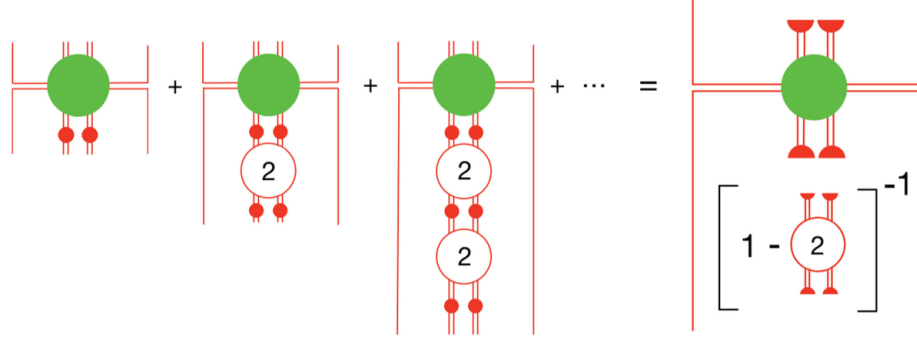


FIG. 16. In these diagrams, the minimum number of \mathcal{O} double-lines that are traversed by a left-right path is two. The blob labeled “2” refers to the same object as in Fig. 11. The top and bottom ends of these diagrams are identified. The green blob with six external double-lines is defined in Fig. 17. On the right side, we represent the sum as a geometric series, where multiplication is vertical.

Next, we consider diagrams such that the minimum number of \mathcal{O} double-lines that are traversed by a left-right path is one. These diagrams are depicted in Fig. 15. In either the Selberg or q -deformed models, they become in the JT limit

$$\begin{aligned} & \int ds_a ds_b e^{-\beta_L s_a^2} e^{-\beta_R s_b^2} \rho(s_L) \rho(s_R) \frac{\langle \mathcal{O}_{aa} \mathcal{O}_{ab} \mathcal{O}_{bb} \mathcal{O}_{ba} \rangle_{\text{disk},c}}{\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk},c}} \\ &= \int ds_a ds_b e^{-\beta_L s_a^2} e^{-\beta_R s_b^2} \rho(s_L) \rho(s_R) (\Gamma_{aa}^\Delta \Gamma_{bb}^\Delta)^{1/2} \\ & \quad \times \left\{ \begin{array}{cc} \Delta & s_a \quad s_b \\ \Delta & s_b \quad s_a \end{array} \right\}, \end{aligned} \quad (9.42)$$

which reproduces the second term in (2.36).

Next, we consider diagrams such that the minimum number of \mathcal{O} double-lines that are traversed by a left-right path is two. These are depicted in Fig. 16. To evaluate these diagrams, we should first solve for the green blob in Fig. 17. The sum of the diagrams in Fig. 16 differs between the q -deformed and Selberg models. We remind the reader that for both the Selberg and q -deformed regulators, the gravitational Feynman rules are replaced by their q -deformed counterparts. Furthermore, in the q -deformed model, each $6j$ symbol is accompanied by a power of ϵ .⁶⁰ In contrast, in the Selberg model, each connected $2n$ -point function is weighted by ϵ^{n-1} .

In the q -deformed model, the sum of the diagrams on the right-hand side of Fig. 17 is

$$\epsilon^2 \begin{array}{c} \Delta \quad \Delta \\ | \quad | \\ \hline \Delta \end{array} + \epsilon^3 \begin{array}{c} \Delta \quad \Delta \\ \diagdown \quad \diagup \\ \hline \Delta \end{array} \quad (9.43)$$

⁶⁰In Sec. VIII C, we referred to ϵ as q_B . Here, we will return to a notation that is consistent with Sec. VII.

We choose to make the factors of ϵ explicit in our graphical representations of products of $6j$ symbols. Furthermore, Fig. 13 implies that

$$\left(1 - \begin{array}{c} \text{blob with 2 double-lines} \end{array} \right)^{-1} = 1 + \epsilon \begin{array}{c} \times \end{array} \quad (9.44)$$

Note that if we multiply the left-hand side of Fig. 17 by the inverse of (9.44) and then identify the top and bottom ends of the diagram, we obtain the left-hand side of Fig. 16. Thus, we need to multiply (9.43) by the inverse of (9.44). In the q -deformed model, the sum of the diagrams in Fig. 16 then becomes

$$\epsilon^2 \begin{array}{c} \Delta \quad \Delta \\ | \quad | \\ \hline \Delta \end{array} \quad (9.45)$$

where the top and bottom ends of this diagram are identified. Equation (9.45) is a graphical representation of the following expression:

$$\begin{aligned} & \int ds_a \rho_q(s_a) ds_b \rho_q(s_b) ds_c \rho_q(s_c) e^{-\beta_L s_a^2 - \beta_R s_c^2} (\Gamma_{aa,q}^\Delta \Gamma_{cc,q}^\Delta)^{1/2} \\ & \quad \times \epsilon^2 \left\{ \begin{array}{cc} \Delta & s_a \quad s_b \\ \Delta & s_b \quad s_a \end{array} \right\}_q \left\{ \begin{array}{cc} \Delta & s_b \quad s_c \\ \Delta & s_c \quad s_b \end{array} \right\}_q. \end{aligned} \quad (9.46)$$

Using the pentagon identity in (G52), this becomes

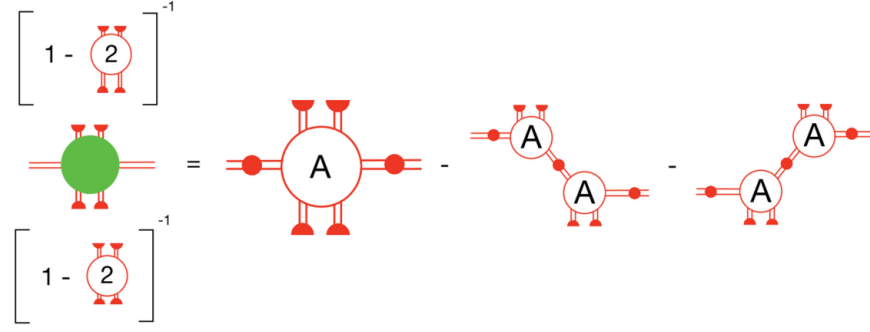


FIG. 17. This equation is used to solve for the green blob with six external double-lines. A blob labeled “A” represents a sum over connected amputated planar diagrams (for the blob with four external double-lines, this is equivalent to summing over 1PI diagrams). The green blob should consist of connected diagrams only, or else the diagrams in Fig. 16 would permit left-right paths that cross fewer than two double-lines. Furthermore, any diagrams that contribute to the “2” blob should be amputated off the bottom two and top two legs of this green blob, because in Fig. 16 we want to make all of the appearances of the “2” blob explicit. That is, there should not be any “2” blobs hidden in the green blob. If it were not for the two subtractions on the right-hand side, then the green blob would be defined to be the sum of connected six-point diagrams with “2” blobs amputated off the top two and bottom two legs. However, we must also exclude from this green blob certain connected diagrams that would permit left-right paths that cross over one double-line (this explains the subtractions on the right-hand side). The top and bottom ends of these diagrams are not identified.

$$\int ds_a \rho_q(s_a) ds_c \rho_q(s_c) e^{-\beta_L s_a^2 - \beta_R s_c^2} (\Gamma_{aa,q}^\Delta \Gamma_{cc,q}^\Delta)^{1/2} \times \epsilon^2 \sum_{m=0}^{\infty} q^{\Delta x} \left\{ \begin{matrix} 2\Delta + m & s_a & s_c \\ \Delta & s_c & s_a \end{matrix} \right\}_q, \quad (9.47)$$

which gives in the JT limit

$$\int ds_a \rho(s_a) ds_c \rho(s_c) e^{-\beta_L s_a^2 - \beta_R s_c^2} (\Gamma_{aa}^\Delta \Gamma_{cc}^\Delta)^{1/2} \times \sum_{m=0}^{\infty} \left\{ \begin{matrix} 2\Delta + m & s_a & s_c \\ \Delta & s_c & s_a \end{matrix} \right\}. \quad (9.48)$$

From (9.47), it is clear that the $\epsilon \rightarrow 1$ and $q \rightarrow 1$ limits commute in the q -deformed model.

We now consider the sum of the diagrams in Fig. 16 using the Selberg regulator. The explicit expression for the amputated blob with six external double-lines in Fig. 17 is given by summing over the four connected chord diagrams with six external lines. However, two of these chord diagrams are canceled by the subtractions on the right-hand side of Fig. 17. The remaining two chord diagrams are both weighted by ϵ^2 , which is in contrast to (9.43), where one diagram is weighted by ϵ^3 and the other by ϵ^2 . We again want to multiply the sum of these two chord diagrams by the inverse of (9.44). Using (A82) and (G52), we have that

$$\begin{array}{c} \Delta \quad \Delta \\ | \quad | \\ \hline \end{array} + \begin{array}{c} \Delta \quad \Delta \\ \diagdown \quad \diagup \\ \hline \end{array} = \sum_{n=0}^{\infty} q^{n\Delta} \left(1 + (-1)^n q^{2n\Delta + \frac{n(n-1)}{2}} \right) \begin{array}{c} \diagup \quad \diagdown \\ | \quad | \\ \hline \end{array} \quad (9.49)$$

Next, we multiply (9.49) by the inverse of (9.44) and identify the top and bottom ends. The sum over the diagrams in Fig. 16 finally becomes

$$\epsilon^2 \sum_{n=0}^{\infty} q^{n\Delta} \frac{1 + (-1)^n q^{2n\Delta + \frac{n(n-1)}{2}}}{1 + \epsilon(-1)^n q^{2n\Delta + \frac{n(n-1)}{2}}} \times \begin{array}{c} 2\Delta + n \\ | \\ \hline \end{array} \quad (9.50)$$

The JT limit corresponds to taking $\epsilon \rightarrow 1$ and $q \rightarrow 1$. For the Selberg model, the $q \rightarrow 1$ limit must be taken first, or else the answer would agree with the q -deformed model. The result is that only the even n terms in the above sum contribute. In the JT limit, the result becomes

$$\int ds_a \rho(s_a) ds_c \rho(s_c) e^{-\beta_L s_a^2 - \beta_R s_c^2} (\Gamma_{aa}^\Delta \Gamma_{cc}^\Delta)^{1/2} \times \sum_{n=0}^{\infty} \left\{ \begin{matrix} 2\Delta + 2n & s_a & s_c \\ \Delta & s_c & s_a \end{matrix} \right\}, \quad (9.51)$$

which reproduces the third term in (2.36).

The next class of 't Hooft diagrams, in which the lowest number of \mathcal{O} double-lines crossed by a left-right path is three, is computed in Appendix F 2.

C. Comments on the pair of pants

Throughout this section, we have explained how double-trumpet correlators in the matrix model may be directly computed from the disk correlators. We expect that our computational techniques generalize from the double-trumpet to topologies with more handles and boundaries. In this subsection we sketch an approach to computing the empty pair of pants.

It helps to represent nontrivial topologies using the plane. For instance, we represented the double-trumpet as a square with one pair of opposite edges identified. For the pair of pants, we may use two hexagons with an appropriate set of edge identifications. See Fig. 18. One should consider only matrix model diagrams which are connected at the level of the \mathcal{O} double-lines (ignoring trivial \mathcal{O} bubble diagrams), and nontrivially wrap all three holes.⁶¹ Otherwise, the diagram (or a disconnected subdiagram) will have already been included in a double trumpet leg. One may enumerate all the possible graphs as follows. First, one should specify the number (greater than or equal to one) of double-lines passing through each of the three blue edges in Fig. 18. Then, on each hexagon, one should draw all possible planar graphs with the specified number of external lines on each edge. We have drawn some examples in Fig. 18. Of course, this procedure will overcount the graphs. For a better count, one should use the appropriate analog of the inverse “two-to-two propagator” mentioned in Sec. V to strip off the appropriate subdiagrams from the external lines before gluing the two hexagons together, so that one is gluing correctly amputated disk diagrams. Note that in the Selberg model, the competition between the vanishing of the total disk amplitudes between nonsymmetrized states and the divergence of the inverse propagator results in only symmetrized

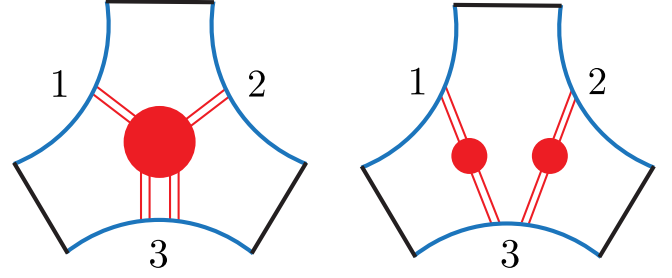


FIG. 18. A representation of the pair of pants as two hexagons. The blue edges are identified in pairs according to the labels 1, 2, and 3. The black edges refer to the boundaries that correspond to the single trace insertions in the matrix integral. By drawing planar graphs on the hexagons and gluing them across the edges, we may construct graphs with the pair of pants topology. Each red blob stands for a sum of planar diagrams with the appropriate number of external double-lines.

states appearing in each glued edge, by identical reasoning to the double trumpet two-point function calculation. Moreover, in the q -deformed model, we expect that the symmetrizers again cancel and a Hagedorn spectrum of states propagates across the blue edge.

We leave a careful study of this counting problem to future work. This matrix model analysis should lead to expressions involving integrals of products of $6j$ symbols. We expect that further calculations like the one in Sec. II C, which covers the simplest irreducible diagram in the pair of pants, can demonstrate the correspondence between certain classes of 't Hooft diagrams and certain geodesics on the pair of pants. Because the pair of pants has infinitely many different closed geodesics while the double-trumpet has only one closed geodesic, we expect the study of the pair of pants (as well as geometries with handles) to be more involved than the analysis presented in this paper.

D. UV divergences in the double-trumpet

In this section we comment on the matrix model interpretation of the problematic UV behavior in the matter partition function on the double-trumpet. We first comment on the Hagedorn temperature in the q -deformed model and then argue more generally that whenever the double-trumpet is undefined, the matrix model saddle is unstable.

1. Hagedorn temperature in the q -deformed model

An interesting outcome of our analysis is that the empty double-trumpet in the q -deformed model does not depend on the parameter ϵ that was introduced by the regularization scheme. We explicitly found that $\mathcal{D}^{(2)}$ and $\mathcal{D}^{(3)}$ do not contribute to the empty double-trumpet, and we conjectured that $\mathcal{D}^{(n)}$ for $n > 3$ also do not contribute. The contribution from $\mathcal{D}^{(1)}$ agrees with the $\epsilon = 0$ model, which in the JT limit is given by (4.42). Equation (4.42) may be interpreted as the double-trumpet path integral of a matter

⁶¹A hole is nontrivially wrapped if every path that begins on the hole and ends on any other hole crosses one or more \mathcal{O} double-lines.

theory minimally coupled to JT gravity. From (4.42), we find that the matter partition function is

$$Z(b) = \frac{1 - e^{-b}}{1 - e^{-b} - e^{-\Delta b}}. \quad (9.52)$$

The matter theory associated to this partition function has many more states than the theory of a free scalar quantized in global AdS₂. In particular, the states counted in $Z_{\text{scalar}}(b)$ span a Fock space that is defined by acting on the vacuum with creation operators a_n^\dagger for $n \geq 0$. The operator a_n^\dagger raises the energy by $\Delta + n$, and all of the creation operators commute. The states that contribute to $Z(b)$ may be enumerated in the same way, except without the $[a_n^\dagger, a_m^\dagger] = 0$ condition. That is, different orderings of the creation operators correspond to different states. Put differently, we may say that $Z(b)$ counts all the words that may be constructed using the letters \mathcal{O} and ∂ such that the rightmost letter is \mathcal{O} . The energy of a word is given by $\Delta n_{\mathcal{O}} + n_{\partial}$, where $n_{\mathcal{O}}$ and n_{∂} are the numbers of \mathcal{O} and ∂ letters.

For sufficiently small (but nonzero) b , the denominator of (9.52) goes to zero, which indicates that there is a Hagedorn temperature. When b is below the inverse Hagedorn temperature, the sum over the ‘t Hooft diagrams that contribute to (9.31) does not converge. Away from the JT limit, we see from (9.29) that a nonconvergent sum is still possible when $q_A + \tilde{q} > 1$. In the remainder of this section, we show that the saddle point we are doing perturbation theory around is perturbatively unstable in this regime. Because the empty double-trumpet does not depend on ϵ , we will for simplicity first analyze the $\epsilon = 0$ model for $q_A < 1$, which was introduced in Eq. (8.35). After integrating out B , the matrix integral becomes

$$\begin{aligned} \mathcal{Z}_{q_A, q_B=0, \tilde{q}} &= \int d^N \lambda \exp \left(\sum_{a < b} \log(\lambda_a - \lambda_b)^2 \right. \\ &\quad - N \sum_a [V_{q_A}(\lambda_a) + V_{c.t.}^{q_A, \tilde{q}}(\lambda_a)] \\ &\quad \left. - \frac{1}{2} \sum_{a, b} \log F^{q_A, \tilde{q}}(\lambda_a, \lambda_b) \right), \end{aligned} \quad (9.53)$$

where we have written the matrix as an integral over the eigenvalues of A . By making the change of variables

$$\lambda = \frac{2x}{\sqrt{1 - q_A}} \quad (9.54)$$

for each eigenvalue λ , the matrix integral becomes

$$\begin{aligned} \mathcal{Z}_{q_A, q_B=0, \tilde{q}} &\propto \int d^N x \exp \left(\sum_{a < b} \log(x_a - x_b)^2 \right. \\ &\quad - N \sum_a [\tilde{V}_{q_A}(x_a) + \tilde{V}_{c.t.}^{q_A, \tilde{q}}(x_a)] \\ &\quad \left. - \frac{1}{2} \sum_{a, b} \log \tilde{F}^{q_A, \tilde{q}}(x_a, x_b) \right), \end{aligned} \quad (9.55)$$

where \tilde{V}_{q_A} was defined in (8.26), $\tilde{V}_{c.t.}^{q_A, \tilde{q}}$ was defined in (8.40), and $\tilde{F}^{q_A, \tilde{q}}$ was defined in (8.36). The density of states $\tilde{\rho}(x)$ defines a saddle of the matrix integral when it extremizes the total potential energy, which is

$$\begin{aligned} V_{\text{total}} &= N \int dx \tilde{\rho}(x) [\tilde{V}_{q_A}(x) + \tilde{V}_{c.t.}^{q_A, \tilde{q}}(x)] \\ &\quad + \frac{1}{2} \int dx_a \tilde{\rho}(x_a) dx_b \tilde{\rho}(x_b) \\ &\quad \times [\log \tilde{F}^{q_A, \tilde{q}}(x_a, x_b) - \log(x_a - x_b)^2]. \end{aligned} \quad (9.56)$$

The function $\tilde{F}^{q_A, \tilde{q}}$ modifies the Coulomb repulsive force between different eigenvalues. Our saddle point of interest is $\tilde{\rho}(x) = N\tilde{\rho}_{q_A, 0}(x)$, which was defined in (8.29). This saddle is stable when the Hessian of V_{total} evaluated for $\tilde{\rho} = N\tilde{\rho}_{q_A, 0}$ is positive-definite. If we vary a single eigenvalue at position x_0 by δx_0 , the density of states changes by

$$\tilde{\rho}(x) \rightarrow \tilde{\rho}(x) - \delta x_0 \delta'(x - x_0). \quad (9.57)$$

More generally, if we vary all of the eigenvalues according to the rule $\delta x = f(x)$ for some smooth function f , then the change in the density of states is

$$\delta \tilde{\rho}(x) = -\frac{d}{dx} [\tilde{\rho}(x) f(x)]. \quad (9.58)$$

Note that the support of $\delta \tilde{\rho}(x)$ is $x \in [-1, 1]$, which is the same as the support of $\tilde{\rho}_{q_A, 0}(x)$. Hence, a small deformation away from the saddle point corresponds to varying the density of states by a function $\delta \tilde{\rho}(x)$ that has support for $x \in [-1, 1]$ and integrates to zero. We may expand $\delta \tilde{\rho}(x)$ in a basis of Chebyshev polynomials,

$$\delta \tilde{\rho}(x) = \sum_{n=1}^{\infty} \frac{\delta c_n}{\sqrt{1 - x^2}} T_n(x), \quad (9.59)$$

where the $n = 0$ term was omitted to ensure that

$$\int_{-1}^1 dx \delta \tilde{\rho}(x) = 0. \quad (9.60)$$

Note that the integral of each term in (9.59) is proportional to $\sqrt{x+1}$ (resp. $\sqrt{1-x}$) near the $x = -1$ (resp. $x = 1$) endpoint. This is consistent with (9.58) and the behavior of $\tilde{\rho}(x)$ near the endpoints. If we expand V_{total} around the

saddle to second order in the deformation and use (9.59), we obtain

$$V_{\text{total}} = \frac{1}{2} \int dx_a \delta \tilde{\rho}(x_a) dx_b \delta \tilde{\rho}(x_b) \times [\log \tilde{F}^{q_A, \tilde{q}}(x_a, x_b) - \log(x_a - x_b)^2]. \quad (9.61)$$

$$= \frac{\pi^2}{2} \sum_{n=1}^{\infty} \delta c_n^2 \left[-\frac{\tilde{q}^n}{(1-q^n)n} + \frac{1}{n} \right], \quad (9.62)$$

where we have used (8.39) as well as

$$\int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} T_n(x) \log(y-x)^2 = -\frac{2\pi}{n} T_n(y), \quad n \geq 1, \quad y \in (-1, 1). \quad (9.63)$$

The Hessian is positive-definite when each term in the sum in (9.62) is positive. Thus, when $q + \tilde{q} > 1$, the saddle becomes unstable. This stability criterion exactly coincides with (9.26).

2. General stability analysis

For the general q -deformed model with nonzero ϵ , the analysis above does not change very much. In fact, for a general ‘t Hooft-scaled single-trace two-matrix model, one can show that the Hessian directly determines the double-trumpet. As a reminder, the two matrices in the ‘t Hooft-scaled model are A and B , and the scale factor a is chosen such that the saddle-point spectrum of A/a has support on $[-1, 1]$. After integrating out B , the total potential energy of the eigenvalues of A takes the form

$$V_{\text{total}} = \sum_{k=1}^{\infty} N^{2-k} \int \left[\prod_{i=1}^k dx_i \tilde{\rho}(x_i) \right] \tilde{V}^{(k)}(x_1, \dots, x_k) \quad (9.64)$$

where $\tilde{\rho}(x)$ is the density of states of the rescaled matrix A/a and $x \in (-1, 1)$. Because this is a ‘t Hooft-scaled model, N does not appear in $\tilde{V}^{(k)}$. If we denote the saddle for $\tilde{\rho}$ by $N\tilde{\rho}_0$ and write $\tilde{\rho} = N\tilde{\rho}_0 + \delta\tilde{\rho}$, then (9.64) takes the form

$$V_{\text{total}} = \sum_{k=2}^{\infty} N^{2-k} \int \left[\prod_{i=1}^k dx_i \delta\tilde{\rho}(x_i) \right] \mathcal{V}^{(k)}(x_1, \dots, x_k), \quad (9.65)$$

where $\delta\tilde{\rho}$ does not appear in the expression for $\mathcal{V}^{(k)}$. The $k=1$ term does not appear because we are expanding around a saddle of V_{total} . In the large N limit the only nonvanishing term in the sum is for $k=2$, because $\delta\tilde{\rho}$ is order one in the large N expansion. Using (9.59) again, we finally have that

$$V_{\text{total}} = \frac{1}{2} \sum_{n,m=1}^{\infty} \delta c_n V_{nm}^{(H)} \delta c_m, \quad (9.66)$$

where

$$V_{nm}^{(H)} = 2 \int dx_1 dx_2 \frac{T_n(x_1)}{\sqrt{1-x_1^2}} \mathcal{V}^{(2)}(x_1, x_2) \frac{T_n(x_2)}{\sqrt{1-x_2^2}} \quad (9.67)$$

is the Hessian. For general N , the range of the δc_n coefficients is restricted by the condition that the total density of states should be nonnegative. However, in the large- N limit, the δc_n coefficients are valued on the entire real line because the perturbation $\delta\tilde{\rho}$ is always subleading compared to the saddle $N\tilde{\rho}_0$. Thus, at large N , the matrix integral becomes Gaussian in the δc_n coefficients, which reflects the fact that the multiboundary correlators are dominated by disks and cylinders.⁶²

From (9.59) we have that

$$\int_{-1}^1 dx \delta\tilde{\rho}(x) T_n(x) = \frac{\pi}{2} \delta c_n, \quad n \geq 1, \quad (9.68)$$

and it follows that at large N ,

$$\left\langle \text{Tr} T_n \left(\frac{A}{a} \right) \text{Tr} T_m \left(\frac{A}{a} \right) \right\rangle = \int dx_1 dx_2 \langle \tilde{\rho}(x_1) \tilde{\rho}(x_2) \rangle T_n(x_1) T_m(x_2) \quad (9.69)$$

$$= \left(\int_{-1}^1 dx_1 N\tilde{\rho}_0(x_1) T_n(x_1) \right) \left(\int_{-1}^1 dx_2 N\tilde{\rho}_0(x_2) T_m(x_2) \right) + \frac{\pi^2}{4} \langle \delta c_n \delta c_m \rangle. \quad (9.70)$$

Extracting the connected part of the above, we have that

$$\begin{aligned} \left\langle \text{Tr} T_n \left(\frac{A}{a} \right) \text{Tr} T_m \left(\frac{A}{a} \right) \right\rangle_{\text{cyl}} &= \frac{\pi^2}{4} \langle \delta c_n \delta c_m \rangle \\ &= \frac{\pi^2}{4} [V^{(H)}]_{nm}^{-1}. \end{aligned} \quad (9.71)$$

We have thus shown that the double-trumpet is well-defined precisely when the matrix model saddle is stable. Thus, the conclusions in Sec. IX D 1 hold for any ϵ . Similarly, although we do not have an explicit expression for the regulated double-trumpet in the Selberg model, we

⁶²In particular, in a one-matrix model, (9.64) only contains terms for $k=1, 2$, and the $k=2$ term is fixed by the Vandermonde determinant, while the $k=1$ term is fixed by the matrix potential. It is known that the cylinder amplitude is universal once the endpoints of the spectrum have been fixed. This result is reproduced by a Gaussian measure for the δc_n coefficients.

know that in the double-scaled limit the double-trumpet correlators look like

$$\begin{aligned} & \langle \text{Tr} \cos(b_1 \sqrt{H}) \text{Tr} \cos(b_2 \sqrt{H}) \rangle_{\text{cyl}} \\ &= \frac{b_1}{4} Z_{\text{scalar}}(b_1) \delta(b_1 - b_2), \end{aligned} \quad (9.72)$$

which means that in the double-scaling limit, the spectrum of the Hessian extends down to zero, leading to an instability.⁶³

X. DISCUSSION

A. Summary of results

In this work, we described various methods for studying two-matrix models dual to JT gravity minimally coupled to matter. Our aim was to demonstrate how the gravitational path integral may be used to determine the ETH ensemble that arises from coarse-graining the CFT data involving heavy operators at large but finite N .

In one approach, we identified an operator equation obeyed by two operators, \mathcal{O} and H , that is analogous to the statement in 1D CFT that the only primary operators (aside from the identity) that appear in the $\mathcal{O}\mathcal{O}$ OPE have dimensions 2Δ plus an even non-negative integer. Together with conformal invariance and associativity of the OPE, this condition entirely fixes the n -point \mathcal{O} functions to be those of a GFF.⁶⁴ We constructed a two-matrix ensemble by imposing this operator equation as a constraint on the matrices that represent the operators. This matrix model should compute the correct holographic n -point functions because we are representing the operators using matrices, and matrix multiplication is associative. This result is further supported by explicit checks of Schwinger-Dyson equations.

In another approach, we described an algorithm for determining a matrix potential that reproduces all of the gravitational disk correlators. To organize the calculation, we introduced a fictitious parameter ϵ such that gravitational Feynman diagrams with more crossings are weighted by more factors of ϵ . The matrix potential can be written in an ϵ expansion. Different schemes for weighting gravitational diagrams with powers of ϵ correspond to different ways to take the double-scaling limit. We discussed two specific schemes: the “Selberg regulator” and the “ q -deformed regulator.” By construction, the disk correlators of the matrix model do not depend on which regulator is used.

⁶³In the Selberg model, the regulated disk density of states is the same as in the q -deformed model, so we can still use (8.19) to relate H and A and set $a = \frac{2}{\sqrt{1-q}}$. In the double-scaling (or $q \rightarrow 1$) limit, we set $e^{-b} = q^n$ with fixed b such that $T_n(\frac{A}{a})$ becomes $\cos(b\sqrt{H})$ (see the discussion leading to (9.31)).

⁶⁴See Appendix D.

