# Matrix Models for Eigenstate Thermalization

Daniel Louis Jafferis,[1] David K. Kolchmeyer,[1,2] Baur Mukhametzhanov,[3,4] and Julian Sonner[5]
[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 28 December 2022; revised 26 March 2023; accepted 31 July 2023; published 20 September 2023)

We develop a class of matrix models that implement and formalize the eigenstate thermalization hypothesis (ETH) and point out that, in general, these models must contain non-Gaussian corrections in order to correctly capture thermal mean-field theory or to capture nontrivial out-of-time-order correlation functions (OTOCs) as well as their higher-order generalizations. We develop the framework of these ETH matrix models and put it in the context of recent studies in statistical physics incorporating higher statistical moments into the ETH ansatz. We then use the ETH matrix model in order to develop a matrix-integral description of Jackiw-Teitelboim (JT) gravity coupled to a single scalar field in the bulk. This particular example takes the form of a double-scaled ETH matrix model with non-Gaussian couplings matching disk correlators and the density of states of the gravitational theory. Having defined the model from the disk data, we present evidence that the model correctly captures the JT + matter theory with multiple boundaries and, conjecturally, at higher genus. This is a shorter companion paper to the work [D. L. Jafferis *et al.*, companion paper, *Jackiw-Teitelboim gravity with matter, generalized eigenstate thermalization hypothesis, and random matrices*, Phys. Rev. D **108**, 066015 (2023).], serving as a guide to the much more extensive material presented there, as well as developing its underpinning in statistical physics.

DOI: 10.1103/PhysRevX.13.031033 Subject Areas: Gravitation, Statistical Physics, String Theory

## I. INTRODUCTION

Theories defined by matrix integrals arise in a variety of physical contexts. Two particularly prominent, but *a priori* unrelated, examples lie at the heart of this paper. On the one hand, such matrix theories arise as the description of ergodic phases of quantum chaotic systems; on the other hand, they also feature prominently in the study of lower-dimensional quantum gravity where they allow a construction of the nonperturbative sum over (smooth) geometries. Progress in lower-dimensional holography in recent years [1–6] has led to the striking realization [7] that these two approaches, in certain contexts, are in fact one and the same. Its clearest expression can perhaps be found in the rewriting of the path integral of two-dimensional Jackiw-Teitelboim (JT) gravity in terms of a matrix integral over boundary Hamiltonians [8]. The potentially profound connection of two-dimensional gravity to the physics of quantum chaos has been further developed in Refs. [9,10] by demonstrating that JT gravity can indeed be described as a quantum ergodic phase. It is clearly an important goal to develop this connection between quantum chaos and quantum gravity further, in particular, by extending it away from pure JT gravity and low dimensions. This work, as well as its longer companion paper, Ref. [11], develops the gravity/chaos correspondence further by devising a matrix-integral description of JT gravity coupled to scalar matter and, furthermore, by incorporating a different paradigm of quantum chaos into the story, namely, that of the eigenstate thermalization hypothesis (ETH). The two-dimensional models we mentioned before may be *exactly* rewritten in terms of matrix descriptions of ergodic phases, but one would expect that *effective* descriptions of gravity in terms of quantum ergodic matrix theories should also be possible in higher dimensions, at least at or after certain timescales. In conventional quantum-chaotic approaches, this timescale would be assumed to be the Thouless time, but by adding specific nonlinearities or constraints to the matrix integral, one may extend the region of validity to earlier times. The ETH matrix model developed in this work and

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

in Ref. [11] offers a general framework of the kind of structure, which we expect to also be relevant in such cases.

This paper, being a shorter companion to the work in Ref. [11], serves two dual purposes. First, we give an accessible and succinct summary of the salient features, results, and methods of the longer paper, with the aim to guide the reader through the more extensive technical material appearing there. Second, we supply a slightly different perspective on the results, emphasizing general aspects relevant to a statistical description of chaotic quantum systems which unifies random-matrix theory (RMT) and eigenstate thermalization into a common framework. We illustrate the approach by describing how it implements thermal mean-field theory (see Sec. III), the simplest example known to us, as well as to JT gravity, which introduces more structure beyond thermal mean-field theory (TMFT).

While our principal object of interest remains the matrix-model description of matter-coupled JT gravity, we believe that the more general structure of the ETH matrix model is of interest for more general quantum chaotic systems. At its heart is the simple observation that an ensemble-averaged description, even of thermal mean-field theory, must contain important non-Gaussian correlations of the form

$$\overline{\mathcal{O}_{a_1b_1}\mathcal{O}_{a_2b_2}\cdots\mathcal{O}_{a_nb_n}}\Big|_{\mathrm{ETH}} = e^{-(n-1)S(\bar{E})}g^{(n)}_{\mathcal{O},a_1b_1\cdots a_nb_n}(E_1,\ldots,E_n) + \text{disconnected}, \tag{1.1}$$

which contain terms that are suppressed by further factors of the microcanonical entropy $S(\bar{E})$ compared to the usual Gaussian correlations evidenced in the standard ETH ansatz. A similar point was made in Ref. [12], which recognized the need for an ansatz that reproduces correct out-of-time-order correlation function (OTOC) dynamics. By $\mathcal{O}_{ab}$, we mean a matrix element of a simple operator in an energy eigenbasis, and the functions $g^{(n)}_{\mathcal{O},a_1b_1\cdots a_nb_n}(E_1,\ldots,E_n)$ are smoothly varying order-one functions of $n$ energies. We will comment on the index structure of $g^{(n)}_{\mathcal{O},a_1b_1\cdots a_nb_n}(E_1,\ldots,E_n)$ later; for now, we note that in order to contribute to the thermal correlation function at leading order, each $b$ index must be equal to one of the $a$ indices. This explains why $g^{(n)}_{\mathcal{O},a_1b_1\cdots a_nb_n}$ is a function of only $n$ energies, rather than $2n$ energies. By "disconnected," we are referring to contributions that depend on functions of fewer than $n$ energy arguments, such as the smooth functions that appear in the standard Gaussian ETH ansatz. This observation comes from recent work on the statistical physics of OTOCs, demonstrating the need for extending ETH to include non-Gaussian contributions (see, e.g., Refs. [12–16]). In this work, we formalize the observation above in terms of a two-matrix model of the form

$$\mathcal{Z}_{\mathrm{ETH}} = \int dHd\mathcal{O}e^{-N\mathrm{Tr}V(H,\mathcal{O})}, \tag{1.2}$$

where $H$ is a matrix defining the statistics of energy levels, while the $\mathcal{O}$ matrix generates the correlations of matrix elements of the additional matter field. Below, we will be more careful about the definition of the correct integration measure and potentials appearing in the exponent. Translated to the potential of the ETH matrix model, the highly suppressed non-Gaussianities of Eq. (1.1) appear, in fact, at $\mathcal{O}(1)$, which is a reflection of the fact that they must contribute at leading order to higher-order correlation functions in order to produce the necessary connected parts. Our ETH matrix model is a further generalization of the ETH matrix model of Refs. [12,17] because it applies to systems outside of the usual thermodynamic (or semiclassical) regime, as we will explain.

The first part of this paper is dedicated to a more careful definition of the ETH matrix model, with general quantum chaotic applications in mind. The second part specializes this structure to the theory of JT gravity with matter, where the additional structure present allows us to go much further in pinning down the full matrix potential $V(H,\mathcal{O})$, guided by the gravitational path integral as well as arguments based on locality and conformal symmetry. Note that the JT-matter gravitational path integral is not well defined all the way to the UV, unlike the pure JT case. We will point out how this is encoded in the dual ETH matrix model as well.

## II. MATRIX MODEL FOR EIGENSTATE THERMALIZATION

An important signature shown by quantum chaotic systems is the presence of distinctive statistical correlations between their energy levels. There are two interrelated frameworks that are usually invoked to quantitatively describe these correlations, namely, RMT and the ETH. The purpose of this section is to describe a generalization of ETH that is powerful enough to incorporate more fine-grained correlations than the usual "Gaussian ansatz" and, at the same time, introduce a random-matrix description of such a generalized ETH ensemble. A generalized ETH ansatz for the matrix elements was introduced previously in Ref. [12] (and elaborated upon in Ref. [18]) with an emphasis on the thermodynamic limit, and we generalize it by letting the energy eigenvalues also come from a random matrix. Our generalization will set the context for our matrix-model description of JT gravity coupled to simple scalar matter, described in Sec. IV, as well as in the longer companion paper [11].

### A. Degaussing ETH

The level correlations implied by ETH are usually stated in the form of an ansatz for the matrix elements of a simple operator in the energy basis [19,20],

$$\langle E_a|\mathcal{O}|E_b\rangle = \overline{\langle E_a|\mathcal{O}|E_a\rangle}\delta_{ab} + e^{-S(\bar{E})/2} f_{\mathcal{O}}(E_a, E_b) R_{ab}, \tag{2.1}$$

where $\overline{\langle E_a|\mathcal{O}|E_a\rangle}$ is a smooth function of $E_a$, $S(\bar{E})$ is the microcanonical entropy evaluated at the average energy $\bar{E} = \frac{1}{2}(E_a + E_b)$, and $f_{\mathcal{O}}(E_a, E_b)$ is a smooth function of both energies. Finally, $R_{ab}$ is a random matrix, traditionally taken to be sampled from a *Gaussian* distribution with zero mean and unit variance. One therefore postulates the existence of an ensemble—we will refer to this as the ETH ensemble—from which matrix elements of simple operators in the energy basis are sampled. The usual ETH ansatz above is equivalent to specifying the first two nontrivial moments of this ensemble, namely, compatible with a purely Gaussian probability distribution.

That this ensemble cannot in fact obey purely Gaussian statistics can be seen by noting that this would imply that all higher-order thermal correlation functions factorize, which is in contradiction with, for example, the existence of a nonvanishing Lyapunov exponent diagnosed from OTOCs in eigenstates [12–14,21,22]. Another argument forcing us to consider non-Gaussian statistical ensembles comes from considering the main application we have in mind in this work, namely, CFTs with holographic duals and, more specifically, JT gravity coupled to matter. Suppose we use bulk gravity Witten-diagram techniques to compute the manifestly crossing-invariant four-point function

$$\langle \mathcal{O}(\tau_1)\mathcal{O}(\tau_2)\mathcal{O}(\tau_3)\mathcal{O}(\tau_4)\rangle = (\tau_{12}\tau_{34})^{-2\Delta} + (\tau_{14}\tau_{23})^{-2\Delta} + (\tau_{13}\tau_{24})^{-2\Delta} \tag{2.2}$$

of the boundary 1D conformal quantum system. We may also consider its generalization to finite temperature,

$$\langle \mathrm{Tr} e^{-\beta H}\mathcal{O}(\tau_1)\mathcal{O}(\tau_2)\mathcal{O}(\tau_3)\mathcal{O}(\tau_4)\rangle = \left(\frac{\beta}{\pi}\sin\left(\frac{\pi}{\beta}\tau_{12}\right)\frac{\beta}{\pi}\sin\left(\frac{\pi}{\beta}\tau_{34}\right)\right)^{-2\Delta} \tag{2.3}$$

$$+\left(\frac{\beta}{\pi}\sin\left(\frac{\pi}{\beta}\tau_{14}\right)\frac{\beta}{\pi}\sin\left(\frac{\pi}{\beta}\tau_{23}\right)\right)^{-2\Delta} \tag{2.4}$$

$$+\left(\frac{\beta}{\pi}\sin\left(\frac{\pi}{\beta}\tau_{13}\right)\frac{\beta}{\pi}\sin\left(\frac{\pi}{\beta}\tau_{24}\right)\right)^{-2\Delta}. \tag{2.5}$$

By crossing invariance, we refer to the invariance of the four-point function under permutations of the external operators. Using the bulk perspective, the expression above results straightforwardly from a computation using gravity Witten diagrams, so long as we work in the semiclassical regime. From the boundary perspective, this expression implies that the OPE between $\mathcal{O}(x)\mathcal{O}(y)$ only receives contributions from double-trace primaries $[\mathcal{O}\mathcal{O}]_n$ of dimension $2\Delta + n$ for $n \in 2\mathbb{Z}_{\geq 0}$ in addition to the identity operator. The latter perspective may not seem to be immediately relevant to the discussion, but we shall come to this statement shortly and put it to good use in the context of thermal mean-field theory and indeed JT gravity, where it will be of great help to constrain the relevant ETH ensemble.

To set the scene, we now ask what the answer would be if we evaluated the correlation function above in the Gaussian ETH ensemble, i.e., using the ansatz (2.1) to define quadratic Wick contractions of the operator $\mathcal{O}$ [23]. We thus compute

$$\overline{\mathrm{Tr}\prod_{j=1}^{4} e^{-\beta_j H}\mathcal{O}}\Bigg|_{\mathrm{Gauss}} = \sum_{a_1\ldots a_4} e^{-\sum_j \beta_j E_{a_j}} F(E_{a_1}, E_{a_2})^{-1} \times F(E_{a_3}, E_{a_4})^{-1}(\delta_{a_1 a_3} + \delta_{a_2 a_4}) \tag{2.6}$$

$$+\sum_a e^{-(\beta_1+\cdots+\beta_4)E_a} F(E_a, E_a)^{-2}. \tag{2.7}$$

Note that, as indicated, we evaluated the above expectation value with respect to the purely Gaussian ETH ensemble, using the Wick contraction

$$\overbrace{\mathcal{O}_{ab}\mathcal{O}_{cd}} = e^{-S(\overline{E})}\,|f_{\mathcal{O}}(E_a, E_b)|^2\,\delta_{ad}\delta_{bc} := F(E_a, E_b)^{-1}\delta_{ad}\delta_{bc} \tag{2.8}$$

following from the Gaussian ETH ansatz above and where, in the last equality, we have introduced a more convenient notation that will be helpful later. The average energy in the exponent reads $\bar{E} = \frac{1}{2}(E_a + E_b)$. We now notice that the third term (2.7) is nonplanar [24], and as such, it is subleading at large $N = \dim H \gg 1$ or, equivalently, in the $e^S$ counting.

Considering only the first two terms that contribute at leading order, however, gives an answer that is not crossing invariant, equivalent to only keeping the first two terms in Eq. (2.2) above. Following Ref. [12], we conclude that the term making the overall result invariant must come from a non-Gaussian connected contribution to the ETH ensemble, which nevertheless contributes at leading order owing to additional Hilbert space sums, i.e., a contribution of the form

$$\overline{\mathcal{O}_{a_1b_1}\mathcal{O}_{a_2b_2}\mathcal{O}_{a_3b_3}\mathcal{O}_{a_4b_4}} \supset e^{-3S(\overline{E})} g^{(4)}_{\mathcal{O},a_1b_1\cdots a_4b_4}(E_1, E_2, E_3, E_4), \tag{2.9}$$

where $\bar{E}$ is now the average of all four energies, and $g_{\mathcal{O}}$ is a smooth function of the individual energies. The entropic suppression factor is dictated by the need to precisely balance the four sums over energy eigenstates in the four-point function (2.6) to result in a leading-order contribution. For a general $n$-point function, this generalizes easily to

$$\overline{\mathcal{O}_{a_1b_1}\mathcal{O}_{a_2b_2}\cdots\mathcal{O}_{a_nb_n}} = e^{-(n-1)S(\bar{E})} g^{(n)}_{\mathcal{O},a_1b_1\cdots a_nb_n}(E_1,\ldots,E_n) + \text{disconnected}. \tag{2.10}$$

In typical expressions of interest, the above averaged matrix elements will be integrated against a suitable set of densities of states, which we will incorporate below into our construction. The notation of the overline $\overline{\cdots}$ indicates purely the expectation value with respect to the ensemble for the matrix elements (where the eigenvalues are fixed to a typical instance), while $\langle\cdots\rangle$ indicates an expectation value that includes the energy integrals, a structure we refer to as the ETH matrix model for reasons that will become clear presently.

Returning to our discussion of the crossing-invariant four-point function above, we now see that the third planar, i.e., leading order in $e^S$, contribution to Eq. (2.2) comes from the quartic non-Gaussianity $g^{(4)}_{\mathcal{O}}(E_1,E_2,E_3,E_4)$, which shows that indeed the ETH ensemble has to contain nontrivial higher statistical moments even to match thermal mean-field theory at the level of the four-point function. At the level of the usual ETH ansatz, these non-Gaussianities are extremely subtle, that is, highly suppressed in entropy [cf. Eq. (2.10)], but their effect on thermal correlation functions is significant and necessary even to ensure basic properties like crossing symmetry.

In the formula (2.1), traditionally referred to as the ETH ansatz, the function $f_{\mathcal{O}}(E_1,E_2) = f_{\mathcal{O}}(\bar{E},\omega)$ is left unspecified, reminding us that it depends on a given physical system. We take the same point of view regarding the higher moments $g^{(n)}_{\mathcal{O}}$, in general. In Sec. IV, when we construct the specific ETH ensembles describing TMFT, as well as matter-coupled JT gravity, we are able to fully determine (at least in principle) these non-Gaussianities.

We now comment on the index structure of $g^{(n)}_{\mathcal{O},a_1b_1\cdots a_nb_n}\times(E_1,\ldots,E_n)$. An important point is that a single-trace matrix model for $\mathcal{O}$ makes the same predictions for thermal correlators in the thermodynamic limit as the Foini-Kurchan (FK) ensemble [12]. Assuming no degeneracy in the spectrum of the Hamiltonian, the energy eigenbasis is only defined up to an energy-dependent rotation, $|E_a\rangle \to e^{i\theta_a}|E_a\rangle$. Because Eq. (2.10) should hold for any choice of the phases $\theta_a$, the terms on the right side should contain Kronecker delta symbols that each involve one $a$ and one $b$ index. Invoking a typicality argument, we also want the ansatz to be invariant under unitary rotations that act within microcanonical Hilbert spaces. In other words, we want

$$\sum_{\tilde{a}_1\tilde{b}_1\cdots\tilde{a}_n\tilde{b}_n} U_{a_1\tilde{a}_1}U^\dagger_{\tilde{b}_1b_1}\cdots U_{a_n\tilde{a}_n}U^\dagger_{\tilde{b}_nb_n}\overline{\mathcal{O}_{\tilde{a}_1\tilde{b}_1}\ldots\mathcal{O}_{\tilde{a}_n\tilde{b}_n}} = \overline{\mathcal{O}_{a_1b_1}\ldots\mathcal{O}_{a_nb_n}}, \tag{2.11}$$

where $U$ is a block-diagonal unitary where each block has size $e^{S(E)}$. To achieve this, we consider a generalized ansatz such that there are exactly $n$ Kronecker symbols per term. In particular, in the ansatz, we exclude symbols involving two $a$ or two $b$ indices. The general index structure is then

$$\delta_{a_1b_{\sigma(1)}}\delta_{a_2b_{\sigma(2)}}\cdots\delta_{a_nb_{\sigma(n)}}, \tag{2.12}$$

where $\sigma\in S_n$ is a permutation of $n$ elements. If $\sigma$ has more than one cycle, then we say that Eq. (2.12) factors into several terms that individually take the form of Eq. (2.12) for lower values of $n$. The generalized ETH ansatz should have the property that a factor with $n$ symbols is uniquely associated to a specific function of $n$ energies [12]. Hence, the generalized ETH ansatz for a single simple operator is completely characterized by a single smooth function of $n$ energies for each $n$. In particular, we have

$$g^{(n)}_{\mathcal{O},a_1b_1\cdots a_nb_n}(E_1,\ldots,E_n) = \sum_{\sigma\in S_n} g^{(n)}_{\mathcal{O}}(E_{\sigma(1)},\ldots,E_{\sigma(n)}) \times \delta_{b_{\sigma(1)}a_{\sigma(2)}}\delta_{b_{\sigma(2)}a_{\sigma(3)}}\cdots\delta_{b_{\sigma(n)}a_{\sigma(1)}}, \tag{2.13}$$

where $\sigma$ is a permutation of $\{1,2,\ldots,n\}$, and $g^{(n)}_{\mathcal{O}}(E_1,E_2,\ldots,E_n)$ is, without loss of generality, a cyclically invariant function of $n$ energies. The generalized ETH ansatz is the minimal modification of the standard Gaussian ETH ansatz that is needed for compatibility with thermal mean-field theory, and it is completely characterized by the functions $g^{(n)}_{\mathcal{O}}(E_1,\ldots,E_n)$ for all $n\in\mathbb{N}$. These functions are theory dependent, and in JT gravity it is easy to deduce them from the expressions for the disk correlators [25]. This structure is naturally produced by a two-matrix model with a single-trace matrix potential, which is why the potential in Eq. (1.2) is written with a single trace.

Having demonstrated the similarities between our ETH matrix model and the FK ansatz, we now emphasize how our matrix model improves the ansatz. Foini and Kurchan assume a thermodynamic limit where the microcanonical ensemble dominates the canonical ensemble. The FK ansatz is motivated from typicality arguments. Simple operators are conjugated by random block-diagonal unitary matrices, where each block corresponds to a microcanonical window. These windows are taken to be arbitrarily small (but still much larger than the level spacing size). Our matrix model applies more generally to systems where correlators receive contributions from many different microcanonical windows, with the relative weights determined by the density of states $\rho(E)$ for the Hamiltonian. In other words, we are interested in systems where the relevant energies are not necessarily parametrically large, but the spectrum is still exponentially dense. Our matrix model allows for correlations between energy eigenvalues in different energy windows, as we explain in more detail in the next section. In the thermodynamic limit, these correlations are not important for computing a thermal correlator (and hence were not considered in Ref. [12]), but they are important in general (such as in JT gravity with matter). Our ansatz is thus motivated by the typicality arguments presented in Ref. [12] combined with the fact that energy levels of chaotic systems exhibit nontrivial correlations across different microcanonical windows, which resonates with the fact that the energy-level statistics is well described by matrix models. A compelling reason to further generalize the FK ansatz to a two-matrix model is that the matrix model naturally and elegantly reproduces the FK ansatz, including the correct sizes of all of the non-Gaussian correlations. Hence, the FK ansatz is both generalized and succinctly summarized by Eq. (1.2).

### B. Matrix-model description

One of the main insights about chaotic quantum systems is the (conjectured and empirically robustly observed) fact that their energy eigenvalues follow correlation laws associated with those of random matrix ensembles [26]. One defines the joint probability distribution of the eigenvalues of a random matrix [27],

$$d\mu[H] = \mu_0^{-1} e^{-\sum_a V(E_a)} \prod_{a<b} |E_a - E_b|^\beta, \qquad (2.14)$$

where $\beta \in \{1, 2, 4\}$ depends on the symmetry class [28] and $\mu_0$ is a normalization factor, whose value we do not need here. In our principal application to JT with matter, we are interested in the case $\beta = 2$, namely, when the matrix model is in the unitary class, but the remaining cases are relevant for more general ETH ensembles as well as other bulk theories; see, e.g., Ref. [29]. In a number of standard applications, the potential $V(E_a)$ is taken to be quadratic, but we are interested in higher-order non-Gaussian generalizations of this, for example, the so-called SSS potential, describing the eigenvalue density of JT gravity in terms of a matrix integral [8]. Integrating this probability density over all but $n$ eigenvalues gives the $n$-level density, or $n$-level correlation function [27],

$$\rho^{(n)}(E_1, \ldots E_n) = \rho(E_1) \cdots \rho(E_n) + \text{connected}. \qquad (2.15)$$

A special case is the spectral density, obtained by marginalizing the distribution (2.14) over all but a single eigenvalue. We often use the notation

$$\rho(E) = e^{S(E)} = e^{S_0} \rho_0(E), \qquad (2.16)$$

meaning that we alternately write the full spectral density as the exponent of the microcanonical entropy at a given energy $E$, or as a factor $e^{S_0}$ times a smooth order-one function $\rho_0(E)$, as is often employed in the SYK or JT context, where $S_0$ has the meaning of the ground-state entropy [4,5,30,31]. More generally, we can think of $e^{S_0}$ as a bookkeeping parameter of entropic factors. Putting together the RMT description of the energy levels of chaotic systems with the generalized ETH of Sec. II A above, it is very natural to describe the overall structure in terms of a two-matrix model where the energy-level statistics are generated by the random Hamiltonian $H$, while the operator matrix element statistics (the generalized ETH) are generated by a second random matrix, which we denote with the same symbol $\mathcal{O}$, by a slight abuse of notation. The matrix integral

$$\mathcal{Z}_{\text{ETH}} = \int d\mu[H] d\mu[\mathcal{O}] e^{-\mathcal{V}[H,\mathcal{O}]} \qquad (2.17)$$

is then seen as the joint probability distribution of energy levels and matrix elements of this ensemble with $d\mu[H]d\mu[\mathcal{O}]$ the yet-to-be-specified appropriate measures for the two random matrices, but which we anticipate will generally be neither quadratic in $H$ nor in $\mathcal{O}$. Viewed in the energy eigenbasis, this is exactly the ETH ensemble we have previously invoked. In this section, we describe the general features of such an ensemble, before constructing a specific instance in Sec. IV, capable of describing matter-coupled JT gravity.

In view of the ETH ansatz and its beyond-Gaussian generalization, it is most natural to define the ETH matrix model in the energy eigenbasis of the Hamiltonian with a single-trace matrix potential

$$\mathcal{Z}_{\text{ETH}} = \int d\mu[H] d\mu[\mathcal{O}] \exp\Bigg( e^{S_0} \sum_{n\geq 2} \sum_{a_1 \cdots a_n} G^{(n)}(E_{a_1}, \ldots, E_{a_n}) \times \mathcal{O}_{a_1 a_2} \mathcal{O}_{a_2 a_3} \cdots \mathcal{O}_{a_n a_1} \Bigg), \qquad (2.18)$$

with the measure for the energy eigenvalues defined to be Eq. (2.14) above, while

$$d\mu[\mathcal{O}] \coloneqq \prod_i d\mathcal{O}_{ii} \prod_{i<j} d\text{Re}\mathcal{O}_{ij} d\text{Im}\mathcal{O}_{ij}. \qquad (2.19)$$

The $G^{(n)}$ functions encode the $g_{\mathcal{O}}^{(n)}$ functions introduced in Eq. (2.13), and their precise relationship can only be determined by actually performing the matrix integral above, which is difficult, in general. The reader may want to convince themselves that this matrix model results in the necessary $e^{S_0}$ scaling shown in Eq. (2.10) if the matrix-model coupling functions $G^{(n)}$ are of $\mathcal{O}(1)$ in $e^{S_0}$. This structure is thus a matrix-integral representation of the generalized ETH above, capable of producing, at the same time, the statistical distribution of energy eigenvalues and corresponding spectral integrals, as well as the statistics of the operator matrix elements $\mathcal{O}_{ab}$.

The coupling functions $G^{(n)}(E_1, \ldots E_n)$ are smooth functions of the energy arguments, and they take on specific functional forms only once a particular theory has been specified. As an example, we indicate below (again, much more detail can be found in Ref. [11]) how these functions can, in principle, be determined systematically for the JT + scalar matrix model from thermal correlation functions.

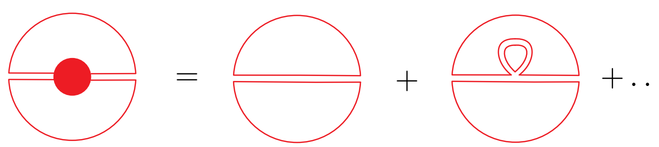
Note that $d\mu[H]$, as defined in Eq. (2.14) above, contains both the Vandermonde factor $\Delta^\beta(E)$ and the exponential of the potential $V(E_a)$. The joint measure can be seen to be the correct one by following the usual Fadeev-Popov procedure [32]. One starts in a general basis where neither $H$ nor $\mathcal{O}$ is diagonal and then transforms to the eigenbasis of the $H$ matrix, sending $H \to UHU^\dagger$ and $\mathcal{O} \to U\mathcal{O}U^\dagger$. In general, $H$ and $\mathcal{O}$ are not simultaneously diagonalizable; therefore, we may reduce to integrating only over the eigenvalues of $H$, but we must integrate over the full Haar measure for Hermitian matrices for the second matrix $\mathcal{O}$.