We showed that in any single-trace two-matrix model, the cylinder (as well as the cylinder two-point function) can be determined directly from the disk amplitudes without explicit knowledge of the potential. Our strategy was to systematically classify ‘t Hooft diagrams with cylinder topology. For the matrix models of interest in this paper, our procedure for determining the cylinder only returns a definite answer if the method for taking the double-scaling limit is known. We obtained formulas using both the q -deformed and Selberg regulators. To the extent that we checked, the Selberg model reproduces the cylinder amplitudes of JT gravity minimally coupled to a scalar field.⁶⁵ The bulk dual of the q -deformed model is not known, but the cylinder amplitude indicates that the partition function of the unknown matter that propagates on the double-trumpet has a Hagedorn temperature.

After removing the regulators, the cylinder amplitudes of the matrix models (or equivalently, the double-trumpet amplitudes of their bulk duals) are formally undefined due to the UV divergence associated to the shrinking cycle in the off-shell gravitational path integral. Hence, the matrix models do not have a well-defined genus expansion in the JT limit.⁶⁶ We explained that this is because the saddle point of the matrix integral is perturbatively unstable. It would be interesting to find a new saddle that the unstable saddle can decay to. We leave this task for future work.

B. Generalization to higher dimensions

The ETH ensembles considered in this work are analogous to the ETH ensembles that one might consider in higher dimensional AdS/CFT.⁶⁷ In a general holographic CFT, if we know the correlation functions, OPE coefficients, and scaling dimensions of the light operators only (including $1/N$ corrections), then we can construct an ensemble by averaging over all of the CFT data involving the heavy operators that is consistent with this information [56–58].⁶⁸ For example, the light-light-heavy structure constants will be constrained by the four-point function of light operators because heavy operators may run in the intermediate channel. These structure constants are

⁶⁵Due to an ambiguity in the measure of the matrix \mathcal{O} , the empty double-trumpet in the Selberg model is treated as part of the data that defines the model, rather than a nontrivial prediction of the model. The double-trumpet two-point function is a nontrivial prediction. For the q -deformed model, both the empty double-trumpet and the two-point function are nontrivial predictions.

⁶⁶Note that the regulated models can have a well-defined genus expansion.

⁶⁷We would like to thank A. Belin, J. de Boer and P. Nayak for important discussions on the issues described in this section.

⁶⁸This ensemble has the property that observables involving states that are far below the black hole threshold are not affected by ensemble averaging [59].

constrained by the requirement that the four-point function of light operators is computed correctly.

Correlators with external heavy operators also lead to crossing-symmetry constraints on the structure constants. Each structure constant can be thought of as a random tensor whose rank is the number of heavy operators. The crossing equation is quadratic in the structure constants. We can impose crossing symmetry as a constraint by squaring the crossing equation and adding it to the potential with a large coefficient Λ . This means that the potential is at most quartic in the tensors (note that in the CFT ensembles proposed by [57,58], non-Gaussianities also play an important role [60,61]). The ensemble will also include a set of variables that represent the scaling dimensions of the heavy operators. These are analogous to the eigenvalues of H in our matrix model. The potential will depend on these eigenvalues in a complicated way through their appearance in the crossing equations. The number of eigenvalues must be scaled to infinity as Λ is sent to infinity; the precise way to do this might be similar to the scaling procedure discussed in Sec. VIB.

To reproduce the correlation functions of light operators at finite temperature, only tensors up to rank 2 given by light-heavy-heavy structure constants will play a significant role. The resulting matrix models will be very similar to those constructed in this paper, except that vector spaces on which they act will be graded by spin, and there will be a matrix for every light operator, including multi-twist operators. In holographic theories, it may be possible to express the ensemble only in terms of the primitive operators, dual to bulk fields, by integrating out the multitwist matrices.

Previous work in $\text{AdS}_3/\text{CFT}_2$ [57,58] has shown that imposing crossing symmetry for four-point functions of only heavy operators is not necessary for the ensemble to have a gravity dual. However, if we constrain the structure constants to ensure that crossing symmetry is obeyed for *all* correlators, then the ensemble is an average over all the CFTs that are consistent with the initially given information on the light operator data. The given light operator data may not correspond to any solution of the crossing equations, in which case we may say that the light operator data belongs to the swampland.

Suppose we use the light operator data from a theory in the landscape to construct an ensemble by squaring the crossing equations, as described above. Suppose also that we impose crossing symmetry for all correlators, so that the solutions to the constraints are actual CFTs. Suppose for simplicity that there is a unique solution to the crossing equations given the light data. One might think that computing ensemble-averaged observables in the $\Lambda \rightarrow \infty$ limit is no easier than solving the entire theory, which is extremely difficult. Nonetheless, it still may be possible to study interesting features of the ensemble for large Λ . To be more precise, let Λ_h refer to the Λ parameter that multiplies

the square of the crossing equation for four external heavy operators.⁶⁹ This parameter should be viewed as a control over how much coarse-graining is being performed.⁷⁰ For $\Lambda_h = \infty$ and fixed N (or fixed central charge), the spectrum of heavy operators will be a sum over delta functions, corresponding to the solution of the crossing equations. For Λ_h large but finite, these delta functions will be smeared by various amounts, resulting in a coarse-grained spectrum. It is unlikely that the distribution of CFT data *uniformly* converges to a sum of delta functions in the $\Lambda_h \rightarrow \infty$ limit. For large but finite Λ_h , we expect that the spectrum of scaling dimensions will still look smooth (as opposed to being a sum of slightly smeared delta functions) for scaling dimensions $\gtrsim \Lambda_h$. Another way to obtain a smooth spectrum is to take the $N \rightarrow \infty$ limit in tandem with the $\Lambda_h \rightarrow \infty$ limit. Because Λ_h is interpreted as a coarse-graining parameter, it is reasonable to expect that the smooth parts of the spectrum can be reproduced by a semiclassical bulk dual. Perhaps there are features in the smooth spectrum for scaling dimensions on the order of Λ_h that can distinguish whether the light CFT data is in the landscape or swampland. It would also be interesting to study this ensemble using the Schwinger-Dyson techniques of Sec. VI.

Note that in the ETH ensemble described in Sec. VI, there are light operators but no light states, and there are heavy states but no heavy operators. The heavy states are the eigenvalues of H , while \mathcal{O} is the light operator. Our constraint is analogous to the statement in 1D CFT that the spectrum of primary operators in the \mathcal{OO} OPE should be that of the generalized free field. Given our checks of the Schwinger-Dyson equations in Sec. VI, we believe that this constraint is sufficient to construct an ensemble that reproduces all of the holographic disk correlators. There is no analog of crossing symmetry of heavy external operators in our model. While we have not explicitly studied the space of solutions to our constraint, we know that solutions exist. For example, the SYK model in the appropriate scaling regime furnishes a solution to the constraint. Our constraint holds as an operator equation in any ensemble-averaged theory whose correlators reproduce the disk correlators of JT gravity minimally coupled to a scalar field. Our ensemble is maximally ignorant in the sense that all theories that obey the constraint are averaged over.

C. Other ETH ensembles

A point we want to emphasize is that it is important to have a principle that determines the ETH ensemble beyond

⁶⁹In general, there may be multiple Λ parameters associated to the multiple constraints we impose.

⁷⁰ Λ_h is analogous to the γ parameter in (2.48) of [62]. The authors of [62] were able to give their coarse-grained/ensemble-averaged theory a gravity description involving interacting end-of-the-world branes.

the requirement that it reproduces gravitational path integral calculations. Without such a principle, we would not be able to learn anything new from the ETH ensemble. One such principle is that the ensemble should be the maximally ignorant ensemble that agrees with the gravitational path integral calculations that we know how to do. In practice, one should fix an ansatz for the ensemble and tune the parameters to match the gravitational path integral results. Different ETH ansätze correspond to different schemes for coarse-graining the CFT data. Our ETH ansatz was simply a two-matrix model with a single-trace potential. It would be interesting to compare our notion of coarse-graining to other notions of coarse-graining that exist in the literature [63,64]. One might also try to connect the ETH matrix model that would arise from such coarse-graining procedures to the Goldstone effective theories of chaos described in [65,66].

Another ETH ensemble was considered in [26,29,67]. In [26], the principle that determines the ensemble was called the assumption of local typicality. This is the statement that simple observables are unaffected by conjugating the simple operators with block-diagonal random unitary matrices that act separately in each microcanonical window. From this principle, one can construct an ansatz for multipoint thermal correlators of simple operators. To leading order in e^S , this ensemble is the same as our two-matrix ensemble. For example, the ETH diagrams in Fig. 2 of [67] are the same as the ‘t Hooft diagrams on the left side of (5.3) after accounting for the fact that the one-point correlators of \mathcal{O} vanish in our model. To leading order in e^S , the diagrams in [29] are isomorphic to the planar ‘t Hooft diagrams in this paper after specializing to the case of having a single simple operator. However, the ensemble constructed from local typicality differs from our matrix ensemble at subleading order in e^S . Using the assumption of local typicality, [26] was able to match certain wormhole contributions to multipoint amplitudes, but without the contribution from the matter determinant. Unlike the principle of local typicality, our ensemble includes an average over the eigenvalues of \mathcal{O} . This averaging appears to be important for reproducing the matter determinant, which is an important aspect of bulk locality. Furthermore, [68,69] studied the matrix elements of simple operators in extremal (or zero-energy) microstates of certain supersymmetric black holes and found that they are approximately described by a Gaussian random matrix. Hence, including the \mathcal{O} operator in the matrix ensemble is natural and well-motivated.

D. Future directions

There are various directions we wish to explore, some of which have already been mentioned. We are interested in studying the fate of the double-trumpet UV divergences in our models by finding a stable saddle that the unstable saddle can decay to. This computation might be easiest in

the Gaussian toy model of Sec. IV, or its q -deformed counterpart in (8.35). In the q -deformed model, we can tune the parameters such that only one eigenvalue of the Hessian has the wrong sign. In this case, there is only one unstable direction.

Another interesting idea is to find a modular-invariant spectrum of primary operators in a CFT_2 using a constrained ensemble constructed analogously to the one that we considered in Secs. III and VI.⁷¹ The scaling dimensions of the primaries in a given spin sector can be represented as eigenvalues in an eigenvalue integral, and the modular crossing equation can be squared and added to the potential with a large coefficient. This question is interesting to study because the Maloney-Witten-Keller [70,71] partition function, which is a sum over $\text{SL}(2, \mathbb{Z})$ images of the vacuum Virasoro character, yields a modular-invariant spectrum that exhibits negativity. By representing the scaling dimensions as eigenvalues in an eigenvalue integral, it is impossible for the spectrum to become negative. The Schwinger-Dyson equations of the constrained model may be useful for finding a modular-invariant spectrum free of negativity.

It would also be interesting to use the q -deformed model of Sec. VIII to study the bulk reconstruction of JT gravity with matter in the same manner as [55]. In particular, [55] noted that in the double-scaled SYK model, one may write an ensemble-averaged correlator as an overlap between a bra and a ket state. These states live in a bulk Hilbert space, and their quantum numbers count the number of chords that pass through a spatial slice. These chord numbers encode the length of the slice as well as the positions of matter particles along the slice. Because the disk correlators of the q -deformed model are also computed by summing over chord diagrams, one can reconstruct the same bulk Hilbert space from the q -deformed model. In the planar limit, the disk correlators (such as (8.21)) exactly agree with the double-scaled SYK model. However, the subleading corrections to the correlators differ between the two models.⁷² It would be interesting to study how bulk reconstruction differs between these two models when subleading corrections are considered. Note that due to these corrections, the bulk to boundary map is no longer an exact isometry. This is an important feature of realistic models of AdS/CFT [72].

Finally, we want to better understand the relationship between the q -deformed and Selberg models of Sec. VIII

⁷¹We thank Tom Hartman for suggesting this idea to us.

⁷²As mentioned previously, finding the subleading corrections in the q -deformed model in the JT limit is challenging due to UV divergences. However, away from the JT limit, we found a parameter regime (9.26) where the double-trumpet was finite. Without a better nonperturbative understanding of the q -deformed model, one should compare the q -deformed model and the double-scaled SYK model in a parameter regime where the matrix model saddle is stable.

and the constrained ensemble of Sec. III. The former models appear to be perturbatively unstable (the most explicit check is performed in Sec. IX D 1) while the constrained model is nonperturbatively well-defined, since the potential in (3.20) is bounded from below.⁷³ We have not been able to explicitly show that the Selberg and q -deformed models resemble (3.20) in the double-scaling/JT limit, although we have not ruled out this possibility. If we cannot find a stable saddle in the Selberg and q -deformed models, then we can instead use the constrained ensemble (3.20) to understand how the UV divergences in the bulk can be regulated by nonperturbative effects. We leave more explicit studies of the constrained model to future work.

ACKNOWLEDGMENTS

We would like to thank Nick Agia, Alex Belin, Noam Chai, Jan de Boer, Anatoly Dymarsky, Lorenz Eberhardt, Akash Goel, Tom Hartman, Clifford Johnson, Zohar Komargodski, Henry Lin, Juan Maldacena, Dalimil Mazáč, Vladimir Narovlansky, Pranjali Nayak, Joaquin Turiaci, and Herman Verlinde for stimulating discussions. This work was performed in part at Aspen Center for Physics, which is supported by National Science Foundation Grant No. PHY-1607611. This work has been partially supported by the DOE through the Grant No. DE-SC0007870, the SNF through Project Grant No. 200020 182513, as well as the NCCR 51NF40-141869 The Mathematics of Physics (SwissMAP). The work of B.M. was supported by a grant from the Simons Foundation (651444, B.M.) and NSF Grant No. PHY-2014071.

APPENDIX A: SPECIAL FUNCTIONS

In this appendix, we introduce the special functions used in the main text. For the properties of Wilson polynomials and Wilson functions we follow [73,74].

1. Wilson polynomials

Wilson polynomials $W_n(x) = W_n(x; a, b, c, d)$ are n th order polynomials of x^2 . They depend on four complex parameters $a = a_1, b = a_2, c = a_3, d = a_4$, such that non-real parameters appear in complex conjugate pairs. They are defined by

$$W_n(x; a, b, c, d) = (a+b)_n(a+c)_n(a+d)_n \times {}_4F_3 \left(\begin{matrix} -n, n+a+b+c+d-1, a+ix, \\ a+b, a+c, a+d \end{matrix}; 1 \right). \quad (\text{A1})$$

They are symmetric in a, b, c, d and obey an orthogonality relation

$$\int_0^\infty \frac{dx}{2\pi} M(x) W_n(x) W_m(x) = r_n \delta_{nm},$$

$$r_n = n! \frac{(n-1 + \sum a_j)_n}{\Gamma(2n + \sum a_j)} \prod_{i < j} \Gamma(a_i + a_j + n). \quad (\text{A2})$$

The measure is

$$M(x) \equiv M(x; a, b, c, d) = \frac{\prod_{j=1}^4 \Gamma(a_j \pm ix)}{\Gamma(\pm 2ix)}. \quad (\text{A3})$$

For a general choice of $a_j \in \mathbb{C}$ there is also a discrete part of the measure [73]. However, it is absent if $\text{Re} a_j > 0$, the case relevant for JT gravity.

We also define rescaled Wilson polynomials⁷⁴ $P_n(x) \equiv P_n(x; a, b, c, d)$ that are orthonormal with respect to the Schwarzian measure

$$P_n(x) = \frac{1}{\sqrt{r_n}} \prod_{j=1}^4 \Gamma(a_j \pm ix)^{1/2} W_n(x), \quad (\text{A4})$$

$$\int_0^\infty \frac{dx}{2\pi \Gamma(\pm 2ix)} P_n(x) P_m(x) = \delta_{nm}. \quad (\text{A5})$$

2. Choice of parameters relevant for JT gravity

The choice of parameters relevant for JT gravity is

$$a = d^* = \Delta_1 + ik_1, \quad b = c^* = \Delta_2 + ik_2. \quad (\text{A6})$$

We therefore define

$$P_n^{\Delta_1, \Delta_2}(x; k_1, k_2) := P_n(x; \Delta_1 \pm ik_1, \Delta_2 \pm ik_2), \quad (\text{A7})$$

where by $\Delta \pm ik$ we mean that P_n depends on both parameters and the order is not important due to symmetry in a, b, c, d . Pictorially, we represent the Wilson polynomial as

⁷³We expect that $V(H)$ can be chosen to be bounded from below. The role of $V(H)$ is to ensure that in the double-scaling limit, the density of states of H becomes, in the saddle-point approximation, the disk density of states of JT gravity. Away from the double-scaling limit, the density of states of H will have compact support. This means that as the double-scaling limit is taken, $V(x)$ can be designed to approach $+\infty$ for $x \rightarrow \pm\infty$.

⁷⁴Strictly speaking, P_n are not polynomials, but we abuse the language slightly and still call them Wilson polynomials.

$$P_n^{\Delta_1, \Delta_2}(x; k_1, k_2) = \begin{array}{c} \Delta_1 \quad \quad \Delta_2 \\ \quad \quad x \\ \quad \quad \diagdown \quad \diagup \\ \quad \quad k_1 \quad k_2 \\ \quad \quad | \\ \Delta_1 + \Delta_2 + n \end{array} . \quad (\text{A8})$$

We often denote the bottom line by n instead of $\Delta_1 + \Delta_2 + n$ for brevity.

3. Wilson function

The Wilson function is a certain analytic continuation of the Wilson polynomial in its degree. It is defined by [73]

$$\mathcal{W}_\lambda(x; a, b, c, d) \equiv \phi_\lambda(x; a, b, c, 1-d) \quad (\text{A9})$$

$$= \frac{1}{\Gamma(a+b)\Gamma(a+c)\Gamma(a+d)} \frac{\Gamma(g+a+i\lambda)}{\Gamma(g-a-i\lambda)\Gamma(g+i\lambda \pm ix)} \quad (\text{A10})$$

$$W(g+a-1+i\lambda; a+ix, a-ix, \tilde{a}+i\lambda, \tilde{b}+i\lambda, \tilde{c}+i\lambda), \quad (\text{A11})$$

where a “very-well poised” hypergeometric series is

$$W(a; b, c, d, e, f) \equiv {}_7F_6 \left(\begin{matrix} a, & 1+\frac{a}{2}, & b, & c, & d, & e, & f \\ \frac{a}{2}, & 1+a-b, & 1+a-c, & 1+a-d, & 1+a-e, & 1+a-f \end{matrix} ; 1 \right). \quad (\text{A12})$$

The “dual” parameters are defined by

$$g = \frac{1}{2} \sum_{i=1}^4 a_i = \frac{1}{2} \sum_{i=1}^4 \tilde{a}_i, \quad (\text{A13})$$

$$\tilde{a}_1 = g - a_4, \quad \tilde{a}_4 = g - a_1, \quad (\text{A14})$$

$$\tilde{a}_2 = g - a_3, \quad \tilde{a}_3 = g - a_2. \quad (\text{A15})$$

Here and below we set $a_1 = a, a_2 = b, a_3 = c, a_4 = d$. We will use a_1, a_2, a_3, a_4 , and a, b, c, d interchangeably.

Two more representations that follow from some non-trivial identities of hypergeometric functions are [see Eq. (3.3) and Proposition 4.4 in [73]]

$$\begin{aligned} \mathcal{W}_\lambda(x; a, b, c, d) &= \frac{\Gamma(d-a)}{\Gamma(a+b)\Gamma(a+c)\Gamma(d \pm ix)\Gamma(\tilde{d} \pm i\lambda)} \\ &\times {}_4F_3 \left(\begin{matrix} a+ix, a-ix, \tilde{a}+i\lambda, \tilde{a}-i\lambda \\ a+b, a+c, a-d+1 \end{matrix} ; 1 \right) \\ &+ (a \leftrightarrow d), \end{aligned} \quad (\text{A16})$$

$$\begin{aligned} \mathcal{W}_\lambda(x, a_1, a_2, a_3, a_4) &= \frac{\Gamma(-2i\lambda)}{\Gamma(g+i\lambda \pm ix) \prod_{i=1}^4 \Gamma(\tilde{a}_i - i\lambda)} \\ &\times {}_4F_3 \left(\begin{matrix} \tilde{a}_1+i\lambda, \tilde{a}_2+i\lambda, \tilde{a}_3+i\lambda, \tilde{a}_4+i\lambda \\ g+i\lambda+ix, g+i\lambda-ix, 1+2i\lambda \end{matrix} ; 1 \right) \\ &+ (\lambda \rightarrow -\lambda), \end{aligned} \quad (\text{A17})$$

The representations (A16) and (A17) are useful for writing the Wilson function as contour integrals, as we will discuss below. Equation (A17) is also useful for deriving asymptotics of the Wilson function (and related $6j$ -symbol) at large x .

The Wilson function has the following properties:

- (i) $\mathcal{W}_\lambda(x; a, b, c, d)$ is symmetric in a, b, c, d .
- (ii) It satisfies a duality $\mathcal{W}_\lambda(x; a, b, c, d) = \mathcal{W}_x(\lambda; \tilde{a}, \tilde{b}, \tilde{c}, \tilde{d})$.

For $\lambda = \pm i(\tilde{a} + n)$ the second term in (A16) vanishes because of $\Gamma(\tilde{a} + i\lambda)$ in the denominator, while the first term reduces to the Wilson polynomial up to gamma-function factors.

4. $\mathfrak{sl}(2, \mathbb{R})$ $6j$ -symbol

The $6j$ -symbol used in the main text is defined by

$$\begin{aligned} &\left\{ \begin{array}{ccc} \Delta_1 & k_1 & x \\ \Delta_2 & k_2 & \lambda \end{array} \right\} \\ &= (\gamma_{1x}^{\Delta_1} \gamma_{2\lambda}^{\Delta_1} \gamma_{1\lambda}^{\Delta_2} \gamma_{2x}^{\Delta_2})^{1/2} \mathcal{W}_\lambda(x; \Delta_1 \pm ik_1, \Delta_2 \pm ik_2) \end{aligned} \quad (\text{A18})$$

$$= \prod_{i=1}^4 (\Gamma(a_i \pm ix) \Gamma(\tilde{a}_i \pm i\lambda))^{1/2} \mathcal{W}_\lambda(x; a, b, c, d), \quad (\text{A19})$$

where in the first line by $\Delta \pm ik$ we again mean that \mathcal{W}_λ depends on both and the order is not important due to symmetry in a, b, c, d . The parameters are chosen as in (A6). We also defined

$$\gamma_{12}^{\Delta} = \Gamma(\Delta \pm ik_1 \pm ik_2). \quad (\text{A20})$$

The dual parameters exchange Δ_1 and Δ_2

$$a = d^* = \Delta_1 + ik_1, \quad b = c^* = \Delta_2 + ik_2, \quad (\text{A21})$$

$$\tilde{a} = \tilde{d}^* = \Delta_2 + ik_1, \quad \tilde{b} = \tilde{c}^* = \Delta_1 + ik_2. \quad (\text{A22})$$

Pictorially, we represent the $6j$ -symbol as

$$\left\{ \begin{array}{ccc} \Delta_1 & k_1 & x \\ \Delta_2 & k_2 & \lambda \end{array} \right\} = \begin{array}{c} \Delta_1 \quad \Delta_2 \\ \diagdown \quad \diagup \\ x \\ \diagup \quad \diagdown \\ k_1 \quad k_2 \\ \diagdown \quad \diagup \\ \lambda \end{array} \quad (\text{A23})$$

This is a $6j$ symbol associated to the triple tensor product $\pi^+ \otimes \pi^P \otimes \pi^+$, where π^+ is a positive discrete series representation and π^P is a principal series representation of $\mathfrak{sl}(2, \mathbb{R})$. See [74] for details.

Using the representation (A17) we can derive the asymptotic at large x . In this limit the hypergeometric function drops out and we have

$$\left\{ \begin{array}{ccc} \Delta & k_1 & x \\ \Delta & k_2 & \lambda \end{array} \right\} \approx \frac{2\pi}{x} e^{-\pi x} x^{2i\lambda} \Gamma(2i\lambda) \prod_{j=1}^4 \left(\frac{\Gamma(\tilde{a}_j - i\lambda)}{\Gamma(\tilde{a}_j + i\lambda)} \right)^{1/2} + (\lambda \rightarrow -\lambda), \quad (x \rightarrow \infty). \quad (\text{A24})$$

5. Properties of the $6j$ -symbol

The Wilson function transform (of type I) is defined in [73]

$$(\mathcal{F}f)(\lambda) = \int_0^\infty \frac{dx}{2\pi} M(x) \mathcal{W}_\lambda(x) f(x). \quad (\text{A25})$$

The Wilson function transform of the Wilson polynomial is the Wilson polynomial with dual parameters up to a sign

$$(\mathcal{F}W_n)(\lambda) = (-1)^n \tilde{W}_n(\lambda), \quad (\text{A26})$$

where $\tilde{W}_n(x; a, b, c, d) := W_n(x; \tilde{a}, \tilde{b}, \tilde{c}, \tilde{d})$. This can be written in terms of the $6j$ -symbol and rescaled polynomials P_n

$$\int_0^\infty \frac{dx}{2\pi\Gamma(\pm 2ix)} \left\{ \begin{array}{ccc} \Delta_1 & k_1 & x \\ \Delta_2 & k_2 & \lambda \end{array} \right\} P_n^{\Delta_1, \Delta_2}(x) = (-1)^n P_n^{\Delta_2, \Delta_1}(\lambda). \quad (\text{A27})$$

Note that in the rhs Δ_1, Δ_2 are exchanged, which is how duality acts on the JT parameters (A22). Since P_n are a full set of orthonormal functions, (A27) implies

$$\left\{ \begin{array}{ccc} \Delta_1 & k_1 & x \\ \Delta_2 & k_2 & \lambda \end{array} \right\} = \sum_{n=0}^{\infty} (-1)^n P_n^{\Delta_1, \Delta_2}(x) P_n^{\Delta_2, \Delta_1}(\lambda). \quad (\text{A28})$$

As pointed out in [73], this is not actually convergent. We assume that it can be thought of as an identity of distributions and converges after integrating with a test function. We represent (A28) pictorially

$$\begin{array}{c} \Delta_1 \quad \Delta_2 \\ \diagdown \quad \diagup \\ x \\ \diagup \quad \diagdown \\ k_1 \quad k_2 \\ \diagdown \quad \diagup \\ \lambda \end{array} = \sum_{n=0}^{\infty} (-1)^n \begin{array}{c} \Delta_1 \quad \Delta_2 \\ \diagdown \quad \diagup \\ x \\ \diagup \quad \diagdown \\ k_1 \quad k_2 \\ \diagdown \quad \diagup \\ \Delta_1 + n \quad \Delta_2 + n \\ \diagdown \quad \diagup \\ \lambda \end{array} \quad (\text{A29})$$

An analogous equation can be written for the resolution of identity

$$2\pi\Gamma(\pm 2ix)\delta(x - \lambda) = \sum_{n=0}^{\infty} P_n^{\Delta_1, \Delta_2}(x) P_n^{\Delta_1, \Delta_2}(\lambda), \quad (\text{A30})$$

$$\begin{array}{c} \Delta_1 \quad \Delta_2 \\ \diagdown \quad \diagup \\ x \\ \diagup \quad \diagdown \\ k_1 \quad k_2 \\ \diagdown \quad \diagup \\ \lambda \end{array} = \sum_{n=0}^{\infty} \begin{array}{c} \Delta_1 \quad \Delta_2 \\ \diagdown \quad \diagup \\ x \\ \diagup \quad \diagdown \\ k_1 \quad k_2 \\ \diagdown \quad \diagup \\ \Delta_1 + n \quad \Delta_2 + n \\ \diagdown \quad \diagup \\ \lambda \end{array} \quad (\text{A31})$$

Here, note that Δ_1, Δ_2 are interchanged in the bottom part of the diagram in comparison to (A29). Furthermore, note that each term appearing in the sum on the right-hand side of (A29) or (A31) may be interpreted as a conformal block (dressed with the Schwarzian mode) that corresponds to the exchange of a primary with dimension $\Delta_1 + \Delta_2 + n$ in a four-point function. For the simple case where $\Delta_1 = \Delta_2$, we verify this statement in Appendix C 1.

The ‘‘pentagon identity’’ can be derived by analytically continuing the formula at the top of p. 33 in [74]

of Wilson polynomials in the rhs, and obtain another useful relation (Theorem 7.5 in [74])

$$\Delta \begin{array}{c} \Delta_1 \quad \Delta_2 \\ | \quad | \\ k_1 \quad x \quad k_2 \\ | \quad | \\ k_4 \quad \lambda \quad k_3 \end{array} = \sum_{n=0}^{\infty} \Delta \begin{array}{c} \Delta_1 \quad \Delta_2 \\ | \quad | \\ k_1 \quad n \quad k_2 \\ | \quad | \\ k_4 \quad \lambda \quad k_3 \\ \Delta_1 \quad \Delta_2 \end{array} . \quad (A32)$$

$$\begin{array}{c} \Delta_1 \quad \Delta_2 \\ | \quad | \\ \text{---} \quad \text{---} \\ | \\ n \end{array} = \begin{array}{c} \Delta_1 \quad \Delta_2 \\ | \quad | \\ \text{---} \quad \text{---} \\ | \\ n \end{array} , \quad (A33)$$

This is essentially the identity resolution (A31) with an additional horizontal line Δ . But written out explicitly it contains a product of two $6j$ -symbols in the lhs and one $6j$ -symbol and two Wilson polynomials P_n in the rhs. One can integrate this with the Wilson polynomial P_n (pictorially, attach it e.g. in the bottom), use orthogonality

where we omitted the energy labels of regions. They are the same on both sides, except for the loop in the lhs which represents an integral with the Schwarzian measure $\int_0^\infty \frac{dx}{2\pi i(\pm 2ix)}$.