As mentioned before, the $G^{(n)}(E_1, \ldots, E_n)$ are continuous functions of the energies, and they determine the matrix model potential, which is organized as an expansion in $\mathcal{O}^n$. We often write the $n$th order term in this expansion as $W^{(n)}[E_{a_1}, \ldots E_{a_n}; \mathcal{O}]$. One chooses $V(H)$ in such a way as to match the leading density of states to the chaotic system of interest [33]. We can now deduce the statistical distribution of energy eigenvalues, as well as of operator matrix elements, by introducing sources for the energy, as well as for moments of the operator $\mathcal{O}$ in the energy eigenbasis in order to generate the various moments of the ETH ensemble—these will, of course, be nothing but our $g^{(n)}_{\mathcal{O},a_1 b_1 \ldots a_n b_n}(E_1, \ldots E_n)$ above. Let us illustrate this for the quadratic and quartic moments before stating the general answer. From the model (2.18), one computes the two-point function of two operators at arbitrary Euclidean separation,

$$\begin{aligned}
&\langle \mathrm{Tr}(e^{-\beta_1 H}\mathcal{O}e^{-\beta_2 H}\mathcal{O})\rangle \\
&= \int d\mu[H]d\mu[\mathcal{O}]\mathrm{tr}(e^{-\beta_1 H}\mathcal{O}e^{-\beta_2 H}\mathcal{O})e^{e^{S_0}\sum_n W^{(n)}[E_{a_1},\ldots E_{a_n};\mathcal{O}]} \\
&= \int dE_a dE_b \rho(E_a)\rho(E_b)e^{-\beta_1 E_a - \beta_2 E_b}\overline{\mathcal{O}_{ab}\mathcal{O}_{ba}} + \cdots \qquad (2.20)
\end{aligned}$$

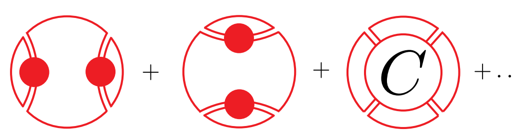
where in the second line, we used the overbar notation to indicate the expectation value evaluated in the $\mathcal{O}$ matrix potential, written in the $H$ eigenbasis. We have consequently performed all energy integrals except for the pair $E_a$, $E_b$ appearing explicitly in $\overline{\mathcal{O}_{ab}\mathcal{O}_{ba}}$. This process has produced a factor of the pair correlation function $\rho^{(2)}(E_a, E_b)$, which at leading order equals the product $\rho(E_a)\rho(E_b)$. The corrections are subleading in the "genus expansion," that is, in powers of $e^{S_0}$. We may denote this equation graphically as

$$\left\langle \mathrm{Tr}\, e^{-\beta H}\mathcal{O}(\tau)\mathcal{O}(0)\right\rangle \ := \qquad (2.21)$$



where we have not drawn any of the in-filling $H$ double lines, which are automatically accounted for by the integrals over the densities of states, as explained above. We may similarly compute the four-point function, which, for simplicity, we only represent graphically,

$$\left\langle \mathrm{Tr}\, e^{-\beta H}\mathcal{O}(\tau_1)\cdots\mathcal{O}(\tau_4)\right\rangle \ := \qquad (2.22)$$




where the third diagram marked with a "$C$" implements the contribution of the interaction terms in the matrix model. Of course, analogous computations apply for higher-order correlations. In general, the noncrossing partitions discussed in Ref. [18] directly correspond to sums over planar 't Hooft diagrams in our ETH matrix model. Each diagram above specifies such a partition. In any given application, one must choose the couplings of the matrix model, as well as the density of states, so as to match the physical correlators of interest. As we will show in Sec. III, we can do this for the JT-matter matrix model, where we fix the full ETH matrix model by matching to the planar correlation functions. We also provide some evidence for our conjecture that the resulting model reproduces the higher topology correlation functions as well.

### 1. ETH and topology

From the perspective of generalized eigenstate thermalization, that is, non-Gaussian ETH, various entropic factors $e^S$ (or equivalently the bookkeeping parameter $e^{S_0}$) are mere kinematical necessities imposed on us by the requirement of producing the correct (e.g., crossing-invariant) four- (and higher-) point functions at leading order; i.e., they arise from the need to compensate the multiple sums over Hilbert space occurring in expressions such as

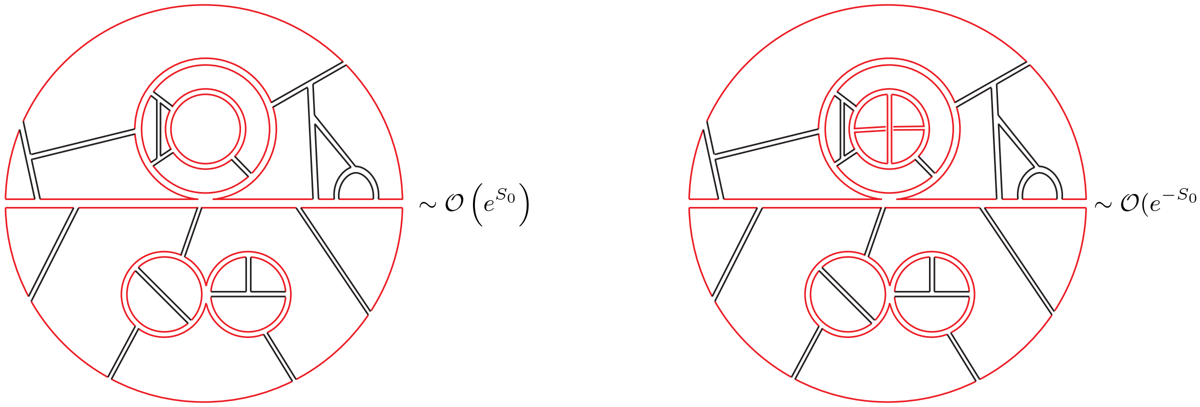



FIG. 1. Topology of the ETH matrix model. The left panel shows a planar diagram contributing to the two-point function of operators (2.20), which, un-normalized, gives an answer of order $e^{S_0}$. Any nonplanar diagram, exemplified by the diagram on the right, contributes at lower order in the $e^{S_0}$ expansion, the example above contributing at $\mathcal{O}(e^{-S_0})$. The order of an arbitrary diagram is determined by its Euler characteristic $e^{\chi S_0}$ when viewing its double-line representation as a triangulation of a 2D Riemann surface. This well-known feature of matrix models is the reason why non-Gaussian contributions in the matrix model action are needed for the ETH matrix model in order to reproduce basic facts even about thermal mean-field theory.

Eq. (2.6). On the other hand, the matrix-model description developed here gives an alternative interpretation in terms of topology. One may arrange the various matrix-model diagrams contributing to a given correlation function in a 't Hooft expansion and organize the powers of $e^{S_0}$ in terms of the topology shared by all diagrams at a given order (see Fig. 1). It is a classic result by t'Hooft [34] that the power of $e^{S_0}$ in a diagram is given by $e^{\chi S_0}$, where $\chi = 2 - 2g - b$ is the Euler characteristic of the surface on which the diagram can be drawn without self-intersections, $g$ is the genus (number of handles) of the surface, and $b$ is the number of boundaries. This is deduced from the ribbon-graph representation by the familiar formula $2 - 2g = V - E + F$ in terms of the number of vertices $V$, edges $E$, and faces $F$. This is, of course, a well-known feature of matrix models (and two-dimensional gravity), but its relation to the non-Gaussianities occurring in extensions of the eigenstate thermalization hypothesis [12] is useful and, to the best of our knowledge, new. Connections between topological expansion and quantum chaos have been described recently [8–10,35–39], largely motivated by the occurrence of so-called "wormhole" solutions contributing to the Euclidean quantum gravity path integral. We can incorporate this topological structure into the quantum chaos and ETH discussion in a natural way. As mentioned above, we define our generalized ETH ansatz using a single-trace matrix potential. Single-trace matrix models admit an interpretation where the physical degrees of freedom that live on the two-dimensional surfaces interact locally. The local nature of JT gravity minimally coupled to a free scalar field further motivates us to consider only single-trace matrix models.

### C. Testable predictions of the ETH matrix model

Having described our ETH matrix model as a generalization of the FK ansatz, we now comment on the testable predictions of our ansatz, which are related. As mentioned above, the index structure of a contribution to an ensemble-averaged product of matrix elements directly and uniquely determines the energy dependence of the expression, once the data that fix the specific form of the ansatz are supplied (namely, the $g^{(n)}$ functions or, equivalently, the $G^{(n)}$ functions). We provide examples of this using the following ensemble-averaged products:

$$\overline{\mathcal{O}_{12}\mathcal{O}_{21}} = e^{-S(\overline{E})} g_{\mathcal{O}}^{(2)}(E_1, E_2), \tag{2.23}$$

$$\overline{\mathcal{O}_{12}\mathcal{O}_{23}\mathcal{O}_{34}\mathcal{O}_{41}} = e^{-3S(\overline{E})} g_{\mathcal{O}}^{(4)}(E_1, E_2, E_3, E_4), \tag{2.24}$$

$$\overline{\mathcal{O}_{12}\mathcal{O}_{21}\mathcal{O}_{34}\mathcal{O}_{43}} = (e^{-S(\frac{E_1+E_2}{2})} g_{\mathcal{O}}^{(2)}(E_1, E_2))(e^{-S(\frac{E_3+E_4}{2})} g_{\mathcal{O}}^{(2)}(E_3, E_4)), \tag{2.25}$$

$$\overline{\mathcal{O}_{12}\mathcal{O}_{21}\mathcal{O}_{13}\mathcal{O}_{31}} = e^{-3S(\overline{E})} g_{\mathcal{O}}^{(4)}(E_1, E_2, E_1, E_3) + e^{-S(\frac{E_1+E_2}{2})} g_{\mathcal{O}}^{(2)}(E_1, E_2) e^{-S(\frac{E_1+E_3}{2})} g_{\mathcal{O}}^{(2)}(E_1, E_3), \tag{2.26}$$

$$\overline{\mathcal{O}_{12}\mathcal{O}_{23}\mathcal{O}_{33}\mathcal{O}_{31}} = e^{-3S(\overline{E})} g_{\mathcal{O}}^{(4)}(E_1, E_2, E_3, E_3), \tag{2.27}$$

$$\overline{\mathcal{O}_{11}\mathcal{O}_{11}\mathcal{O}_{11}\mathcal{O}_{11}} = 6e^{-3S(E_1)} g_{\mathcal{O}}^{(4)}(E_1, E_1, E_1, E_1) + 3e^{-2S(E_1)} g_{\mathcal{O}}^{(2)}(E_1, E_1) g_{\mathcal{O}}^{(2)}(E_1, E_1). \tag{2.28}$$

To be concrete, all of the above ensemble-averaged products could be evaluated in a given chaotic theory by averaging over some couplings of the model. The first two lines above are used to determine the $g^{(n)}$ functions that fix the ansatz. The remaining lines are predictions of the ansatz for the statistics of the matrix elements. In the language of matrix models, general correlators are built out of connected correlators. Furthermore, note that the factorized contribution in Eq. (2.28) dominates over the unfactorized one in the thermodynamic limit, where $S(E_1)$ becomes parametrically large [18] [the same is true in Eq. (2.26)]. The further corrections to Eq. (2.28) from handle contributions to ‘t Hooft diagrams begin at order $e^{-4S(E_1)}$. These are, in principle, determined by the $g^{(n)}$ functions but, in practice, are hard to explicitly compute (see Sec. IV F for some progress on nondisk ‘t Hooft diagrams). Thus, for systems away from the thermodynamic limit, our ansatz makes a nontrivial prediction for the suppressed contributions to ensemble-averaged matrix elements, such as the fourth moment of $\mathcal{O}_{11}$, Eq. (2.28). In the language of Refs. [12,18], the fourth moment (2.28) corresponds to a noncrossing partition and is expected to contribute to thermal correlators at leading order. However, because we have chosen to present a model with an $\mathcal{O} \to -\mathcal{O}$ symmetry, the terms in Eq. (2.28) make only an $e^{-S(E_1)}$ contribution to thermal correlators.

An ansatz that applies away from the thermodynamic limit is more suitable for numeric tests. The recent numerical study [40] was able to observe the factorization in Eq. (2.26) as well as the suppression of Eq. (2.28) in the large-system-size limit. It would also be interesting to probe the energy dependence of the suppressed contributions of Eq. (2.28), which are measurable at finite system size.

In the later sections, we test our model against TMFT and JT gravity with matter, which are dual to the SYK model in the appropriate regimes. It would be interesting to numerically test our ansatz against the SYK model outside of these regimes, as in Ref. [41].

## III. EXAMPLE 1: THERMAL MEAN-FIELD THEORY

As a simple example, we give a matrix model description of a TMFT, in which correlators factorize into two-point functions. An example of this is the generalized free field arising in the semiclassical limit $G_N \to 0$ of JT gravity coupled to a free massive scalar field, which we will consider in the next section. However, the discussion in this section is more general and applies to TMFT with an arbitrary two-point correlator and an arbitrary density of states.

Thermal mean-field theory is determined by its two-point correlator

$$\mathcal{G}_\beta(\tau) = \frac{1}{Z(\beta)} \mathrm{Tr} e^{-\beta H} \mathcal{O}(\tau)\mathcal{O}(0) \tag{3.1}$$

$$= \int_{-\infty}^{\infty} d\omega e^{\tau\omega} e^{-\frac{\beta}{2}\omega} \hat{\mathcal{G}}_\beta(\omega). \tag{3.2}$$

In the second line, we defined the Laplace transform $\hat{\mathcal{G}}_\beta(\omega)$. The integral over $\omega$ is presumed to converge for $0 < \tau < \beta$, which implies that $\hat{\mathcal{G}}_\beta(\omega)$ must decay sufficiently rapidly at large $\omega$: $\hat{G}_\beta(\omega) \lesssim \# e^{-(\beta/2)|\omega|}, |\omega| \to \infty$. In Eq. (3.2), we introduced an explicit factor $e^{-(\beta/2)\omega}$, such that the KMS condition takes a simple form

$$\text{KMS}: \ \mathcal{G}_\beta(\tau) = \mathcal{G}_\beta(\beta - \tau) \Leftrightarrow \hat{\mathcal{G}}_\beta(\omega) = \hat{\mathcal{G}}_\beta(-\omega). \tag{3.3}$$

The four-point correlator in TMFT factorizes into two-point functions

$$\mathcal{G}_\beta^{(4)}(\tau_1, \ldots, \tau_4) = \frac{1}{Z(\beta)} \mathrm{Tr} e^{-\beta H} \mathcal{O}(\tau_1) \ldots \mathcal{O}(\tau_4) \tag{3.4}$$

$$= \mathcal{G}_\beta(\tau_{12})\mathcal{G}_\beta(\tau_{34}) + \mathcal{G}_\beta(\tau_{14})\mathcal{G}_\beta(\tau_{23}) + \mathcal{G}_\beta(\tau_{13})\mathcal{G}_\beta(\tau_{24}), \tag{3.5}$$

where $\tau_{ij} = \tau_i - \tau_j$ and we assume $\beta > \tau_1 > \tau_2 > \tau_3 > \tau_4 > 0$. Similarly, higher-point correlators are given by a sum of all Wick contractions.

In order to discuss the matrix model for TMFT, it is convenient to first write the correlators in the energy basis,

$$\mathcal{G}_\beta(\tau) = \frac{1}{Z(\beta)} \int dE_1 dE_2 \rho(E_1)\rho(E_2) e^{-(\beta-\tau)E_1 - \tau E_2} \langle |\mathcal{O}_{E_1E_2}|^2 \rangle, \tag{3.6}$$

$$\mathcal{G}_\beta^{(4)}(\tau_1, \ldots, \tau_4) = \frac{1}{Z(\beta)} \int \prod_{j=1}^{4} (dE_j \rho(E_j) e^{-\beta_j E_j}) \times \langle \mathcal{O}_{E_1E_2}\mathcal{O}_{E_2E_3}\mathcal{O}_{E_3E_4}\mathcal{O}_{E_4E_1} \rangle, \tag{3.7}$$

where

$$\beta_1 = \beta - \tau_{14}, \quad \beta_2 = \tau_{12}, \quad \beta_3 = \tau_{23}, \quad \beta_4 = \tau_{34}. \tag{3.8}$$

The integrals over energies are performed with the density of states $\rho(E)$, which, of course, depends on the particular system and should be given in addition to the two-point function to define TMFT. In particular, the partition function is given by the usual expression

$$Z(\beta) = \int dE\rho(E)e^{-\beta E}. \tag{3.9}$$

We work in the thermodynamic limit. This could be either a large volume limit or, more generally, a large number of degrees of freedom, while keeping the temperature fixed. In holographic systems, such as JT gravity, this corresponds to the semiclassical gravity limit $G_N \to 0$. The average energy of the system in the thermodynamic limit is large, typically proportional to the volume or the number of degrees of freedom. In practice, this means that energy integrals, such as Eqs. (3.6), (3.7), and (3.9), are dominated by large energies $E_j$, while the differences are much smaller, $E_j \gg |E_i - E_k|$.

For example, the partition function is dominated by the saddle energy determined by the usual thermodynamic relation

$$S'(E) = \beta \Rightarrow E = E(\beta), \tag{3.10}$$

where $\rho(E) = e^{S(E)}$. Including fluctuations around the saddle, we have

$$Z(\beta) \approx e^{S(E(\beta))-\beta E(\beta)} \int d\omega e^{S''(E)\frac{\omega^2}{2}} \tag{3.11}$$

$$= e^{S(E(\beta))-\beta E(\beta)} \int d\omega e^{-\frac{\beta^2}{C}\frac{\omega^2}{2}} \tag{3.12}$$

$$= \frac{\sqrt{2\pi C}}{\beta} e^{S(E(\beta))-\beta E(\beta)}, \tag{3.13}$$

where we used the standard thermodynamic relation $S''(E) = (dT/dE)(d/dT)S'(E) = -(1/CT^2)$ and $C = (dE/dT)$ is the heat capacity. In the thermodynamic limit, the heat capacity $C$ is typically proportional to the volume or number of degrees of freedom; therefore, $C \to \infty$. Similarly, further non-Gaussian fluctuations are suppressed.

### A. Two-point function

Now, we consider Eqs. (3.6), (3.7), and higher-point correlators in TMFT. Given the two-point function (3.2), we determine the correlators of $\mathcal{O}$ in the energy basis. Then, we discuss the matrix model that computes these correlators.

We start with the two-point function. It turns out that

$$\langle |\mathcal{O}_{E_1E_2}|^2\rangle = \frac{1}{\rho(E)}\hat{\mathcal{G}}_{\beta(E)}(\omega), \tag{3.14}$$

where the inverse temperature is determined by the thermodynamic relation $\beta(E) = S'(E)$ and

$$E = \frac{E_1 + E_2}{2}, \qquad \omega = E_1 - E_2. \tag{3.15}$$

This result is valid in the thermodynamic limit. Let us check that it holds. We insert Eq. (3.14) into Eq. (3.6). We switch to coordinates $E$, $\omega$ and use

$$\rho(E_1)\rho(E_2) = e^{S(E+\frac{\omega}{2})+S(E-\frac{\omega}{2})} \tag{3.16}$$

$$= e^{2S(E)}\left(1 + O\left(\frac{\omega^2}{C}\right)\right), \tag{3.17}$$

where $C \gg 1$ is the heat capacity. Corrections in the last equation are vanishing in the thermodynamic limit. Together, we have

$$\frac{1}{Z(\beta)}\int dE_1dE_2\rho(E_1)\rho(E_2)e^{-(\beta-\tau)E_1-\tau E_2}\langle|\mathcal{O}_{E_1E_2}|^2\rangle \tag{3.18}$$

$$\approx \frac{1}{Z(\beta)}\int dE\rho(E)e^{-\beta E}\int_{-\infty}^{\infty} d\omega e^{\omega(\tau-\frac{\beta}{2})}\hat{\mathcal{G}}_{\beta(E)}(\omega) \tag{3.19}$$

$$\approx \int_{-\infty}^{\infty} d\omega e^{\omega(\tau-\frac{\beta}{2})}\hat{\mathcal{G}}_{\beta}(\omega) \tag{3.20}$$

$$= \mathcal{G}_\beta(\tau). \tag{3.21}$$

In the second equality, the integral over $E$ is computed by the saddle approximation, setting $\beta = S'(E)$.

### B. Four-point function

Now, we compute the four-point correlator (3.6). In the energy basis, it turns out to be

$$\langle\mathcal{O}_{E_1E_2}\mathcal{O}_{E_2E_3}\mathcal{O}_{E_3E_4}\mathcal{O}_{E_4E_1}\rangle = \langle|\mathcal{O}_{E_1E_2}|^2\rangle\langle|\mathcal{O}_{E_3E_4}|^2\rangle\left(\frac{\delta(E_1-E_3)}{\rho(E_1)} + \frac{\delta(E_2-E_4)}{\rho(E_2)}\right) \tag{3.22}$$

$$+\langle\mathcal{O}_{E_1E_2}\mathcal{O}_{E_2E_3}\mathcal{O}_{E_3E_4}\mathcal{O}_{E_4E_1}\rangle_c, \tag{3.23}$$

where the disconnected part is determined by Eq. (3.14), while the connected part is

$$\langle\mathcal{O}_{E_1E_2}\mathcal{O}_{E_2E_3}\mathcal{O}_{E_3E_4}\mathcal{O}_{E_4E_1}\rangle_c = (\hat{\mathcal{G}}(\omega_1)\hat{\mathcal{G}}(\omega_2)\hat{\mathcal{G}}(\omega_3)\hat{\mathcal{G}}(\omega_4))^{1/2}\frac{\delta(E_1+E_3-E_2-E_4)}{\rho(E)^3}, \tag{3.24}$$

where

$$E = \frac{1}{4}\sum_{j=1}^{4} E_j, \qquad \omega_1 = E_1 - E_2, \qquad \omega_2 = E_2 - E_3,$$
$$\omega_3 = E_3 - E_4, \qquad \omega_4 = E_4 - E_1. \tag{3.25}$$

Let us check that by inserting these expressions into Eq. (3.7), we obtain the TMFT four-point correlator (3.5). Each of the three terms in Eqs. (3.22) and (3.23) corresponds to the three terms in Eq. (3.7), respectively. The idea of the computation is to change coordinates $E_j$ to $E, \omega_j$,

$$dE_1 dE_2 dE_3 dE_4 = dE d\omega_1 d\omega_2 d\omega_3 \tag{3.26}$$

$$= \delta(\omega_1 + \omega_2 + \omega_3 + \omega_4) dE d\omega_1 d\omega_2 d\omega_3 d\omega_4. \tag{3.27}$$

We also have

$$E_1 = E + \frac{2\omega_1 + \omega_2 - \omega_4}{4}, \quad E_2 = E + \frac{2\omega_2 + \omega_3 - \omega_1}{4}, \tag{3.28}$$

$$E_3 = E + \frac{2\omega_3 + \omega_4 - \omega_2}{4}, \quad E_4 = E + \frac{2\omega_4 + \omega_1 - \omega_3}{4}. \tag{3.29}$$

The entropies in $\rho(E_j) = e^{S(E_j)}$ are expanded around the average value $E$ to linear order in $\omega_j$, Further corrections are suppressed in the thermodynamic limit. For example,

$$\rho(E_1) \approx e^{S(E) + S'(E)\frac{2\omega_1 + \omega_2 - \omega_4}{4}} \left(1 + O\left(\frac{\omega_j^2}{C}\right)\right). \tag{3.30}$$

The integral over the average energy $E$ will then set $S'(E) = \beta$ as usual, and similarly for other entropic factors.

Putting it all together, we have, for the first term in Eq. (3.22),

$$\frac{1}{Z(\beta)} \int \prod_{j=1}^{4} (dE_j \rho(E_j) e^{-\beta_j E_j}) \langle |\mathcal{O}_{E_1 E_2}|^2 \rangle \langle |\mathcal{O}_{E_3 E_4}|^2 \rangle \frac{\delta(E_1 - E_3)}{\rho(E_1)} \tag{3.31}$$

$$\approx \frac{1}{Z(\beta)} \int dE \rho(E) e^{-\beta E} \int_{-\infty}^{\infty} \prod_{j=1}^{4} d\omega_j \delta(\omega_1 + \omega_2) \delta(\omega_3 + \omega_4) \hat{\mathcal{G}}_{\beta(E)}(\omega_1) \hat{\mathcal{G}}_{\beta(E)}(\omega_3) \tag{3.32}$$

$$\exp\left\{-(\beta + \beta_1)\frac{2\omega_1 + \omega_2 - \omega_4}{4} - \beta_2 \frac{2\omega_2 + \omega_3 - \omega_1}{4} - \beta_3 \frac{2\omega_3 + \omega_4 - \omega_2}{4} - \beta_4 \frac{2\omega_4 + \omega_1 - \omega_3}{4}\right\} \tag{3.33}$$

$$\approx \int_{-\infty}^{\infty} d\omega_1 d\omega_3 e^{(\tau_{12} - \frac{\beta}{2})\omega_1 + (\tau_{34} - \frac{\beta}{2})\omega_3} \hat{\mathcal{G}}_{\beta}(\omega_1) \hat{\mathcal{G}}_{\beta}(\omega_3) \tag{3.34}$$

$$= \mathcal{G}_{\beta}(\tau_{12}) \mathcal{G}_{\beta}(\tau_{34}). \tag{3.35}$$

In the second equality, we computed the $E$ integral by saddle approximation and therefore substituted $\beta(E)$ by $\beta$.

The other two terms in Eqs. (3.22) and (3.23) are computed similarly to above and give the other two terms in Eq. (3.5). In particular, note that the connected part of the correlator (3.24) gives the "crossed" term $\mathcal{G}_\beta(\tau_{13})\mathcal{G}_\beta(\tau_{24})$.

### C. Semiclassical limit of JT gravity

An example of TMFT arises as a semiclassical limit $G_N \to 0$ of JT gravity coupled to a free scalar field—usually called the generalized free field (GFF). JT gravity will be discussed in more detail in Sec. IV.

The two-point function of GFF, briefly discussed in Sec. II, is

$$\mathcal{G}_\beta(\tau) = \left(\frac{\pi/\beta}{\sin(\pi\tau/\beta)}\right)^{2\Delta}. \tag{3.36}$$

The Laplace transform is

$$\hat{\mathcal{G}}_\beta(\omega) = \frac{(2\pi/\beta)^{2\Delta - 1}}{2\pi\Gamma(2\Delta)} \Gamma\left(\Delta \pm \frac{i\beta}{2\pi}\omega\right). \tag{3.37}$$

TMFT correlators with this particular form of the two-point function arise in JT gravity in the semiclassical limit due to [42]

$$\frac{\Gamma(\Delta \pm i\sqrt{E_1} \pm i\sqrt{E_2})}{\Gamma(2\Delta)} \approx \frac{1}{\rho(E)} \frac{(2\sqrt{E})^{2\Delta - 1}}{2\pi\Gamma(2\Delta)} \Gamma\left(\Delta \pm \frac{i}{2\sqrt{E}}\omega\right), \tag{3.38}$$

where "±" in the lhs means a product of four gamma functions for all choices of signs. We defined $E = (E_1 + E_2/2), \omega = E_1 - E_2$. The semiclassical limit [43] corresponds to $E_1, E_2 \gg |E_1 - E_2|$, and the semiclassical density of states is

$$\rho(E) \approx \frac{1}{(2\pi)^2} e^{2\pi\sqrt{E}}. \tag{3.39}$$

To obtain the semiclassical limit of the four-point and higher correlators in JT gravity, we need the limit of the so-called $sl(2,\mathbb{R})$ $6j$ symbol. In Appendix C of Ref. [11], we show that

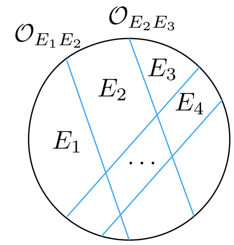
$$\begin{Bmatrix} \Delta & \sqrt{E_1} & \sqrt{E_2} \\ \Delta & \sqrt{E_3} & \sqrt{E_4} \end{Bmatrix} \approx \frac{\delta(E_1+E_3-E_2-E_4)}{\rho(E)}, \qquad E = \frac{1}{4}\sum_{j=1}^{4} E_j, \tag{3.40}$$

which is, again, in the limit $E_j \gg |E_i - E_k|$.

### D. Higher TMFT correlators

Higher correlators in a general TMFT can also be written in the energy basis. It turns out that they are computed by the so-called "chord diagrams" that we summarize below, which can be understood as a semiclassical limit of JT gravity with matter, where the exact correlators are also computed by chord diagrams (see Sec. II in Ref. [11]). Regardless, given the answer below, we can check that the chord diagrams simply compute the TMFT Wick contractions.

A given $n$-point correlator can be written in the energy basis as a sum over chord diagrams,

$$\langle \mathcal{O}_{E_1E_2}\mathcal{O}_{E_2E_3}\ldots\mathcal{O}_{E_nE_1}\rangle =$$ 

$$+ \quad \ldots \tag{3.41}$$

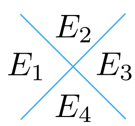
In the rhs, we sum over all possible chord diagrams, where all operators are connected pairwise. Each diagram, of course, corresponds to the particular Wick contraction in TMFT. However, the chord diagram contains more information. Each chord diagram splits the interior of the disc into regions, to each of which we assign an energy parameter $E_j$. Then, we write a formula for each diagram according to the following rules.

(i) For each matrix element $\mathcal{O}_{E_1E_2}$, we write a factor

$$\left(\frac{1}{\rho(E)}\hat{\mathcal{G}}_{\beta(E)}(\omega)\right)^{1/2}, \tag{3.42}$$

where $E = (E_1 + E_2/2), \omega = E_1 - E_2$.

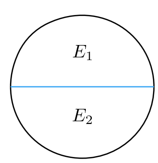
(ii) For each intersection of chords, we write



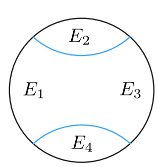
$$= \frac{\delta(E_1+E_3-E_2-E_4)}{\rho(E)}, \qquad E = \frac{1}{4}\sum_{j=1}^{4} E_j. \tag{3.43}$$

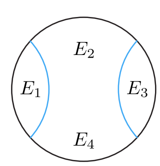
(iii) Each energy $E_j$, if any, that is not adjacent to the boundary is integrated over with the density of states $\rho(E_j)$.

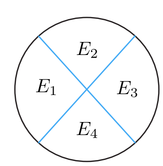
For example, the two-point and four-point functions that we already discussed, Eqs. (3.14) and (3.23), correspond to the diagrams

$$\langle \mathcal{O}_{E_1E_2}\mathcal{O}_{E_2E_1}\rangle =$$ 

$$= \frac{1}{\rho(E)}\hat{\mathcal{G}}_{\beta(E)}(\omega) \tag{3.44}$$

$$\langle \mathcal{O}_{E_1E_2}\mathcal{O}_{E_2E_3}\mathcal{O}_{E_3E_4}\mathcal{O}_{E_4E_1}\rangle =$$ 

$+$ 

$+$ 

(3.45)

$$= \frac{1}{\rho(E)^2}(\hat{\mathcal{G}}_{\beta(E)}(\omega_1)\hat{\mathcal{G}}_{\beta(E)}(\omega_2)\hat{\mathcal{G}}_{\beta(E)}(\omega_3)\hat{\mathcal{G}}_{\beta(E)}(\omega_4))^{1/2} \tag{3.46}$$

$$\left(\frac{\delta(E_1-E_3)}{\rho(E_1)} + \frac{\delta(E_2-E_4)}{\rho(E_2)} + \frac{\delta(E_1+E_3-E_2-E_4)}{\rho(E)}\right). \tag{3.47}$$

In the four-point function, we use the fact that, in the thermodynamic limit at large energies, we can approximate $\rho(E_1)...\rho(E_4) \approx \rho(E)^4$. In the first two diagrams for the four-point function, we also introduce the energies for the regions in a redundant way at the cost of a delta function. Higher correlators can be computed in a similar manner.