The Yang-Baxter equation (2.15) can be derived using (A29) and (A33)

$$\begin{array}{c} \diagup \quad \diagdown \\ | \\ \text{---} \end{array} = \sum_{n=0}^{\infty} (-1)^n \begin{array}{c} \diagup \quad \diagdown \\ | \\ \text{---} \quad \text{---} \\ | \\ n \end{array} = \sum_{n=0}^{\infty} (-1)^n \begin{array}{c} \diagup \quad \diagdown \\ | \\ \text{---} \quad \text{---} \\ | \\ n \end{array} \quad (A34)$$

$$= \sum_{n=0}^{\infty} (-1)^n \begin{array}{c} \text{---} \quad \text{---} \\ | \\ \diagup \quad \diagdown \\ | \\ n \end{array} = \begin{array}{c} \diagup \quad \diagdown \\ | \\ \text{---} \end{array} . \quad (A35)$$

6. Contour integral representation

A useful representation of the $6j$ -symbol was discussed in Appendix B of [16]

$$\left\{ \begin{array}{ccc} \Delta_1 & k_1 & k_2 \\ \Delta_2 & k_3 & k_4 \end{array} \right\} = e^{i\phi/2} \int_{-i\infty}^{i\infty} \frac{du}{2\pi i} f(u), \quad (A36)$$

where

$$f(u) = \Gamma\left(u - \frac{i}{2}k_{1+3} \pm ik_2\right) \Gamma\left(u + \frac{i}{2}k_{1+3} \pm ik_4\right) \frac{\Gamma(\Delta_1 - u - \frac{i}{2}k_{1-3})\Gamma(\Delta_2 - u + \frac{i}{2}k_{1-3})}{\Gamma(\Delta_1 + u + \frac{i}{2}k_{1-3})\Gamma(\Delta_2 + u - \frac{i}{2}k_{1-3})}, \quad (A37)$$

$$e^{i\phi} = \frac{\Gamma(\Delta_1 + ik_1 \pm ik_2)\Gamma(\Delta_1 - ik_3 \pm ik_4)\Gamma(\Delta_2 - ik_1 \pm ik_4)\Gamma(\Delta_2 + ik_3 \pm ik_2)}{\Gamma(\Delta_1 - ik_1 \pm ik_2)\Gamma(\Delta_1 + ik_3 \pm ik_4)\Gamma(\Delta_2 + ik_1 \pm ik_4)\Gamma(\Delta_2 - ik_3 \pm ik_2)}. \quad (A38)$$

We also used a short notation $k_{1+3} = k_1 + k_3, k_{1-3} = k_1 - k_3$. This integral is related to Eq. (B.28) in [16] by a shift of u . The integrand contains semi-infinite sequences of poles. The contour is such that poles of $\Gamma(u + \dots)$ are to the left of the contour and poles of $\Gamma(-u + \dots)$ are to the right. Closing the contour to the right and picking up the poles one obtains a sum of two ${}_4F_3$'s related to the representation (A16) of the Wilson function.

The integral simplifies significantly in a special case $\Delta_1 = \Delta_2 = \Delta, k_3 = -k_1$. In this case the phase factor is absent $e^{i\phi} = 1$ and (after relabeling)

$$\left\{ \begin{array}{ccc} \Delta & k & k_1 \\ \Delta & k & k_2 \end{array} \right\} = \int_{\epsilon-i\infty}^{\epsilon+i\infty} \frac{du}{2\pi i} \Gamma(u \pm ik_1) \Gamma(u \pm ik_2) \times \frac{\Gamma(\Delta - u \pm ik)}{\Gamma(\Delta + u \pm ik)}. \quad (\text{A39})$$

We also used that the $6j$ -symbol is an even function of k_j . For the computation of the geodesic on the pair of pants in Sec. II C, it is convenient to take the Fourier transform of (A39) in k, k_1, k_2 . We note that

$$\Gamma(u \pm ik) = \Gamma(2u) \int_0^\infty db \frac{2 \cos(bk)}{(2 \cosh \frac{b}{2})^{2u}} \quad (\text{A40})$$

$$\frac{\Gamma(x+ik)}{\Gamma(y+ik)} = \frac{1}{\Gamma(y-x)} \int_0^\infty db e^{-ibk} e^{-\frac{b}{2}(x+y-1)} \left(2 \sinh \frac{b}{2} \right)^{y-x-1}. \quad (\text{A41})$$

It is somewhat easier to derive the inverse Fourier of these formulas. Then one has integrals of gamma-functions that can be computed by deforming the contour and picking up residues at poles of the gamma functions. Applying these expressions to $\Gamma(u \pm ik_1), \Gamma(u \pm ik_2), \frac{\Gamma(\Delta-u+ik)}{\Gamma(\Delta+u+ik)}, \frac{\Gamma(\Delta-u-ik)}{\Gamma(\Delta+u-ik)}$ we find from (A39)

$$\left\{ \begin{array}{ccc} \Delta & k & k_1 \\ \Delta & k & k_2 \end{array} \right\} = \int_0^\infty db_1 db_2 db db' \cos(b_1 k_1) \cos(b_2 k_2) e^{ik(b-b')} \quad (\text{A42})$$

$$\frac{e^{-(\Delta-\frac{1}{2})(b+b')}}{\sinh \frac{b}{2} \sinh \frac{b'}{2}} \int_{-i\infty}^{i\infty} \frac{du}{2\pi i} \left(\frac{\sinh \frac{b}{2} \sinh \frac{b'}{2}}{\cosh \frac{b_1}{2} \cosh \frac{b_2}{2}} \right)^{2u} = \frac{1}{2} \int_0^\infty db_1 db_2 db db' \cos(b_1 k_1) \cos(b_2 k_2) \times e^{ik(b-b')} e^{-(\Delta-\frac{1}{2})(b+b')} \quad (\text{A43})$$

$$\delta \left(\sinh \frac{b}{2} \sinh \frac{b'}{2} - \cosh \frac{b_1}{2} \cosh \frac{b_2}{2} \right). \quad (\text{A44})$$

The u -integral resulted in a delta function. Now we change variables to $b_\pm = b \pm b'$ and integrate out b_+ using the delta function. We have

$$\int_0^\infty db db' = \frac{1}{2} \int_{-\infty}^\infty db_- \int_{|b_-|}^\infty db_+, \quad (\text{A45})$$

$$\delta \left(\sinh \frac{b}{2} \sinh \frac{b'}{2} - \cosh \frac{b_1}{2} \cosh \frac{b_2}{2} \right) = \frac{4}{\sinh \frac{B}{2}} \delta(b_+ - B), \quad (\text{A46})$$

where B is defined by

$$\cosh \frac{B}{2} = \cosh \frac{b_-}{2} + 2 \cosh \frac{b_1}{2} \cosh \frac{b_2}{2}. \quad (\text{A47})$$

Note that this implies $B > |b_-|$, assuming $B > 0$. So B is in the integration range $b_+ \in (|b_-|, \infty)$. Therefore integrating out b_+ leads to

$$\left\{ \begin{array}{ccc} \Delta & k & k_1 \\ \Delta & k & k_2 \end{array} \right\} = 2 \int_0^\infty db_1 db_2 db_- \cos(b_1 k_1) \cos(b_2 k_2) \times \cos(b_- k) \frac{e^{-(\Delta-\frac{1}{2})B}}{\sinh \frac{B}{2}}. \quad (\text{A48})$$

Finally, we give yet another integral representation of the $6j$ -symbol that looks a bit simpler than (A36), though we do not use it in the main text. It is

$$\left\{ \begin{array}{ccc} \Delta_1 & k_1 & x \\ \Delta_2 & k_2 & \lambda \end{array} \right\} = \left(\frac{\gamma_{2x}^{\Delta_2} \gamma_{2\lambda}^{\Delta_1}}{\gamma_{1x}^{\Delta_1} \gamma_{1\lambda}^{\Delta_2}} \right)^{1/2} \int_{-i\infty}^{i\infty} \frac{du}{2\pi i} \frac{\Gamma(\Delta_1 \pm ix + u) \Gamma(\Delta_2 \pm i\lambda + u)}{\Gamma(\Delta_1 + \Delta_2 \pm ik_2 + u)} \times \Gamma(\pm ik_1 - u). \quad (\text{A49})$$

Closing the contour to the right or, equivalently, using formula (B.18) in [16], we obtain the sum of two ${}_4F_3$'s that is related to the representation of the Wilson function (A17). The integrand doesn't obey all symmetries of the $6j$ -symbol, so applying symmetries of the $6j$ -symbol we can get different integral representations.

7. q -deformation

We now discuss the q -deformed version of Wilson polynomials and functions with $0 < q < 1$. First, the q -Pochhammer symbol is defined by

$$(A; q)_k \equiv \prod_{n=0}^{k-1} (1 - Aq^n),$$

$$(A_1, \dots, A_p; q)_k \equiv \prod_{i=1}^p (A_i; q)_k. \quad (\text{A50})$$

Many of the q -deformed functions that we discuss below can be reduced back to the undeformed case in the limit $q \rightarrow 1$ using

$$\frac{1-q^a}{1-q} \rightarrow a, \quad \frac{(q^a; q)_n}{(1-q)^n} \rightarrow (a)_n = \prod_{j=0}^{n-1} (a+j). \quad (\text{A51})$$

The q -hypergeometric series (often called “basic hypergeometric”) is defined by⁷⁵

$${}_{r+1}\phi_r \left(\begin{matrix} A_1, \dots, A_{r+1} \\ B_1, \dots, B_r \end{matrix}; q, x \right) = \sum_{n=0}^{\infty} \frac{(A_1, \dots, A_{r+1}; q)_n}{(B_1, \dots, B_r; q)_n} \frac{x^n}{(q; q)_n}. \quad (\text{A52})$$

Using (A51), in the limit $q \rightarrow 1$ the basic hypergeometric series reduces to the usual hypergeometric function

$${}_{r+1}\phi_r \left(\begin{matrix} q^{a_1}, \dots, q^{a_{r+1}} \\ q^{b_1}, \dots, q^{b_r} \end{matrix}; q, x \right) \rightarrow {}_{r+1}F_r \left(\begin{matrix} a_1, \dots, a_{r+1} \\ b_1, \dots, b_r \end{matrix}; x \right). \quad (\text{A53})$$

Another useful function is the q -Gamma function

$$\Gamma_q(x) = (1-q)^{1-x} \frac{(q; q)_{\infty}}{(q^x; q)_{\infty}}. \quad (\text{A54})$$

In the limit $q \rightarrow 1$ this becomes the gamma function $\Gamma(x)$.

8. Askey-Wilson polynomial

The Askey-Wilson polynomial is defined by

$$\begin{aligned} p_n(\cos \theta; A, B, C, D|q) \\ = A^{-n} (AB, AC, AD; q)_n \\ \times {}_4\phi_3 \left(\begin{matrix} q^{-n}, q^{n-1}ABCD, Ae^{i\theta}, Ae^{-i\theta} \\ AB, AC, AD \end{matrix}; q, q \right). \end{aligned} \quad (\text{A55})$$

They are degree n polynomials in $\cos \theta$. They are also symmetric in A, B, C, D . Setting $A = q^a, B = q^b, C = q^c, D = q^d$ and taking $q \rightarrow 1$ we recover the Wilson polynomial (A1) up to an overall factor

$$\begin{aligned} p_n \left(\frac{q^{ix} + q^{-ix}}{2}; q^a, q^b, q^c, q^d|q \right) \\ = (1-q)^{3n} W_n(x; a, b, c, d) + \dots, \quad q \rightarrow 1. \end{aligned} \quad (\text{A56})$$

Here, we also introduced a variable x by

$$e^{i\theta} = q^{-ix}, \quad \theta = x \log \frac{1}{q}. \quad (\text{A57})$$

We will sometimes denote the parameters by $A_1 = A, A_2 = B, A_3 = C, A_4 = D$. Askey-Wilson polynomials satisfy an orthogonality relation. Assuming $|A_j| < 1$ (otherwise there might be an extra contribution from a discrete set of points) we have

$$\int_0^\pi \frac{d\theta}{2\pi} M_q(\theta) p_n(\cos \theta) p_m(\cos \theta) = h_n \delta_{nm}, \quad (\text{A58})$$

$$h_n = \frac{1}{1-ABCDq^{2n-1}} \frac{(ABCDq^{n-1}; q)_{\infty}}{(q^{n+1}; q)_{\infty}} \prod_{i < j} \frac{1}{(A_i A_j q^n; q)_{\infty}}, \quad (\text{A59})$$

where the integration measure is

$$M_q(\theta) = \frac{(e^{\pm 2i\theta}; q)_{\infty}}{\prod_{j=1}^4 (A_j e^{\pm i\theta}; q)_{\infty}} \quad (\text{A60})$$

$$= \frac{(1-q)^{2 \sum_{j=1}^4 a_j - 6}}{(q; q)_{\infty}^6} \frac{\prod_{j=1}^4 \Gamma_q(a_j \pm ix)}{\Gamma_q(\pm 2ix)}. \quad (\text{A61})$$

As before, $\theta = x \log \frac{1}{q}$ and $A_j = q^{a_j}$.

We also define orthonormal rescaled Askey-Wilson polynomials

$$\begin{aligned} P_n(x; a_1, a_2, a_3, a_4|q) \\ = |\log q|^{1/2} (1-q)(q; q)_{\infty} \left(h_n \prod_{j=1}^4 (q^{a_j \pm ix}; q)_{\infty} \right)^{-1/2} \\ \times p_n \left(\frac{q^{ix} + q^{-ix}}{2}; q^{a_1}, q^{a_2}, q^{a_3}, q^{a_4}|q \right). \end{aligned} \quad (\text{A62})$$

They are orthonormal with respect to the q -deformed Schwarzian measure

$$\int_0^{\pi/|\log q|} dx \rho_q(x) P_n(x) P_m(x) = \delta_{nm}, \quad (\text{A63})$$

where

$$\rho_q(x) = \frac{1}{2\pi \Gamma_q(\pm 2ix)}. \quad (\text{A64})$$

As in the undeformed case, the parameters relevant in JT gravity in the main text are chosen to be

$$a = d^* = \Delta_1 + ik_1, \quad b = c^* = \Delta_2 + ik_2 \quad (\text{A65})$$

and we define

$$P_n^{\Delta_1, \Delta_2}(x; k_1, k_2|q) := P_n(x; \Delta_1 \pm ik_1, \Delta_2 \pm ik_2|q). \quad (\text{A66})$$

⁷⁵For the more general case ${}_p\phi_p$ see e.g. [75].

9. Askey-Wilson function

The Askey-Wilson function is defined by

$$\mathcal{W}_\Lambda(X; A, B, C, D|q) \equiv \phi_\Lambda(X; A, B, C, q/D|q) \quad (\text{A67})$$

$$\frac{(AB, AC, AD, GA^{-1}\Lambda^{-1}, G\Lambda X^{\pm 1}; q)_\infty}{(GA\Lambda; q)_\infty} \quad (\text{A68})$$

$${}_8W_7(q^{-1}GA\Lambda; AX, AX^{-1}, \tilde{A}\Lambda, \tilde{B}\Lambda, \tilde{C}\Lambda; q, \tilde{D}\Lambda^{-1}), \quad (\text{A69})$$

where the very well-poised ${}_8\phi_7$ series is defined by

$${}_8W_7(A; A_1, \dots, A_5; q, z) = {}_8\phi_7\left(\begin{matrix} A, qA^{1/2}, -qA^{1/2}, A_1, \dots, A_5 \\ A^{1/2}, -A^{1/2}, qA/A_1, \dots, qA/A_5 \end{matrix}; q, z\right). \quad (\text{A70})$$

The “dual” parameters are defined as

$$G = (ABCD)^{1/2} = (\tilde{A}\tilde{B}\tilde{C}\tilde{D})^{1/2}, \quad (\text{A71})$$

$$\tilde{A} = G/D, \quad \tilde{D} = G/A, \quad (\text{A72})$$

$$\tilde{B} = G/C, \quad \tilde{C} = G/B. \quad (\text{A73})$$

Another representation that generalizes (A16) is [74]⁷⁶

$$\begin{aligned} \mathcal{W}_\Lambda(X; A, B, C, D|q) & \quad (\text{A74}) \\ &= \frac{(AB, AC, DX^{\pm 1}, \tilde{D}\Lambda^{\pm 1}; q)_\infty}{(D/A; q)_\infty} \\ & \times {}_4\phi_3\left(\begin{matrix} AX, AX^{-1}, \tilde{A}\Lambda, \tilde{A}\Lambda^{-1} \\ AB, AC, qA/D \end{matrix}; q, q\right) \\ & + (A \leftrightarrow D), \end{aligned} \quad (\text{A75})$$

One can also write down a representation that generalizes (A17), see [76]. The Askey-Wilson function \mathcal{W}_Λ is symmetric in A, B, C, D and satisfies “duality”

$$\mathcal{W}_\Lambda(X; A, B, C, D|q) = \mathcal{W}_X(\Lambda; \tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}|q). \quad (\text{A76})$$

It is also symmetric in X, Λ (as clear from (A74))

$$\mathcal{W}_{\Lambda^{\pm 1}}(X^{\pm 1}; A, B, C, D|q) = \mathcal{W}_\Lambda(X; A, B, C, D|q) \quad (\text{A77})$$

for any choice of signs. To recover the Wilson function (A16), we set $A_j = q^{a_j}$, $X = q^{-ix}$, $\Lambda = q^{-i\lambda}$ and take $q \rightarrow 1$

⁷⁶See formula (8.15) in [74]. Our definition differs from that paper by $D \rightarrow q/D$, such that in our convention the Askey-Wilson function is symmetric in A, B, C, D .

$$\begin{aligned} \mathcal{W}_{q^{-i\lambda}}(q^{-ix}; q^a, q^b, q^c, q^d|q) \\ = (1-q)^{5-2\sum_j a_j} (q; q)_\infty^5 \mathcal{W}_\lambda(x; a, b, c, d) + \dots \end{aligned} \quad (\text{A78})$$

10. q -deformed $6j$ -symbol

We define the q -deformed $6j$ -symbol by

$$\begin{aligned} \left\{ \begin{matrix} \Delta_1 & k_1 & x \\ \Delta_2 & k_2 & \lambda \end{matrix} \right\}_q &= |\log q| (1-q)^{-6+2\sum_j a_j} (q; q)_\infty^{-5} \\ & \times \prod_{j=1}^4 (\Gamma_q(a_j \pm ix) \Gamma_q(\tilde{a}_j \pm i\lambda))^{1/2} \\ & \times \mathcal{W}_{q^{-i\lambda}}(q^{-ix}; q^a, q^b, q^c, q^d|q) \quad (\text{A79}) \\ &= |\log q| (1-q)^2 (q; q)_\infty^3 \\ & \times \prod_{j=1}^4 [(q^{a_j \pm ix}; q)_\infty (q^{\tilde{a}_j \pm i\lambda}; q)_\infty]^{-1/2} \\ & \times \mathcal{W}_{q^{-i\lambda}}(q^{-ix}; q^a, q^b, q^c, q^d|q), \end{aligned} \quad (\text{A80})$$

where the parameters a_j are chosen as in (A65). In the limit $q \rightarrow 1$ we recover the undeformed $6j$ -symbol

$$\left\{ \begin{matrix} \Delta_1 & k_1 & x \\ \Delta_2 & k_2 & \lambda \end{matrix} \right\}_q \rightarrow \left\{ \begin{matrix} \Delta_1 & k_1 & x \\ \Delta_2 & k_2 & \lambda \end{matrix} \right\}. \quad (\text{A81})$$

The definition of the $6j$ -symbol looks unpleasant, but it satisfies many beautiful identities. For example, a q -deformed analog of (A28) is [77]

$$\begin{aligned} \left\{ \begin{matrix} \Delta_1 & k_1 & x \\ \Delta_2 & k_2 & \lambda \end{matrix} \right\}_q &= \sum_{n=0}^{\infty} (-1)^n q^{n(\Delta_1 + \Delta_2) + \frac{n(n-1)}{2}} \\ & \times P_n^{\Delta_1, \Delta_2}(x; k_1, k_2|q) \\ & \times P_n^{\Delta_2, \Delta_1}(\lambda; k_1, k_2|q). \end{aligned} \quad (\text{A82})$$

a. Orthogonality relation

The Askey-Wilson function defines the Askey-Wilson transform [76] and satisfies an orthogonality relation that we express in terms of the $6j$ -symbols (though we don’t use it in the main text)⁷⁷

$$\begin{aligned} \int_0^{\pi/|\log q|} dx \rho_q(x) \frac{\Gamma_q(\frac{1}{2} \pm B \pm ix)}{\sqrt{C_q}} \left\{ \begin{matrix} \Delta_1 & k_1 & x \\ \Delta_2 & k_2 & \lambda \end{matrix} \right\}_q \\ \times \left\{ \begin{matrix} \Delta_1 & k_1 & x \\ \Delta_2 & k_2 & \lambda' \end{matrix} \right\}_q + \text{discrete} \end{aligned} \quad (\text{A83})$$

⁷⁷These formulas are adopted from Sec. 8.3 in [74]. We set the parameter t in that paper to be $t = q^{B+1/2}$.

$$= \frac{\delta(\lambda - \lambda')}{\rho_q(\lambda)} \frac{\sqrt{\tilde{C}_q}}{\Gamma_q(\frac{1}{2} \pm \tilde{B} \pm i\lambda)}, \quad (\text{A84})$$

where

$$C_q = \Gamma_q\left(\frac{1}{2} \pm (B + \Delta_1) \pm ik_1\right) \Gamma_q\left(\frac{1}{2} \pm (B + \Delta_2) \pm ik_2\right). \quad (\text{A85})$$

The dual parameters \tilde{B}, \tilde{C}_q are defined by

$$\tilde{B} = -B - \Delta_1 - \Delta_2, \quad \tilde{C}_q = C_q|_{B \rightarrow \tilde{B}}. \quad (\text{A86})$$

The discrete part is

$$\text{discrete} = \frac{\Gamma_q(\Delta_1 \pm (B + \frac{1}{2}) \pm ik_1) \Gamma_q(\Delta_2 \pm (B + \frac{1}{2}) \pm ik_2)}{\sqrt{C}} \quad (\text{A87})$$

$$\sum_{k > B+1/2} \left\{ \begin{matrix} \Delta_1 & k_1 & x_k \\ \Delta_2 & k_2 & \lambda \end{matrix} \right\}_q \left\{ \begin{matrix} \Delta_1 & k_1 & x_k \\ \Delta_2 & k_2 & \lambda' \end{matrix} \right\}_q \Gamma_q(\Delta_1 \pm ik_1 \pm ix_k)^{-1} \Gamma_q(\Delta_2 \pm ik_2 \pm ix_k)^{-1} \quad (\text{A88})$$

$$\frac{(q^{-1/2-B+\Delta_1 \pm ik_1}, q^{-1/2-B+\Delta_2 \pm ik_2}; q)_k}{(q^{1/2-B-\Delta_1 \pm ik_1}, q^{1/2-B-\Delta_2 \pm ik_2}; q)_k} \frac{1 - q^{2k-2B-1}}{|\log q|} q^{-4k(\Delta_1 + \Delta_2 + B) + k^2 - k}, \quad (\text{A89})$$

where $x_k = i(B + 1/2 - k)$.

b. Physical interpretation of B

The parameter B can be thought of as the magnetic field. Indeed, the full density of states in the limit $q \rightarrow 1$ is

$$\lim_{q \rightarrow 1} \rho_q(x) \Gamma_q\left(\frac{1}{2} \pm B \pm ix\right) = \rho(x) \Gamma\left(\frac{1}{2} \pm B \pm ix\right) \quad (\text{A90})$$

$$= \frac{2x \sinh(2\pi x)}{\cos(2\pi B) + \cosh(2\pi x)}. \quad (\text{A91})$$

This is the density of states of a nonrelativistic particle moving on a hyperbolic plane in the magnetic field B [19,20].

To recover the formulas in the Schwarzian limit we take $B \rightarrow i\infty$. In this limit

$$\frac{\Gamma(\frac{1}{2} \pm B \pm ix)}{\sqrt{C}} \approx \frac{\sqrt{\tilde{C}}}{\Gamma(\frac{1}{2} \pm \tilde{B} \pm ix)} \approx e^{-\pi i(\Delta_1 + \Delta_2)}, \quad (\text{A92})$$

where $C = \lim_{q \rightarrow 1} C_q$. And the orthogonality relation (A83) becomes (2.16).

c. Contour integral representation

A generalization of the integral representation (A49) is

$$\left\{ \begin{matrix} \Delta_1 & k_1 & x \\ \Delta_2 & k_2 & \lambda \end{matrix} \right\}_q = \frac{|\log q|}{1-q} \left(\frac{\gamma_{2x,q}^{\Delta_2} \gamma_{2\lambda,q}^{\Delta_1}}{\gamma_{1x,q}^{\Delta_1} \gamma_{1\lambda,q}^{\Delta_2}} \right)^{1/2} \frac{\Gamma_q(1+2ik_1) \Gamma_q(-2ik_1)}{\Gamma(1+2ik_1) \Gamma(-2ik_1)} \quad (\text{A93})$$

$$\int_{-i\infty}^{i\infty} \frac{du}{2\pi i} q^{u-ik_1} \frac{\Gamma_q(\Delta_1 \pm ix + u) \Gamma_q(\Delta_2 \pm i\lambda + u)}{\Gamma_q(\Delta_1 + \Delta_2 \pm ik_2 + u)} \times \frac{\Gamma(\pm ik_1 + u + 1)}{\Gamma_q(\pm ik_1 + u + 1)} \Gamma(\pm ik_1 - u), \quad (\text{A94})$$

where

$$\gamma_{12,q}^{\Delta} = \Gamma_q(\Delta \pm ik_1 \pm ik_2). \quad (\text{A95})$$

This representation can be derived from Eq. (B.23) in [13].

APPENDIX B: FOUR-POINT FUNCTION ON THE DISK

In this section we check that the relative coefficients between the three terms in (2.13) are indeed as written. The first two terms are computed by the same path integral up to relabeling of β_j and must clearly enter with the same coefficient. Let us check the relative coefficient between the second and third terms. We note that in the limit $\beta_2 \rightarrow 0$ they must be equal

$$(B1)$$

Setting $\beta_2 = 0$, we can integrate over s_2 in the corresponding amplitudes and (B1) holds thanks to the identities

$$\begin{aligned} & \int_0^\infty ds_2 \rho(s_2) (\Gamma_{12}^\Delta \Gamma_{23}^\Delta)^{1/2} \frac{\delta(s_2 - s_4)}{\rho(s_2)} \\ &= \int_0^\infty ds_2 \rho(s_2) (\Gamma_{12}^\Delta \Gamma_{23}^\Delta)^{1/2} \left\{ \begin{matrix} \Delta & s_1 & s_2 \\ \Delta & s_3 & s_4 \end{matrix} \right\} \\ &= (\Gamma_{14}^\Delta \Gamma_{43}^\Delta)^{1/2}. \end{aligned} \quad (B2)$$

The second identity is the $n = 0$ case of (A27).