We now finally turn to the matrix potential computing the same correlators. We will use the chord diagram results discussed here in order to find the potential of the matrix model.

### E. TMFT matrix model

Now, we write down the matrix model describing the TMFT correlators with an arbitrary two-point function $\hat{\mathcal{G}}_\beta(\omega)$ and a density of states $\rho(E) = e^{S(E)}$. We consider the matrix potential for the operator $\mathcal{O}_{ab}$ written in the energy basis,

$$V(R) = N\left(\frac{1}{2}\sum_{ab} g_{ab}|R_{ab}|^2 + \frac{1}{4}\sum_{abcd} g_{abcd}R_{ab}R_{bc}R_{cd}R_{da}\right), \tag{3.48}$$

$$g_{ab} = 2\sinh^2\frac{S'(\overline{E})\omega}{2}, \tag{3.49}$$

$$g_{abcd} = \frac{1}{2}\frac{\delta(E_a - E_c)}{\rho(E_a)} + \frac{1}{2}\frac{\delta(E_b - E_d)}{\rho(E_b)} - \frac{\delta(E_a + E_c - E_b - E_d)}{\rho(E)}, \tag{3.50}$$

where $\bar{E} = (E_a + E_b/2), \omega = E_a - E_b, E = \frac{1}{4}(E_a + E_b + E_c + E_d)$, and $\beta(\bar{E}) = S'(\bar{E})$ is the inverse temperature corresponding to the energy $\bar{E}$. The matrix $R_{ab}$ is a rescaling of $\mathcal{O}_{ab}$,

$$\mathcal{O}_{ab} \coloneqq \frac{(\hat{\mathcal{G}}_{\beta(\bar{E})}(\omega))^{1/2}}{\rho(\bar{E}))} R_{ab}. \tag{3.51}$$

With this rescaling, the exact (planar) two-point function is $\langle |R_{ab}|^2\rangle_{\text{disk}} = 1$; see Eq. (3.14).

Several comments are in order. First, the matrix potential, Eqs. (3.48)–(3.50), is written in the thermodynamic limit. This is consistent with the fact that the TMFT correlators considered in the previous subsections are valid in the thermodynamic limit.

Second, in addition to the matrix potential for $\mathcal{O}$ above, one must add a potential for the Hamiltonian $V(H)$, whose role is to set the correct density of states $\rho(E)$. In this particular example, where we work in the thermodynamic limit, the only effect of the density of states on the correlators is to set the relation between the temperature and entropy in a standard way, $\beta = S'(\bar{E})$. This case was manifest in the computation of TMFT correlators in Secs. III A and III B. Similarly, in the case of TMFT, the relevant dependence of matrix model couplings $g_{ab}, g_{abcd}$ on energies is within a microcanonical window, again due to the thermodynamic limit. This is not always true in more general ETH matrix models. For example, in JT gravity, the dependence of both the density of states and matrix model couplings on a wide range of energies, far outside of any microcanonical window, is important.

Furthermore, the matrix potential, Eqs. (3.48)–(3.50), is of the general form (2.18) in the following sense. The potential, Eqs. (3.48)–(3.50), is written in the thermodynamic limit. To properly define it and to be able to use it in computations, we need to back away from the thermodynamic limit. When we do this "regularization," we expect that the delta functions in the quartic interaction (3.50) must be smeared and become smooth functions of the energies. Therefore, once we do this regularization, the potential becomes of the general form (2.18), with the couplings $g_{ab}, g_{abcd}$ given by smooth functions of energies.

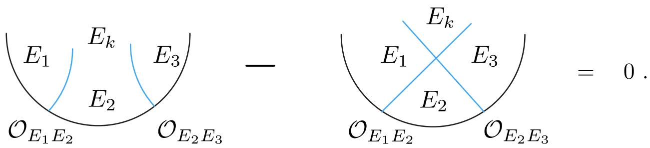
The thermodynamic limit discussed above is analogous to the double-scaling limit of matrix models; e.g., see Refs. [44,45]. In the latter case, one tunes the couplings of the matrix model to zoom into the edge of the spectrum. In the thermodynamic limit, we tune the couplings to zoom into a microcanonical window somewhere in the middle of the spectrum. In both cases, to compute the correlation functions, we must first work away from the double-scaling or thermodynamic limit. Then, at large $N$, we can compute the sum over all planar t'Hooft diagrams, including loops. After resumming the loops, we can take the double-scaling or thermodynamic limit. On the other hand, if we try to work directly with the double-scaling or thermodynamic limit of the potential, we would find that individual loop diagrams diverge, simply because the energy spectrum is unbounded in the double-scaling or thermodynamic limit. All such divergences cancel once we resum all loops.

The reader might wonder how we could possibly guess the potential in Eqs. (3.48)–(3.50). There are several arguments. One, perhaps convoluted, argument is that the potential above arises as a semiclassical ($G_N \to 0$) limit of the matrix model holographically dual to JT gravity coupled to a free scalar. This case is discussed in detail in the companion paper [11].

More directly, the potential can be understood as follows. In our search for the matrix model, we attempt to find a matrix $\mathcal{O}_{E_1E_2}$ that approximates a TMFT operator $\mathcal{O}$. However, in TMFT, the operator $\mathcal{O}$ is not completely random, and it satisfies certain constraints. In particular, TMFT operators at different Euclidean times commute inside correlators, for example, $\langle \text{Tr}e^{-\beta H}\mathcal{O}(\tau_1)\mathcal{O}(\tau_2)\mathcal{O}(\tau_3) \times \mathcal{O}(\tau_4)\rangle_\beta = \langle \text{Tr}e^{-\beta H}\mathcal{O}(\tau_2)\mathcal{O}(\tau_1)\mathcal{O}(\tau_3)\mathcal{O}(\tau_4)\rangle_\beta$. It turns out that, in the energy basis, this implies

$$\int dE_2\,\rho(E_2)\left(\frac{\delta(E_2-E_k)}{\rho(E_2)}-\frac{\delta(E_1+E_3-E_2-E_k)}{\rho(E)}\right)\times\left(R_{E_1E_2}R_{E_2E_3}-\frac{\delta(E_1-E_3)}{\rho(E_1)}\right)=0, \tag{3.52}$$

where $E=\frac{1}{4}(E_1+E_3+E_2+E_k)$. This equation holds inside TMFT correlation functions. Pictorially, we can express the constraint as




(3.53)

This is an operator identity that holds inside correlation functions, so the top of the diagram is filled with the other operators in the correlator. The first term $\delta(E_2-E_k)/\rho(E_2)$ leaves the operators intact, while the role of the second term $\delta(E_1+E_3-E_2-E_k)/\rho(E)$ is to exchange them. This interpretation is correct inside correlation functions except for the contraction $\langle R_{E_1E_2}R_{E_2E_3}\rangle$. Therefore, we explicitly subtract it by writing $R_{E_1E_2}R_{E_2E_3}-[\delta(E_1-E_3)/\rho(E_1)]$.

Let us check that Eq. (3.52) is indeed true. The average of the constraint itself is trivially zero since

$$\langle R_{E_1E_2}R_{E_2E_3}\rangle=\frac{\delta(E_1-E_3)}{\rho(E_1)}. \tag{3.54}$$

More nontrivially, we can multiply Eq. (3.52) by $R_{E_3E_4}R_{E_4E_1}$ and compute the correlator

$$\int dE_2\rho(E_2)\left(\frac{\delta(E_2-E_k)}{\rho(E_2)}-\frac{\delta(E_1+E_3-E_2-E_k)}{\rho(E)}\right) \tag{3.55}$$

$$\left\langle\left(R_{E_1E_2}R_{E_2E_3}-\frac{\delta(E_1-E_3)}{\rho(E_1)}\right)R_{E_3E_4}R_{E_4E_1}\right\rangle. \tag{3.56}$$

Using Eq. (3.47), this is equivalent to

$$\int dE_2\,\rho(E_2)\left(\frac{\delta(E_2-E_k)}{\rho(E_2)}-\frac{\delta(E_1+E_3-E_2-E_k)}{\rho(E_{123k})}\right) \tag{3.57}$$

$$\left(\frac{\delta(E_2-E_4)}{\rho(E_2)}+\frac{\delta(E_1+E_3-E_2-E_4)}{\rho(E_{1234})}\right)=0, \tag{3.58}$$

where $E_{1234}=\frac{1}{4}(E_1+E_2+E_3+E_4)$. This is straightforward to check. The only subtlety is the product of second terms in both brackets. For that term, one has to use an identity $\rho(E_1+E_3-E_4)/\rho(E_1+E_3/2)^2\approx\frac{1}{\rho(E_4)}$ that holds in the thermodynamic limit. Higher correlators can be checked in a similar manner.

Now, the reason that we discuss the constraint (3.52) is that a randomly chosen finite-dimensional matrix $\mathcal{O}_{E_1E_2}$ does not obey such constraints. Thus, we need to somehow impose it in the matrix model. One way to do that is to write a potential

$$V(R)=N|\text{constraint}|^2, \tag{3.59}$$

where $N$ is the size of the matrix. Then, in the thermodynamic limit $N\to\infty$, the constraint is imposed since the measure of the matrix model is $e^{-V(R)}$. It is straightforward to compute the square of the constraint (3.52),

$$\int dE_1dE_3dE_k\rho(E_1)\rho(E_3)\rho(E_k) \tag{3.60}$$

$$\int dE_2\rho(E_2)\left(\frac{\delta(E_2-E_k)}{\rho(E_2)}-\frac{\delta(E_1+E_3-E_2-E_k)}{\rho(E)}\right)\times\left(R_{E_1E_2}R_{E_2E_3}-\frac{\delta(E_1-E_3)}{\rho(E_1)}\right) \tag{3.61}$$

$$\int dE_4\rho(E_4)\left(\frac{\delta(E_4-E_k)}{\rho(E_4)}-\frac{\delta(E_1+E_3-E_4-E_k)}{\rho(E)}\right)\times\left(R_{E_1E_4}R_{E_4E_3}-\frac{\delta(E_1-E_3)}{\rho(E_1)}\right)^*. \tag{3.62}$$

There is actually a constraint (3.52) for any $E_1, E_3, E_k$, so we include all of them by integrating over $E_1, E_3, E_k$. Note that $(R_{E_1E_4}R_{E_4E_3})^*=R_{E_3E_4}R_{E_4E_1}$, so the quartic term has the correct structure appearing in Eq. (3.48). A straightforward computation shows that Eq. (3.62) produces the potential (3.48) up to terms that do not depend on $R$.

Regardless of how we derived the potential (3.48), we can take it as given and study whether it is consistent with the correlation functions in TMFT. In the companion paper [11], we study the generalization of Eq. (3.48) to JT gravity coupled to a scalar field and derive certain Schwinger-Dyson equations that relate the correlators $\langle\mathcal{O}_{E_1E_2}\mathcal{O}_{E_2E_3}\ldots\rangle$ to the potential. Namely, in Ref. [11], we consider Schwinger-Dyson equations of the form

$$\int d\mu[\mathcal{O}]\frac{\partial}{\partial\mathcal{O}_{ab}}(\mathcal{O}_{ab}e^{-V(\mathcal{O})})=0, \tag{3.63}$$

$$\int d\mu[\mathcal{O}]\frac{\partial}{\partial\mathcal{O}_{ad}}(\mathcal{O}_{ab}\mathcal{O}_{bc}\mathcal{O}_{cd}e^{-V(\mathcal{O})})=0. \tag{3.64}$$

For example, for the quartic potential (3.48), the first Schwinger-Dyson equation gives

$$\langle R_{ab}R_{ba}\rangle=\frac{1}{N}g_{ab}^{-1}-g_{ab}^{-1}\sum_{cd}g_{abcd}\langle R_{ab}R_{bc}R_{cd}R_{da}\rangle, \tag{3.65}$$

where we used the rescaled matrix $R_{ab}$ instead of $\mathcal{O}_{ab}$, Eq. (3.51). In Ref. [11], we showed that Schwinger-Dyson equations (3.63) and (3.64) are indeed satisfied. This automatically implies that these equations are also satisfied in the TMFT case since the latter is just the semiclassical $G_N \to 0$ limit of JT gravity. We do not repeat that analysis here and refer the interested reader to Ref. [11] for details. One can also consider generalizations of Eqs. (3.63) and (3.64) with more insertions of $\mathcal{O}_{ab}$, though we did not do so.

## IV. EXAMPLE 2: JT + MATTER MATRIX MODEL

### A. Matter correlation functions in JT gravity

Let us now turn to our main application of the structure we defined above, namely, the matrix model description of JT gravity coupled to a scalar field. The theory is defined via the action

$$S_{\rm JT} = -S_0\chi - \frac{1}{2}\int_{\mathcal{M}} \sqrt{g}\phi(R+2) - \int_{\partial\mathcal{M}} \sqrt{h}\phi(K-1) + S_{\rm m}, \tag{4.1}$$

with the matter action

$$S_{\rm m} = \frac{1}{2}\int_{\mathcal{M}} \sqrt{g}(g^{ab}\partial_a\varphi\partial_b\varphi + m^2\varphi^2). \tag{4.2}$$

Here, $\chi$ is the Euler number of the two-manifold $\mathcal{M}$ over which the Lagrangian density is integrated, $\phi$ is the dilaton field that forms part of the definition of the two-dimensional gravity theory under study, while $\varphi$ is an additional matter scalar field of mass $m$. The term integrated over the boundary of the manifold $\partial\mathcal{M}$ contains a Gibbons-Hawking-York term in the form of the integrated extrinsic curvature $K$ as well as a boundary cosmological constant. Without the addition of the matter scalar $\varphi$, this theory has been shown to be equivalent to a matrix model [8], while here (see also the longer companion paper [11]), we generalize this to the theory including the matter field.

We start by laying out a few useful facts and computations in the theory above. All expressions quoted here can be obtained by using a convenient set of Feynman rules, developed in Ref. [46] and reviewed in Ref. [11]. First, the inclusion of the Euler number in the action means that amplitudes have an expansion in the topology of the manifold $\mathcal{M}$, which is used to calculate them gravitationally.

We start with the leading order, where the topology is that of a disk. At disk level, the density of states of JT gravity is given by

$$\rho_0(E)dE = \frac{1}{2\pi^2}\sinh(2\pi\sqrt{E})dE = \frac{1}{\pi^2}s\,\sinh(2\pi s)ds, \tag{4.3}$$

where we have defined the variable $s^2 = E$.