APPENDIX C: SEMICLASSICAL LIMIT OF FOUR-POINT CORRELATOR

1. Wilson polynomials and conformal blocks

In this section we show that in the semiclassical and zero temperature limit $\beta \gg \tau_i$

$$\begin{aligned} & \frac{1}{Z(\beta)} \int_0^\infty \prod_{j=1}^4 (ds_j \rho(s_j) e^{-\beta_j s_j^2}) (\Gamma_{12}^\Delta \Gamma_{23}^\Delta \Gamma_{34}^\Delta \Gamma_{41}^\Delta)^{1/2} \\ & \times P_n^{\Delta\Delta}(s_2; s_1, s_3) P_n^{\Delta\Delta}(s_4; s_1, s_3) \end{aligned} \quad (C1)$$

$$\approx \frac{(2\Delta)_n^2}{n!(4\Delta + n - 1)_n} (\beta_2 \beta_4)^{-2\Delta} g_{2\Delta+n}(z), \quad (C2)$$

where $z = \frac{\tau_{12}\tau_{34}}{\tau_{13}\tau_{24}}$. The rhs is the conformal block of the double-trace operator $2\Delta + n$. We also assume

$$\begin{aligned} \beta_1 &= \beta - \tau_{14}, & \beta_2 &= \tau_{12}, & \beta_3 &= \tau_{23}, & \beta_4 &= \tau_{34}; \\ \beta &> \tau_1 > \tau_2 > \tau_3 > \tau_4 > 0 \end{aligned} \quad (C3)$$

Using the definition of Wilson polynomials in Sec. A we have

$$\frac{1}{Z(\beta)} \frac{(2\Delta)_n^2}{n!(4\Delta + n - 1)_n} \frac{\Gamma(4\Delta + 2n)}{\Gamma(2\Delta)^2 \Gamma(2\Delta + n)^2} \int_0^\infty \prod_{j=1}^4 (ds_j \rho(s_j) e^{-\beta_j s_j^2}) \gamma_{12}^\Delta \gamma_{23}^\Delta \gamma_{34}^\Delta \gamma_{41}^\Delta \quad (C4)$$

$$\frac{(2\Delta + is_1 \pm is_3)_n^2}{\Gamma(2\Delta + n \pm is_1 \pm is_3)} {}_4F_3 \left(\begin{matrix} -n, 4\Delta + n - 1, \Delta + is_1 \pm is_2 \\ 2\Delta, 2\Delta + is_1 \pm is_3 \end{matrix}; 1 \right) {}_4F_3 \left(\begin{matrix} -n, 4\Delta + n - 1, \Delta + is_1 \pm is_4 \\ 2\Delta, 2\Delta + is_1 \pm is_3 \end{matrix}; 1 \right), \quad (C5)$$

where $\gamma_{nm}^\Delta = \Gamma(\Delta \pm is_n \pm is_m)$. To take the semiclassical limit we change variables

$$s_j^2 = s_1^2 + \omega_j, \quad j = 2, 3, 4 \quad (C6)$$

and take the limit $s_1^2 \gg \omega_j$. More carefully, we should first restore the factors of gravitational coupling $\gamma = \frac{\bar{\phi}}{8\pi G_N}$ that we previously set to $2\gamma = 1$. This is easily done by dimensional analysis. The coupling γ has dimensions of

length, k^2 has dimensions of energy, and β has dimensions of length. We will not do this explicitly, but instead think of it as the high-energy limit $s_1^2 \gg \omega_j$.

Using that

$$\Gamma(a \pm is) \approx 2\pi |s|^{2a-1} e^{-\pi|s|}, \quad |s| \rightarrow \infty, \quad (C7)$$

we find

$$\frac{1}{Z(\beta)} \frac{(2\Delta)_n^2}{n!(4\Delta + n - 1)_n} \frac{\Gamma(4\Delta + 2n)}{\Gamma(2\Delta)^2 \Gamma(2\Delta + n)^2} \int_0^\infty ds_1 \rho(s_1) e^{-\beta s_1^2} (2s_1)^{4\Delta-3} \quad (C8)$$

$$\int_{-\infty}^\infty \frac{d\omega_2 d\omega_3 d\omega_4}{(2\pi)^3} e^{-\beta_2 \omega_2 - \beta_3 \omega_3 - \beta_4 \omega_4} e^{i\frac{\omega_3}{2s_1}} \Gamma\left(\Delta \pm i\frac{\omega_2}{2s_1}\right) \Gamma\left(\Delta \pm i\frac{\omega_4}{2s_1}\right) \Gamma\left(\Delta \pm i\frac{\omega_{32}}{2s_1}\right) \Gamma\left(\Delta \pm i\frac{\omega_{34}}{2s_1}\right) \quad (C9)$$

$$\frac{(-1)^n (2\Delta - i\frac{\omega_3}{2s_1})^2}{\Gamma(2\Delta + n \pm i\frac{\omega_3}{2s_1})^3} {}_3F_2 \left(\begin{matrix} -n, 4\Delta + n - 1, \Delta - i\frac{\omega_2}{2s_1} \\ 2\Delta, 2\Delta - i\frac{\omega_3}{2s_1} \end{matrix}; 1 \right) {}_3F_2 \left(\begin{matrix} -n, 4\Delta + n - 1, \Delta - i\frac{\omega_4}{2s_1} \\ 2\Delta, 2\Delta - i\frac{\omega_3}{2s_1} \end{matrix}; 1 \right) \quad (C10)$$

where $\omega_{ij} = \omega_i - \omega_j$. The integral over s_1 is computed by a saddle approximation. The value of the saddle is determined from $\rho(s_1)e^{-\beta s_1^2} \sim e^{2\pi s_1 - \beta s_1^2}$ and is given by $s_1 = \frac{\pi}{\beta}$. Then $\int_0^\infty ds_1 \rho(s_1) e^{-\beta s_1^2} = Z(\beta)$ is the partition function, while in the rest of the integral we simply set $s_1 = \frac{\pi}{\beta}$

$$\frac{(2\Delta)_n^2}{n!(4\Delta + n - 1)_n} \frac{\Gamma(4\Delta + 2n)}{\Gamma(2\Delta)^2 \Gamma(2\Delta + n)^2} \left(\frac{2\pi}{\beta}\right)^{4\Delta-3} \quad (C11)$$

$$\int_{-\infty}^{\infty} \frac{d\omega_2 d\omega_3 d\omega_4}{(2\pi)^3} e^{-\beta_2 \omega_2 - \beta_3 \omega_3 - \beta_4 \omega_4} e^{\frac{\beta}{2} \omega_3} \Gamma\left(\Delta \pm i \frac{\beta}{2\pi} \omega_2\right) \Gamma\left(\Delta \pm i \frac{\beta}{2\pi} \omega_4\right) \Gamma\left(\Delta \pm i \frac{\beta}{2\pi} \omega_{32}\right) \Gamma\left(\Delta \pm i \frac{\beta}{2\pi} \omega_{34}\right) \quad (C12)$$

$$\frac{(-1)^n (2\Delta - i \frac{\beta}{2\pi} \omega_3)_n^2}{\Gamma(2\Delta + n \pm i \frac{\beta}{2\pi} \omega_3)} {}_3F_2\left(\begin{matrix} -n, 4\Delta + n - 1, \Delta - i \frac{\beta}{2\pi} \omega_2 \\ 2\Delta, 2\Delta - i \frac{\beta}{2\pi} \omega_3 \end{matrix}; 1\right) {}_3F_2\left(\begin{matrix} -n, 4\Delta + n - 1, \Delta - i \frac{\beta}{2\pi} \omega_4 \\ 2\Delta, 2\Delta - i \frac{\beta}{2\pi} \omega_3 \end{matrix}; 1\right) \quad (C13)$$

Now we take the zero temperature limit $\beta \rightarrow \infty$. Again using the asymptotics of gamma functions at large argument, we find

$$\frac{(2\Delta)_n^2}{n!(4\Delta + n - 1)_n} \frac{\Gamma(4\Delta + 2n)}{\Gamma(2\Delta)^2 \Gamma(2\Delta + n)^2} \quad (C14)$$

$$\int_{-\infty}^{\infty} d\omega_2 d\omega_3 d\omega_4 e^{-\beta_2 \omega_2 - \beta_3 \omega_3 - \beta_4 \omega_4} e^{\frac{\beta}{2}(\omega_3 + |\omega_3| - |\omega_2| - |\omega_4| - |\omega_{32}| - |\omega_{34}|)} \quad (C15)$$

$$|\omega_3|^{-4\Delta+1} |\omega_2 \omega_4 \omega_{32} \omega_{34}|^{2\Delta-1} {}_2F_1\left(\begin{matrix} -n, 4\Delta + n - 1 \\ 2\Delta \end{matrix}; \frac{\omega_2}{\omega_3}\right) {}_2F_1\left(\begin{matrix} -n, 4\Delta + n - 1 \\ 2\Delta \end{matrix}; \frac{\omega_4}{\omega_3}\right) \quad (C16)$$

One can show that at large β

$$e^{\frac{\beta}{2}(\omega_3 + |\omega_3| - |\omega_2| - |\omega_4| - |\omega_{32}| - |\omega_{34}|)} \approx \begin{cases} 1, & \text{if } 0 < \omega_2 < \omega_3, 0 < \omega_4 < \omega_3 \\ 0, & \text{otherwise} \end{cases}. \quad (C17)$$

Therefore we find

$$\frac{(2\Delta)_n^2}{n!(4\Delta + n - 1)_n} \frac{\Gamma(4\Delta + 2n)}{\Gamma(2\Delta)^2 \Gamma(2\Delta + n)^2} \quad (C18)$$

$$\int_0^\infty d\omega_3 e^{-\beta_3 \omega_3} \omega_3^{-4\Delta+1} \quad (C19)$$

$$\int_0^{\omega_3} d\omega_2 e^{-\beta_2 \omega_2} (\omega_2 \omega_{32})^{2\Delta-1} F(-n, 4\Delta + n - 1, 2\Delta, \omega_2/\omega_3) \quad (C20)$$

$$\int_0^{\omega_3} d\omega_4 e^{-\beta_4 \omega_4} (\omega_4 \omega_{34})^{2\Delta-1} F(-n, 4\Delta + n - 1, 2\Delta, \omega_4/\omega_3). \quad (C21)$$

After rescaling $\omega_2 = x\omega_3, \omega_4 = y\omega_3$ and using that

$$\int_0^1 dx e^{-ax} [x(1-x)]^{2\Delta-1} F(-n, 4\Delta + n - 1, 2\Delta, x) = \sqrt{\pi} e^{-a/2} a^{-2\Delta+\frac{1}{2}} \Gamma(2\Delta) I_{2\Delta+n-\frac{1}{2}}(a/2) \quad (C22)$$

we have

$$\frac{(2\Delta)_n^2}{n!(4\Delta + n - 1)_n} \frac{\pi \Gamma(4\Delta + 2n)}{\Gamma(2\Delta + n)^2} \quad (C23)$$

$$(\beta_2 \beta_4)^{-2\Delta+\frac{1}{2}} \int_0^\infty d\omega_3 e^{-(\beta_3 + \frac{\beta_2 + \beta_4}{2})\omega_3} I_{2\Delta+n-\frac{1}{2}}(\beta_2 \omega_3/2) I_{2\Delta+n-\frac{1}{2}}(\beta_4 \omega_3/2) \quad (C24)$$

$$= \frac{(2\Delta)_n^2}{n!(4\Delta + n - 1)_n} (\beta_2 \beta_4)^{-2\Delta} g_{2\Delta+n}(z), \quad (C25)$$

where $z = \frac{\beta_2\beta_4}{(\beta_2+\beta_3)(\beta_3+\beta_4)} = \frac{\tau_{12}\tau_{34}}{\tau_{13}\tau_{24}}$ and $g_h(z) = z^h F(h, h, 2h, z)$. We checked this integral numerically. This is indeed the expected answer.

2. Semiclassical limit of the $6j$ -symbol

It is also interesting to take the semiclassical limit of the $6j$ -symbol and check that after an appropriate integration over energies it gives the crossed GFF Wick contraction. We consider

$$\begin{aligned} & \begin{Bmatrix} \Delta & k_1 & k_2 \\ \Delta & k_3 & k_4 \end{Bmatrix} \\ &= (\gamma_{12}\gamma_{23}\gamma_{34}\gamma_{41})^{1/2} \mathcal{W}_{k_4}(k_2; \Delta \pm ik_1, \Delta \pm ik_3). \end{aligned} \quad (C26)$$

We set

$$k_j^2 = k_1^2 + \omega_j, \quad k_j \approx k_1 + \frac{\omega_j}{2k_1} \equiv k_1 + \nu_j, \quad j=2,3,4 \quad (C27)$$

and take $k_1^2 \gg \omega_j$. It is convenient to use the expression for the Wilson function (A16) giving

$$\begin{aligned} & \mathcal{W}_{k_4}(k_2; \Delta \pm ik_1, \Delta \pm ik_3) \\ &= \frac{\Gamma(-2ik_1)}{\Gamma(2\Delta + ik_1 \pm ik_3)\Gamma(\Delta - ik_1 \pm ik_2)\Gamma(\Delta - ik_1 \pm ik_4)} \end{aligned} \quad (C28)$$

$${}_4F_3\left(\begin{matrix} \Delta + ik_1 \pm ik_2, \Delta + ik_1 \pm ik_4 \\ 2\Delta + ik_1 \pm ik_3, 2ik_1 + 1 \end{matrix}; 1\right) + \text{c.c.} \quad (C29)$$

At large k_1 the hypergeometric function simplifies significantly ($k_{ij} = k_i - k_j$)

$$\begin{aligned} & {}_4F_3\left(\begin{matrix} \Delta + ik_1 \pm ik_2, \Delta + ik_1 \pm ik_4 \\ 2\Delta + ik_1 \pm ik_3, 2ik_1 + 1 \end{matrix}; 1\right) \\ & \approx {}_2F_1\left(\begin{matrix} \Delta + ik_{12}, \Delta + ik_{14} \\ 2\Delta + ik_{13} \end{matrix}; 1\right) \end{aligned} \quad (C30)$$

$$= \frac{\Gamma(2\Delta + ik_{13})\Gamma(ik_{2+4-1-3})}{\Gamma(\Delta + ik_{23})\Gamma(\Delta + ik_{43})}, \quad (C31)$$

where $k_{i+j} = k_i + k_j$ etc. Also using that

$$\begin{aligned} \Gamma(\Delta + ix) &\approx \sqrt{2\pi}|x|^{\Delta-\frac{1}{2}} e^{-\frac{\pi}{2}|x|} e^{i\text{sgn}(x)\frac{\pi}{2}(\Delta-\frac{1}{2})+ix\log\frac{|x|}{\sigma}}, \\ x &\in \mathbb{R}, \quad |x| \gg 1, \end{aligned} \quad (C32)$$

we find

$$\begin{aligned} & \mathcal{W}_{k_4}(k_2; \Delta \pm ik_1, \Delta \pm ik_3) \\ & \approx \frac{e^{2\pi k_1}}{2\pi} (2k_1)^{1-4\Delta} e^{\frac{\pi}{2}(\nu_2+\nu_3+\nu_4)} \end{aligned} \quad (C33)$$

$$\left(\frac{(2k_1)^{i\nu_{2+4-3}} \Gamma(i\nu_{2+4-3})}{\Gamma(\Delta + i\nu_2)\Gamma(\Delta - i\nu_{34})\Gamma(\Delta + i\nu_4)\Gamma(\Delta - i\nu_{32})} + \text{c.c.} \right). \quad (C34)$$

At large k_1 it is highly oscillating unless $\nu_{2+4-3} \equiv \nu_2 + \nu_4 - \nu_3 = 0$. Near this value we substitute $\nu_3 = \nu_2 + \nu_4$ in the gamma functions in the denominator, expand the gamma function in the numerator and find

$$\begin{aligned} & \mathcal{W}_{k_4}(k_2; \Delta \pm ik_1, \Delta \pm ik_3) \\ & \approx \frac{e^{2\pi k_1} (2k_1)^{1-4\Delta} e^{\frac{\pi}{2}(\nu_2+\nu_3+\nu_4)} 2\sin(\nu_{2+4-3}\log(2k_1))}{2\pi\Gamma(\Delta \pm i\nu_2)\Gamma(\Delta \pm i\nu_4)} \nu_{2+4-3} \end{aligned} \quad (C35)$$

$$\approx \frac{e^{2\pi k_1} (2k_1)^{1-4\Delta} e^{\frac{\pi}{2}(\nu_2+\nu_3+\nu_4)}}{\Gamma(\Delta \pm i\nu_2)\Gamma(\Delta \pm i\nu_4)} \delta(\nu_{2+4-3}). \quad (C36)$$

In the last line we used that $\lim_{t \rightarrow \infty} \frac{2\sin(tx)}{x} = 2\pi\delta(x)$. Also taking the limit of gamma functions in the prefactor in (C26) we find a rather simple expression

$$\begin{Bmatrix} \Delta & k_1 & k_2 \\ \Delta & k_3 & k_4 \end{Bmatrix} \approx (2\pi)^2 \frac{e^{-2\pi k}}{2k} \delta(k_1 + k_3 - k_2 - k_4) \quad (C37)$$

$$\approx \frac{\delta(k_1 + k_3 - k_2 - k_4)}{\rho(k)}, \quad (C38)$$

where we restored the original variables and defined the average momentum $k = \frac{1}{4}(k_1 + k_2 + k_3 + k_4)$. To reiterate, we take the limit where all k_j are large, but all differences $k_i - k_j$ are finite. It is straightforward to use the above formula to show that appropriate integration over energies gives the crossed GFF Wick contraction in the four-point correlator, as expected.

APPENDIX D: OPERATOR ALGEBRA OF THE GENERALIZED FREE FIELD

The purpose of this appendix is to show that the condition (3.10), together with associativity of the OPE and conformal symmetry, guarantees that all n -point correlators of \mathcal{O} agree with those of the bosonic generalized free field. The upshot is that the operator algebra of \mathcal{O} in the generalized free field theory is completely characterized by the operator equation (3.10) and conformal invariance. All other operator equations must be implied by (3.10) and conformal invariance.

As mentioned in the main text, (3.10) implies that the OPE takes the form (3.9), which we repeat here (in a slightly modified form):

$$\begin{aligned} \mathcal{O}(\tau_1)\mathcal{O}(\tau_2) &= \frac{1}{\tau_{12}^{\Delta}} + \sum_{n=0}^{\infty} \tau_{12}^{2n} [\mathcal{O}\mathcal{O}]_n \Big|_{\tau=\frac{\tau_1+\tau_2}{2}} \\ &+ \text{descendants.} \end{aligned} \quad (D1)$$

We have chosen to expand the operators about their midpoint. Written in this way, all of the powers of τ_{12} are even. In particular, one can show that the descendants only contribute even powers of τ_{12} . If we compute the Lorentzian commutator $[\mathcal{O}(t_1), \mathcal{O}(t_2)]$, all of the terms on the right-hand side of (D1) drop out except for the identity contribution, which appears on the right-hand side of (3.10). If any other primaries with dimensions not in $2\Delta + 2\mathbb{Z}_{\geq 0}$ appeared in the \mathcal{OO} OPE, these primaries would contribute to the right side of (3.10), so by imposing (3.10) as a constraint on the algebra we are demanding that the OPE takes the form (D1).

Let $F_n(\tau_1, \tau_2, \dots, \tau_{2n})$ be the $2n$ -point function of \mathcal{O} in a theory where the OPE (D1) holds, and let $G_n(\tau_1, \tau_2, \dots, \tau_{2n})$ be the $2n$ -point function of \mathcal{O} in the generalized free field theory. We will inductively show that $F_n = G_n$ for all $n \in \mathbb{N}$.⁷⁸ The base case $n = 1$ is trivial. Next, suppose that the result has been shown for $n \leq m - 1$ with $m \in \mathbb{N}$. Define

$$H_m(\tau_1, \tau_2, \dots, \tau_{2m}) := F_m(\tau_1, \tau_2, \dots, \tau_{2m}) - G_m(\tau_1, \tau_2, \dots, \tau_{2m}). \quad (\text{D2})$$

This function should be defined via analytic continuation from the domain $\tau_1 > \tau_2 > \dots > \tau_{2m}$. One can easily see that H_m is nonsingular for $\tau_1 \rightarrow \tau_2$ because the identity contributions to the $\mathcal{O}(\tau_1)\mathcal{O}(\tau_2)$ OPE cancel. Hence, we can analytically continue to $\tau_1 < \tau_2$. Furthermore, H_m is invariant under $\tau_1 \leftrightarrow \tau_2$ because only even powers of τ_{12} appear in (D1). Hence, the $\tau_1 \rightarrow \tau_3$ limit is nonsingular because the $\tau_2 \rightarrow \tau_3$ limit is nonsingular. One can thus show that H_m is entire in τ_1 . Using the same logic, one can show that H_m is an entire function of all of the τ variables. Furthermore, H_m goes to zero as $\tau_1 \rightarrow \infty$ due to conformal symmetry. Hence, H_m vanishes identically.

APPENDIX E: TWO-POINT FUNCTION ON THE DOUBLE-TRUMPET

In this section we compute the two-point correlator on the double-trumpet. We use the boundary particle formalism of [19]. It is convenient to work in the coordinates $ds^2 = d\rho^2 + \cosh^2 \rho d\tau^2$. We first compute the JT path integral on a piece of the disk in the Fig. 19 bounded by four points: $A_1 = (\rho_1, \tau_1)$, $A_2 = (\rho_2, \tau_2)$, $B_1 = (0, \tau_1)$, $B_2 = (0, \tau_2)$. The wiggly line connecting A_1, A_2 is the boundary of physical length u . The embedding of this boundary segment into the hyperbolic disk is integrated over with the Schwarzian action. The other four lines connecting points A_1, A_2, B_1, B_2 are geodesics. We construct the path integral for $g(u; A_1, A_2)$ out of the Hartle-Hawking wave function (region bounded by the

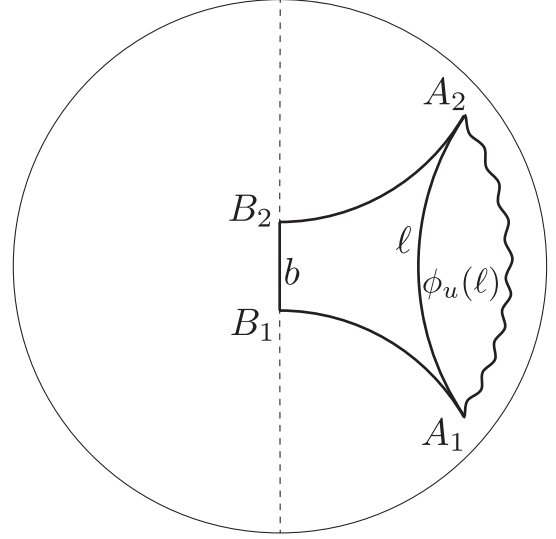


FIG. 19. Geometry relevant for the computation of $g(u; A_1, A_2)$.

boundary and the geodesic ℓ) and the hyperbolic quadrangle $A_1 A_2 B_2 B_1$

$$g(u; A_1, A_2) = e^{\phi_b A} \phi_u(\ell). \quad (\text{E1})$$

The factor $e^{\phi_b A}$, where A is the area of the quadrangle $A_1 A_2 B_2 B_1$, is the JT path integral on the quadrangle. An exercise in hyperbolic geometry shows that in the limit when A_1, A_2 are close to the boundary ($\rho_1, \rho_2 \rightarrow \infty$) this area is

$$A \approx \pi - 2(e^{-\rho_1} + e^{-\rho_2}) \coth \frac{\tau_2 - \tau_1}{2} + \dots \quad (\text{E2})$$

Also multiplying by $\phi_b = \frac{\gamma}{\epsilon} \equiv \gamma q$, rescaling the radial coordinate⁷⁹ $e^{-\rho} = \frac{1}{q} e^{-w}$ and setting $\gamma = 1/2$ we get a contribution

$$e^{\phi_b A} \approx e^{\pi q - (e^{-w_1} + e^{-w_2}) \coth \frac{\tau_2 - \tau_1}{2}}. \quad (\text{E3})$$

The first term πq cancels out if one adds appropriate corner terms in the JT action.⁸⁰ In any case it doesn't depend on any of the coordinates. The second term will be important in the computation below.

The second factor in (E1) is the Hartle-Hawking wave function

$$\phi_u(\ell) = e^{-\ell/2} \int_0^\infty ds \rho(s) e^{-us^2} 2K_{2is}(2e^{-\ell/2}) \quad (\text{E4})$$

⁷⁹This is analogous to the rescaling $z = \frac{\gamma}{q}$ in [19]. It is necessary for the particle in a magnetic field to be equivalent to the Schwarzian theory.

⁸⁰Such terms are necessary for locality, see [78] for details.

⁷⁸We thank Dalimil Mazáč for providing this proof to us.

$$\equiv e^{-\ell/2} \int_0^\infty ds \rho(s) e^{-us^2} \hat{\phi}_s(\ell). \quad (\text{E5})$$

In the energy basis it obeys the orthogonality relations

$$\int_{-\infty}^\infty d\ell \hat{\phi}_s(\ell) \hat{\phi}_{s'}(\ell) = \frac{\delta(s-s')}{\rho(s)}, \quad (\text{E6})$$

$$\int_0^\infty ds \rho(s) \hat{\phi}_s(\ell) \hat{\phi}_s(\ell') = \delta(\ell - \ell'). \quad (\text{E7})$$

In (E1) the (renormalized) geodesic distance ℓ between the points A_1, A_2 is $e^\ell = e^{w_1+w_2} \sinh^2 \frac{\tau_2 - \tau_1}{2}$. Putting everything together we have

$$g(u; A_1, A_2) = \exp \left(-(e^{-w_1} + e^{-w_2}) \coth \frac{\tau_2 - \tau_1}{2} \right) \frac{e^{-\frac{w_1+w_2}{2}}}{\sinh \frac{\tau_2 - \tau_1}{2}} \times \int_0^\infty ds \rho(s) e^{-us^2} 2K_{2is} \left(\frac{2e^{-\frac{w_1+w_2}{2}}}{\sinh \frac{\tau_2 - \tau_1}{2}} \right). \quad (\text{E8})$$

1. Trumpet partition function

To make sure our formulas are correct, we first compute the trumpet partition function. We set $w_1 = w_2 = w$, $\tau_2 = \tau_1 + b \equiv \tau + b$ and integrate over w . The correct measure of integration is⁸¹ $\frac{1}{b} d\tau dw e^w \sinh \frac{b}{2}$, where $1/b$ is from gauge fixing $U(1)$ symmetry. The τ integral is trivial $\int_0^b \frac{d\tau}{b} = 1$ and the rest gives

$$Z_{tr}(\beta, b) = \int_{-\infty}^\infty dw \left(e^w \sinh \frac{b}{2} \right) \cdot e^{-2e^{-w} \coth \frac{b}{2}} \frac{e^{-w}}{\sinh \frac{b}{2}} \times \int_0^\infty ds \rho(s) e^{-\beta s^2} 2K_{2is} \left(\frac{2e^{-w}}{\sinh \frac{b}{2}} \right) \quad (\text{E9})$$

$$= \int_0^\infty ds e^{-\beta s^2} \frac{\cos(bs)}{\pi}, \quad (\text{E10})$$

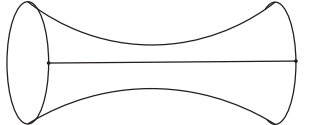
where we used that

⁸¹The measure which naturally arises in the boundary particle is induced by the metric $ds^2 = d\rho^2 + \cosh^2 \rho d\tau^2$, which gives $\sqrt{g} d\rho d\tau = \cosh \rho d\rho d\tau \approx q \frac{1}{2} e^w dw d\tau$. This differs from the correct measure by a factor $2 \sinh \frac{b}{2}$. This factor is explained by the fact that the boundary particle is equivalent to the Schwarzian theory with the path integral measure $\prod_u \frac{d\tau(u)}{\tau'(u)}$, while the correct measure on the double-trumpet is $2 \sinh \frac{b}{2} \prod_u \frac{d\tau(u)}{\tau'(u)}$. The factor $2 \sinh \frac{b}{2}$ can be understood from considering Eq. (2.25) in [79]. We first map it to the double-trumpet coordinate by $\tan \frac{\phi}{2} = e^{-\tau}$. Then two solutions of (2.25) in [79] become $\Psi_\pm = e^{\pm\tau/2} / \sqrt{\tau'}$. Under Euclidean time evolution $u \rightarrow u + 2\pi$, we have $\tau \rightarrow \tau + b$ and $\Psi_\pm \rightarrow e^{\pm b/2} \Psi_\pm$. Then we compute the trace of evolution operator around the Euclidean circle $\text{Tr}(-1)^F U = e^{b/2} - e^{-b/2} = 2 \sinh \frac{b}{2}$, where the minus sign is from $(-1)^F$.

$$\int_0^\infty \frac{dx}{x} e^{-2x \cosh \frac{b}{2}} 2K_{2is}(2x) = \frac{1}{\rho(s)} \frac{\cos(bs)}{\pi}. \quad (\text{E11})$$

2. Two-point function without 1-loop determinant

The two-point function on the double-trumpet without the matter determinant can be computed by putting together two factors of $g(u; A_1, A_2)$ from the previous subsection, for the left and right trumpets. We also insert $e^{-\Delta \ell_{LR}}$, where $e^{\ell_{LR}} = e^{w_L + w_R} \cosh^2 \frac{\tau_L - \tau_R}{2}$ is the (renormalized) geodesic distance between the two operators. See Fig. 20. We have

$$\langle \text{Tr} e^{-\beta_L H} \mathcal{O} \text{Tr} e^{-\beta_R H} \mathcal{O} \rangle \supset \text{Diagram} \quad (\text{E12})$$


$$= \int_0^\infty db \int_0^b d\tau \int_{-\infty}^\infty dw_L dw_R \int_0^\infty ds_L ds_R \rho(s_L) \times \rho(s_R) e^{-\beta_L s_L^2 - \beta_R s_R^2} \quad (\text{E13})$$

$$\times e^{-2e^{-w_L} \coth \frac{b}{2}} 2K_{2is_L} \left(\frac{2e^{-w_L}}{\sinh \frac{b}{2}} \right) \times e^{-2e^{-w_R} \coth \frac{b}{2}} 2K_{2is_R} \left(\frac{2e^{-w_R}}{\sinh \frac{b}{2}} \right) \quad (\text{E14})$$

$$\times \sum_{n=-\infty}^\infty \left(\frac{e^{-(w_L + w_R)}}{\cosh^2 \frac{\tau + nb}{2}} \right)^\Delta, \quad (\text{E15})$$

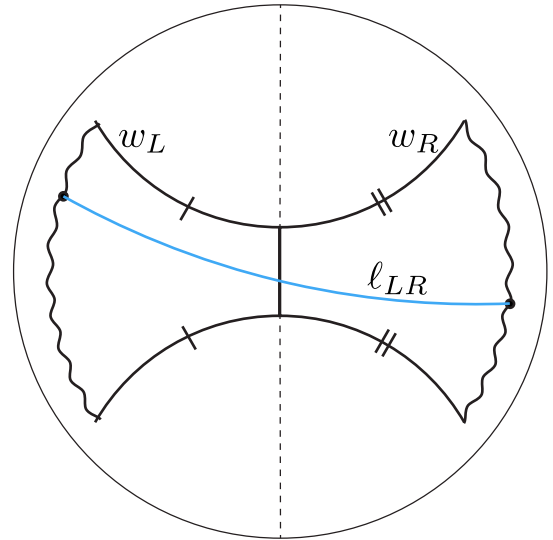


FIG. 20. Geometry for the computation of the two-point function on the double-trumpet. The blue line is the geodesic of (renormalized) length ℓ_{LR} connecting operators on the left and right boundaries.

where the sum over n is the sum over windings of the geodesic connecting two operators. The integral over τ gives

$$\int_0^b d\tau \sum_{n=-\infty}^{\infty} \left(\cosh \frac{\tau + nb}{2} \right)^{-2\Delta} = \int_{-\infty}^{\infty} d\tau \left(\cosh \frac{\tau}{2} \right)^{-2\Delta} \quad (\text{E16})$$

$$= \frac{2^{2\Delta} \Gamma(\Delta)^2}{\Gamma(2\Delta)}. \quad (\text{E17})$$

To compute w_L , w_R integrals we change variables to $x_L = \frac{e^{-w_L}}{\sinh \frac{b}{2}}$ and similarly for w_R . And use that

$$\int_0^{\infty} \frac{dx}{x} x^{\Delta} e^{-2x \cosh \frac{b}{2}} 2K_{2is}(2x) = 2^{-2\Delta} \Gamma(\Delta \pm 2is)^{1/2} \left(\cosh \frac{b}{4} \right)^{-2\Delta+1} \Phi_s(b), \quad (\text{E18})$$

$$\Phi_s(b) = \frac{2^{2\Delta} \Gamma(\Delta)}{\Gamma(2\Delta)} \Gamma(\Delta \pm 2is)^{1/2} \times F\left(\frac{1}{2} + 2is, \frac{1}{2} - 2is, \Delta + \frac{1}{2}, -\sinh^2 \frac{b}{4}\right). \quad (\text{E19})$$

This is a generalization of (E11). We introduced a function $\Phi_s(b)$ which is essentially a Jacobi function that defines Jacobi transform, e.g. see [80]. It satisfies orthogonality relations

$$\int_0^{\infty} db \left(\sinh \frac{b}{4} \right)^{2\Delta} \left(\cosh \frac{b}{4} \right)^{2(1-\Delta)} \Phi_{s_L}(b) \Phi_{s_R}(b) = \frac{\delta(s_L - s_R)}{\rho(s_L)}, \quad (\text{E20})$$

$$\int_0^{\infty} ds \rho(s) \Phi_s(b) \Phi_s(b') = \frac{\delta(b - b')}{\left(\sinh \frac{b}{4} \right)^{2\Delta} \left(\cosh \frac{b}{4} \right)^{2(1-\Delta)}}. \quad (\text{E21})$$

Using above integrals we have

$$\frac{\Gamma(\Delta)^2}{\Gamma(2\Delta)} \int_0^{\infty} ds_L ds_R \rho(s_L) \rho(s_R) \times e^{-\beta_L s_L^2 - \beta_R s_R^2} (\Gamma(\Delta \pm 2is_L) \Gamma(\Delta \pm 2is_R))^{1/2} \quad (\text{E22})$$

$$\times \int_0^{\infty} db \left(\sinh \frac{b}{4} \right)^{2\Delta} \left(\cosh \frac{b}{4} \right)^{2(1-\Delta)} \Phi_{s_L}(b) \Phi_{s_R}(b). \quad (\text{E23})$$

The b integral here turns out to be the orthogonality relation for the Jacobi transform, so we find

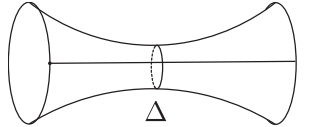
$$\langle \text{Tr} e^{-\beta_L H} \mathcal{O} \text{Tr} e^{-\beta_R H} \mathcal{O} \rangle = \frac{\Gamma(\Delta)^2}{\Gamma(2\Delta)} \int_0^{\infty} ds \rho(s) e^{-(\beta_L + \beta_R) s^2} \Gamma(\Delta \pm 2is). \quad (\text{E24})$$

3. Two-point function with 1-loop determinant

To include the 1-loop determinant in our computation, we use the formula (2.24) for the determinant at fixed b and insert it in the computation in the previous subsection. We start with the $n = 1$ term in (2.24) corresponding to the operator Δ propagating on the closed geodesic. Instead of the integral (E23) we need to compute

$$\int_0^{\infty} db \frac{e^{-\Delta' b}}{1 - e^{-b}} \left(\sinh \frac{b}{4} \right)^{2\Delta} \left(\cosh \frac{b}{4} \right)^{2(1-\Delta)} \Phi_{s_L}(b) \Phi_{s_R}(b) = \left\{ \begin{array}{ccc} \Delta' & s_L & s_R \\ \Delta & s_R & s_L \end{array} \right\}. \quad (\text{E25})$$

We did not find a derivation of this formula, but we checked that it holds numerically. For the $n = 1$ term we need this integral with $\Delta' = \Delta$. Assuming it is correct, we have a contribution to the two-point function

$$\langle \text{Tr} e^{-\beta_L H} \mathcal{O} \text{Tr} e^{-\beta_R H} \mathcal{O} \rangle \supset \text{Diagram} \quad (\text{E26})$$


$$= \frac{\Gamma(\Delta)^2}{\Gamma(2\Delta)} \int_0^{\infty} ds_L ds_R \rho(s_L) \rho(s_R) e^{-(\beta_L s_L^2 + \beta_R s_R^2)} \quad (\text{E27})$$

$$\times (\Gamma(\Delta \pm 2is_L) \Gamma(\Delta \pm 2is_R))^{1/2} \left\{ \begin{array}{ccc} \Delta & s_L & s_R \\ \Delta & s_R & s_L \end{array} \right\}. \quad (\text{E28})$$

For the $n = 2$ term in (2.24), we expand it as $\frac{e^{-2\Delta b}}{(1 - e^{-b})(1 - e^{-2b})} = \sum_{m=0}^{\infty} \frac{e^{-(2\Delta+2m)b}}{1 - e^{-b}}$ and use (E25) with $\Delta' = 2\Delta + 2m$. Higher-order terms $n \geq 3$ can be similarly computed as a sum over contributions of higher-trace operators.

APPENDIX F: MORE 'T HOOFT DIAGRAMS ON THE DOUBLE-TRUMPET

In this appendix, we compute 't Hooft diagrams for the empty double-trumpet and two-point function on the double-trumpet, where a right-left path crosses three lines.

1. Empty double-trumpet

To compute $\mathcal{D}^{(3)}$, we consider diagrams that allow for left-right paths that cross over three \mathcal{O} double-lines, but we exclude those diagrams that were already counted in $\mathcal{D}^{(1)}$

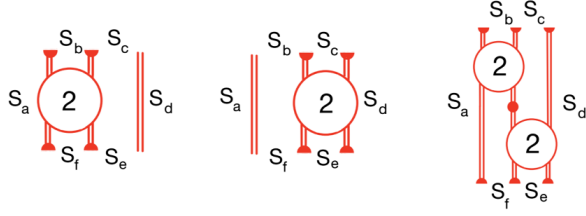


FIG. 21. On the left we graphically represent $\mathcal{X}_1^{(2)}$. In the center we graphically represent $\mathcal{X}_2^{(2)}$. We may think of $\mathcal{X}_1^{(2)}$ and $\mathcal{X}_2^{(2)}$ as operators that act on a function of s_f and s_e to produce a function of s_b and s_c . On the right, we represent the product $\mathcal{X}_1^{(2)}\mathcal{X}_2^{(2)}$.

and $\mathcal{D}^{(2)}$. It is convenient to think of each diagram as an operator that acts on functions of two energies. Before we can enumerate the diagrams that contribute to $\mathcal{D}^{(3)}$, we need to define a few operators. The first two operators we consider are defined as follows:

$$\begin{aligned} [\mathcal{X}_1^{(2)}]_{ad}(s_b, s_c; s_f, s_e) &= \mathcal{X}_{ac}^{(2)}(s_b, s_f) \frac{\delta(s_c - s_e)}{\rho(s_c)}, \\ [\mathcal{X}_2^{(2)}]_{ad}(s_b, s_c; s_f, s_e) &= \mathcal{X}_{bd}^{(2)}(s_c, s_e) \frac{\delta(s_b - s_f)}{\rho(s_b)}. \end{aligned} \quad (\text{F1})$$

For convenience, we will often write $\mathcal{X}_1^{(2)}$ and $\mathcal{X}_2^{(2)}$ to refer to these operators.⁸² We depict these operators in Fig. 21. The product of $\mathcal{X}_i^{(2)}$ and $\mathcal{X}_j^{(2)}$ (with $i, j \in \{1, 2\}$) is defined as follows:

$$\begin{aligned} [\mathcal{X}_j^{(2)}\mathcal{X}_i^{(2)}]_{ad}(s_b, s_c; s_f, s_e) &= \int_0^\infty ds_1 \rho(s_1) ds_2 \rho(s_2) [\mathcal{X}_j^{(2)}]_{ad}(s_b, s_c; s_1, s_2) \\ &\quad \times [\mathcal{X}_i^{(2)}]_{ad}(s_1, s_2; s_f, s_e). \end{aligned} \quad (\text{F2})$$

We use this product to define the operators \mathcal{B} and \mathcal{P} as follows:

$$\mathcal{B} = \sum_{n=0}^{\infty} (\mathcal{X}_1^{(2)} + \mathcal{X}_2^{(2)})^n \quad (\text{F3})$$

$$\mathcal{P} = \mathcal{B} - 1 - \sum_{n=1}^{\infty} ((\mathcal{X}_1^{(2)})^n + (\mathcal{X}_2^{(2)})^n). \quad (\text{F4})$$

The motivation for defining \mathcal{B} and \mathcal{P} in this way will be clear from Fig. 24. The next two operators we introduce, \mathcal{A} and \mathcal{F} , are respectively defined from the connected planar six-point and four-point functions of \mathcal{O} . See Fig. 22. The last operator we need, $\mathcal{X}^{(3)}$, is defined in Fig. 23.

⁸²All of the subsequent operators we define will depend on the same set of energies, which we will no longer explicitly write unless necessary.

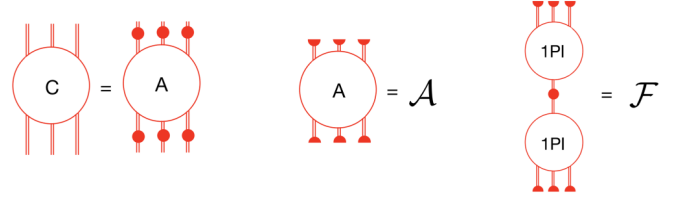


FIG. 22. On the left we relate the sum over connected planar six-point diagrams (represented by the “C” blob) to the sum of amputated six-point diagrams (represented by the “A” blob). The operator \mathcal{A} is defined by multiplying the sum of amputated six-point diagrams by a factor of $\sqrt{\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}}}$ for each external double-line, where s_a and s_b label the two energies appearing in the double-line. Each factor of $\sqrt{\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}}}$ is represented by a red semicircle. The operator \mathcal{F} is defined in terms of the 1PI planar four-point function as shown on the right. Note that \mathcal{A} and \mathcal{F} do not depend on the normalization of \mathcal{O} .

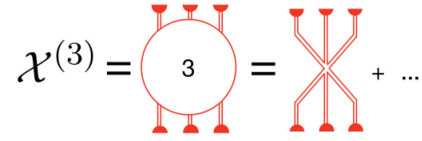


FIG. 23. The blob with a “3” corresponds to the sum over all connected planar six-point diagrams such that any left-right path through the blob crosses over more than three double-lines. We define $\mathcal{X}^{(3)}$ as the product of this blob with a factor of $\sqrt{\langle \mathcal{O}_{ab} \mathcal{O}_{ba} \rangle_{\text{disk}}}$ for each external leg, where s_a and s_b are the two energies associated with the leg. These factors are represented by red semicircles. The simplest contribution to the blob labeled “3” is a tree six-point vertex (in this case, there are no left-right paths that can go through the blob, because we only let left-right paths cross over double-lines, and all of the double-lines are external. So it is vacuously true that all left-right paths through the blob cross over more than three double-lines).

The product of any two operators is defined as in (F2). The integrals over s_1 and s_2 correspond to integrating over the energies that appear in the closed index loops that are formed when a diagram in the first operator is attached to a diagram in the second operator. Note that when either $\mathcal{X}_1^{(2)}$ or $\mathcal{X}_2^{(2)}$ appear in a product, at most one closed index loop is formed, which reflects the fact that delta functions appear in (F1). Having defined $\mathcal{X}_1^{(2)}$, $\mathcal{X}_2^{(2)}$, \mathcal{B} , \mathcal{P} , \mathcal{A} , \mathcal{F} , and $\mathcal{X}^{(3)}$, we note that the explicit formulas for all of these operators are known *a priori*, except for $\mathcal{X}^{(3)}$. This is because $\mathcal{X}_1^{(2)}$ and $\mathcal{X}_2^{(2)}$ follow from (9.35), while \mathcal{A} and \mathcal{F} follow directly from the connected planar correlators of \mathcal{O} , which are known from the (regulated) gravitational Feynman rules. To determine $\mathcal{X}^{(3)}$, we use the following relation:

$$\mathcal{A} = \mathcal{F} + \mathcal{P} + \mathcal{B}\mathcal{X}^{(3)}\mathcal{B}[1 - \mathcal{X}^{(3)}\mathcal{B}]^{-1}, \quad (\text{F5})$$

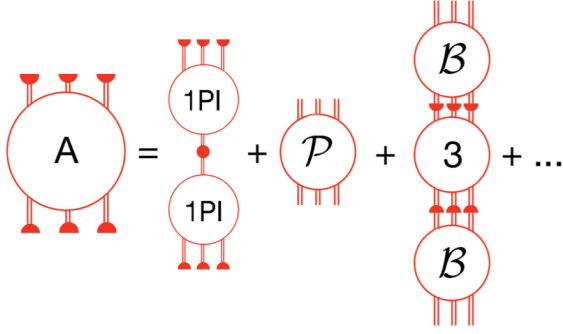


FIG. 24. To solve for $\mathcal{X}^{(3)}$, we systematically classify the diagrams that contribute to the amputated connected six-point function, which is contained in \mathcal{A} and appears on the left-hand side. The diagrams that contribute to \mathcal{F} (first term on the right-hand side) all allow for left-right paths that cross over only one double-line. The remaining diagrams on the right side have the property that any left-right path crosses over three or more double-lines. We organize these diagrams by explicitly showing all the locations where a left-right path can cross over three double-lines. The remaining parts of the diagrams are encapsulated in $\mathcal{X}^{(3)}$. The blobs labeled \mathcal{B} and \mathcal{P} should be replaced by sums over products of the first two diagrams in Fig. 21, according to (F3) and (F4). The ... includes further terms in the geometric series in (F5). Note that \mathcal{P} is defined similarly to \mathcal{B} , except that \mathcal{P} only contains connected diagrams. Given that the left-hand side is a sum of connected diagrams, only connected diagrams may appear on the right side.

which is depicted and explained in Fig. 24. After using (F5) to solve for $\mathcal{X}^{(3)}$, we may directly compute the sum of all diagrams that contribute to $\mathcal{D}^{(3)}$. The basic building blocks we use to enumerate these diagrams are $\mathcal{X}_1^{(2)}$, $\mathcal{X}_2^{(2)}$, and $\mathcal{X}^{(3)}$. Note that \mathcal{B} counts all of the diagrams that may be enumerated using $\mathcal{X}_1^{(2)}$ and $\mathcal{X}_2^{(2)}$ only. The result is

$$\mathcal{D}_{ad}^{(3)} = \text{Tr} \sum_{n=1}^{\infty} \frac{1}{n} ([\mathcal{X}_1^{(2)} + \mathcal{X}_2^{(2)}]^n - [\mathcal{X}_1^{(2)}]^n - [\mathcal{X}_2^{(2)}]^n) + \text{Tr} \sum_{n=1}^{\infty} \frac{1}{n} [\mathcal{X}^{(3)} \mathcal{B}]^n. \quad (\text{F6})$$

The trace in (F6) corresponds to setting $s_b = s_f$ and $s_c = s_e$ in Fig. 21 and then integrating over these energies with measure $d\rho(s)$. We are left with a function of s_a and s_d that corresponds to $\mathcal{D}_{ad}^{(3)}$. On the first line of (F6), we enumerate all of the diagrams that may be built out of $\mathcal{X}_1^{(2)}$ and $\mathcal{X}_2^{(2)}$. The subtractions ensure that we only count connected diagrams. On the second line, we organize the diagrams by the number of $\mathcal{X}^{(3)}$ factors that appear. The $\frac{1}{n}$ factors ensure that there is no overcounting and also supply the correct symmetry factors to the diagrams that have a \mathbb{Z}_n symmetry.

Using (F5), we have that

$$\mathcal{A} - \mathcal{F} - \mathcal{P} + \mathcal{B} = \mathcal{B}[1 - \mathcal{X}^{(3)}\mathcal{B}]^{-1} := \mathcal{R}, \quad (\text{F7})$$

from which it follows that

$$1 - \mathcal{R}^{-1}\mathcal{B} = \mathcal{X}^{(3)}\mathcal{B}. \quad (\text{F8})$$

The second line of (F6) becomes

$$\begin{aligned} \text{Tr} \sum_{n=1}^{\infty} \frac{1}{n} [\mathcal{X}^{(3)}\mathcal{B}]^n &= -\text{Tr} \log \mathcal{R}^{-1}\mathcal{B} \\ &= \text{Tr} \log \mathcal{R} - \text{Tr} \log \mathcal{B}, \end{aligned} \quad (\text{F9})$$

so that (F6) becomes

$$\begin{aligned} \mathcal{D}_{ad}^{(3)} &= \text{Tr} \log(1 - \mathcal{X}_1^{(2)}) + \text{Tr} \log(1 - \mathcal{X}_2^{(2)}) \\ &\quad + \text{Tr} \log \mathcal{R}, \end{aligned} \quad (\text{F10})$$

where we have used (F3) to substitute for \mathcal{B} . Our method for computing $\mathcal{D}^{(3)}$ is valid for any single-trace, two-matrix model. Given the connected planar correlators which are fixed by the disk amplitudes of the model, we may determine $\mathcal{X}_1^{(2)}$, $\mathcal{X}_2^{(2)}$, and \mathcal{R} and thus $\mathcal{D}^{(3)}$.

We now obtain explicit expressions for $\mathcal{D}^{(3)}$ using the q -deformed and Selberg regulators. We first define the operators \mathcal{S}_1^q and \mathcal{S}_2^q as follows:

$$[\mathcal{S}_1^q]_{ad}(s_b, s_c; s_f, s_e) \equiv \left\{ \begin{array}{ccc} \Delta & s_b & s_c \\ \Delta & s_f & s_a \end{array} \right\}_q \frac{\delta(s_c - s_e)}{\rho_q(s_c)}, \quad (\text{F11})$$

$$[\mathcal{S}_2^q]_{ad}(s_b, s_c; s_f, s_e) \equiv \frac{\delta(s_b - s_f)}{\rho_q(s_c)} \left\{ \begin{array}{ccc} \Delta & s_c & s_d \\ \Delta & s_e & s_b \end{array} \right\}_q. \quad (\text{F12})$$

Note that the Yang-Baxter equation is

$$\mathcal{S}_1^q \mathcal{S}_2^q \mathcal{S}_1^q = \mathcal{S}_2^q \mathcal{S}_1^q \mathcal{S}_2^q. \quad (\text{F13})$$

Using the regulated gravitational Feynman rules to determine the disk amplitudes, we find that

$$\mathcal{A} = \mathcal{F} + \epsilon^2 \mathcal{S}_2^q \mathcal{S}_1^q + \epsilon^2 \mathcal{S}_1^q \mathcal{S}_2^q + \epsilon^\# \mathcal{S}_1^q \mathcal{S}_2^q \mathcal{S}_1^q, \quad (\text{F14})$$

where $\# = 3$ for the q -deformed regulator and $\# = 2$ for the Selberg regulator, and

$$[\mathcal{F}]_{ad}(s_b, s_c; s_f, s_e) = \epsilon^2 \left\{ \begin{array}{ccc} \Delta & s_b & s_c \\ \Delta & s_d & s_a \end{array} \right\}_q \left\{ \begin{array}{ccc} \Delta & s_e & s_f \\ \Delta & s_a & s_d \end{array} \right\}_q. \quad (\text{F15})$$

From (F1) and the q -analog of (9.35), we have that

$$\mathcal{X}_1^{(2)} = \frac{\epsilon \mathcal{S}_1^q}{1 + \epsilon \mathcal{S}_1^q}, \quad \mathcal{X}_2^{(2)} = \frac{\epsilon \mathcal{S}_2^q}{1 + \epsilon \mathcal{S}_2^q}, \quad (\text{F16})$$

and it follows using (F4) that

$$\mathcal{B} - \mathcal{P} = [1 - \mathcal{X}_1^{(2)}]^{-1} + [1 - \mathcal{X}_2^{(2)}]^{-1} - 1, \quad (\text{F17})$$

$$= 1 + \epsilon \mathcal{S}_1^q + \epsilon \mathcal{S}_2^q. \quad (\text{F18})$$

Using (F7) and (F10), we finally have

$$\begin{aligned} \mathcal{D}^{(3)} = & -\text{Tr} \log(1 + \epsilon \mathcal{S}_1^q) - \text{Tr} \log(1 + \epsilon \mathcal{S}_2^q) \\ & + \text{Tr} \log(1 + \epsilon \mathcal{S}_1^q + \epsilon \mathcal{S}_2^q + \epsilon^2 \mathcal{S}_2^q \mathcal{S}_1^q \\ & + \epsilon^2 \mathcal{S}_1^q \mathcal{S}_2^q + \epsilon^\# \mathcal{S}_1^q \mathcal{S}_2^q \mathcal{S}_1^q). \end{aligned} \quad (\text{F19})$$

The first line of (F19) only depends on the spectrum of \mathcal{S}_1^q and \mathcal{S}_2^q , which is independent of the two energies that $\mathcal{D}^{(3)}$ could depend on. We now argue that the second line of (F19) is also independent of these energies.

To compute the second line of (F19), we will construct two convenient orthonormal bases for the space of functions of two energies that the operators \mathcal{S}_1^q and \mathcal{S}_2^q act on. We will refer to basis vectors of the first (resp. second) basis as $|n, m\rangle_1$ (resp. $|n, m\rangle_2$), where $n, m \in \mathbb{Z}_{\geq 0}$. The basis vector $|s\rangle$ in the space of functions of one energy is normalized such that

$$\langle s_1 | s_2 \rangle = \frac{\delta(s_1 - s_2)}{\rho_q(s_1)}. \quad (\text{F20})$$

Then, we define $|n, m\rangle_1$ by

$$\begin{aligned} (\langle s_f | \otimes \langle s_e |) |n, m\rangle_1 \\ = P_n^{\Delta, \Delta}(s_f; s_a, s_e | q) P_m^{2\Delta + n, \Delta}(s_e; s_a, s_d | q), \end{aligned} \quad (\text{F21})$$

and $|n, m\rangle_2$ is defined by

$$\begin{aligned} (\langle s_f | \otimes \langle s_e |) |n, m\rangle_2 \\ = P_n^{\Delta, 2\Delta + m}(s_f; s_a, s_d | q) P_m^{\Delta, \Delta}(s_e; s_f, s_d | q). \end{aligned} \quad (\text{F22})$$

Using (A63), we have

$${}_1\langle n, m | \tilde{n}, \tilde{m} \rangle_1 = {}_2\langle n, m | \tilde{n}, \tilde{m} \rangle_2 = \delta_{n, \tilde{n}} \delta_{m, \tilde{m}}, \quad (\text{F23})$$

while (G41) ensures that these bases are complete. Our two bases are eigenbases of \mathcal{S}_1^q and \mathcal{S}_2^q ,

$$\begin{aligned} \mathcal{S}_1^{q, \epsilon} |n, m\rangle_1 &= (-1)^n q^{\frac{n(n-1)}{2}} q^{2\Delta n} |n, m\rangle_1, \\ \mathcal{S}_2^{q, \epsilon} |n, m\rangle_2 &= (-1)^m q^{\frac{m(m-1)}{2}} q^{2\Delta m} |n, m\rangle_2, \end{aligned} \quad (\text{F24})$$

which shows that \mathcal{S}_1^q and \mathcal{S}_2^q are Hermitian.

The operators \mathcal{S}_1^q and \mathcal{S}_2^q , subject to the relation (F13), generate an algebra. A casimir of this algebra is given by

$$\mathcal{C} := (\mathcal{S}_1^q \mathcal{S}_2^q \mathcal{S}_1^q)^2. \quad (\text{F25})$$

Using (G52) and (A82), we find that

$$\begin{aligned} \mathcal{C} |n, m\rangle_1 &= \mathcal{C}_{n, m} |n, m\rangle_1, \quad \mathcal{C} |n, m\rangle_2 = \mathcal{C}_{n, m} |n, m\rangle_2, \\ \mathcal{C}_{n, m} &:= q^{(n+m)(6\Delta-1+n+m)}. \end{aligned} \quad (\text{F26})$$

In particular, the casimir depends only on the sum of n and m . Because n and m are non-negative integers, each value of the casimir corresponds to a finite-dimensional representation of the algebra. Hence, in either of the two bases we introduced, any operator in the algebra takes a block-diagonal form, where each block is finite dimensional. To compute the second line of (F19), we need to compute the log of the determinant of the operator appearing in parentheses. It suffices to compute the determinant of this operator in each individual block. The second line of (F19) thus takes the form of an infinite sum, where each term in the sum corresponds to a single block. Let us consider the matrix elements of \mathcal{S}_1^q and \mathcal{S}_2^q in the $|n, m\rangle_1$ basis for concreteness. The matrix elements of \mathcal{S}_1^q are determined by its eigenvalues, which do not depend on the two energies that $\mathcal{D}^{(3)}$ could depend on. To determine the matrix elements of \mathcal{S}_2^q , we use (G55), which implies that

$$|n, m\rangle_1 = \sum_{y=0}^{n+m} S_q^{\Delta, \Delta, \Delta, n+m}(y, n) |n + m - y, y\rangle_2. \quad (\text{F27})$$

Hence, the matrix elements of \mathcal{S}_2^q in the basis that diagonalizes \mathcal{S}_1^q do not depend on the two energies that $\mathcal{D}^{(3)}$ could have depended on. It follows that $\mathcal{D}_{ad}^{(3)}$ is independent of s_a, s_d and does not contribute to the empty double-trumpet.

Having explicitly shown that $\mathcal{D}^{(2)}$ and $\mathcal{D}^{(3)}$ do not contribute to the empty double-trumpet, we conjecture that $\mathcal{D}^{(n)}$ does not contribute for all $n \geq 2$. The formula for $\mathcal{D}^{(n)}$ should involve determinants of operators acting on the space of functions of n energies, and we expect that these operators are part of the algebra generated by the appropriate analogs of \mathcal{S}_1^q and \mathcal{S}_2^q . An argument similar to the one given above will then show that $\mathcal{D}_{ad}^{(n)}$ does not depend on s_a or s_d .

2. Two-point function on the double-trumpet

Continuing from the end of Sec. IX B, we now consider diagrams such that the minimum number of \mathcal{O} double-lines that are crossed by a left-right path is three. This is the last class of diagrams that we explicitly compute. In this case, the analog of the green blob in Fig. 16 will have eight external double-lines. Six of the double-lines will wrap

around the double-trumpet, while the other two double-lines go to the AdS boundaries where they represent an insertion of \mathcal{O} . The green blob should only include connected diagrams. However, there are connected diagrams that allow for left-right paths that cross over only one or two double-lines. These diagrams must be subtracted from the sum over all connected eight-point diagrams (this subtraction is analogous to the subtraction on the right-hand side of Fig. 16). In Figs. 25–27, we enumerate all of the connected diagrams that should be subtracted. We then define the green blob in Fig. 28. In Fig. 29, we enumerate all the diagrams that we wish to resum.