#### 1. Disk correlation functions

Turning now to correlation functions [47], we focus on the two-point function, which at disk level reads

$$\langle {\rm Tr}\, e^{-\beta H}\mathcal{O}(\tau)\mathcal{O}(0)\rangle_{\rm disk} = e^{S_0}\int_0^\infty ds_1 ds_2 \rho(s_1)\rho(s_2)e^{-(\beta_1 s_1^2+\beta_2 s_2^2)}\Gamma_{12}^{\Delta}. \tag{4.4}$$

Here, $\beta_1 = \tau, \beta_2 = \beta - \tau$, and we have also introduced the short-hand notation

$$\Gamma_{12}^{\Delta} \coloneqq \frac{\Gamma(\Delta \pm is_1 \pm is_2)}{\Gamma(2\Delta)} = \Gamma(2\Delta)^{-1}\Gamma(\Delta+is_1+is_2)\Gamma(\Delta+is_1-is_2)\Gamma(\Delta-is_1+is_2)\Gamma(\Delta-is_1-is_2). \tag{4.5}$$

Then, we can express the four-point function as

$$\langle {\rm Tr} e^{-\beta H}\mathcal{O}(\tau_1)\ldots\mathcal{O}(\tau_4)\rangle_{\rm disk} = e^{S_0}\int_0^\infty \prod_{i=1}^4 (ds_i\rho(s_i)e^{-\beta_i s_i^2})(\Gamma_{12}^{\Delta}\Gamma_{23}^{\Delta}\Gamma_{34}^{\Delta}\Gamma_{41}^{\Delta})^{1/2}\left(\frac{\delta(s_1-s_3)}{\rho(s_1)} + \frac{\delta(s_2-s_4)}{\rho(s_2)} + \begin{Bmatrix} \Delta & s_1 & s_2 \\ \Delta & s_3 & s_4 \end{Bmatrix}\right). \tag{4.6}$$

The third term in this expression uses the bracket notation for the $6j$ symbol of the $\mathfrak{sl}(2,\mathbb{R})$ algebra. Expressions for any higher-point functions can be obtained (see Refs. [11,46]), but the two- and four-point functions reviewed above shall suffice for present purposes. However, we need to consider correlation functions on geometries of different topology, starting with the so-called "double trumpet" [8], which is topologically a cylinder.

#### 2. Direct match to JT + matter theory

Let us proceed somewhat naively to determine the couplings of the ETH matrix model describing matter-coupled JT gravity. In fact, the answers we arrive at are correct, but more machinery is needed to justify them properly, such as an understanding of how to obtain them in a double-scaling limit. This and other issues will be discussed in the remainder of this paper, including studying the theory on higher topologies, i.e., away from the disk level.

We start by matching the disk density of states,

$$\langle \mathrm{Tr} e^{-\beta H} \rangle = e^{S_0}\rho_0(E) = e^{S_0}\sinh(2\pi\sqrt{E}), \quad (4.7)$$

which involves choosing the potential $V$ in Eq. (2.14) appropriately. Computing the full disk two-point function in the matrix model, as in Eq. (2.20), leads to the identification

$$\overline{\mathcal{O}_{ab}\mathcal{O}_{ba}} = e^{-S_0}\frac{\Gamma(\Delta \pm i\sqrt{E_a} \pm i\sqrt{E_b})}{\Gamma(2\Delta)} \coloneqq e^{-S_0}\Gamma^{\Delta}_{ab}. \quad (4.8)$$

In an analogous fashion, one can determine the quartic non-Gaussianity of the ETH matrix model for JT + matter. The computation is performed in Ref. [11] and leads to the identification

$$e^{-3S(\overline{E})}g^{(4)}_{\mathcal{O}}(E_a, E_b, E_c, E_d) = e^{-3S_0}(\Gamma^{\Delta}_{ab}\Gamma^{\Delta}_{bc}\Gamma^{\Delta}_{cd}\Gamma^{\Delta}_{da})^{1/2}\begin{Bmatrix} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{Bmatrix}. \quad (4.9)$$

In the companion paper, we describe a systematic procedure determining the full $\mathcal{O}$ potential in an expansion working in the number of "bulk" line crossings.

### 3. Double-trumpet correlation functions

We now include the scalar one-loop determinant in the otherwise "empty" double trumpet; that is, we consider the effect of the scalar field on the trumpet absent any explicit insertions of the $\mathcal{O}$ operator at the boundary,

$$\langle \mathrm{Tr} e^{-\beta_L H}\mathrm{Tr} e^{-\beta_R H}\rangle = \int_0^\infty db b Z_{\mathrm{tr}}(\beta_L, b) Z_{\mathrm{tr}}(\beta_R, b) Z_{\mathrm{scalar}}(b). \quad (4.10)$$

In other words, it differs from the usual JT double trumpet only by the inclusion of $Z_{\mathrm{scalar}}(b)$. The factors $Z_{\mathrm{tr}}(\beta_R, b)$ are given by the trumpet partition function for a geometry characterized by an asymptotically $\mathrm{AdS}_2$ boundary of length $\beta$ and a geodesic boundary of length $b$ in the interior. To obtain the explicit expression $Z_{\mathrm{tr}}(\beta_R, b) = (1/2\sqrt{\pi\beta})e^{-(b^2/4\beta)}$, one integrates over the fluctuations of the asymptotic boundary weighted by the Schwarzian measure [8], giving the one-loop exact result [52] we quoted. In the companion paper, we show how to evaluate the scalar determinant in several different ways, useful for various different points of view. One can show that

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

$$= 1 + \sum_{\Delta' \in \mathcal{S}}\frac{e^{-\Delta' b}}{1-e^{-b}} \quad (4.11)$$

$$= \exp\left(\sum_{w=1}^{\infty}\frac{e^{-wb\Delta}}{w(1-e^{-wb})}\right). \quad (4.12)$$

One important feature, apparent in all representations, is the presence of a UV divergence for $b \to 0$. This divergence is reproduced from our matrix model, in the double-scaling limit. The first expression above makes apparent the relation to the spectrum of conformal primary operators and their descendants; for example, the $n = 1$ term can be recognized as a sum over the primary state of dimension $\Delta$ together with all of its descendants. Similarly, the $n = 2$ term sums over the double-trace operators with dimensions $2\Delta + 2m$, together with all descendants. Analogous interpretations continue to hold for all higher $n$. In the last expression in Eq. (4.11), $\mathcal{S}$ is defined to be the set of scaling dimensions of all primary operators in the bosonic generalized free field (other than the identity). The $1/1 - e^{-b}$ factor may be expanded in a geometric series, which is associated with the descendants. The second formula is naturally related to the Selberg trace formula for the heat kernel of a scalar operator on the double trumpet, that is, as a sum over (multiply wound) primitive geodesics on this particular hyperbolic manifold.

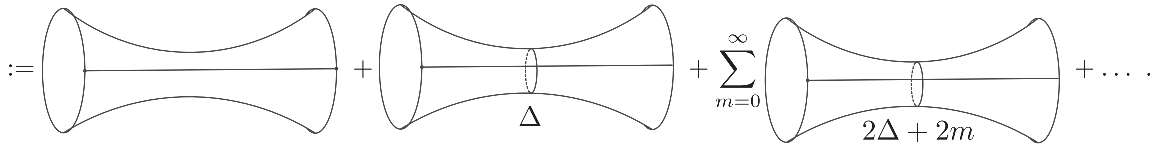
### 4. Double-trumpet two-point function

We next determine the two-point correlation function of the JT-matter theory, with an $\mathcal{O}$ insertion on each boundary. In the companion paper (see Appendix D of Ref. [11]), it is shown that the result arranges itself into a sum over terms corresponding to the structure present in the expansion (4.11) above. Focusing explicitly on the first three such contributions, we obtain

$$\langle \mathrm{Tr} e^{-\beta_L H}\mathcal{O}\mathrm{Tr} e^{-\beta_R H}\mathcal{O}\rangle_{\mathrm{cyl}} \quad (4.13)$$

$$= \int_0^\infty ds_L ds_R \rho(s_L)\rho(s_R) e^{-(\beta_L s_L^2 + \beta_R s_R^2)}(\Gamma^{\Delta}_{LL}\Gamma^{\Delta}_{RR})^{1/2} \quad (4.14)$$

$$\times\left(\frac{\delta(s_L - s_R)}{\rho(s_L)} + \begin{Bmatrix} \Delta & s_L & s_R \\ \Delta & s_R & s_L \end{Bmatrix} + \sum_{m=0}^{\infty}\begin{Bmatrix} 2\Delta + 2m & s_L & s_R \\ \Delta & s_R & s_L \end{Bmatrix} + \dots, \quad (4.15)$$




(4.16)

where $\Gamma_{LL}^{\Delta} = [\Gamma(\Delta)^2/\Gamma(2\Delta)]\Gamma(\Delta + is_L)\Gamma(\Delta - is_L)$, and analogously for $\Gamma_{RR}^{\Delta}$ in terms of $s_R$. In the expression above, the first term results from a simple propagation of an $\mathcal{O}$ line along a geodesic connecting the two boundary $\mathcal{O}$ insertions through the bulk. The second takes the form of a geodesic $\mathcal{O}$ propagation through the bulk, this time crossed by an $\mathcal{O}$ bulk loop along a closed geodesic, the intersection of boundary-to-boundary geodesic and bulk loop giving rise to the $6j$ symbol labeled by the scaling dimension $\Delta$, as shown. Finally, the third term shows the analogous process but now involving the sum over double traces of dimension $2\Delta + 2m$ in the loop. The higher contributions we omitted correspond, in a similar fashion, to the $n \geq 3$ terms in the expansion of the scalar one-loop determinant (4.11) above.

### 5. *Higher-genus correlation functions*

The general higher-genus, $n$-boundary geometry can be assembled via a pants decomposition of a genus-$g$ Riemann surface with $n$ geodesic boundaries glued to $n$ trumpet geometries. In the companion paper, we accumulate evidence that an arbitrary $\mathcal{O}$ $n$-point correlation function at genus $g$ can be computed as follows.

Let us first ignore the contribution to the correlator from the determinant of the scalar field. Then, the correlator is computed by summing over all geodesics that connect the $\mathcal{O}$ operators on the asymptotic boundaries, with a weighting of $e^{-\Delta\ell}$ for each geodesic, where $\ell$ is a renormalized length. These geodesic configurations are classified by their topologies. Each topology is represented by a set of lines drawn on the surface with no voluntary intersections. These lines divide the surface into subregions, and each subregion is characterized by the number of boundaries $n$ and the genus $g$. Each boundary of each subregion is labeled by an $s$ parameter. For every subregion, we should include a factor of $\langle\rho(s_1)\ldots\rho(s_n)\rangle_{g,n}$, which we define to be the inverse Laplace transform of $Z_{g,n}(\beta_1, \ldots, \beta_n)$, which is the path integral defined in Eq. (127) of Ref. [8]. For every intersection of two bulk lines, we include a factor of a $6j$ symbol that depends on the four adjacent $s$ parameters. An asymptotic AdS boundary with Euclidean length $\beta$ is assigned a factor of $e^{-\beta s^2}$. Finally, one should integrate over all of the $s$ parameters from 0 to $\infty$. For disk topologies, these Feynman rules become the rules of Ref. [46], which we have reviewed in Ref. [11].

To include the contribution from the scalar-field matter determinant, we also should sum over all ways of drawing closed geodesics on the spacetime. The rules for an intersection of two geodesics are still as above. The scaling dimensions assigned to the closed geodesics should take all values in $\mathcal{S}$, which was defined in Eq. (4.11). Finally, wherever there is a simple closed geodesic, we include the following function of its two adjacent $s$ parameters:

$$\int_0^\infty db\, b \frac{e^{-\Delta' b}}{1 - e^{-b}} \frac{2}{b}\cos(bs_1)\frac{2}{b}\cos(bs_2), \tag{4.17}$$

where $\Delta'$ is the dimension of the primary operator propagating on the closed geodesic. This ensures that the closed geodesic is weighted by $e^{-\Delta' b}/1 - e^{-b}$ in the moduli space integral.

The evidence we have for the above conjecture comes from our computation of the double-trumpet two-point function as well as a computation on the pair of pants, presented in Ref. [11]. See also Refs. [35,53,54], which discuss the two-point function on the disk with a handle.

## B. Matching to the ETH matrix model

We have now assembled all the data needed to fully specify the free coupling functions of the general ETH matrix model. This can be viewed alternatively as an exercise demonstrating the usefulness of the general ETH matrix model and the generalized ETH, or as an extension of the JT matrix model [8] to include a scalar field. The task at hand is to specify the higher-order coupling functions $G^{(n)}(E_a, \ldots E_n)$ in the ETH matrix model in Eq. (2.18). We may achieve this in two different ways [55].

(i) We use the nearly conformal invariance and locality of the JT-matter theory in order to constrain the potential nonlinearities directly. The leading nonlinear couplings of the ETH matrix model are given as a "constraint-squared" type potential. Roughly speaking, we translate the observation around Eq. (2.2) into the constraint

$$[\mathcal{O}(t_1), \mathcal{O}(t_2)] = \begin{cases} -\frac{2i\sin(\pi\Delta)}{|t_1 - t_2|^{2\Delta}}\operatorname{sign}(t_1 - t_2) & \Delta \notin \mathbb{Z}_{\geq 1} \\ \frac{2\pi i}{(2\Delta-1)!}(-1)^{\Delta-1}\delta^{(2\Delta-1)}(t_{12}) & \Delta \in \mathbb{Z}_{\geq 1}. \end{cases} \tag{4.18}$$

This constraint shows that the commutator is proportional to a c-number or, in other words, the unit operator. This constraint finds its representation in the matrix model as a specific non-Gaussianity, as we shall see.

(ii) We deduce the higher-order couplings by matching correlation functions to the gravity predictions at disk level (determined in Sec. IV A) and conjecture that the so-defined matrix model correctly predicts all higher-genus contributions. This is the analogue of the statement that in the single-matrix model of Ref. [8], the disk density of states determines the full matrix potential and thus the higher-genus amplitudes. In particular, we match away from the double-scaling limit to a suitably $q$-deformed version of the gravity amplitudes before taking the $q \to 1$ limit that recovers the original amplitudes.

(iii) In addition, whichever strategy we use in order to determine the $\mathcal{O}$ potential, we need to add counterterms to the $H$ potential in order to (re)adjust the disk density of states $\rho_0(E)$ to the desired form—for example, the $\sinh(2\pi\sqrt{E})$ behavior of JT gravity with matter. That this is needed can already be seen at the Gaussian level for the $\mathcal{O}$ statistics: In this case, we can directly integrate out the $\mathcal{O}$ matrix to obtain

$$\mathcal{Z}_{\mathrm{ETH}} = \int dH e^{-V_{\mathrm{SSS}}(H) - \tilde{V}(H)}, \qquad (4.19)$$

where we denoted the single-trace $H$ potential of Ref. [8] by the initials of its three authors, and

$$\tilde{V}(H) = \sum_a V_{\mathrm{ct}}(E_a) + \frac{1}{2}\sum_{a,b} \log F(E_a, E_b), \quad (4.20)$$

written in the energy eigenbasis. Here, $V_{\mathrm{ct}}$ is the counterterm potential we are trying to determine, and $F(E_a, E_b)$ is the coefficient of the quadratic term in the $\mathcal{O}$ matrix. It is clear that integrating out has changed the naive disk density of states, which we can compensate by a judicial choice of $V_{\mathrm{ct}}(H)$. An important result of Ref. [11] is that this can be achieved with a single-trace counterterm.