Next, note that (F3)–(F5) imply that

$$\mathcal{A} - \mathcal{F} = \mathcal{P} + \mathcal{B}([1 - \mathcal{X}^{(3)}\mathcal{B}]^{-1} - 1) \quad (\text{F28})$$

$$= -1 - \frac{\mathcal{X}_1^{(2)}}{1 - \mathcal{X}_1^{(2)}} - \frac{\mathcal{X}_2^{(2)}}{1 - \mathcal{X}_2^{(2)}} + \mathcal{B}[1 - \mathcal{X}^{(3)}\mathcal{B}]^{-1} \quad (\text{F29})$$

$$= -1 - \frac{\mathcal{X}_1^{(2)}}{1 - \mathcal{X}_1^{(2)}} - \frac{\mathcal{X}_2^{(2)}}{1 - \mathcal{X}_2^{(2)}} + \frac{1}{\mathcal{B}^{-1} - \mathcal{X}^{(3)}} \quad (\text{F30})$$

$$= -1 - \frac{\mathcal{X}_1^{(2)}}{1 - \mathcal{X}_1^{(2)}} - \frac{\mathcal{X}_2^{(2)}}{1 - \mathcal{X}_2^{(2)}} + \frac{1}{1 - \mathcal{X}_1^{(2)} - \mathcal{X}_2^{(2)} - \mathcal{X}^{(3)}} \quad (\text{F31})$$

from which it follows that

$$\begin{aligned} & \frac{1}{1 - \mathcal{X}_1^{(2)} - \mathcal{X}_2^{(2)} - \mathcal{X}^{(3)}} \\ &= 1 + \frac{\mathcal{X}_1^{(2)}}{1 - \mathcal{X}_1^{(2)}} + \frac{\mathcal{X}_2^{(2)}}{1 - \mathcal{X}_2^{(2)}} + \mathcal{A} - \mathcal{F}. \end{aligned} \quad (\text{F32})$$

Recalling the graphical representations of $\mathcal{X}_1^{(2)}$ and $\mathcal{X}_2^{(2)}$ in Fig. 21 as well as the representation of $\mathcal{X}^{(3)}$ in Fig. 23, it follows that the left-hand side of (F32) is graphically represented by

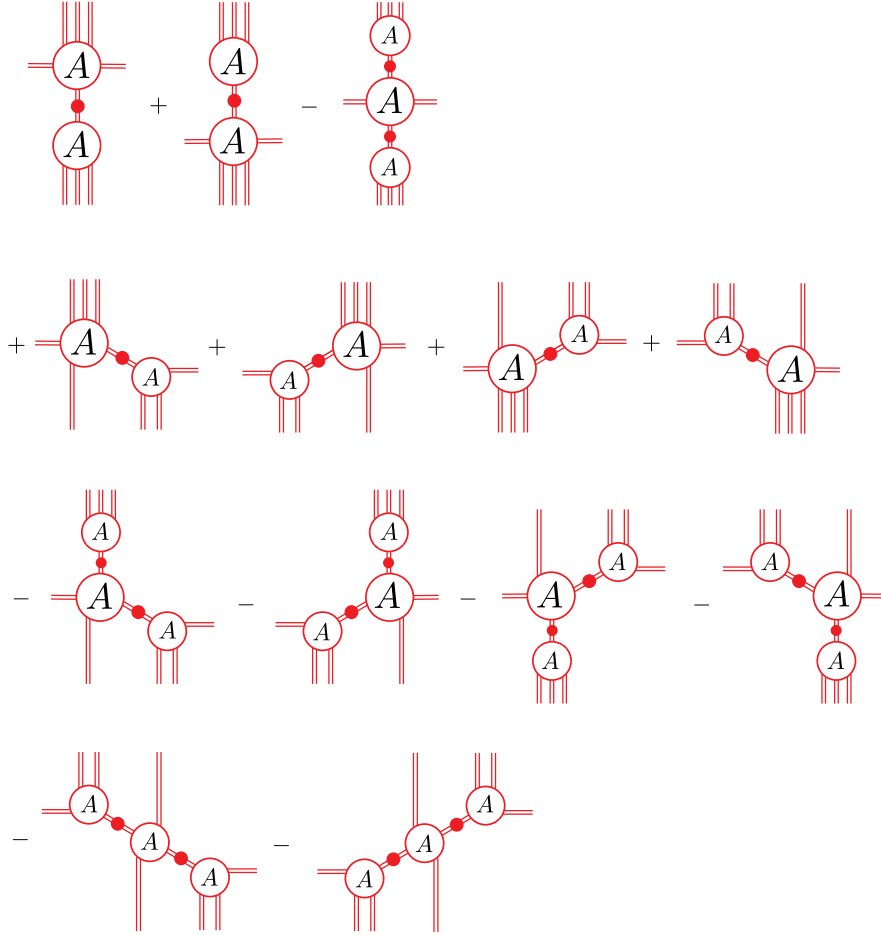


FIG. 25. In this figure, Figs. 26 and 27, we enumerate all of the connected eight-point diagrams that admit a “shortcut” from the left to the right side of the diagram. A “shortcut” is defined as a left-right path that crosses over one or two double-lines. In keeping with the conventions in Fig. 22, a blob with an “A” in it refers to a sum over connected planar amputated diagrams.

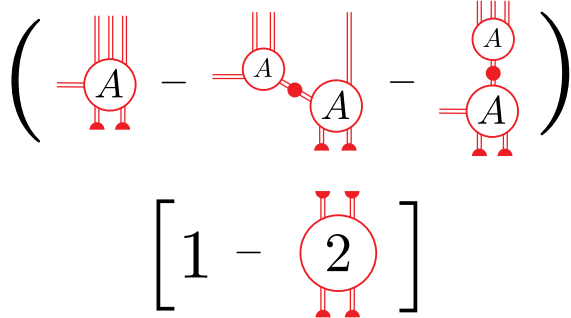


FIG. 26. In this figure, Figs. 25 and 27, we enumerate all of the connected eight-point diagrams that admit a “shortcut” from the left to the right side of the diagram. A blob with an “A” in it refers to a sum over connected planar amputated diagrams. The three factors that are arranged vertically are meant to be multiplied. The “2” blob was defined in Fig. 11.

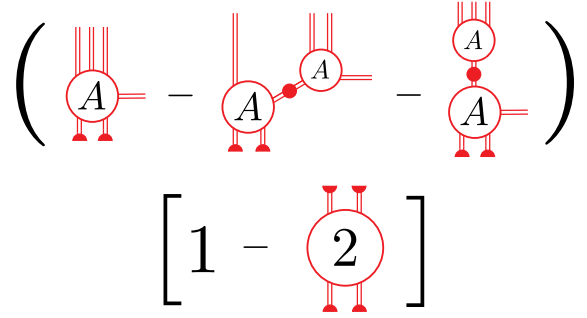


FIG. 27. In this figure, Figs. 25, and 26, we enumerate all of the connected eight-point diagrams that admit a “shortcut” from the left to the right side of the diagram. A blob with an “A” in it refers to a sum over connected planar amputated diagrams. The three factors that are arranged vertically are meant to be multiplied. The “2” blob was defined in Fig. 11.

$$\left[1 - \begin{array}{c} \text{blob 2} \\ \text{blob 2} \\ \text{blob 3} \end{array} \right]^{-1} \quad (\text{F33})$$

which appears in Fig. 29.

We now evaluate the right-hand side of Fig. 28 using the q -deformed regulator. There are 27 connected chord diagrams with eight external lines. These contribute to the “A” blob. Of these chord diagrams, 21 are subtracted. The right-hand side of Fig. 28 becomes the sum over the following six chord diagrams:

$$\begin{aligned} & \epsilon^3 \begin{array}{c} | \quad | \quad | \\ \hline | \quad | \quad | \end{array} + \epsilon^4 \left(\begin{array}{c} \diagdown \quad | \quad | \\ \hline | \quad \diagup \quad | \end{array} + \begin{array}{c} | \quad \diagdown \quad | \\ \hline | \quad | \quad \diagup \end{array} \right) \\ & + \epsilon^5 \left(\begin{array}{c} \diagdown \quad \diagup \\ \hline \diagup \quad \diagdown \end{array} + \begin{array}{c} \diagdown \quad \diagup \\ \hline \diagup \quad \diagdown \end{array} \right) + \epsilon^6 \begin{array}{c} \diagdown \quad \diagup \\ \hline \diagup \quad \diagdown \end{array} \end{aligned} \quad (\text{F34})$$

This is a symmetric sum over six permutations. To obtain the sum over the diagrams in Fig. 29, we should multiply (F34) by the inverse of (F32) and identify the top and bottom ends of the diagrams (or equivalently, take a trace). In the q -deformed model, (F32) becomes

$$\begin{aligned}
 & \left[1 - \text{blob}_2 - \text{blob}_2 - \text{blob}_3 \right]^{-1} \\
 & \quad \text{blob}_8 = \text{blob}_A - \text{blob}_B \\
 & \left[1 - \text{blob}_2 - \text{blob}_2 - \text{blob}_3 \right]^{-1}
 \end{aligned}$$

FIG. 28. We define the green blob with eight external double-lines. First, we subtract all diagrams that permit shortcuts from the sum over connected amputated eight-point diagrams. The blue blob is defined to be the sum over all the diagrams in Figs. 25–27. Furthermore, we must amputate off factors of the “2” and “3” blobs from the top and bottom three legs. The “3” blob is defined in Fig. 23. This figure defines the green blob in analogy to Fig. 17. The top and bottom ends of this diagram are not identified.

$$\begin{aligned}
 & \left(\begin{array}{c} | \\ | \\ | \end{array} \right) + \epsilon \left(\begin{array}{c} \diagup \diagdown \\ | \end{array} + \begin{array}{c} | \\ \diagdown \diagup \end{array} \right) \\
 & + \epsilon^2 \left(\begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} \right) + \epsilon^3 \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \\ \diagup \diagdown \end{array}
 \end{aligned} \tag{F35}$$

Multiplying (F34) by the inverse of (F35) has the effect of undoing the symmetrization over the three lines in (F34). The final result for Fig. 29 is

$$\epsilon^3 \begin{array}{c} | \\ | \\ | \end{array} \tag{F36}$$

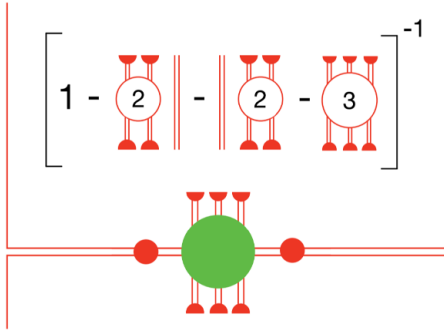


FIG. 29. We enumerate the sum over all diagrams such that the minimum number of double-lines crossed by a left-right path is three. The green blob is defined in Fig. 28. The top and bottom ends of this diagram are identified.

After using the pentagon identity (9.47) and taking the JT limit, this becomes

$$\begin{aligned}
 & \int ds_a \rho(s_a) ds_c \rho(s_c) e^{-\beta_L s_a^2 - \beta_R s_c^2} (\Gamma_{aa}^\Delta \Gamma_{cc}^\Delta)^{1/2} \\
 & \times \sum_{n,m=0}^{\infty} \left\{ \begin{array}{cc} 3\Delta + n + m & s_a \quad s_c \\ \Delta & s_c \quad s_a \end{array} \right\}.
 \end{aligned} \tag{F37}$$

Finally, we turn to the Selberg regulator. Using the Selberg regulator, (F32) becomes

$$\begin{aligned}
& \left| \begin{array}{c} | \\ | \\ | \end{array} \right| + \epsilon \left(\begin{array}{c} \diagdown \diagup \\ | \end{array} + \begin{array}{c} | \\ \diagdown \diagup \end{array} \right) \\
& + \epsilon^2 \left(\begin{array}{c} \diagdown \diagup \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \\ \diagdown \diagup \end{array} \right)
\end{aligned} \tag{F38}$$

which differs from (F35) by the weighting assigned to the last diagram.

We now consider the right-hand side of Fig. 28 using the Selberg regulator. Note that each chord diagram that contributes to the “A” blob is weighted by ϵ^3 . Furthermore, every diagram in Fig. 25 is weighted by ϵ^3 . In Figs. 26 and 27, the terms inside the round brackets are weighted by ϵ^2 . The factor inside the square brackets may be evaluated from (9.44). To simplify the remainder of this computation, we now set $q = 1$. From (9.44), we have that

$$1 - \text{blob} = \left(1 + \epsilon \begin{array}{c} \diagdown \diagup \end{array} \right)^{-1} = \frac{1}{1 - \epsilon^2} \left(1 - \epsilon \begin{array}{c} \diagdown \diagup \end{array} \right) \tag{F39}$$

where in the last equality we used (2.14). It follows that the diagrams in Fig. 26 evaluate to

$$\frac{2\epsilon^4}{1 + \epsilon} \left(\begin{array}{c} \diagdown \diagup \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} \right) \tag{F40}$$

while those in Fig. 27 evaluate to

$$\frac{2\epsilon^4}{1 + \epsilon} \left(\begin{array}{c} \diagdown \diagup \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} \right). \tag{F41}$$

The right-hand side of Fig. 28 then becomes the sum over the following chord diagrams:

$$\begin{aligned}
& \epsilon^3 \left(\begin{array}{c} | \\ | \\ | \end{array} + \begin{array}{c} \diagdown \diagup \\ | \end{array} + \begin{array}{c} | \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \\ \diagdown \diagup \end{array} \right) \\
& + \left(\epsilon^3 - \frac{2\epsilon^4}{1 + \epsilon} \right) \left(\begin{array}{c} \diagdown \diagup \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagdown \diagup \\ \diagup \diagdown \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \\ \diagdown \diagup \end{array} \right).
\end{aligned} \tag{F42}$$

Next, we must multiply the inverse of (F38) with (F42) and then take a trace. Evaluating the inverse of (F38) is easier when $q = 1$. The inverse of (F38) is

$$c_1 \begin{array}{c} | \\ | \\ | \end{array} + c_2 \begin{array}{c} \diagdown \diagup \\ | \end{array} + c_3 \begin{array}{c} | \\ \diagdown \diagup \end{array} + c_4 \begin{array}{c} \diagdown \diagup \\ \diagdown \diagup \end{array} + c_5 \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} + c_6 \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \\ \diagdown \diagup \end{array} \tag{F43}$$

where the coefficients c_j are

$$\begin{aligned}
 c_1 &= \frac{2\epsilon(-\epsilon^3 + \epsilon + 1) + 1}{(\epsilon - 1)^2(\epsilon(\epsilon(6\epsilon + 7) + 4) + 1)} \\
 c_2 &= c_3 = \frac{\epsilon((\epsilon - 2)\epsilon(\epsilon + 1) - 1)}{(\epsilon - 1)^2(\epsilon(\epsilon(6\epsilon + 7) + 4) + 1)} \\
 c_4 &= c_5 = \frac{\epsilon^3(\epsilon + 2)}{(\epsilon - 1)^2(\epsilon(\epsilon(6\epsilon + 7) + 4) + 1)} \\
 c_6 &= -\frac{\epsilon^2(2\epsilon^2 + 1)}{(\epsilon - 1)^2(\epsilon(\epsilon(6\epsilon + 7) + 4) + 1)}.
 \end{aligned} \tag{F44}$$

We now multiply (F43) with the final line of (F42) and then take the $\epsilon \rightarrow 1$ limit (we will come back to the rest of (F42) later). We then take a trace by identifying the top and bottom ends of the diagrams. The result is

$$\begin{aligned}
 & \frac{1}{3!} \left(\text{diagram 1} - \text{diagram 2} \right) + \frac{1}{3!} \left(\text{diagram 3} - \text{diagram 4} \right) \\
 & + \frac{1}{3!} \left(\text{diagram 5} - \text{diagram 6} \right) + \frac{1}{3!} \left(\text{diagram 7} - \text{diagram 8} \right).
 \end{aligned} \tag{F45}$$

Because we have identified the top and bottom ends of each diagram, it turns out that (F45) evaluates to zero. In fact, each of the bracketed terms separately vanish.

We now multiply (F43) with the sum over the six terms appearing in brackets in (F42). In the JT limit, the result finally becomes

$$\begin{aligned}
 & \frac{1}{3!} \left(\text{diagram 9} + \text{diagram 10} + \text{diagram 11} \right. \\
 & \left. + \text{diagram 12} + \text{diagram 13} + \text{diagram 14} \right)
 \end{aligned} \tag{F46}$$

Now we would like to compute (F46), where we identify the top and bottom parts of the diagram. We start with the first three terms

$$\text{tr} \text{diagram 9} = \int_0^\infty dk_1 dk_2 \rho(k_1) \rho(k_2) \frac{|k_1| |k_2|}{|k_1| |k_2|} = \sum_{n,m=0}^\infty \left\{ \begin{matrix} 3\Delta + n + m \\ \Delta \end{matrix} \begin{matrix} s_a & s_b \\ s_b & s_a \end{matrix} \right\}, \tag{F47}$$

$$\text{tr} \text{diagram 10} = \text{tr} \text{diagram 11} = \sum_{n,m=0}^\infty (-1)^n \left\{ \begin{matrix} 3\Delta + n + m \\ \Delta \end{matrix} \begin{matrix} s_a & s_b \\ s_b & s_a \end{matrix} \right\}. \tag{F48}$$

The trace here means that we identify top and bottom parts of the diagram and integrate over k_1, k_2 with the density of states. In both equations we used the pentagon identity (A32) twice and orthogonality of Wilson polynomials. In the second equation we also used (A27).

The last three terms in (F46) are slightly more tricky to compute. The fourth term gives

$$\text{tr} \text{---} \text{---} \text{---} = \sum_{n=0}^{\infty} \text{tr} \text{---} \text{---} \text{---} = \sum_{n,m=0}^{\infty} (-1)^m \text{tr} \text{---} \text{---} \text{---} \quad (\text{F49})$$

$$= \sum_{n,m=0}^{\infty} (-1)^m \text{tr} \text{---} \text{---} \text{---} . \quad (\text{F50})$$

In the first equality we used the pentagon identity. Then we used an expansion of the $6j$ -symbol into Wilson polynomials (A28). And finally, we computed the trace by connecting top and bottom lines.

To compute (F50) we note that it is an overlap between two products of Wilson polynomials

$$\begin{array}{c} k_1 \\ \diagup \\ s_a \quad n \quad s_b \\ \diagdown \\ m \end{array} = P_n(k_1; \Delta \pm i s_a, \Delta \pm i k_2) P_m(k_2; 2\Delta + n \pm i s_a, \Delta \pm i s_b) , \quad (\text{F51})$$

$$\begin{array}{c} k_1 \\ \diagup \\ s_a \quad n \quad s_b \\ \diagdown \\ m \end{array} = P_m(k_1; \Delta \pm i s_a, 2\Delta + n \pm i s_b) P_n(k_2; \Delta \pm i k_1, \Delta \pm i s_b) . \quad (\text{F52})$$

These two products can be thought of as bases in the space of functions of two variables k_1, k_2 . Since both of them are bases, they can be related to each other by a linear transformation. Indeed, such a transformation can be derived by analytically continuing the Theorem 7.6 of [74]

$$\begin{array}{c} k_1 \\ \diagup \\ s_a \quad n \quad s_b \\ \diagdown \\ m \end{array} = \sum_{j=0}^{n+m} c_j R_j \begin{array}{c} k_1 \\ \diagup \\ s_a \quad j \quad s_b \\ \diagdown \\ n+m-j \end{array} , \quad (\text{F53})$$

where

$$c_j = \sqrt{\mathcal{N}_j} \binom{n+m}{j} \frac{(2\Delta)_n (2\Delta)_m}{(2\Delta)_j} \frac{(6\Delta + n + m - 1)_j}{(4\Delta + j - 1)_j (4\Delta + 2j)_{n+m-j}}, \quad (\text{F54})$$

$$\mathcal{N}_j = \frac{j!(n+m-j)!}{n!m!} \frac{(6\Delta+n+m+j-1)_{n+m-j}}{(6\Delta+2n+m-1)_m} \frac{(4\Delta+j-1)_j}{(4\Delta+n-1)_n} \quad (\text{F55})$$

$$\frac{\Gamma(4\Delta + 2n) \Gamma(4\Delta + n + m + j) \Gamma(2\Delta + n + m - j) \Gamma(2\Delta + j)^2}{\Gamma(4\Delta + 2j) \Gamma(4\Delta + 2n + m) \Gamma(2\Delta + m) \Gamma(2\Delta + n)^2} \quad (\text{F56})$$

and R_j is the Racah polynomial defined on page 1 of [74]

$$R_j \equiv R_j(n; 2\Delta - 1, 2\Delta - 1, -n - m - 1, 4\Delta + n + m - 1) \quad (\text{F57})$$

$$= {}_4F_3 \left(\begin{matrix} -j, -n, 4\Delta + j - 1, 4\Delta + n - 1 \\ 2\Delta, -n - m, 6\Delta + n + m - 1 \end{matrix}; 1 \right). \quad (\text{F58})$$

In (F50) the top and bottom lines are identified and carry the index m . This means that we are interested in the term $j = n$ in the rhs of (F53). In this case some of the expressions simplify. In particular, $\mathcal{N}_n = 1$. We find

$$\begin{array}{c} m \\ | \\ \diagup \quad \diagdown \\ | \quad | \\ \diagdown \quad \diagup \\ | \\ n \end{array} = (c_j R_j)|_{j=n} \quad (F59)$$

$$= \left(n + \frac{m}{n}\right) \frac{(2\Delta)_n (6\Delta + n + m - 1)_n}{(4\Delta + 2n)_m (4\Delta + n - 1)_n} {}_4F_3 \left(\begin{matrix} -n, -n, 4\Delta + n - 1, 4\Delta + n - 1 \\ 2\Delta, -n - m, 6\Delta + n + m - 1 \end{matrix}; 1 \right). \quad (\text{F60})$$

Finally, we find

$$\text{tr} \text{---} \text{---} \text{---} = \text{tr} \text{---} \text{---} \text{---} = \sum_{n,m=0}^{\infty} (-1)^m c_n R_n \left\{ \begin{matrix} 3\Delta + n + m & s_a & s_b \\ \Delta & s_b & s_a \end{matrix} \right\}. \quad (\text{F61})$$

A similar computation shows

$$\text{tr} \begin{array}{c} \diagup \quad \diagdown \\ | \\ \hline \diagdown \quad \diagup \\ | \end{array} = \sum_{n,m=0}^{\infty} (-1)^{n+m} c_n R_n \left\{ \begin{array}{cc} 3\Delta + n + m & s_a \quad s_b \\ \Delta & s_b \quad s_a \end{array} \right\}. \quad (\text{F62})$$

Combining the above results we have

$$\frac{1}{3!} \text{tr} \left(\begin{array}{c} | \\ | \\ | \\ | \end{array} + \begin{array}{c} \diagup \\ | \\ | \\ | \end{array} + \begin{array}{c} | \\ \diagdown \\ | \\ | \end{array} + \begin{array}{c} \diagup \\ \diagdown \\ | \\ | \end{array} + \begin{array}{c} \diagdown \\ \diagup \\ | \\ | \end{array} + \begin{array}{c} \diagup \\ \diagdown \\ \diagup \\ \diagdown \end{array} \right) \quad (\text{F63})$$

$$= \frac{1}{3!} \sum_{n,m=0}^{\infty} [1 + 2(-1)^n + 2(-1)^m c_n R_n + (-1)^{n+m} c_n R_n] \left\{ \begin{array}{cc} 3\Delta + n + m & s_a \quad s_b \\ \Delta & s_b \quad s_a \end{array} \right\} \quad (\text{F64})$$

$$= \frac{1}{3!} \sum_{N=0}^{\infty} \left\{ \begin{array}{cc} 3\Delta + N & s_a \quad s_b \\ \Delta & s_b \quad s_a \end{array} \right\} \sum_{n=0}^N [1 + 2(-1)^n + 2(-1)^m c_n R_n + (-1)^{n+m} c_n R_n] \Big|_{m=N-n}. \quad (\text{F65})$$

We would like to compare this result with JT gravity. The next term in (2.36) (that we didn't write there) is

$$\sum_{n,m=0}^{\infty} \left\{ \begin{array}{cc} 3\Delta + 2n + 3m & s_a \quad s_b \\ \Delta & s_b \quad s_a \end{array} \right\} = \sum_{N=0}^{\infty} \left\{ \begin{array}{cc} 3\Delta + N & s_a \quad s_b \\ \Delta & s_b \quad s_a \end{array} \right\} \sum_{n,m=0}^{\infty} \delta_{2n+3m,N}. \quad (\text{F66})$$

To match this with (F65) we need

$$\frac{1}{3!} \sum_{n=0}^N [1 + 2(-1)^n + 2(-1)^m c_n R_n + (-1)^{n+m} c_n R_n] \Big|_{m=N-n} = \sum_{n,m=0}^{\infty} \delta_{2n+3m,N}, \quad (\text{F67})$$

where $c_n R_n$ is defined in (F60). Incredibly, this identity is indeed true. We didn't find its derivation, but checked it numerically for several values of N .

APPENDIX G: REPRESENTATION THEORY OF $\mathcal{U}_q(\mathfrak{su}(1,1))$

In this appendix, we introduce the $\mathcal{U}_q(\mathfrak{su}(1,1))$ algebra and derive important relations involving the various 6j symbols that are encountered throughout this paper. We mostly follow the discussion in [74], although our conventions differ. Any results that are not proven here are attributable to [74]. The main purpose of this section is to derive the identities (G52) and (G55).

1. The $\mathcal{U}_q(\mathfrak{su}(1,1))$ algebra

The $\mathcal{U}_q(\mathfrak{su}(1,1))$ algebra is a q -deformation of the universal enveloping algebra of $\mathfrak{su}(1,1)$. The generators K , K^{-1} , E , and F obey the relations

$$KK^{-1} = 1 = K^{-1}K, \quad KE = q^{1/2}EK, \quad KF = q^{-1/2}FK, \quad EF - FE = \frac{K^2 - K^{-2}}{q^{1/2} - q^{-1/2}}, \quad (\text{G1})$$

and the Casimir is given by

$$\Omega = \frac{q^{-1/2}K^2 + q^{1/2}K^{-2} - 2}{(q^{-1/2} - q^{1/2})^2} + EF = \frac{q^{-1/2}K^{-2} + q^{1/2}K^2 - 2}{(q^{-1/2} - q^{1/2})^2} + FE. \quad (\text{G2})$$

This algebra is a Hopf $*$ -algebra, and the comultiplication operator Δ is defined as follows:

$$\Delta(K) = K \otimes K, \quad \Delta(E) = K \otimes E + E \otimes K^{-1}, \quad \Delta(K^{-1}) = K^{-1} \otimes K^{-1}, \quad \Delta(F) = K \otimes F + F \otimes K^{-1}. \quad (\text{G3})$$

Note that this algebra is not cocommutative, which means that the tensor product of two representations depends on their ordering. The adjoint operation is defined by

$$K^\dagger = K, \quad E^\dagger = -F, \quad F^\dagger = -E, \quad (K^{-1})^\dagger = K^{-1}. \quad (\text{G4})$$

We may define an automorphism of the algebra as follows:

$$\tilde{K} = K^{-1}, \quad \tilde{K}^{-1} = K, \quad \tilde{E} = F, \quad \tilde{F} = E. \quad (\text{G5})$$

The tilde generators also obey (G1), (G4), and (G2) for the same Ω . The coproduct is not invariant under (G5). If X represents a generator of the algebra such that

$$\Delta(X) = \sum_i a_i \otimes b_i, \quad (\text{G6})$$

then

$$\Delta(\tilde{X}) = \sum_i \tilde{b}_i \otimes \tilde{a}_i. \quad (\text{G7})$$

We are interested in three unitary representations: the positive discrete series, the negative discrete series, and the principal series.

a. Discrete series

The positive discrete series π_k^+ is labeled by $k > 0$. The representation space is $\ell^2(\mathbb{Z}_{\geq 0})$ with orthonormal basis $\{e_n\}_{n \in \mathbb{Z}_{\geq 0}}$. The action is given by

$$\begin{aligned} \pi_k^+(K)e_n &= q^{\frac{k+n}{2}}e_n, & \pi_k^+(K^{-1})e_n &= q^{-\frac{k+n}{2}}e_n \\ (q^{-1/2} - q^{1/2})\pi_k^+(E)e_n &= q^{-\frac{1}{4} - \frac{k}{2} - \frac{n}{2}}\sqrt{(1 - q^{n+1})(1 - q^{2k+n})}e_{n+1} \\ (q^{-1/2} - q^{1/2})\pi_k^+(F)e_n &= -q^{\frac{1}{4} - \frac{k}{2} - \frac{n}{2}}\sqrt{(1 - q^n)(1 - q^{2k+n-1})}e_{n-1} \\ (q^{-1/2} - q^{1/2})^2\pi_k^+(\Omega)e_n &= (q^{k-\frac{1}{2}} + q^{\frac{1}{2}-k} - 2)e_n. \end{aligned} \quad (\text{G8})$$

The positive discrete series representation becomes the negative discrete series representation under the automorphism in (G5). That is, if X is a generator of the algebra, then it is represented by $\pi_k^-(X)$ in the negative discrete series representation, where

$$\pi_k^-(X) = \pi_k^+(\tilde{X}). \quad (\text{G9})$$

b. Principal series

The principal unitary series representations $\pi_{\rho,\epsilon}^P$ are labeled by $0 \leq \rho \leq \frac{\pi}{|\log q|}$ and $\epsilon \in [0, 1)$, where $(\rho, \epsilon) \neq (0, \frac{1}{2})$. At times we may find it convenient to allow ϵ to take values outside of this range with the understanding that ϵ is defined modulo unit shifts. The representation space is $\ell^2(\mathbb{Z})$ with orthonormal basis $\{e_n\}_{n \in \mathbb{Z}}$. The action is given by

$$\begin{aligned} \pi_{\rho,\epsilon}^P(K)e_n &= q^{\frac{n+\epsilon}{2}}e_n, & \pi_{\rho,\epsilon}^P(K^{-1})e_n &= q^{-\frac{n+\epsilon}{2}}e_n, \\ (q^{-1/2} - q^{1/2})\pi_{\rho,\epsilon}^P(E)e_n &= q^{-\frac{1}{4} - \frac{n}{2} - \frac{\epsilon}{2}}\sqrt{(1 - q^{n+\epsilon+i\rho+\frac{1}{2}})(1 - q^{n+\epsilon-i\rho+\frac{1}{2}})}e_{n+1} \\ (q^{-1/2} - q^{1/2})\pi_{\rho,\epsilon}^P(F)e_n &= -q^{\frac{1}{4} - \frac{n}{2} - \frac{\epsilon}{2}}\sqrt{(1 - q^{n+\epsilon+i\rho-\frac{1}{2}})(1 - q^{n+\epsilon-i\rho-\frac{1}{2}})}e_{n-1} \\ (q^{-1/2} - q^{1/2})^2\pi_{\rho,\epsilon}^P(\Omega)e_n &= (q^{i\rho} + q^{-i\rho} - 2)e_n \end{aligned} \quad (\text{G10})$$

If we define the automorphism in (G5) to act on e_n , ρ , and ϵ as follows:

$$\tilde{\rho} = \rho, \quad \tilde{\epsilon} = -\epsilon, \quad \tilde{e}_n = (-1)^n e_{-n}, \quad (\text{G11})$$

then (G10) is invariant under the automorphism.