Having outlined the general procedure, let us now expand on the two approaches to determine the $\mathcal{O}$ potential in some more detail.

### C. Double-scaling limit

While we did not specify this explicitly, in a typical application, the ETH ansatz (2.1) is invoked for random matrices $R_{ab}$ of finite dimension, and thus for a locally finite Hilbert space, so that $e^{S(E)} < \infty$. In fact, for the main application we have in mind, namely, 2D gravity, we are interested in the case where the number of eigenvalues, that is, the dimension of Hilbert space, is scaled to infinity. As has been studied extensively in past applications of matrix models to the theory of 2D gravity (see, e.g., the review [45]), in order to pass to the limit of smooth fluctuating surfaces, this must be accompanied by a rescaling of one or several coupling parameters, leaving appropriate ratios finite. In fact, in matching matrix-model correlation functions to the JT + matter (as well as in the semiclassical limit, i.e., thermal mean-field theory discussed above) expressions, we implicitly assume that such a double-scaling limit has been taken. This double-scaling procedure leaves a theory with a continuous spectrum of eigenvalues supported on a noncompact cut in the complex energy plane, which can be matched to that of JT gravity [Eq. (4.3), below] for $E \in [0, \infty)$. While the direct-matching procedure may therefore seem a bit *ad hoc* at first, we in fact establish these results more carefully by introducing two classes of finite regulated matrix models, the so-called “$q$-deformed” and “Selberg models.” Both of these involve introducing a regulator, which renders the Hilbert-space dimension finite for $q < 1$ and is chosen so that, in the limit $q \to 1$, we recover the gravity correlation functions, i.e., the matching in Eqs. (4.7)–(4.9). In all cases, we fix the (two-)matrix model using only disk data and then proceed to show that it continues to correctly capture topologically nontrivial correlators of JT + matter. We return to the regulated models and their double-scaling limits in Sec. IV E after discussing the gravitational correlators both at disk level and higher genus that our ETH-matrix model for JT + matter is designed to reproduce.

### D. Constraint-squared potential

In this section, we argue that by imposing a constraint on the matrices $H$ and $\mathcal{O}$, we can construct a matrix model that correctly computes the disk correlators of JT gravity minimally coupled to a scalar field. Our argument is not rigorous, but it is modeled after a rigorous result in 1D CFT, which states that if $\mathcal{O}$ is a primary with dimension $\Delta$ and if the spectrum of primary operators appearing in the $\mathcal{O}\mathcal{O}$ OPE is that of a bosonic generalized free field (GFF) with dimension $\Delta$, then all of the correlators of $\mathcal{O}$ are exactly those of the GFF [56]. JT gravity with matter on the disk admits a semiclassical limit in which the correlators become those of a GFF, and one might expect that even away from the semiclassical limit, there is some condition that can be placed on the operators $\mathcal{O}$ and $H$ that is only obeyed by gravitational correlators computed using the Feynman rules discussed towards the end of Sec. IV A. If such a condition exists, we may deduce it as follows. First, note that the $\mathcal{O}\mathcal{O}$ OPE in the GFF takes the form

$$\mathcal{O}(\tau)\mathcal{O}(0) = \frac{1}{\tau^{2\Delta}} + \sum_{n=0}^{\infty} \tau^{2n}[\mathcal{O}\mathcal{O}]_n + \text{descendants}. \qquad (4.21)$$

Aside from the identity, the primary operators above have dimensions $2\Delta + 2n$ for $n$ a non-negative integer. Equation (4.21) implies Eq. (4.18), which we repeat here:

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

In a 1D CFT, Eq. (4.18) also implies Eq. (4.21) because any contribution to Eq. (4.21) with a different power of $\tau$ would make an additional nontrivial contribution to the right side of Eq. (4.18). Hence, a quadratic operator equation for $\mathcal{O}$ is enough to guarantee that the spectrum of primaries appearing in the $\mathcal{O}\mathcal{O}$ OPE is that of a GFF, which in turn guarantees that the correlators of $\mathcal{O}$ are those of a GFF. We emphasize that the proof of this assumes the usual bootstrap axioms of conformal invariance and OPE associativity.

Next, we consider JT gravity away from the semiclassical limit and search for an operator equation for $\mathcal{O}$ and $H$ that is consistent with the Feynman rules described in Sec. IV A and is quadratic in $\mathcal{O}$. We could only find one such operator equation [57]:

$$e^{-S_0}\sum_b \begin{Bmatrix} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{Bmatrix}\left[\frac{\mathcal{O}_{ab}\mathcal{O}_{bc}}{e^{-S_0}\sqrt{\Gamma^{\Delta}_{ab}\Gamma^{\Delta}_{bc}}} - \frac{\delta(s_a-s_c)}{e^{S_0}\rho(s_a)}\right] = \left[\frac{\mathcal{O}_{ad}\mathcal{O}_{dc}}{e^{-S_0}\sqrt{\Gamma^{\Delta}_{ad}\Gamma^{\Delta}_{dc}}} - \frac{\delta(s_a-s_c)}{e^{S_0}\rho(s_a)}\right]. \quad (4.22)$$

If we insert the right side of Eq. (4.22) into any correlator, the Feynman rules of Sec. IV A dictate that we should omit diagrams where a bulk line connects the two adjacent $\mathcal{O}$ insertions. If we insert the left side of Eq. (4.22) into a correlator, the diagrams that contribute are the same, except with an additional crossing of the two bulk lines that end on the adjacent $\mathcal{O}$ insertions. (This means that the two lines could intersect twice. Such a double crossing can then be undone due to an orthogonality relation of the $6j$ symbols.)

We believe that Eq. (4.22) is the appropriate generalization of Eq. (4.18) to nearly conformal CFTs. Because of the constraining power of Eq. (4.18) in 1D CFTs, we expect that Eq. (4.22) is also highly constraining. Recall that the aforementioned rigorous 1D CFT result rests on three assumptions: the operator equation (4.18), associativity of the OPE, and conformal invariance. We explained above that Eq. (4.18) should generalize to Eq. (4.22) in a nearly conformal CFT, and thus we impose Eq. (4.22) as a constraint in a matrix model. If we define

$$M^d_{ac} \coloneqq \frac{1}{2}\sum_b\left(e^{-S_0}\begin{Bmatrix} \Delta & s_a & s_b \\ \Delta & s_c & s_d \end{Bmatrix} - \frac{\delta(s_b-s_d)}{e^{S_0}\rho(s_b)}\right) \times\left(\frac{\mathcal{O}_{ab}\mathcal{O}_{bc}}{e^{-S_0}\sqrt{\Gamma^{\Delta}_{ab}\Gamma^{\Delta}_{bc}}} - \frac{\delta(s_a-s_c)}{e^{S_0}\rho(s_a)}\right), \quad (4.23)$$

then the ensemble is defined by the following matrix integral:

$$\int dH d\mathcal{O}\exp\left(-\mathrm{Tr}V(H) - \frac{\Lambda}{2}\sum_{a,c,d}|M^d_{ac}|^2\right), \quad (4.24)$$

where $\Lambda$ is a large parameter that enforces $M^d_{ac}=0$, or Eq. (4.22), as a constraint. Associativity of the OPE is guaranteed by the fact that we are representing the operators using matrices, and matrix multiplication is associative. Of course, we do not expect this matrix model to respect conformal invariance, but it should reproduce a nearly conformally invariant theory due to the use of $6j$ symbols in constructing the matrix potential. In Ref. [11], we show that many Schwinger-Dyson equations of this matrix model are solved by the desired gravitational disk correlators, and we also discuss how Eq. (4.24) may be defined away from the strict double-scaling limit. Note that the matrix potential of the $\mathcal{O}$ integral, which is responsible for the correlations of the generalized ETH ansatz, is bounded from below. The potential $V(H)$ is chosen such that the eigenvalue distribution of $H$ to leading order in $e^{S_0}$ matches the disk density of states of JT gravity.

The question of higher-genus correlators in double-scaled matrix models dual to JT gravity with matter is subtle, and we explore it in two separately defined models, which we introduce below.

### E. Iterative procedure to determine the matrix model

We now discuss a matching calculation that allows one to determine the choices of the coupling functions $G^{(n)}$, introduced in Eq. (2.18), that result in the matrix model computing the correct disk correlators. A careful treatment requires us to regulate the gravitational disk amplitudes such that the energies (or, equivalently, the $s$ parameters) are integrated over a finite range. This is because loop 't Hooft diagrams in the matrix model [such as on the right-hand side of Eq. (2.21)] involve energy integrals, and a finite spectrum guarantees that these integrals converge. To achieve this, we note that the special functions appearing in Sec. IV A admit $q$ deformations,

$$\rho(s)\to\rho_q(s), \quad \Gamma^{\Delta}_{12}\to\Gamma^{\Delta}_{12,q}, \begin{Bmatrix} \Delta_1 & s_1 & s_2 \\ \Delta_2 & s_3 & s_4 \end{Bmatrix} \to \begin{Bmatrix} \Delta_1 & s_1 & s_2 \\ \Delta_2 & s_3 & s_4 \end{Bmatrix}_q, \quad (4.25)$$

in such a way that the usual JT + matter rules are recovered in the limit $q\to 1$. Precise definitions are provided in Ref. [11]. The important point is that the $s$ parameter is integrated only in the range $[0,(\pi/|\log q|)]$. These special functions obey a nontrivial Yang-Baxter relation, which is proved in Ref. [11], which ensures that the $q$-deformed Feynman rules are well defined. Note that these $q$-deformed Feynman rules do not obey the constraint (4.22), nor do they appear to obey any other constraint.

Next, it helps to organize the calculation by introducing a fictitious parameter $\epsilon$ such that higher-point gravitational amplitudes are weighted by higher powers of $\epsilon$ (roughly speaking). We determine the couplings order by order in $\epsilon$, and at the end of the calculation, we set $\epsilon \to 1$ (as well as $q \to 1$) so that the gravitational amplitudes return to their original values.

We consider two schemes for introducing the parameter $\epsilon$. One way, called the "Selberg regulator," is to weight each connected $2n$-point function [given by $g_{\mathcal{O}}^{(2n)}$, introduced in Eq. (2.13)] by a factor of $\epsilon^{n-1}$. The other way, called the "$q$-deformed regulator," is to weight each $6j$ symbol by $\epsilon$. The key point is that with either regulator, to a finite order in $\epsilon$, only finitely many $g_{\mathcal{O}}^{(n)}$ functions are nonzero. Hence, we can match these connected correlators using Eq. (2.18), where only finitely many $G^{(n)}$ are nonzero. By working to higher orders in $\epsilon$, we can naturally construct a series representation for $G^{(n)}$. The models constructed using the Selberg and $q$-deformed regulators are, respectively, referred to as the Selberg and $q$-deformed matrix models. The JT limit is defined to be the $\epsilon \to 1, q \to 1$ limit [58].

The $q$-deformed regulator is designed to match the Feynman rules of the double-scaled SYK model [7,59]. In the Majorana SYK model with $N$ Majorana fermions, interacting via a random coupling involving a subset of $p$ degrees of freedom at a time, one scales $N, p \to \infty$, such that

$$\lambda = \frac{2p^2}{N}, \qquad q = e^{-\lambda} \tag{4.26}$$

are held finite. In addition to the Hamiltonian, one may define another operator $\mathcal{O}$ that takes the same form of the Hamiltonian but with an independent set of random couplings. The number of fermions appearing in $\mathcal{O}$ sets the scaling dimension $\Delta$. Correlators of $H$ and $\mathcal{O}$ match those of the $q$-deformed regulator.

The Selberg regulator is designed to treat all of the gravitational Feynman diagram contributions to a single connected $n$-point function $g_{\mathcal{O}}^{(n)}$ on an equal footing. In Ref. [11], we present highly nontrivial results that indicate that the double trumpet of the Selberg model agrees with Eq. (4.10).

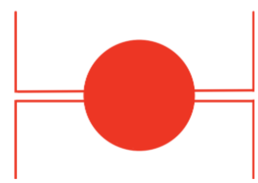
### F. Cylinder amplitudes in the matrix model

We partially outline our techniques for computing cylinder amplitudes in the $q$-deformed and Selberg matrix models. While these models produce the same disk amplitudes in the JT limit, they return different results for the matter determinant on the double trumpet. The model calculation we choose to present here concerns the double-trumpet two-point function, with one $\mathcal{O}$ inserted into each trace. This quantity is computed in the matrix model by summing over all 't Hooft diagrams with cylinder topology and two external double lines associated with the $\mathcal{O}$ matrix. We can classify these diagrams systematically, and each class of diagrams may be computed from our knowledge of the disk correlators.

Our first class of diagrams may be summed by setting $b = a$ in Eq. (4.8) and then integrating the remaining energy $E_a$ using the disk density of states:

$$\int_0^\infty dE_a e^{S_0} \rho_0(E_a) \overline{\mathcal{O}_{aa}\mathcal{O}_{aa}}, \tag{4.27}$$

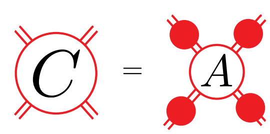
which is graphically represented as follows:



,

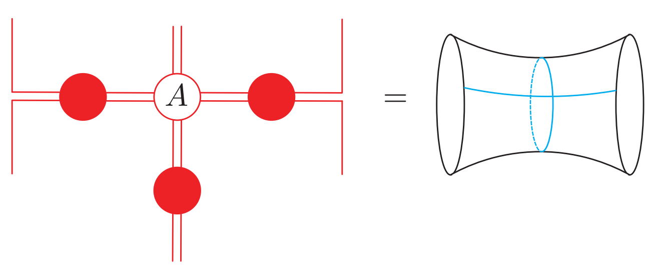
where the top and bottom ends of the diagram are identified so that all of the single lines shown are connected, reflecting the fact that there is a single integral in Eq. (4.27). It is already known [35] that Eq. (4.27) is equal to the double-trumpet two-point function without the matter determinant contribution, and we have shown that this is naturally associated with a sum over a class of 't Hooft diagrams.

The remaining 't Hooft diagrams represent nontrivial contributions of the matter determinant on the double trumpet. For instance, let us define a blob with an "$A$" to represent a sum over four-point, planar, amputated 't Hooft diagrams:




. (4.28)

We can compute another class of 't Hooft diagrams as follows:




, (4.29)

where on the left side the top and bottom of the diagram are again identified, and we have used Eqs. (4.8) and (4.9) to produce an explicit analytic expression that is graphically represented on the right side. The expression is obtainable from the gravitational Feynman rules (see the end of Sec. IV A). Explicitly, we are computing

$$e^{2S_0}\int_0^\infty dE_a dE_b \rho_0(E_a)\rho_0(E_b)e^{-3S(\overline{E})}g_{\mathcal{O}}^{(4)} \times (E_a, E_b, E_a, E_b)(\overline{\mathcal{O}_{ab}\mathcal{O}_{ba}})^{-1}. \quad (4.30)$$

The presence of the $(\overline{\mathcal{O}_{ab}\mathcal{O}_{ba}})^{-1}$ reflects the fact that a red blob on the topmost vertical double line in Eq. (4.29) is missing, to avoid overcounting ‘t Hooft diagrams.