2. Clebsch-Gordan coefficients

We now consider the Clebsch-Gordan coefficients that we will need.

a. $\pi^+ \otimes \pi^+$

Consider the tensor product representation $\pi_{k_1}^+ \otimes \pi_{k_2}^+$, which is spanned by $e_{n_1} \otimes e_{n_2}$ for $n_1, n_2 \in \mathbb{Z}_{\geq 0}$. A state of definite K weight is given by

$$|p\rangle \equiv \sum_{n=0}^p v_n^p e_n \otimes e_{p-n}, \quad (\text{G12})$$

where $p \in \mathbb{Z}_{\geq 0}$, and we wish to choose v_n^p such that $|p\rangle$ is an eigenvector of the casimir Ω . The correct choice is to take

$$\begin{aligned}
v_n^p(x, k_1, k_2) = & -(q^{\frac{1}{2}-k_1-k_2})^n q^{pk_1} \sqrt{\frac{(q^{2k_2}; q)_p}{(q^{2k_1+2k_2}; q)_p}} \sqrt{\frac{(q^{2k_1}, q^{-p}; q)_n}{(q, q^{1-p-2k_2}; q)_n}} \\
& \times \sqrt{\frac{(q^{2k_1}, q^{2k_1+2k_2-1}, q^{-p}; q)_x (q^{px-\frac{x(x-1)}{2}})(1-q^{2k_1+2k_2-1+2x})}{(q, q^{2k_1+2k_2+p}, q^{2k_2}; q)_x (1-q^{2k_1+2k_2-1})(-q^{2k_1})^x}} \\
& \times R_n(\mu(x); q^{2k_1-1}, q^{2k_2-1}, p|q),
\end{aligned} \tag{G13}$$

where $x \in \{0, 1, \dots, p\}$, and R_n is a dual q -Hahn polynomial, defined on page 450 of [81]. Note that the quantities inside the square roots are positive. With v_n^p chosen as in (G13), $|p\rangle$ in (G12) is a vector in the $\pi_{k_3}^+$ representation, where $k_3 = k_1 + k_2 + x$.

The Clebsch-Gordan coefficients $v_n^p(x, k_1, k_2)$ obey the following orthogonality relations:

$$\sum_{x=0}^p v_m^p(x, k_1, k_2) v_n^p(x, k_1, k_2) = \delta_{nm}, \quad n, m \in \{0, 1, \dots, p\} \tag{G14}$$

$$\begin{aligned}
\sum_{n=0}^p v_n^p(x_1, k_1, k_2) v_n^p(x_2, k_1, k_2) &= \delta_{x_1 x_2}, \\
x_1, x_2 \in \{0, 1, \dots, p\},
\end{aligned} \tag{G15}$$

and it follows that the decomposition of $\pi_{k_1}^+ \otimes \pi_{k_2}^+$ is given by

$$\pi_{k_1}^+ \otimes \pi_{k_2}^+ = \bigoplus_{j=0}^{\infty} \pi_{k_1+k_2+j}^+. \tag{G16}$$

b. $\pi^+ \otimes \pi^p$

Consider the tensor product representation $\pi_{\Delta}^+ \otimes \pi_{\rho, \epsilon}^p$, whose decomposition into irreps contains both principal series and positive discrete series representations [74]. We are only interested in the principal series representations. The space $\pi_{\Delta}^+ \otimes \pi_{\rho, \epsilon}^p$ is spanned by $e_{n_1} \otimes e_{n_2}$ for $n_1 \in \mathbb{Z}_{\geq 0}$ and $n_2 \in \mathbb{Z}$. A state of definite K weight is given by

$$|\tau, p\rangle = \sum_{n \geq 0} f_n^p(\tau, \rho, \epsilon, \Delta) e_n \otimes e_{p-n}, \tag{G17}$$

where $p \in \mathbb{Z}$, and we wish to define $f_n^p(\tau, \rho, \epsilon, \Delta)$ such that $|\tau, p\rangle$ is an eigenvector of Ω that belongs to the principal series representation $\pi_{\tau, \Delta+\epsilon}^p$. The Clebsch-Gordan coefficient is given by

$$\begin{aligned}
f_n^p(\tau, \rho, \epsilon, \Delta) = & (-1)^n \sqrt{\frac{|\log q|}{2\pi}} \frac{1}{(q; q)_n (q^{2\Delta}; q)_n (q^{\frac{1}{2}-p-\epsilon \pm i\rho}; q)_n} \frac{(q^{\pm 2i\tau}, q, q^{\frac{1}{2}-p-\epsilon \pm i\rho}, q^{2\Delta}; q)_{\infty}}{(q^{\frac{1}{2}-p-\Delta-\epsilon \pm i\tau}, q^{\Delta \pm i\rho \pm i\tau}; q)_{\infty}} \\
& \times p_n \left(\frac{q^{i\tau} + q^{-i\tau}}{2}, q^{\frac{1}{2}-p-\Delta-\epsilon}, q^{\Delta \pm i\rho} | q \right),
\end{aligned} \tag{G18}$$

where p_n is a continuous dual q -Hahn polynomial, defined on page 429 of [81]. Because $|\tau_1, p\rangle$ and $|\tau_2, p\rangle$ must be orthogonal for $\tau_1 \neq \tau_2$, we have that

$$\sum_{n \geq 0} f_n^p(\tau_1, \rho, \epsilon, \Delta) f_n^p(\tau_2, \rho, \epsilon, \Delta) = \delta(\tau_1 - \tau_2), \tag{G19}$$

And the normalization of (G18) is chosen to make (G19) consistent with the orthogonality relation of the continuous dual q -Hahn polynomials provided in [74].

c. $\pi^p \otimes \pi^-$

The decomposition of $\pi^p \otimes \pi^-$ may be deduced from the decomposition of $\pi^+ \otimes \pi^p$ and the automorphism (G5). In particular, $\pi_{\rho, \epsilon}^p \otimes \pi_{\Delta}^-$ is spanned by $e_{n_1} \otimes e_{n_2}$ with $n_1 \in \mathbb{Z}$ and $n_2 \in \mathbb{Z}_{\geq 0}$. If we define

$$|\tau, p\rangle = \sum_{n \geq 0} (-1)^n f_n^{-p}(\tau, \rho, -\epsilon, \Delta) e_{p+n} \otimes e_n, \tag{G20}$$

then $|\tau, p\rangle$ belongs to a $\pi_{\tau, \epsilon-\Delta}^p$ representation.

d. $\pi^- \otimes \pi^-$

The decomposition of $\pi^- \otimes \pi^-$ may be deduced from the decomposition of $\pi^+ \otimes \pi^+$ and the automorphism (G5). In particular, $\pi_{k_1}^- \otimes \pi_{k_2}^-$ is spanned by $e_{n_1} \otimes e_{n_2}$ for $n_1, n_2 \in \mathbb{Z}_{\geq 0}$. If we define

$$|p\rangle = \sum_{n=0}^p v_{p-n}^p(x, k_2, k_1) e_n \otimes e_{p-n}, \quad (\text{G21})$$

then $|p\rangle$ is a vector in the $\pi_{k_1+k_2+x}^-$ representation.

3. $6j$ symbols

In this section, we study three different $6j$ symbols that may be built out of the above Clebsch-Gordan coefficients. We now briefly review the general definition of a $6j$ symbol. Let j_i denote a unitary irrep of an algebra, where $i \in \{1, 2, \dots, 6\}$. To define a $6j$ symbol, consider the triple tensor product $j_1 \otimes j_2 \otimes j_3$. One way to decompose this into irreps is to first take the tensor product of $j_1 \otimes j_2$ and then tensor the result with j_3 . Let us then project the result onto the space of states transforming in a j_4 representation (call this space \mathcal{S}_4). The space \mathcal{S}_4 may be written as a direct sum of orthogonal subspaces, where each subspace corresponds to a representation that appears in the $j_1 \otimes j_2$ tensor product (let j_5 refer to such a representation). Thus an orthonormal basis vector of \mathcal{S}_4 may be written as $|j_4 m_4; j_5\rangle_{12}$, where m_4 labels a basis vector in the j_4 irrep. Alternatively, we can repeat the entire process in the other channel by taking the $j_2 \otimes j_3$ tensor product first. Let j_6 refer to a representation in the tensor product of $j_2 \otimes j_3$. Thus a different orthonormal basis of \mathcal{S}_4 is given by vectors of the form $|j_4 m_4; j_6\rangle_{23}$. The $6j$ symbol is defined to be the overlap

$$U_{j_5 j_6}^{j_1 j_2 j_3 j_4} \equiv {}_{12}\langle j_4 m_4; j_5 | j_4 m_4; j_6 \rangle_{23}, \quad (\text{G22})$$

which does not depend on m_4 . The $6j$ symbol is represented graphically in Fig. 30. Equation (G22) defines a unitary matrix because it is an overlap of orthonormal basis vectors. In particular,

$$\begin{aligned} & \sum_{j_6} U_{j_5 j_6}^{j_1 j_2 j_3 j_4} (U_{j_5 j_6}^{j_1 j_2 j_3 j_4})^* \\ &= \delta_{j_5 \tilde{j}_5}, \sum_{j_5} (U_{j_5 j_6}^{j_1 j_2 j_3 j_4})^* U_{j_5 \tilde{j}_6}^{j_1 j_2 j_3 j_4} = \delta_{j_6 \tilde{j}_6}. \end{aligned} \quad (\text{G23})$$

The $6j$ symbols that we discuss are real.

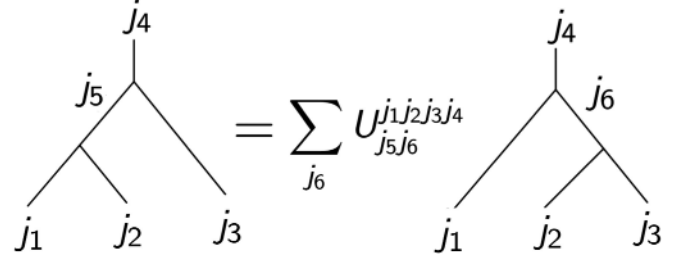


FIG. 30. A diagrammatic illustration of the definition of the $6j$ symbol. Each junction of three lines represents a Clebsch-Gordan coefficient.

In general, $6j$ symbols may be constructed from Clebsch-Gordan coefficients. Let us define the Clebsch-Gordan coefficients $C_{m, m_1, m_2}^{j, j_1, j_2}$ such that

$$|jm\rangle = \sum_{m_1, m_2} C_{m, m_1, m_2}^{j, j_1, j_2} |j_1 m_1\rangle \otimes |j_2 m_2\rangle. \quad (\text{G24})$$

The coefficients obey an orthogonality relation:

$$\sum_{m_1, m_2} C_{m, m_1, m_2}^{j, j_1, j_2} \bar{C}_{\tilde{m}, \tilde{m}_1, \tilde{m}_2}^{\tilde{j}, \tilde{j}_1, \tilde{j}_2} = \delta_{m, \tilde{m}} \delta_{j, \tilde{j}}. \quad (\text{G25})$$

The Clebsch-Gordan coefficients and $6j$ symbols are related as follows:

$$\begin{aligned} & \sum_{m_5} C_{m_4, m_5, m_3}^{j_4, j_5, j_3} C_{m_5, m_1, m_2}^{j_5, j_1, j_2} \\ &= \sum_{j_6, m_6} U_{j_5 j_6}^{j_1 j_2 j_3 j_4} C_{m_4, m_1, m_6}^{j_4, j_1, j_6} C_{m_6, m_2, m_3}^{j_6, j_2, j_3}. \end{aligned} \quad (\text{G26})$$

Using (G25), we have

$$\begin{aligned} & \sum_{m_5, m_2, m_3} C_{m_4, m_5, m_3}^{j_4, j_5, j_3} C_{m_5, m_1, m_2}^{j_5, j_1, j_2} C_{m_6, m_2, m_3}^{j_6, j_2, j_3} \\ &= U_{j_5 j_6}^{j_1 j_2 j_3 j_4} C_{m_4, m_1, m_6}^{j_4, j_1, j_6}. \end{aligned} \quad (\text{G27})$$

We now derive another identity that we will use later. We will start with (G26) and multiply by three new Clebsch-Gordan coefficients on each side to obtain

$$\begin{aligned} & \sum_{m_5, m_1, m_2, m_3, m_8, m_9} C_{m_4, m_5, m_3}^{j_4, j_5, j_3} C_{m_5, m_1, m_2}^{j_5, j_1, j_2} C_{m_8, m_7, m_1}^{j_8, j_7, j_1} C_{m_9, m_8, m_2}^{j_9, j_8, j_2} C_{m_{10}, m_9, m_3}^{j_{10}, j_9, j_3} \\ &= \sum_{j_6, m_6, m_1, m_2, m_3, m_8, m_9} U_{j_5 j_6}^{j_1 j_2 j_3 j_4} C_{m_4, m_1, m_6}^{j_4, j_1, j_6} C_{m_6, m_2, m_3}^{j_6, j_2, j_3} C_{m_8, m_7, m_1}^{j_8, j_7, j_1} C_{m_9, m_8, m_2}^{j_9, j_8, j_2} C_{m_{10}, m_9, m_3}^{j_{10}, j_9, j_3}. \end{aligned} \quad (\text{G28})$$

Next, we use (G27) to simplify the m_1, m_2, m_8 sum on the left and the m_2, m_3, m_9 sum on the right. The result is

$$\begin{aligned} & \sum_{m_5, m_3, m_9} U_{j_8 j_5}^{j_7 j_1 j_2 j_9} C_{m_{10}, m_9, m_3}^{j_{10}, j_9, j_3} C_{m_9, m_7, m_5}^{j_9, j_7, j_5} C_{m_4, m_5, m_3}^{j_4, j_5, j_3} \\ &= \sum_{j_6, m_6, m_1, m_8} U_{j_5 j_6}^{j_1 j_2 j_3 j_4} U_{j_9 j_6}^{j_8 j_2 j_3 j_{10}} C_{m_{10}, m_8, m_6}^{j_{10}, j_8, j_6} C_{m_8, m_7, m_1}^{j_8, j_7, j_1} C_{m_4, m_1, m_6}^{j_4, j_1, j_6}. \end{aligned} \quad (\text{G29})$$

Next, we use (G27) again to simplify the m_5, m_3, m_9 sum on the left and the m_6, m_1, m_8 sum on the right. The result is

$$\begin{aligned} & U_{j_8 j_5}^{j_7 j_1 j_2 j_9} U_{j_9 j_4}^{j_7 j_5 j_3 j_{10}} C_{m_{10}, m_7, m_4}^{j_{10}, j_7, j_4} \\ &= \sum_{j_6} U_{j_5 j_6}^{j_1 j_2 j_3 j_4} U_{j_9 j_6}^{j_8 j_2 j_3 j_{10}} U_{j_8 j_4}^{j_7 j_1 j_6 j_{10}} C_{m_{10}, m_7, m_4}^{j_{10}, j_7, j_4} \end{aligned} \quad (\text{G30})$$

which becomes

$$U_{j_8 j_5}^{j_7 j_1 j_2 j_9} U_{j_9 j_4}^{j_7 j_5 j_3 j_{10}} = \sum_{j_6} U_{j_5 j_6}^{j_1 j_2 j_3 j_4} U_{j_9 j_6}^{j_8 j_2 j_3 j_{10}} U_{j_8 j_4}^{j_7 j_1 j_6 j_{10}}. \quad (\text{G31})$$

a. $\pi^+ \otimes \pi^+ \otimes \pi^+$

Consider the triple tensor product $\pi_{k_1}^+ \otimes \pi_{k_2}^+ \otimes \pi_{k_3}^+$. Our convention for the $6j$ symbol is given in Fig. 30 with the following identifications:

$$\begin{aligned} j_1 &= \pi_{k_1}^+, \quad j_2 = \pi_{k_2}^+, \quad j_3 = \pi_{k_3}^+, \\ j_4 &= \pi_{k_1+k_2+k_3+N}^+, \quad j_5 = \pi_{k_1+k_2+x_{12}}^+, \quad j_6 = \pi_{k_2+k_3+x_{23}}^+, \\ U_{j_5 j_6}^{j_1 j_2 j_3 j_4} &= S_q^{k_1, k_2, k_3, N}(x_{12}, x_{23}), \quad \sum_{j_6} = \sum_{x_{23}=0}^N, \end{aligned} \quad (\text{G32})$$

where the $6j$ symbol $S_q^{k_1, k_2, k_3, N}(x_{12}, x_{23})$ is defined as follows:

$$\begin{aligned} S_q^{k_1, k_2, k_3, N}(x_{12}, x_{23}) &= q^{k_2(N-x_{12}-x_{23})} \begin{bmatrix} N \\ x_{23} \end{bmatrix}_q \times {}_4\phi_3 \left(\begin{matrix} q^{-1+2k_1+2k_2+x_{12}}, q^{-1+2k_2+2k_3+x_{23}}, q^{-x_{12}}, q^{-x_{23}} \\ q^{2k_2}, q^{-1+2k_1+2k_2+2k_3+N}, q^{-N} \end{matrix}; q, q \right) \\ &\times \frac{(q^{-1+2k_1+2k_2+2k_3+N}; q)_{x_{23}} (q^{2k_3}; q)_{N-x_{12}} (q^{2k_2}; q)_{x_{12}}}{\sqrt{(q, q^{2k_1}, q^{2k_2}, q^{2k_1+2k_2+x_{12}-1}; q)_{x_{12}} (q, q^{2k_3}, q^{2(k_1+k_2+x_{12})}, q^{2k_1+2k_2+2k_3+N+x_{12}-1}; q)_{N-x_{12}}}} \\ &\times \frac{\sqrt{(q, q^{2k_2}; q)_{x_{23}} (q, q^{2k_1}, q^{2k_1+2k_2+2k_3+N+x_{23}-1}; q)_{N-x_{23}}}}{\sqrt{(q^{2(k_2+k_3+x_{23})}; q)_{N-x_{23}} (q^{2k_3}, q^{2k_2+2k_3+x_{23}-1}; q)_{x_{23}}}}, \end{aligned} \quad (\text{G33})$$

and the q -binomial coefficient is defined by

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{(q; q)_n}{(q; q)_k (q; q)_{n-k}}. \quad (\text{G34})$$

In particular, we have that for $n_1, n_2, n_3 \in \mathbb{Z}_{\geq 0}$ and $n_1 + n_2 + n_3 \geq N$,

$$\begin{aligned} & v_{n_1}^{n_1+n_2}(x_{12}, k_1, k_2) v_{n_1+n_2-x_{12}}^{n_1+n_2-x_{12}+n_3}(N-x_{12}, k_1+k_2+x_{12}, k_3) \\ &= \sum_{x_{23}=0}^{\min(N, n_2+n_3)} S_q^{k_1, k_2, k_3, N}(x_{12}, x_{23}) v_{n_2}^{n_2+n_3}(x_{23}, k_2, k_3) \\ &\times v_{n_1}^{n_1+n_2+n_3-x_{23}}(N-x_{23}, k_1, k_2+k_3+x_{23}). \end{aligned} \quad (\text{G35})$$

Each Clebsch-Gordan coefficient above corresponds to a junction in Fig. 30. Also, (G23) becomes

$$\sum_{x_{23}=0}^N S_q^{k_1, k_2, k_3, N}(x_{12}, x_{23}) S_q^{k_3, k_2, k_1, N}(x_{23}, \tilde{x}_{12}) = \delta_{x_{12}, \tilde{x}_{12}}. \quad (\text{G36})$$

b. $\pi^- \otimes \pi^- \otimes \pi^-$

The $\pi^- \otimes \pi^- \otimes \pi^-$ $6j$ symbol is related to the $\pi^+ \otimes \pi^+ \otimes \pi^+$ $6j$ symbol via the automorphism (G5). We may make the following substitutions in Fig. 30,

$$\begin{aligned} j_1 &= \pi_{k_1}^-, \quad j_2 = \pi_{k_2}^-, \quad j_3 = \pi_{k_3}^-, \\ j_4 &= \pi_{k_1+k_2+k_3+N}^-, \quad j_5 = \pi_{k_1+k_2+x_{12}}^-, \quad j_6 = \pi_{k_2+k_3+x_{23}}^-, \\ U_{j_5 j_6}^{j_1 j_2 j_3 j_4} &= S_q^{k_3, k_2, k_1, N}(x_{23}, x_{12}), \quad \sum_{j_6} = \sum_{x_{23}=0}^N. \end{aligned} \quad (\text{G37})$$

c. $\pi^+ \otimes \pi^+ \otimes \pi^P$

Consider the triple tensor product $\pi_{\Delta_1}^+ \otimes \pi_{\Delta_2}^+ \otimes \pi_{s_a, \epsilon}^P$. Our convention for the $6j$ symbol is given in Fig. 30 with the following identifications,

$$j_1 = \pi_{\Delta_1}^+, \quad j_2 = \pi_{\Delta_2}^+, \quad j_3 = \pi_{s_a, \epsilon}^P, \quad j_4 = \pi_{s_c, \Delta_1 + \Delta_2 + \epsilon}^P, \quad j_5 = \pi_{\Delta_1 + \Delta_2 + x}^+, \quad j_6 = \pi_{s, \Delta_2 + \epsilon}^P$$

$$U_{j_5 j_6}^{j_1 j_2 j_3 j_4} = (-1)^{x+1} \sqrt{\rho_q(s)} P_x^{\Delta_2, \Delta_1}(s; s_a, s_c | q), \quad \sum_{j_6} = \int_0^{\frac{\pi}{|\log q|}} ds. \quad (\text{G38})$$

While the tensor product of $\pi_{\Delta_2}^+$ and $\pi_{s_a, \epsilon}^P$ does contain positive discrete series representations, the label j_6 only runs over principal series representations because the tensor product of a positive discrete series representation and $\pi_{\Delta_1}^+$ cannot produce $\pi_{s_c, \Delta_1 + \Delta_2 + \epsilon}^P$.

To be precise, the following identity holds:

$$\sum_{x=0}^{n_1+n_2} v_{n_1}^{n_1+n_2}(x, \Delta_1, \Delta_2) f_{n_1+n_2-x}^{n_1+n_2+n_3-x}(s_c, s_a, \epsilon, \Delta_1 + \Delta_2 + x) (-1)^{1+x} \sqrt{\rho_q(s)} P_x^{\Delta_2, \Delta_1}(s; s_a, s_c | q)$$

$$= f_{n_2}^{n_2+n_3}(s, s_a, \epsilon, \Delta_2) f_{n_1}^{n_1+n_2+n_3}(s_c, s, \epsilon + \Delta_2, \Delta_1), \quad n_1, n_2 \in \mathbb{Z}_{\geq 0}, \quad n_3 \in \mathbb{Z}. \quad (\text{G39})$$

Using (A63), this becomes

$$v_{n_1}^{n_1+n_2}(x, \Delta_1, \Delta_2) f_{n_1+n_2-x}^{n_1+n_2+n_3-x}(s_c, s_a, \epsilon, \Delta_1 + \Delta_2 + x)$$

$$= \int_0^{\frac{\pi}{|\log q|}} ds (-1)^{1+x} \sqrt{\rho_q(s)} P_x^{\Delta_2, \Delta_1}(s; s_a, s_c | q) f_{n_2}^{n_2+n_3}(s, s_a, \epsilon, \Delta_2) f_{n_1}^{n_1+n_2+n_3}(s_c, s, \epsilon + \Delta_2, \Delta_1),$$

$$0 \leq x \leq n_1 + n_2. \quad (\text{G40})$$

Each Clebsch-Gordan coefficient in (G40) is represented by a junction in 30. Unitarity of the $6j$ symbol also implies the completeness relation⁸³:

$$\sum_{n=0}^{\infty} P_n^{\Delta_2, \Delta_1}(s_b; s_c, s_a | q) P_n^{\Delta_2, \Delta_1}(s_d; s_c, s_a | q) = \frac{\delta(s_b - s_d)}{\rho_q(s_b)}. \quad (\text{G41})$$

Using the graphical notation in (A8), the completeness relation is given in (A31).

$$\mathbf{d.} \quad \pi^P \otimes \pi^- \otimes \pi^-$$

The $6j$ symbol for $\pi_{s_a, \epsilon}^P \otimes \pi_{\Delta_1}^- \otimes \pi_{\Delta_2}^-$ may be inferred from the $6j$ symbol for $\pi_{\Delta_2}^+ \otimes \pi_{\Delta_1}^+ \otimes \pi_{s_a, -\epsilon}^P$. In particular, it follows from (G39) that

$$(-1)^{n_2} f_{n_2}^{n_2-n_1}(s, s_a, -\epsilon, \Delta_1) (-1)^{n_3} f_{n_3}^{n_2+n_3-n_1}(s_c, s, -\epsilon + \Delta_1, \Delta_2)$$

$$= \sum_{x=0}^{n_2+n_3} v_{n_3}^{n_2+n_3}(x, \Delta_2, \Delta_1) (-1)^{n_2+n_3-x} f_{n_2+n_3-x}^{n_2+n_3-x-n_1}(s_c, s_a, -\epsilon, k_1 + k_2 + x) [-\sqrt{\rho_q(s)} P_x^{\Delta_1, \Delta_2}(s; s_a, s_c | q)],$$

$$n_2, n_3 \in \mathbb{Z}_{\geq 0}, \quad n_1 \in \mathbb{Z}. \quad (\text{G42})$$

Thus, the $6j$ symbol is represented by Fig. 30 with the following identifications,

$$j_1 = \pi_{s_a, \epsilon}^P, \quad j_2 = \pi_{\Delta_1}^-, \quad j_3 = \pi_{\Delta_2}^-, \quad j_4 = \pi_{s_c, \epsilon - \Delta_1 - \Delta_2}^P, \quad j_5 = \pi_{s, \epsilon - \Delta_1}^P, \quad j_6 = \pi_{\Delta_1 + \Delta_2 + x}^-$$

$$U_{j_5 j_6}^{j_1 j_2 j_3 j_4} = -\sqrt{\rho_q(s)} P_x^{\Delta_1, \Delta_2}(s; s_a, s_c | q), \quad \sum_{j_6} = \sum_{x \geq 0}. \quad (\text{G43})$$

⁸³This is equivalent to a special case ($\alpha = 0$) of (4.1.11) in [82]. See also Eqs. (3.8) and (10.1) in [83].

$$\pi_{s_c, \epsilon + \Delta_1 - \Delta_2}^P \begin{array}{c} \pi_{s_b, \epsilon + \Delta_1}^P \\ \pi_{\Delta_1}^+ \quad \pi_{s_a, \epsilon}^P \quad \pi_{\Delta_2}^- \end{array} = \int_0^{\frac{\pi}{|\log q|}} ds_d R_{s_a, s_c, \epsilon}^{\Delta_1 \Delta_2}(s_b, s_d) \begin{array}{c} \pi_{s_c, \epsilon + \Delta_1 - \Delta_2}^P \\ \pi_{s_d, \epsilon - \Delta_2}^P \\ \pi_{\Delta_1}^+ \quad \pi_{s_a, \epsilon}^P \quad \pi_{\Delta_2}^- \end{array} + \dots$$

FIG. 31. A diagrammatic representation of the $6j$ symbol in (G44). On the right-hand side, we have only depicted π^P representations running through the channel. The \dots contain contributions from negative discrete series representations that also run in the channel.

$$\begin{aligned} \begin{array}{c} j_7 \\ j_5 \quad j_6 \quad j_7 \\ j_1 \quad j_2 \quad j_3 \quad j_4 \end{array} &= \sum_{j_8} U_{j_5 j_8}^{j_5 j_3 j_4 j_7} \begin{array}{c} j_7 \\ j_5 \quad j_8 \quad j_7 \\ j_1 \quad j_2 \quad j_3 \quad j_4 \end{array} = \sum_{j_8 j_9} U_{j_5 j_8}^{j_5 j_3 j_4 j_7} U_{j_8 j_9}^{j_8 j_2 j_3 j_7} \begin{array}{c} j_7 \\ j_9 \quad j_8 \quad j_7 \\ j_1 \quad j_2 \quad j_3 \quad j_4 \end{array} \\ \begin{array}{c} j_7 \\ j_5 \quad j_6 \quad j_7 \\ j_1 \quad j_2 \quad j_3 \quad j_4 \end{array} &= \sum_{j_{10}} U_{j_5 j_{10}}^{j_5 j_2 j_3 j_6} \begin{array}{c} j_7 \\ j_6 \quad j_{10} \quad j_7 \\ j_1 \quad j_2 \quad j_3 \quad j_4 \end{array} = \sum_{j_{10} j_{11}} U_{j_5 j_{10}}^{j_5 j_2 j_3 j_6} U_{j_{10} j_{11}}^{j_{10} j_1 j_2 j_7} \begin{array}{c} j_7 \\ j_{11} \quad j_{10} \quad j_7 \\ j_1 \quad j_2 \quad j_3 \quad j_4 \end{array} \\ &= \sum_{j_{10} j_{11} j_{12}} U_{j_5 j_{10}}^{j_5 j_2 j_3 j_6} U_{j_{10} j_{11}}^{j_{10} j_1 j_2 j_7} U_{j_{11} j_{12}}^{j_{11} j_3 j_4 j_7} \begin{array}{c} j_7 \\ j_{12} \quad j_{11} \quad j_7 \\ j_1 \quad j_2 \quad j_3 \quad j_4 \end{array} \end{aligned}$$

FIG. 32. A graphical illustration of the pentagon identity. We may exploit the fact that there are various ways to take a tensor product of four representations to derive a nontrivial relation involving $6j$ symbols.

e. $\pi^+ \otimes \pi^P \otimes \pi^-$

Consider the triple tensor product $\pi_{\Delta_1}^+ \otimes \pi_{s_a, \epsilon}^P \otimes \pi_{\Delta_2}^-$ projected onto the subspace of states that transform in the $\pi_{s_c, \epsilon + \Delta_1 - \Delta_2}^P$ representation. In the two channels, we will project onto the $\pi_{s_b, \Delta_1 + \epsilon}^P$ and $\pi_{s_d, \epsilon - \Delta_2}^P$ channels. The $6j$ symbol is given by

$$R_{s_a, s_c, \epsilon}^{\Delta_1, \Delta_2}(s_b, s_d) = \sqrt{\frac{(q^{1/2 \pm \epsilon \pm i s_a}, q^{\frac{1}{2} \pm (\Delta_1 - \Delta_2 + \epsilon) \pm i s_c}; q)_\infty}{(q^{\frac{1}{2} \pm (\epsilon + \Delta_1) \pm i s_b}, q^{\frac{1}{2} \pm (\Delta_2 - \epsilon) \pm i s_d}; q)_\infty}} \times \begin{Bmatrix} \Delta_2 & s_b & s_c \\ \Delta_1 & s_d & s_a \end{Bmatrix}_q \sqrt{\rho_q(s_b) \rho_q(s_d)}, \quad (\text{G44})$$

and is diagrammatically represented in Fig. 31. Note that the tensor product of $\pi^P \otimes \pi^-$ contains both π^P and π^- representations, and tensoring π^+ with either of these produces a π^P representation. Thus, to derive an orthogonality identity analogous to (G23), we must consider both π^P and π^- representations in the channel on the right-hand side of Fig. 31. However, we have only explicitly considered the π^P representations. The discrete series representations will not play a role in our derivation of the pentagon identity, which we carry out in the next section.

4. Pentagon identity

We now derive a pentagon identity⁸⁴ for the tensor product $\pi^+ \otimes \pi^P \otimes \pi^- \otimes \pi^-$. The pentagon identity may be derived by tensoring four representations together and applying the relation in Fig. 30 in different ways, as indicated in Fig. 32. Equality of the two lines in Fig. 32 indicates that the $6j$ symbols must obey

$$U_{j_6 j_8}^{j_5 j_3 j_4 j_7} U_{j_5 j_9}^{j_1 j_2 j_8 j_7} = \sum_{j_{10}} U_{j_5 j_{10}}^{j_1 j_2 j_3 j_6} U_{j_6 j_9}^{j_1 j_{10} j_4 j_7} U_{j_{10} j_8}^{j_2 j_3 j_4 j_9}. \quad (\text{G45})$$

To obtain the identity of our interest, we make the substitutions

$$\begin{aligned} j_1 &= \pi_{\Delta_1}^+, & j_2 &= \pi_{s_a, \epsilon}^P, & j_3 &= \pi_{\Delta_2}^-, & j_4 &= \pi_{\Delta_3}^-, \\ j_5 &= \pi_{s_b, \epsilon + \Delta_1}^P, & j_6 &= \pi_{s_c, \epsilon + \Delta_1 - \Delta_2}^P, & j_7 &= \pi_{s_d, \epsilon + \Delta_1 - \Delta_2 - \Delta_3}^P, \\ j_8 &= \pi_{\Delta_2 + \Delta_3 + x}^-, & j_9 &= \pi_{s_e, \epsilon - \Delta_2 - \Delta_3}^P, & j_{10} &= \pi_{s_f, \epsilon - \Delta_2}^P, \end{aligned} \quad (\text{G46})$$

so the $6j$ symbols are

⁸⁴See [84] for an explanation of the pentagon identity, which is also called the Biedenharn-Elliott identity.

$$\begin{aligned}
U_{j_6 j_8}^{j_5 j_3 j_4 j_7} &= -\sqrt{\rho_q(s_c)} P_x^{\Delta_2 \Delta_3}(s_c; s_b, s_d | q), & U_{j_5 j_9}^{j_1 j_2 j_8 j_7} &= R_{s_a, s_d, \epsilon}^{\Delta_1, \Delta_2 + \Delta_3 + x}(s_b, s_e), \\
U_{j_5 j_{10}}^{j_1 j_2 j_3 j_6} &= R_{s_a, s_c, \epsilon}^{\Delta_1 \Delta_2}(s_b, s_f), & U_{j_6 j_9}^{j_1 j_{10} j_4 j_7} &= R_{s_f, s_d, \epsilon - \Delta_2}^{\Delta_1 \Delta_3}(s_c, s_e), & U_{j_{10} j_8}^{j_2 j_3 j_4 j_9} &= -\sqrt{\rho_q(s_f)} P_x^{\Delta_2 \Delta_3}(s_f; s_a, s_e | q)
\end{aligned} \quad (\text{G47})$$

and the pentagon identity becomes

$$\begin{aligned}
&\sqrt{\rho_q(s_c)} P_x^{\Delta_2 \Delta_3}(s_c; s_b, s_d | q) R_{s_a, s_d, \epsilon}^{\Delta_1, \Delta_2 + \Delta_3 + x}(s_b, s_e) \\
&= \int_0^{\frac{\pi}{|\log q|}} ds_f R_{s_a, s_c, \epsilon}^{\Delta_1 \Delta_2}(s_b, s_f) R_{s_f, s_d, \epsilon - \Delta_2}^{\Delta_1 \Delta_3}(s_c, s_e) \sqrt{\rho_q(s_f)} P_x^{\Delta_2 \Delta_3}(s_f; s_a, s_e | q),
\end{aligned} \quad (\text{G48})$$

and using (G44) and simplifying, this becomes

$$\begin{aligned}
&\sqrt{\frac{(q^{\frac{1}{2} \pm (\Delta_2 + \Delta_3 - \epsilon) \pm i s_e}; q)_\infty}{(q^{\frac{1}{2} \pm (\Delta_2 + \Delta_3 - \epsilon + x) \pm i s_e}; q)_\infty}} \sqrt{\frac{(q^{\frac{1}{2} \pm (\Delta_1 - \Delta_2 - \Delta_3 + \epsilon - x) \pm i s_d}; q)_\infty}{(q^{\frac{1}{2} \pm (\Delta_1 - \Delta_2 - \Delta_3 + \epsilon) \pm i s_d}; q)_\infty}} P_x^{\Delta_2 \Delta_3}(s_c; s_b, s_d | q) \left\{ \begin{matrix} (\Delta_2 + \Delta_3 + x) & s_b & s_d \\ \Delta_1 & s_e & s_a \end{matrix} \right\}_q \\
&= \int_0^{\frac{\pi}{|\log q|}} ds_f \rho_q(s_f) \left\{ \begin{matrix} \Delta_2 & s_b & s_c \\ \Delta_1 & s_f & s_a \end{matrix} \right\}_q \left\{ \begin{matrix} \Delta_3 & s_c & s_d \\ \Delta_1 & s_e & s_f \end{matrix} \right\}_q P_x^{\Delta_2 \Delta_3}(s_f; s_a, s_e | q).
\end{aligned} \quad (\text{G49})$$

Next, we use the identity

$$\frac{(q^{\frac{1}{2} \pm (k-x) \pm i s}; q)_\infty}{(q^{\frac{1}{2} \pm k \pm i s}; q)_\infty} = q^{-x^2 + 2kx} \quad (\text{G50})$$

and conclude that

$$\begin{aligned}
&q^{x \Delta_1} P_x^{\Delta_2 \Delta_3}(s_c; s_b, s_d | q) \left\{ \begin{matrix} (\Delta_2 + \Delta_3 + x) & s_b & s_d \\ \Delta_1 & s_e & s_a \end{matrix} \right\}_q \\
&= \int_0^{\frac{\pi}{|\log q|}} ds_f \rho_q(s_f) \left\{ \begin{matrix} \Delta_2 & s_b & s_c \\ \Delta_1 & s_f & s_a \end{matrix} \right\}_q \left\{ \begin{matrix} \Delta_3 & s_c & s_d \\ \Delta_1 & s_e & s_f \end{matrix} \right\}_q P_x^{\Delta_2 \Delta_3}(s_f; s_a, s_e | q).
\end{aligned} \quad (\text{G51})$$

Using the completeness relation (G41), this becomes

$$\begin{aligned}
&\sum_{x=0}^{\infty} q^{x \Delta_1} P_x^{\Delta_2 \Delta_3}(s_c; s_b, s_d | q) \left\{ \begin{matrix} (\Delta_2 + \Delta_3 + x) & s_b & s_d \\ \Delta_1 & s_e & s_a \end{matrix} \right\}_q P_x^{\Delta_2 \Delta_3}(s_f; s_a, s_e | q) \\
&= \left\{ \begin{matrix} \Delta_2 & s_b & s_c \\ \Delta_1 & s_f & s_a \end{matrix} \right\}_q \left\{ \begin{matrix} \Delta_3 & s_c & s_d \\ \Delta_1 & s_e & s_f \end{matrix} \right\}_q.
\end{aligned} \quad (\text{G52})$$

Using the graphical notation in (A8), this becomes

$$\Delta_1 \text{ --- } \begin{array}{c} \Delta_2 \\ | \\ \Delta_3 \end{array} = \sum_{x=0}^{\infty} q^{x\Delta_1} \begin{array}{c} \Delta_2 \quad \Delta_3 \\ \diagdown \quad \diagup \\ \Delta_2 + \Delta_3 + x \\ | \\ \Delta_2 \quad \Delta_3 \end{array} \quad (\text{G53})$$

5. Final identity

The last identity that we will need follows from (G31), with the following identifications:

$$\begin{aligned} j_1 &= \pi_{\Delta_1}^-, & j_2 &= \pi_{\Delta_2}^-, & j_3 &= \pi_{\Delta_3}^-, & j_4 &= \pi_{\Delta_1 + \Delta_2 + \Delta_3 + N}^-, \\ j_5 &= \pi_{\Delta_1 + \Delta_2 + x}^-, & j_6 &= \pi_{\Delta_2 + \Delta_3 + y}^-, & j_7 &= \pi_{s_a, \epsilon}^P, & j_8 &= \pi_{s_b, \epsilon - \Delta_1}^P, \\ j_9 &= \pi_{s_c, \epsilon - \Delta_1 - \Delta_2}^P, & j_{10} &= \pi_{s_d, \epsilon - \Delta_1 - \Delta_2 - \Delta_3}^P, & \sum_{j_6} &= \sum_{y=0}^N. \end{aligned} \quad (\text{G54})$$

Then, (G31) becomes

$$P_x^{\Delta_1, \Delta_2}(s_b; s_a, s_c | q) P_{N-x}^{\Delta_1 + \Delta_2 + x, \Delta_3}(s_c; s_a, s_d | q) = \sum_{y=0}^N S_q^{\Delta_3, \Delta_2, \Delta_1, N}(y, x) P_{N-y}^{\Delta_1, \Delta_2 + \Delta_3 + y}(s_b; s_a, s_d | q) P_y^{\Delta_2, \Delta_3}(s_c; s_b, s_d | q). \quad (\text{G55})$$

Using the graphical notation introduced in (A8), this identity becomes

$$\begin{array}{c} \Delta_1 + \Delta_2 + \Delta_3 + N \\ | \\ \Delta_1 + \Delta_2 + x \\ \diagdown \quad \diagup \\ \Delta_1 \quad \Delta_2 \quad \Delta_3 \end{array} = \sum_{y=0}^N S_q^{\Delta_3, \Delta_2, \Delta_1, N}(y, x) \begin{array}{c} \Delta_1 + \Delta_2 + \Delta_3 + N \\ | \\ \Delta_1 + \Delta_2 + y \\ \diagdown \quad \diagup \\ \Delta_1 \quad \Delta_2 \quad \Delta_3 \end{array} \quad (\text{G56})$$

-
- [1] J. M. Deutsch, Quantum statistical mechanics in a closed system, *Phys. Rev. A* **43**, 2046 (1991).
 [2] M. Srednicki, Chaos and quantum thermalization, *Phys. Rev. E* **50**, 888 (1994).
 [3] L. Foini and J. Kurchan, Eigenstate thermalization hypothesis and out of time order correlators, *Phys. Rev. E* **99**, 042139 (2019).

- [4] J. Sonner and M. Vielma, Eigenstate thermalization in the Sachdev-Ye-Kitaev model, *J. High Energy Phys.* **11** (2017) 149.
 [5] C. Murthy and M. Srednicki, Bounds on Chaos from the Eigenstate Thermalization Hypothesis, *Phys. Rev. Lett.* **123**, 230606 (2019).
 [6] T. Anous and J. Sonner, Phases of scrambling in eigenstates, *SciPost Phys.* **7**, 003 (2019).

- [7] P. Nayak, J. Sonner, and M. Vielma, Extended eigenstate thermalization and the role of FZZT branes in the Schwarzsian theory, *J. High Energy Phys.* **03** (2020) 168.
- [8] J. Wang, M. H. Lamann, J. Richter, R. Steinigeweg, A. Dymarsky, and J. Gemmer, Eigenstate Thermalization Hypothesis and Its Deviations from Random-Matrix Theory beyond the Thermalization Time, *Phys. Rev. Lett.* **128**, 180601 (2022).
- [9] P. Saad, S. H. Shenker, and D. Stanford, JT gravity as a matrix integral, [arXiv:1903.11115](#).
- [10] E. Witten, Matrix models and deformations of JT gravity, *Proc. R. Soc. A* **476**, 20200582 (2020).
- [11] H. Maxfield and G. J. Turiaci, The path integral of 3D gravity near extremality; or, JT gravity with defects as a matrix integral, *J. High Energy Phys.* **01** (2021) 118.
- [12] G. Penington, S. H. Shenker, D. Stanford, and Z. Yang, Replica wormholes and the black hole interior, [arXiv:1911.11977](#).
- [13] M. Berkooz, M. Isachenkov, V. Narovlansky, and G. Torrents, Towards a full solution of the large N double-scaled SYK model, *J. High Energy Phys.* **03** (2019) 079.
- [14] D. Mazac and M. F. Paulos, The analytic functional bootstrap. Part I: 1D CFTs and 2D S-matrices, *J. High Energy Phys.* **02** (2019) 162.
- [15] D. Mazac and M. F. Paulos, The analytic functional bootstrap. Part II. Natural bases for the crossing equation, *J. High Energy Phys.* **02** (2019) 163.
- [16] T. G. Mertens, G. J. Turiaci, and H. L. Verlinde, Solving the Schwarzian via the conformal bootstrap, *J. High Energy Phys.* **08** (2017) 136.
- [17] S. J. Suh, Dynamics of black holes in Jackiw-Teitelboim gravity, *J. High Energy Phys.* **03** (2020) 093.
- [18] J. Maldacena, D. Stanford, and Z. Yang, Conformal symmetry and its breaking in two dimensional nearly anti-de-Sitter space, *Prog. Theor. Exp. Phys.* **2016**, 12C104 (2016).
- [19] Z. Yang, The quantum gravity dynamics of near extremal black holes, *J. High Energy Phys.* **05** (2019) 205.
- [20] A. Kitaev and S. J. Suh, Statistical mechanics of a two-dimensional black hole, *J. High Energy Phys.* **05** (2019) 198.
- [21] D. Bagrets, A. Altland, and A. Kamenev, Sachdev–Ye–Kitaev model as Liouville quantum mechanics, *Nucl. Phys.* **B911** 191 (2016).
- [22] J. Liu, E. Perlmutter, V. Rosenhaus, and D. Simmons-Duffin, d -dimensional SYK, AdS loops, and $6j$ symbols, *J. High Energy Phys.* **03** (2019) 052.
- [23] P. Buser, *Geometry and Spectra of Compact Riemann Surfaces*, Modern Birkhäuser Classics (Birkhäuser, 2010).
- [24] P. Saad, Late time correlation functions, baby universes, and ETH in JT gravity, [arXiv:1910.10311](#).
- [25] L. V. Iliesiu, M. Mezei, and G. Sárosi, The volume of the black hole interior at late times, [arXiv:2107.06286](#).
- [26] A. Blommaert, Dissecting the ensemble in JT gravity, [arXiv:2006.13971](#).
- [27] D. Gaiotto, D. Mazac, and M. F. Paulos, Bootstrapping the 3d Ising twist defect, *J. High Energy Phys.* **03** (2014) 100.
- [28] I. Heemskerk, J. Penedones, J. Polchinski, and J. Sully, Holography from conformal field theory, *J. High Energy Phys.* **10** (2009) 079.
- [29] J. Pollack, M. Rozali, J. Sully, and D. Wakeham, Eigenstate Thermalization and Disorder Averaging in Gravity, *Phys. Rev. Lett.* **125**, 021601 (2020).
- [30] M. Srednicki, Thermal fluctuations in quantized chaotic systems, *J. Phys. A* **29**, L75 (1996).
- [31] A. Goel, L. V. Iliesiu, J. Kruthoff, and Z. Yang, Classifying boundary conditions in JT gravity: From energy-branes to α -branes, *J. High Energy Phys.* **04** (2021) 069.
- [32] A. Goel and H. Verlinde, Towards a string dual of SYK, [arXiv:2103.03187](#).
- [33] L. Susskind, Entanglement and chaos in de Sitter space holography: An SYK example, *J. Hologr. Appl. Phys.* **1**, 1 (2021).
- [34] L. Susskind, Scrambling in double-scaled SYK and de Sitter space, [arXiv:2205.00315](#).
- [35] H. Lin and L. Susskind, Infinite temperature’s not so hot, [arXiv:2206.01083](#).
- [36] P. Gao, D. L. Jafferis, and D. K. Kolchmeyer, An effective matrix model for dynamical end of the world branes in Jackiw-Teitelboim gravity, *J. High Energy Phys.* **01** (2022) 038.
- [37] S. Sachdev and J. Ye, Gapless Spin-Fluid Ground State in a Random Quantum Heisenberg Magnet, *Phys. Rev. Lett.* **70**, 3339 (1993).
- [38] A. Kitaev, A simple model of quantum holography 1, in *Talk at KITP* (2015), <http://online.kitp.ucsb.edu/online/entangled15/kitaev/>.
- [39] A. Kitaev, A simple model of quantum holography 2, in *Talk at KITP* (2015), <http://online.kitp.ucsb.edu/online/entangled15/kitaev2/>.
- [40] J. Maldacena and D. Stanford, Remarks on the Sachdev–Ye–Kitaev model, *Phys. Rev. D* **94**, 106002 (2016).
- [41] C. V. Johnson, Nonperturbative Jackiw-Teitelboim gravity, *Phys. Rev. D* **101**, 106023 (2020).
- [42] C. V. Johnson, Explorations of nonperturbative Jackiw-Teitelboim gravity and supergravity, *Phys. Rev. D* **103**, 046013 (2021).
- [43] C. V. Johnson, Jackiw-Teitelboim supergravity, minimal strings, and matrix models, *Phys. Rev. D* **103**, 046012 (2021).
- [44] C. V. Johnson, Low energy thermodynamics of JT gravity and supergravity, [arXiv:2008.13120](#).
- [45] C. V. Johnson, F. Rosso, and A. Svesko, Jackiw-Teitelboim supergravity as a double-cut matrix model, *Phys. Rev. D* **104**, 086019 (2021).
- [46] C. V. Johnson, Consistency conditions for non-perturbative completions of JT gravity, [arXiv:2112.00766](#).
- [47] C. V. Johnson, Quantum Gravity Microstates from Fredholm Determinants, *Phys. Rev. Lett.* **127**, 181602 (2021).
- [48] B. Post, J. van der Heijden, and E. Verlinde, A universe field theory for JT gravity, *J. High Energy Phys.* **05** (2022) 118.
- [49] A. Altland, B. Post, J. Sonner, J. van der Heijden, and E. Verlinde, Quantum chaos in 2D gravity, [arXiv:2204.07583](#).
- [50] J. Ambjorn, J. Jurkiewicz, and Y. M. Makeenko, Multiloop correlators for two-dimensional quantum gravity, *Phys. Lett. B* **251**, 517 (1990).
- [51] E. Brezin and A. Zee, Universality of the correlations between eigenvalues of large random matrices, *Nucl. Phys.* **B402**, 613 (1993).

- [52] D. Schmidlin, Bilinear generating function for the Chebyshev polynomials of the second kind (2015), https://www.researchgate.net/publication/271443041_Bilinear_Generating_Function_for_the_Chebyshev_Polynomials_of_the_Second_Kind.
- [53] E. Brezin and S. Hikami, Vertices from replica in a random matrix theory, *J. Phys. A* **40**, 3545 (2007).
- [54] A. Blommaert and J. Kruthoff, Gravity without averaging, *SciPost Phys.* **12**, 073 (2022).
- [55] H. W. Lin, The bulk Hilbert space of double scaled SYK, [arXiv:2208.07032](https://arxiv.org/abs/2208.07032).
- [56] J. Cotler and K. Jensen, AdS_3 gravity and random CFT, *J. High Energy Phys.* **04** (2021) 033.
- [57] A. Belin and J. de Boer, Random statistics of OPE coefficients and Euclidean wormholes, *Classical Quantum Gravity* **38**, 164001 (2021).
- [58] J. Chandra, S. Collier, T. Hartman, and A. Maloney, Semiclassical 3D gravity as an average of large- c CFTs, [arXiv:2203.06511](https://arxiv.org/abs/2203.06511).
- [59] J.-M. Schlenker and E. Witten, No ensemble averaging below the black hole threshold, [arXiv:2202.01372](https://arxiv.org/abs/2202.01372).
- [60] A. Belin, J. de Boer, and D. Liska, Non-Gaussianities in the statistical distribution of heavy OPE coefficients and wormholes, *J. High Energy Phys.* **06** (2022) 116.
- [61] T. Anous, A. Belin, J. de Boer, and D. Liska, OPE statistics from higher-point crossing, *J. High Energy Phys.* **06** (2022) 102.
- [62] A. Blommaert and M. Usatyuk, Microstructure in matrix elements, [arXiv:2108.02210](https://arxiv.org/abs/2108.02210).
- [63] J. Chandra and T. Hartman, Coarse graining pure states in AdS/CFT , [arXiv:2206.03414](https://arxiv.org/abs/2206.03414).
- [64] N. Engelhardt and A. C. Wall, Coarse graining holographic black holes, *J. High Energy Phys.* **05** (2019) 160.
- [65] A. Altland and J. Sonner, Late time physics of holographic quantum chaos, *SciPost Phys.* **11**, 034 (2021).
- [66] A. Altland, D. Bagrets, P. Nayak, J. Sonner, and M. Vielma, From operator statistics to wormholes, *Phys. Rev. Res.* **3**, 033259 (2021).
- [67] S. Pappalardi, L. Foini, and J. Kurchan, Eigenstate thermalization hypothesis and free probability, [arXiv:2204.11679](https://arxiv.org/abs/2204.11679).
- [68] H. W. Lin, J. Maldacena, L. Rozenberg, and J. Shan, Holography for people with no time, [arXiv:2207.00407](https://arxiv.org/abs/2207.00407).
- [69] H. W. Lin, J. Maldacena, L. Rozenberg, and J. Shan, Looking at supersymmetric black holes for a very long time, [arXiv:2207.00408](https://arxiv.org/abs/2207.00408).
- [70] A. Maloney and E. Witten, Quantum gravity partition functions in three dimensions, *J. High Energy Phys.* **02** (2010) 029.
- [71] C. A. Keller and A. Maloney, Poincare series, 3D gravity and CFT spectroscopy, *J. High Energy Phys.* **02** (2015) 080.
- [72] C. Akers, N. Engelhardt, D. Harlow, G. Penington, and S. Vardhan, The black hole interior from non-isometric codes and complexity, [arXiv:2207.06536](https://arxiv.org/abs/2207.06536).
- [73] W. Groenevelt, The Wilson function transform, *Int. Math. Res. Not.* **2003**, 2779 (2003).
- [74] W. Groenevelt, Wilson function transforms related to Racah coefficients, *Acta Appl. Math.* **91**, 133 (2005).
- [75] G. Gasper, M. Rahman, C. U. Press, G. Rota, B. Doran, P. Flajolet *et al.*, *Basic Hypergeometric Series*, Encyclopedia of Mathematics and Its Applications (Cambridge University Press, 2004).
- [76] E. Koelink and J. Stokman, The Askey-Wilson function transform, *Int. Math. Res. Not.* **2001**, 05 (2000).
- [77] J. V. Stokman, An expansion formula for the Askey-Wilson function, *J. Approx. Theory* **114**, 308 (2002).
- [78] D. Harlow and D. Jafferis, The factorization problem in Jackiw-Teitelboim gravity, *J. High Energy Phys.* **02** (2020) 177.
- [79] D. Stanford and E. Witten, Fermionic localization of the Schwarzian theory, *J. High Energy Phys.* **10** (2017) 008.
- [80] T. H. Koornwinder, *Jacobi Functions and Analysis on Noncompact Semisimple Lie Groups* (Springer Netherlands, Dordrecht, 1984), pp. 1–85.
- [81] R. Koekoek, T. Koornwinder, P. Lesky, and R. Swarttouw, *Hypergeometric Orthogonal Polynomials and Their q -Analogues*, Springer Monographs in Mathematics (Springer Berlin, Heidelberg, 2010).
- [82] Q. M. Tariq, Some generating functions and projection formula for the associated Askey-Wilson polynomials, Ph.D. thesis, Carleton University, 1997.
- [83] R. A. Askey, M. Rahman, and S. Suslov, On a general q -Fourier transformation with nonsymmetric kernels, *J. Comput. Appl. Math.* **68**, 55 (1996).
- [84] J. Baez, The Biedenharn-Elliott identity, <https://math.ucr.edu/home/baez/qg-fall2000/qg10.2.html>.