The right-hand side of Eq. (4.29) exactly computes the second term in Eq. (4.15), and Eq. (4.27) computes the first term in Eq. (4.15). It is reasonable to expect that our strategy of sewing together amputated disk diagrams into cylinder diagrams will produce all of the terms in Eq. (4.15). However, as explained in Ref. [11], the analytic expressions for the remaining classes of ‘t Hooft diagrams depend on how the double-scaling limit is taken or, equivalently, on whether the Selberg or $q$-deformed regulator is used.

Using the Selberg regulator, we have explicitly reproduced the third term in Eq. (4.15), as well as the next term that corresponds to $n = 3$ in Eq. (4.11). We conjecture that all of the contributions from Eq. (4.11) are matched in the Selberg matrix model. In the $q$-deformed matrix model, the results point to a different matter determinant. Instead of Eq. (4.11), the $q$-deformed model predicts that the matter determinant is instead

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

which curiously has a Hagedorn temperature owing to the fact that the sum fails to converge for sufficiently small $b$.

Regardless of whether the matter determinant is Eq. (IV F) or Eq. (4.11), the double trumpet, Eq. (4.10), is ill defined due to the behavior of the integrand at small $b$. This implies that the gravitational theory is not UV complete, although the Selberg and $q$-deformed matrix models are operationally well defined. As we explain in Ref. [11], the ill-defined cylinder amplitude implies that the saddle point that would define the genus expansion is perturbatively unstable.

## V. DISCUSSION

In this paper, together with its companion, Ref. [11], we have described a way of combining the eigenstate thermalization hypothesis (including its generalization in Ref. [12]) with certain matrix integrals into one joint framework. This method allows one to interpret the ETH ansatz as arising from a joint probability distribution describing the statistics of energy levels as well as matrix elements in one unified framework. In this work, we focused on two-matrix integrals because we were interested in the case of one particular operator in addition to the energy eigenvalues, but it should be clear that adding further operators is possible and will lead to multimatrix integrals. The structure of these multimatrix integrals is related to free probability theory, which was invoked in the ETH context previously by Ref. [18]. It would be interesting to investigate this connection further.

We have further expanded on what is presumably the simplest instance of our ETH matrix model, namely, the case of thermal mean-field theory. It is striking that even in this simplest context the matrix model is strongly non-Gaussian, even though, as seen in the energy eigenbasis, all non-Gaussian contributions are entropically suppressed [60]. This is of course a necessary feature of ETH matrix models, in general, and it is compatible with the statistical physics approach of Ref. [12].

More generally, quantum chaotic systems are not expected to be described exactly by a matrix integral; rather, they approach such behavior at late times. Usually the timescale at which a matrix theory description accurately captures a quantum chaotic Hamiltonian is referred to as the Thouless or ergodic timescale. In fact, a more precise definition of this timescale also demands that the statistics of both matrix elements as well as eigenvalues approach those of Gaussian random matrices (see Refs. [15,16] for an in-depth discussion). In this work, we formulated matrix models that apply at (much) earlier timescales precisely by incorporating non-Gaussian statistics into the joint probability distribution of matrix elements and eigenvalues. We extend the applicability of the ETH matrix model beyond the Thouless time by adding more information about the physical system in the form of non-Gaussian terms in the joint potential, a perspective that is made very clear in the constraint matrix model approach we outlined above. In order to generalize this idea, one should introduce a matrixization timescale $t_M$, beyond which a system is well described by an ETH matrix model. By introducing further constraints along the lines of Eq. (4.22), one obtains ETH models whose $t_M$ is pushed further and further towards early times. This procedure is very much in the spirit of an effective matrix model (see Ref. [11]), whose region of validity can be extended by adding more and more UV information, paralleling the procedure one would follow in effective field theory. It is intriguing that certain systems, such as pure JT gravity or JT gravity with matter, as well as the double-scaled SYK model, exhibit a matrixization timescale $t_M$ that formally tends to zero. It would be interesting to ask under what conditions this can happen more generally and, furthermore, whether there may be lower bounds on $t_M$ in higher-dimensional (holographic) theories.

For a more extensive discussion of open questions and for future directions related to our ETH matrix model, the reader is referred to the companion paper [11].

## ACKNOWLEDGMENTS

We would like to thank Nick Agia, Alex Belin, Noam Chai, Jan de Boer, Anatoly Dymarsky, Lorenz Eberhardt, Laura Foini, Akash Goel, Tom Hartman, Clifford Johnson, Zohar Komargodski, Henry Lin, Juan Maldacena, Dalimil Mazáč, Vladimir Narovlansky, Pranjal Nayak, Silvia Pappalardi, 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 Grant No. DE-SC0007870, SNF through Project Grant No. 200020_182513, as well as the NCCR Grant No. 51NF40-141869, Mathematics of Physics (SwissMAP). The work of B. M. was supported by a grant from the Simons Foundation (No. 651444) and NSF Grant No. PHY-2014071.

[1] A. Kitaev, *A Simple Model of Quantum Holography 1*, Talk at KITP, April 7, 2015, http://online.kitp.ucsb.edu/online/entangled15/kitaev.

[2] A. Kitaev, *A Simple Model of Quantum Holography 2*, Talk at KITP, May 27, 2015, http://online.kitp.ucsb.edu/online/entangled15/kitaev2.

[3] A. Kitaev and S. J. Suh, *The Soft Mode in the Sachdev-Ye-Kitaev Model and Its Gravity Dual*, J. High Energy Phys. 05 (2018) 183.

[4] 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).

[5] K. Jensen, *Chaos in* $AdS_2$ *Holography*, Phys. Rev. Lett. **117**, 111601 (2016).

[6] J. Engelsöy, T. G. Mertens, and H. Verlinde, *An Investigation of* $AdS_2$ *Backreaction and Holography*, J. High Energy Phys. 07 (2016) 139.

[7] J. S. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S. H. Shenker, D. Stanford, A. Streicher, and M. Tezuka, *Black Holes and Random Matrices*, J. High Energy Phys. 05 (2017) 118.

[8] P. Saad, S. H. Shenker, and D. Stanford, *JT Gravity as a Matrix Integral*, arXiv:1903.11115.

[9] A. Altland and J. Sonner, *Late Time Physics of Holographic Quantum Chaos*, SciPost Phys. **11**, 034 (2021).

[10] A. Altland, B. Post, J. Sonner, J. van der Heijden, and E. Verlinde, *Quantum Chaos in 2D Gravity*, SciPost Phys. **15**, 064 (2023).

[11] D. L. Jafferis, D. K. Kolchmeyer, B. Mukhametzhanov, and J. Sonner, companion paper, *Jackiw-Teitelboim gravity with matter, generalized eigenstate thermalization hypothesis, and random matrices*, Phys. Rev. D **108**, 066015 (2023).

[12] L. Foini and J. Kurchan, *Eigenstate Thermalization Hypothesis and Out of Time Order Correlators*, Phys. Rev. E **99**, 042139 (2019).

[13] J. Sonner and M. Vielma, *Eigenstate Thermalization in the Sachdev-Ye-Kitaev Model*, J. High Energy Phys. 11 (2017) 149.

[14] C. Murthy and M. Srednicki, *Bounds on Chaos from the Eigenstate Thermalization Hypothesis*, Phys. Rev. Lett. **123**, 230606 (2019).

[15] 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).

[16] A. Dymarsky, *Bound on Eigenstate Thermalization from Transport*, Phys. Rev. Lett. **128**, 190601 (2022).

[17] L. Foini and J. Kurchan, *Eigenstate Thermalization and Rotational Invariance in Ergodic Quantum Systems*, Phys. Rev. Lett. **123**, 260601 (2019).

[18] S. Pappalardi, L. Foini, and J. Kurchan, *Eigenstate Thermalization Hypothesis and Free Probability*, Phys. Rev. Lett. **129**, 170603 (2022).

[19] J. M. Deutsch, *Quantum Statistical Mechanics in a Closed System*, Phys. Rev. A **43**, 2046 (1991).

[20] M. Srednicki, *Chaos and Quantum Thermalization*, Phys. Rev. E **50**, 888 (1994).

[21] T. Anous and J. Sonner, *Phases of Scrambling in Eigenstates*, SciPost Phys. **7**, 003 (2019).

[22] P. Nayak, J. Sonner, and M. Vielma, *Extended Eigenstate Thermalization and the Role of FZZT Branes in the Schwarzian Theory*, J. High Energy Phys. 03 (2020) 168.

[23] For simplicity, let us focus on the case where the operator has a vanishing one-point function so that $\overline{\langle E_a|\mathcal{O}|E_a\rangle} = 0$. The result can obviously be generalized to incorporate a nontrivial one-point function as well.

[24] Here, nonplanar is used in the sense of matrix-model perturbation theory: We draw all index contractions using 't Hooft double-line notation, resulting in a ribbon graph whose topology is called "planar" whenever it can be drawn on a sphere, and nonplanar otherwise.

[25] In JT gravity minimally coupled to a free scalar field, $n$ will only take on even values, owing to the $\mathcal{O} \to -\mathcal{O}$ symmetry. To simplify later formulas, we assume that $\mathcal{O}$ has this symmetry in all further expressions.

[26] O. Bohigas, M.-J. Giannoni, and C. Schmit, *Characterization of Chaotic Quantum Spectra and Universality of Level Fluctuation Laws*, Phys. Rev. Lett. **52**, 1 (1984).

[27] M. L. Mehta, *Random Matrices* (Elsevier, New York, 2004).

[28] We show formulas for the three so-called classical ensembles for $\beta = 1$ (GOE), $\beta = 2$ (GUE), and $\beta = 4$ (GSE).

[29] D. Stanford and E. Witten, *JT Gravity and the Ensembles of Random Matrix Theory*, Adv. Theor. Math. Phys. **24**, 1475 (2020).

[30] S. Sachdev, *Bekenstein-Hawking Entropy and Strange Metals*, Phys. Rev. X **5**, 041025 (2015).

[31] J. Maldacena and D. Stanford, *Remarks on the Sachdev-Ye-Kitaev Model*, Phys. Rev. D **94**, 106002 (2016).

[32] M. Marino, *Les Houches Lectures on Matrix Models and Topological Strings*, arXiv:hep-th/0410165.

[33] Note that a matrix model will result in a continuous density of states. We may view this as a coarse-grained density of state of a general quantum chaotic system, or alternatively, we can take this density as part of the definition of the model, as will be the case in the application to minimally coupled JT gravity.
[34] G. 't Hooft, *A Planar Diagram Theory for Strong Interactions*, Nucl. Phys. B **72**, 461 (1974).
[35] P. Saad, *Late Time Correlation Functions, Baby Universes, and ETH in JT Gravity*, arXiv:1910.10311.
[36] J. Pollack, M. Rozali, J. Sully, and D. Wakeham, *Eigenstate Thermalization and Disorder Averaging in Gravity*, Phys. Rev. Lett. **125**, 021601 (2020).
[37] J. Cotler and K. Jensen, $AdS_3$ *Gravity and Random CFT*, J. High Energy Phys. 04 (2021) 033.
[38] J. Cotler and K. Jensen, *Gravitational Constrained Instantons*, Phys. Rev. D **104**, 081501 (2021).
[39] A. Altland, D. Bagrets, P. Nayak, J. Sonner, and M. Vielma, *From Operator Statistics to Wormholes*, Phys. Rev. Res. **3**, 033259 (2021).
[40] S. Pappalardi, F. Fritzsch, and T. Prosen, *General Eigenstate Thermalization via Free Cumulants in Quantum Lattice Systems*, arXiv:2303.00713.
[41] N. Hunter-Jones, J. Liu, and Y. Zhou, *On Thermalization in the SYK and Supersymmetric SYK Models*, J. High Energy Phys. 02 (2018) 142.
[42] The reader not familiar with the corresponding results in JT may wish to return to this after examining Eqs. (4.4) and (4.6) in Sec. IV.
[43] In our conventions, we set $(\bar{\phi}_r/4G_N) = 1$, where $\bar{\phi}_r$ is the boundary value of the dilaton with dimensions of length. To restore $G_N$, one simply rescales all energies $E_j \to E_j(\bar{\phi}_r/4G_N)$. Therefore, the semiclassical limit $G_N \to 0$ corresponds to the high-energy limit.
[44] P. Di Francesco, P. H. Ginsparg, and J. Zinn-Justin, *2-D Gravity and Random Matrices*, Phys. Rep. **254**, 1 (1995).
[45] P. H. Ginsparg and G. W. Moore, *Lectures on 2-D Gravity and 2-D String Theory, in Theoretical Advanced Study Institute (TASI 92): From Black Holes and Strings to Particles*, arXiv:hep-th/9304011.
[46] T. G. Mertens, G. J. Turiaci, and H. L. Verlinde, *Solving the Schwarzian via the Conformal Bootstrap*, J. High Energy Phys. 08 (2017) 136.
[47] More details can be found in the original references studying disk correlators of the Schwarzian, Refs. [46,48–51].
[48] D. Bagrets, A. Altland, and A. Kamenev, *Sachdev–Ye–Kitaev Model as Liouville Quantum Mechanics*, Nucl. Phys. B **911**, 191 (2016).
[49] Z. Yang, *The Quantum Gravity Dynamics of Near Extremal Black Holes*, J. High Energy Phys. 05 (2019) 205.
[50] A. Kitaev and S. J. Suh, *Statistical Mechanics of a Two-Dimensional Black Hole*, J. High Energy Phys. 05 (2019) 198.
[51] S. J. Suh, *Dynamics of Black Holes in Jackiw-Teitelboim Gravity*, J. High Energy Phys. 03 (2020) 093.
[52] D. Stanford and E. Witten, *Fermionic Localization of the Schwarzian Theory*, J. High Energy Phys. 10 (2017) 008.
[53] L. V. Iliesiu, M. Mezei, and G. Sárosi, *The Volume of the Black Hole Interior at Late Times*, J. High Energy Phys. 07 (2022) 073.
[54] A. Blommaert, *Dissecting the Ensemble in JT Gravity*, J. High Energy Phys. 09 (2022) 075.
[55] We refer here to directly deducing the double-scaled matrix model. In the longer companion paper, we also present two classes of regulated matrix models, which recover the results here in the double-scaled limit.
[56] See Ref. [11] for the proof. We thank Dalimil Mazáč for his invaluable assistance.
[57] If our conjecture in Sec. IV A for the Feynman rules at higher genus is correct, then higher-genus correlators also obey this constraint.
[58] In the Selberg model, the $q \to 1$ limit must precede the $\epsilon \to 1$ limit in order for the double-trumpet amplitudes to come out as desired.
[59] 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.
[60] In Ref. [41], the Gaussian ETH was numerically verified in the SYK model. It would be interesting if one could measure the entropically suppressed non-Gaussianities in a more sensitive study.